SIMPLICITY IN RELATIONAL STRUCTURES AND ITS APPLICATION TO PERMUTATION CLASSES

Robert Brignall

**A Thesis Submitted for the Degree of PhD
at the
University of St. Andrews**

2007

Simplicity in Relational Structures and its Application to Permutation Classes

Robert Brignall

October 25, 2007

DECLARATIONS

I Robert Brignall hereby certify that this thesis, which is approximately 40,000 words, has been written by me, that it is a record of work carried out by me, and that it has not been submitted in any previous application for a higher degree.

Signed: Date:

I was admitted as a research student in September 2004 and as a candidate for the degree of PhD in September 2004; the higher investigation of which this is a record was carried out in the University of St Andrews between September 2004 and July 2007.

Signed: Date:

I hereby certify that the candidate has fulfilled the conditions of the Resolution and Regulations appropriate for the degree of Ph.D. in the University of St Andrews and that the candidate is qualified to submit this thesis in application to that degree.

Signature of Supervisor: Date:

In submitting this thesis to the University of St Andrews I understand that I am giving permission for it to be made available in accordance with the regulations of the University Library for the time being in force, subject to any copyright vested in the work not being affected thereby. I also understand that the title and abstract will be published, and that a copy of the work may be supplied to any bona fide library or research worker, that my thesis will be electronically available for personal or research use, and that the library has the right to migrate my thesis into new electronic forms as required to ensure continued access to the thesis. I have obtained any third-party copyright permissions that may be required in order to allow such access and migration.

Signed: Date:

Acknowledgements

I would like to thank my supervisor, Nik Ruškuc, for introducing me to the area of permutation patterns and for his support throughout this process. Additionally, particular thanks go to Vince Vatter for the many hours of fruitful discussions we have had, and without whom this thesis would be very different and almost certainly less rewarding.

Thanks must also go to my various other collaborators – Sophie Huczynska, Mike Atkinson and Michael Albert – for their time, and to Steve Waton and Steve Linton for their contributions to my knowledge of the area. Also, the anonymous referees of the papers on which this thesis is based have provided enriching comments.

To those with whom I have shared an office – Nelson Silva, Steve Waton, Andreas Distler and Victor Maltcev – thanks for making my time here enjoyable. I am also grateful to the many staff and students of the School of Mathematics and Statistics for their support and friendship, and to countless other friends and acquaintances both in St Andrews and elsewhere.

This thesis has been greatly enhanced by the music of J.S. Bach.

Final thanks, however, must go to my parents, my grandparents, brother and sister, and to Gwen, for her unending support, company and patience.

I acknowledge generous funding from EPSRC.

ABSTRACT

The simple relational structures form the units, or atoms, upon which all other relational structures are constructed by means of the substitution decomposition. This decomposition appears to have first been introduced in 1953 in a talk by Fraïssé, though it did not appear in an article until a paper by Gallai in 1967. It has subsequently been frequently rediscovered from a wide variety of perspectives, ranging from game theory to combinatorial optimization.

Of all the relational structures — a set which also includes graphs, tournaments and posets — permutations are receiving ever increasing amounts of attention. A *simple permutation* is one that maps every nontrivial contiguous set of indices to a set of indices that is never contiguous. Simple permutations and intervals of permutations are important in biomathematics, while permutation classes — downsets under the pattern containment order — arise naturally in settings ranging from sorting to algebraic geometry.

We begin by studying simple permutations themselves, though always aim to establish this theory within the broader context of relational structures. We first develop the technology of "pin sequences", and prove that every sufficiently long simple permutation must contain either a long horizontal or parallel alternation, or a long pin sequence. This gives rise to a simpler unavoidable substructures result, namely that every sufficiently long simple permutation contains a long alternation or oscillation.

Erdős, Fried, Hajnal and Milner showed in 1972 that every tournament could be extended to a simple tournament by adding at most two additional points. We prove analogous results for permutations, graphs, and posets, noting that in these three cases we may need to extend a structure by adding $\lceil (n+1)/2 \rceil$ points in the case of permutations and posets, and $\log_2(n+1)$ points in the graph case.

The importance of simple permutations in permutation classes has been well established in recent years. We extend this knowledge in a variety of ways, first by showing that, in a permutation class containing only finitely many simple permutations, every subset defined by properties belonging to a finite "query-complete set" is enumerated by an algebraic generating function. Such properties include being an even or alternating permutation, or avoiding generalised (blocked or barred) permutations. We further indicate that membership of a permutation class containing only finitely many simple permutations can be computed in linear time.

Using the decomposition of simple permutations, we establish, by representing pin sequences as a language over an eight-letter alphabet, that it is decidable if a permutation class given by a finite basis contains only finitely many simple permutations. We also discuss possible approaches to the same question for other relational structures, in particular the difficulties that arise for graphs. The pin sequence technology provides a further result relating to the wreath product of two permutation classes, namely that $\mathcal{C} \wr \mathcal{D}$ is finitely based whenever $\mathcal{D}$ does not admit arbitrarily long pin sequences. As a partial converse, we also exhibit a number of explicit examples of wreath products that are not finitely based.

Contents

Introduction

This thesis consists of two parts: In Part I we study the structure of simple permutations in the context of relational structures, while in Part II we apply this structural knowledge of simplicity to permutation classes. This division reflects the fact that the study of permutations — and particularly simple permutations — lies in an area of research extending beyond the subject of permutation classes. However, that these two topics are covered under the single title of this thesis reflects the importance in studying simple permutations for the further understanding of permutation classes. Many of the major permutation-based results in this thesis may be found published or available as preprints [28, 29, 30, 31].

In Chapter 1 we begin by introducing permutations and the containment partial order. We then take a more broad view by defining the general construction of relational structures, and demonstrate how permutations, graphs, tournaments and posets may all be described in this language. We then commence our discussion in Sections 1.4 and 1.5 of intervals, simplicity and the substitution decomposition in the context of relational structures, at each stage also translating back to the permutation case.

In Chapter 2 we introduce the new technology of pin sequences and show how sufficiently long simple permutations must contain either a long proper pin sequence, or a long wedge or parallel alternation. We also introduce the language of pins, a necessary prerequisite for the decidability result of Chapter 7. We close the chapter with a speculative discussion on possible analogues of this decomposition theory for graphs.

Motivated by a result of Erdős, Fried, Hajnal and Milner in 1972 for tournaments, Chapter 3 considers the problem of embedding a given relational structure inside a larger simple structure. We demonstrate that a general approach may be used relying on the substitution decomposition, but that the outcome for each type of relational structure may be somewhat unique. To demonstrate this, we look at the simple extensions of permutations, graphs, tournaments and posets.

Much emphasis has been placed in recent years in developing optimal algorithms for computing intervals and the substitution decomposition. In Chapter 4 we review a recent paper by Bergeron, Chauve, Montgolfier and Raffinot who give a linear-time algorithm to compute the intervals in a given permutation. It follows directly from this work that the permutation substitution decomposition may be computed in linear time. We also review some algorithmic results in the case of graphs.

Permutation classes have been intensively studied in recent years, and in Chapter 5 we review some of the results in this area, manifested primarily in constructions between permutation classes, their enumeration and special properties including partial well order and atomicity. Permutation classes containing only finitely many simple permutations have received particular attention, and we cover the most important results concerning these.

One particular property of permutation classes containing only finitely many simple permutations is that they are enumerated by algebraic generating functions. By means of "finite query-complete sets of properties", we show in Chapter 6 that many different subsets of such permutation classes are also enumerated by algebraic generating functions. We close the chapter with some further enumerative results coming from the decomposition of simple permutations in Chapter 2, and note how, using the linear-time substitution decomposition algorithm of Chapter 4, we may establish in linear time whether a given permutation lies in a specified class known to contain only finitely many simple permutations.

Chapter 7 answers affirmatively the natural question arising from the studies of Chapters 5 and 6: is it decidable if a permutation class given by a finite basis contains infinitely many simple permutations? This is done using the decomposition results of Chapter 2, in particular showing that the language of pins lying within a specified class forms a regular language, and hence its infinitude is decidable.

Finally, in Chapter 8, using the technology of pin sequences in a slightly different context we derive a general sufficient result concerning the basis elements of the wreath product between two finitely based permutation classes, relying on whether one of the permutation classes contains arbitrarily long pin sequences or not. In the case where a given class contains arbitrarily long pin sequences, we demonstrate in a number of cases wreath products which are not finitely based. This suggests that the finite basis result is, to some extent, necessary, though we also present some evidence to the contrary.

Part I

Simplicity

Chapter 1

Preliminaries

Expressing an object in terms of smaller, indecomposable objects, is a goal aimed at in a wide variety of subject areas. The first example one finds in mathematics is the Fundamental Theorem of Arithmetic, which demonstrates how any positive integer greater than 1 may be written uniquely (up to ordering) as a product of prime factors. It is a property that is not true for elements of an arbitrary collection, however; take for example the elements of a ring, which in general are not uniquely factorisable (unless the ring is specifically shown to be a Unique Factorisation Domain). When a given collection of objects can be uniquely factorised, emphasis is often placed on the study of the prime or indecomposable elements, as it is these which form the "building blocks" of the collection.

One such family of objects is the family of relational structures – objects governed by a given set of relations – whose most notable members include graphs, tournaments, permutations and posets. Their "factorisation" is relatively straightforward, and will be referred to as the "substitution decomposition", though is known also as the modular decomposition, disjunctive decomposition and X-join. The elemental building blocks of this decomposition are the "simple" structures. This term is used primarily in the context of permutations, while in other contexts these structures are called prime or indecomposable (note in particular that "simple" usually has a different meaning in the context of graphs).

The notion of substitution decomposition dates back at least to a 1953 talk of Fraïssé, but only the abstract of this talk [55] survives. The first article using the substitution decomposition seems to be Gallai [58] (for an English translation, see [59]), who applied them particularly to the study of transitive orientations of graphs. Some work on the substitution decomposition in the general context can be found in Möhring [92]. It has proved to be a useful technique in a wide variety of settings, ranging from game theory to combinatorial optimisation (see Möhring [94] or Möhring and Radermacher [95] for extensive references).

Our relational structure of choice is the permutation. It has sufficient complexity to be worthy of extended study, but also is easily represented graphically. In this setting, much of the motivation for studying the substitution decomposition is for the purposes of enumeration, particularly of permutation classes, and Part II is primarily dedicated to demonstrating the enumerative consequences of this study.

Adapting the permutation-specific theory we will develop to other relational structures is not necessarily obvious; much of the theory depends, as we have indicated, on the graphical representation of permutations, and so, for example, finding a graph-theoretic analogue will not follow immediately. Thus throughout Part I we will discuss the success (or otherwise) of existing attempts in this avenue.

1.1 Permutations, Containment and Order Isomorphism

We begin by introducing the terms we need to study permutations; the definition of a general relational structure will follow after this is established. For $n \in \mathbb{N}$ denote by $[n]$ the set $\{1, 2, \ldots, n\}$, and for $i \leq j$ let $[i, j]$ correspond to the set $\{i, i+1, \ldots, j\}$. We may sometimes also refer to open or half-open segments, for example $(i, j]$ denotes the set $\{i + 1, i + 2, \ldots, j\}$.

In our context, a *permutation* π of length n is an ordering of the elements of $[n]$. For example, $\pi = 918572364$ is a permutation of length 9. Two particular families of permutations to which we will refer relatively often are the *increasing permutations* denoted by $\iota_n = 12 \cdots n$, and the *decreasing permutations* $\delta_n = n(n-1) \cdots 1$.

For $i \in [n]$ denote by $\pi(i)$ the image of the number i under π, and, by extension, $\pi([i, j])$ corresponds to the image of the segment $[i, j]$. The pair $(i, \pi(i))$ represents a *point* of π, and in this pair i is the *position* and $\pi(i)$ the *value* of the point. Viewing π as a set of points immediately indicates the graphical interpretation which will prove invaluable in our forthcoming study. We will, however, postpone this viewpoint momentarily while we introduce some further definitions.

Two finite sequences of the same length, $\alpha = a_1a_2\cdots a_n$ and $\beta = b_1b_2\cdots b_n$, are said to be *order isomorphic* if, for all i, j, we have $a_i < a_j$ if and only if $b_i < b_j$. As such, each sequence of distinct real numbers is order isomorphic to a unique permutation. For a sequence α and set of permutations $\mathcal{C}$, with a slight abuse of notation we will sometimes write statements like "$\alpha \in \mathcal{C}$", meaning "the permutation order isomorphic to α lies in $\mathcal{C}$.". Similarly, any given subsequence (or *pattern*) of a permutation π is order isomorphic to a smaller permutation, σ say, and such a subsequence is called a *copy* of σ in π. We may also say that π *contains* σ (or, in some texts, π *involves* σ) and write $\sigma \leq \pi$. If, on the other hand, π does not contain a copy of some given σ, then π is said to *avoid* σ. For example, $\pi = 918572346$ contains $\sigma = 51342$ because of the subsequence 91572 ($= \pi(1)\pi(2)\pi(4)\pi(5)\pi(6)$), but avoids $\tau = 3142$.

The pattern containment order forms a partial order on the set of all permutations, and in Part II we will be looking at sets of permutations closed under taking subpermutations. A book introducing the study of these permutation patterns has been written by Bóna [22].

1.2 Graphical Representation and Symmetries

As mentioned above, we may think of a permutation π as a set of points $(i, \pi(i))$, and immediately we can form a graphical representation. We can go further, however, and give 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

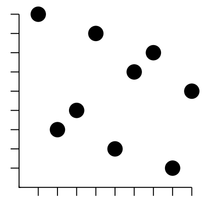

Figure 1.1: The plot of the permutation $\pi = 934826715$.

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 will simply say that a point set is order isomorphic to a permutation. See Figure 1.1 for an example. Steve Waton's PhD thesis [118] extends this graphical interpretation of containment to consider the sets of permutations that can be drawn by taking points lying on a given geometrical shape.

This geometric viewpoint indicates several of the symmetries of pattern containment. The maps $(x, y) \mapsto (-x, y)$, $(x, y) \mapsto (x, -y)$ and $(x, y) \mapsto (y, x)$, when applied to generic point sets, correspond to "reversing", "complementing" and "inverting" permutations respectively. Formally, the *reverse* of a permutation π of length n is the permutation obtained by reading the sequence of symbols of π in reverse order, i.e. from right to left. For each $i \in [n]$, the ith component of the *complement* of π is assigned value $n + 1 - \pi(i)$, while the *inverse* of π is denoted π^{-1} and is defined by $\pi^{-1}(j) = i$, where $j = \pi(i)$. For example, the reverse of $\pi = 934826715$ is 517628439, its complement is 176284395 and $\pi^{-1} = 852396741$.

Of these three symmetries, one of the reverse or complement mappings, together with the inverse mapping generate the dihedral group with eight elements. It is clear to check, either graphically or otherwise, that each of these symmetries preserves pattern containment (for example, $\sigma \leq \pi$ if and only if $\sigma^{-1} \leq \pi^{-1}$). That these are the only symmetries is less immediate but follows directly from the work on permutation reconstruction by Smith [111].

1.3 Relational Structures

The most general objects we will consider are the relational structures, which we now introduce as a precursor to handling simplicity and the substitution decomposition. For any set A, a *k-ary relation* R is a subset of A^k. An ordered sequence of relations over A is then called a *relational structure*.

More specifically, define a *relational language*, $\mathcal{L}$, to be a set of *relational symbols* R together with positive integers n_R denoting the arity of the symbols R. A relational structure $\mathcal{A}$ whose relational symbols are those of $\mathcal{L}$ is then defined by its *ground set* $\mathrm{dom}(\mathcal{A})$ and a set of subsets $R^{\mathcal{A}} \subseteq \mathrm{dom}(\mathcal{A})^{n_R}$ for each $R \in \mathcal{L}$. Such a structure will also be called an *$\mathcal{L}$-structure*. If, for example, $(a_1, \ldots, a_{n_R}) \in R^{\mathcal{A}}$ then we write $R^{\mathcal{A}}(a_1, \ldots, a_{n_R})$, and $R^{\mathcal{A}}$ is an n_R-ary relation.

We will be working primarily with relational structures whose ground sets are finite, though many of these principles may be applied to infinite relational structures. In particular, the substitution decomposition is readily extended to include infinite structures, as shown in [95].

We now briefly review how some well-known objects may be viewed as relational structures.

Permutations. A permutation π on n points may be viewed as the relational structure $\mathcal{A}_\pi$ with ground set $\mathrm{dom}(\mathcal{A}_\pi) = [n]$, on a language containing two binary linear relations, $\mathcal{L} = \{<, \prec, n_< = 2, n_\prec = 2\}$. The first relation, $<^{\mathcal{A}_\pi}$, is the normal ordering on $[n]$, while $i \prec^{\mathcal{A}_\pi} j$ if and only if $\pi(i) < \pi(j)$. For example, $\pi = 934826715$ corresponds to the relational structure $\mathcal{A}_\pi$ on $[9]$ with

$$1 < 2 < 3 < 4 < 5 < 6 < 7 < 8 < 9$$

and

$$8 \prec 5 \prec 2 \prec 3 \prec 9 \prec 6 \prec 7 \prec 4 \prec 1.$$

Graphs. A graph G is a relational structure $\mathcal{A}_G$ on the language $\mathcal{L} = \{E, n_E = 2\}$, where E is a binary symmetric relation, $\mathrm{dom}(\mathcal{A}_G) = V(G)$, and $E^{\mathcal{A}_G}(x, y)$ if and only if $x \sim y$ in G. The analogue to containment in graphs is the notion of the *induced subgraph*:[1] an induced subgraph of G is a graph formed on any subset of vertices from G, with $x \sim y$ in the subgraph if and only if $x \sim y$ in G.

Tournaments. A *tournament* is a complete oriented graph. A tournament T therefore corresponds to the relational structure $\mathcal{A}_T$ on the language $\mathcal{L} = \{\rightarrow, n_{\rightarrow} = 2\}$, but where $\rightarrow$ is now a trichotomous binary relation, i.e. for each $x, y \in \mathrm{dom}(\mathcal{A}_T) = V(T)$, precisely one of $x = y$, $x \rightarrow^{\mathcal{A}_T} y$ or $y \rightarrow^{\mathcal{A}_T} x$ is true. The name "tournament" derives from its use to denote a competition where every pair of players x, y must meet each other in a match, the outcome being either that x wins, denoted $y \rightarrow x$, or that x loses, denoted $x \rightarrow y$. The containment order on tournaments is not surprisingly the same as graphs; an *induced subtournament* of a tournament T is a tournament formed on any subset of vertices of T with $x \rightarrow y$ in the subtournament if and only if $x \rightarrow y$ in T.

Posets. By definition, a poset is a relational structure on the language containing a single binary relation, $<$, which is reflexive, antisymmetric and transitive. The *comparability graph* $G(P, <)$ of a poset $(P, <)$ is a graph with vertex set P, and edge $p \sim q$ if and only if either $p < q$ or $q < p$. Conversely, if G is a comparability graph for some poset $(P, <)$, then the order $<$ is called a *transitive orientation* of (the edges of) G. This connection between posets and graphs arises in a number of combinatorial problems – see Möhring [93] for a survey.

1.4 Intervals and Simplicity

Before we can discuss the substitution decomposition, we must first define how we can find "factors" of a given relational structure, and hence define the elemental relational structures – those structures with no nontrivial factors.

[1]This is sometimes called the "vertex induced subgraph", to distinguish from edge induced subgraphs.

Following Földes [54], we say that a set $X \subseteq \mathrm{dom}(\mathcal{A})$ is an *interval* if for every $R \in \mathcal{L}$ and n_R-tuple $(x_1, \ldots, x_{n_R}) \in \mathrm{dom}(\mathcal{A})^{n_R} \setminus X^{n_R}$, with at least one $x_i \in X$, then

$$R^{\mathcal{A}}(x_1, \ldots, x_{n_R}) \iff R^{\mathcal{A}}(x_1, \ldots, x_{i-1}, y, x_{i+1}, \ldots, x_{n_R}) \text{ for all } y \in X.$$

Informally, an interval corresponds to a subset X of the ground set $\mathrm{dom}(\mathcal{A})$ for which every pair of elements of X have exactly the same relations with the elements of $\mathrm{dom}(\mathcal{A}) \setminus X$. Accordingly, every singleton set $\{x\} \subseteq \mathrm{dom}(\mathcal{A})$ is an interval, as is all of $\mathrm{dom}(\mathcal{A})$. Every other interval is said to be a *proper* interval, and a structure is *simple* if it has no proper intervals.

Simplicity has, to some extent, been studied for relational structures in general, for example, by Földes [54] and Schmerl and Trotter [107]. Much greater attention has, however, been diverted to particular structures, the most pertinent of which we will now review.

Permutations. In the permutation case, an *interval* of π corresponds to a set of contiguous indices $I = [a, b]$ such that the set of values $\pi(I) = \{\pi(i) : i \in I\}$ is also contiguous. Intervals are clearly identified in the plot of a permutation as a set of points enclosed in an axis-parallel rectangle, with no points lying in the regions above, below, to the left or to the right (see Figure 1.2 for an example). Intervals of permutations are interesting in their own right and have applications to biomathematics, particularly to genetic algorithms for sequencing problems, and modelling the genomes of prokaryotes as permutations allows the matching of gene sequences.[2] See Corteel, Louchard, and Pemantle [37] for extensive references.

It then follows that a *simple* permutation is one whose only intervals are of length 0, 1 and n. Figure 1.3 shows three simple permutations of length 12. Note that the eight order-isomorphism preserving symmetries also preserve intervals, and hence simplicity. The number of simple permutations of length $n = 1, 2, \ldots$ is $1, 2, 0, 2, 6, 46, 338, 2926, 28146, \ldots$ (sequence A111111 of [110]), the first few being 1, 12, 21, 2413 and 3142. We will look at the asymptotics of this sequence in Subsection 1.4.2.

[2] In these contexts, the term "common interval" is used, indicating a segment upon which two or more permutations agree; we will encounter this definition again in Chapter 4.

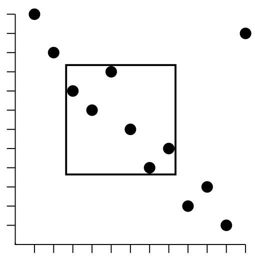

Figure 1.2: An interval in a permutation.

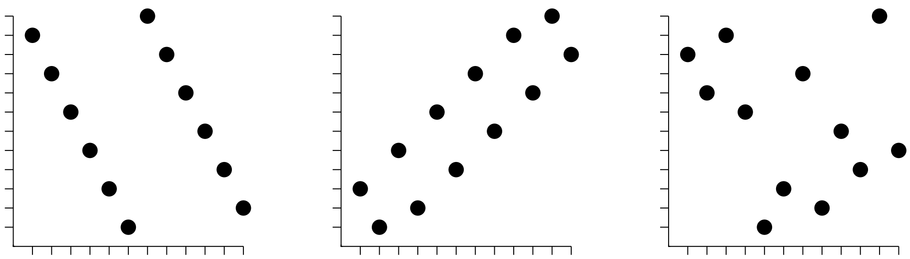

Figure 1.3: The plots of three simple permutations of length 12.

Graphs. An interval in a graph[3] is a set of vertices $X \subseteq V(G)$ such that $N(v) \setminus X = N(w) \setminus X$ for all $v, w \in X$, where $N(v)$ denotes the neighbourhood of v in G. A graph on n vertices therefore has several trivial intervals ($\emptyset$, $V(G)$, and the singletons); a graph with no nontrivial intervals is then often called *prime* or *indecomposable* (the word simple meaning something completely different in this context). These graphs have been the subject of considerable study, see for example Ehrenfeucht, Harju, and Rozenberg [47], Ille [71], and Sabidussi [105]. A survey of indecomposability and the substitution decomposition in graphs can be found in Brandstädt, Le, and Spinrad [27].

Tournaments. An interval in a tournament 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 [40] survey the algebraic and combinatorial results concerning simple tourna-

[3]These are also called autonomous sets, blocks, bound sets, clans, closed sets, clumps, committees, congruences, convex sets, externally related sets, factors, modules, parties solidaires, partive sets, stable sets, and strong intervals.

ments.

Posets. An interval of a poset $(P, <)$ corresponds to a set $A \subseteq P$ which for every $p \in P \setminus A$ satisfies one of $p < A$, $p > A$ or p is incomparable to every point of A. Intervals in a poset correspond to "convex" intervals in its related comparability graph. A subset of $B \subseteq P$ is called $(P, <)$-*convex* if the set $\{r \in P :$ there exist $p, q \in B$ such that $p < r < q\}$ is a subset of B. The following lemma is then easily deduced:

Lemma 1.1 (Buer and Möhring [32]). *Given a poset $(P, <)$, the set of intervals of $(P, <)$ is equal to the set of $(P, <)$-convex intervals of $G(P, <)$.*

1.4.1 Interacting Intervals

In the general context of relational structures, intervals interact with each other in a pleasing way. Two intervals are said to *overlap* if neither interval is contained in the other and their intersection is nontrivial.

Proposition 1.2. *For any two overlapping intervals I and J of the $\mathcal{L}$-structure $\mathcal{A}$,*

(a) $I \cap J$ is an interval of $\mathcal{A}$ (Földes [54, Proposition 1]),

(b) $I \cup J$ is an interval of $\mathcal{A}$ (Földes [54, Proposition 2]), and

(c) $I \setminus J$ is an interval of $\mathcal{A}$.

Proof. We will prove only Case (c) in the case where $\mathcal{L}$ consists solely of a k-ary relation R ($k \geq 2$); the result for a general language $\mathcal{L}$ follows immediately. If I and J are overlapping intervals of $\mathcal{A}$, we must show that if $R^{\mathcal{A}}(x_1, x_2, \ldots, x_k)$ with $x_1 \in I \setminus J$ and not all of $x_2, \ldots, x_k$ lie in $I \setminus J$, then $R^{\mathcal{A}}(y, x_2, \ldots, x_k)$ for any $y \in I \setminus J$.

Since I is an interval and $x_1, y \in I$, we are finished if, for some $i \in [2, k]$, x_i lies in $\mathrm{dom}(\mathcal{A}) \setminus I$, so suppose that every $x_i \in I \cap J$. Since I and J overlap, there exists at least one $z \in J \setminus I$, and so $R^{\mathcal{A}}(x_1, x_2, x_3, \ldots, x_k)$ implies $R^{\mathcal{A}}(x_1, z, x_3, \ldots, x_k)$ because J is an interval. We can now obtain $R^{\mathcal{A}}(y, z, x_3, \ldots, x_k)$, and thus $R^{\mathcal{A}}(y, x_2, x_3, \ldots, x_k)$ as required. $\square$

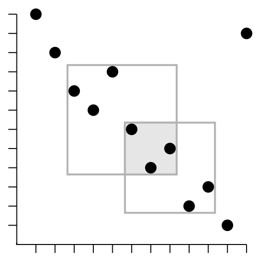

Figure 1.4: Two intervals and their intersection.

For two sets X and Y, let $X \triangle Y$ denote the *symmetric difference* of X and Y, namely $(X \cup Y) \setminus (X \cap Y)$. Providing a relational structure $\mathcal{A}$ is defined by a language consisting only of binary symmetric relations and relations with arity at least 3, then the symmetric difference of two intersecting intervals is also an interval.

Proposition 1.3 (Möhring and Radermacher [95, Theorem 4.1.1])**.** *Let $\mathcal{A}$ be an $\mathcal{L}$-structure for which $n_R \geq 2$ for every $R \in \mathcal{L}$. Then if I and J are overlapping intervals, $I \triangle J$ is also an interval if every binary relation $R \in \mathcal{L}$ is symmetric.*

In the permutation case, Proposition 1.3 clearly does not apply. However, Proposition 1.2 is easily seen by considering the graphical representation, as in Figure 1.4.

1.4.2 Asymptotics

The asymptotic enumeration of simple structures has been studied variously for permutations, tournaments, graphs, and indeed in a more general setting. We will presently review the problem for permutations and graphs, with a view to showing that although both these structures fall within the category of relational structures, the solutions are significantly different (although the approach is essentially identical). One the one hand, the dominant term in the asymptotic enumeration of simple permutations is $n!/e^2$ (a fraction $1/e^2$ of the total number of permutations of length n), while on the other hand almost all graphs are indecomposable.

This difference indicates the caveat that must be added when attempting to study relational structures in their full generality: that certain results do hold for every structure (e.g.

the substitution decomposition), but many other results are only true in certain cases. We will encounter further differences as we progress through this study of simplicity – first in the difficulties of adapting the permutation-specific simple decomposition to the graph case in Chapter 2, and then again in the widely varying bounds on simple extensions in Chapter 3.

Graphs. Let us begin with the graph case, which turns out to be fairly straightforward. Let the random variable X_k denote the number of intervals of size k in a random graph G on n vertices. The probability that a given set of k vertices is an interval is $\frac{2^{n-k}}{2^{\binom{n}{2}}}$, since each of the $n-k$ vertices outside the interval must look at every vertex inside the interval in the same way. As there are $\binom{n}{k}$ ways of choosing the set of k vertices, we have

$$\mathbb{E}[X_k] = \frac{\binom{n}{k}2^{n-k}}{2^{\binom{n}{2}}}.$$

Thus the probability that G is decomposable may be bounded above by the sum of the expected number of proper intervals, i.e. it is bounded by $\mathbb{E}[X_2 + X_3 + \cdots + X_{n-1}]$. By linearity of expectation, this yields

$$\Pr(G \text{ is decomposable}) \leq \frac{2^n}{2^{\binom{n}{2}}} \sum_{k=2}^{n-1} \frac{\binom{n}{k}}{2^k}.$$

Observing that the sum is the binomial expansion of $(1+\frac{1}{2})^n$ less the first two and final terms, we obtain

$$\Pr(G \text{ is decomposable}) \leq \frac{2^n}{2^{\binom{n}{2}}} \left(\left(\frac{3}{2}\right)^n - 1 - \frac{n}{2} - \frac{1}{2^n} \right) \to 0 \text{ as } n \to \infty,$$

and hence almost all graphs are indecomposable. Möhring [91] shows this is also true for several other cases, including tournaments, posets and structures defined on single asymmetric relations. For the tournament version, see also Erdős, Fried, Hajnal and Milner [51].

Permutations. Proceeding as we did with graphs, let the random variable X_k denote the number of intervals of size k in a random permutation π of length n. An interval of length k may be viewed as a mapping from a contiguous set of positions to a contiguous set of values. The set of positions must begin at one of the first $n-k+1$ positions of π, and at the same time the lowest point in the set of values must be one of the lowest $n-k+1$ values of π. Of the $\binom{n}{k}$ sets of values to which the contiguous set of positions may be mapped, only one maps to the chosen contiguous set of values. Thus we have

$$\mathbb{E}[X_k] = \frac{(n-k+1)^2}{\binom{n}{k}} = \frac{(n-k+1)(n-k+1)!k!}{n!}.$$

Already we can see some difficulties may arise; whereas in the graph case it was clear that the denominator (being an exponential in n^2) would always dominate the numerator, here we see that this will not always hold. In particular, $\mathbb{E}[X_2] = \frac{2(n-1)}{n} \to 2$ as $n \to \infty$, implying in fact that, asymptotically, we expect to find two intervals of size two in a random permutation. Seeking the asymptotics of the other terms in $\sum_{k=2}^{n-1} \mathbb{E}[X_k]$, we consider the cases $k=3$, $k=4$, $k=n-2$ (assuming $n \geq 4$) and $k=n-1$ separately:

$$\begin{aligned}
\mathbb{E}[X_3] &= \frac{6(n-2)}{n(n-1)} \leq \frac{6}{n} \to 0 \\
\mathbb{E}[X_4] &= \frac{4!(n-3)}{n(n-1)(n-2)} \leq \frac{24}{n^2} \to 0 \\
\mathbb{E}[X_{n-2}] &= \frac{3 \cdot 3!}{n(n-1)} \leq \frac{24}{n^2} \to 0 \\
\mathbb{E}[X_{n-1}] &= \frac{4}{n} \to 0.
\end{aligned}$$

The remaining terms form a partial sum, which converges providing $\frac{\mathbb{E}[X_{k+1}]}{\mathbb{E}[X_k]} < 1$. Simplifying this equation gives $2k^2-(3n+1)k+n^2+n+1 > 0$, a quadratic in k, which yields two roots. The smaller of these satisfies $0 < k^- \leq n$, the larger $k^+ > n$. Thus for $k \leq k^-$, $\mathbb{E}[X_k]$ is decreasing, while for $k^- < k < n$, $\mathbb{E}[X_k]$ is increasing, and hence $\mathbb{E}[X_k] \leq 24/n^2$ for $4 \leq k \leq n-2$. Thus

$$\sum_{k=4}^{n-2} \mathbb{E}[X_k] \leq (n-5)\frac{24}{n^2} \leq \frac{24}{n} \to 0.$$

Subsequently, the only term of $\sum_{k=2}^{n-1} \mathbb{E}[X_k]$ which is non-zero in the limit $n \to \infty$ is $k=2$. Ignoring larger intervals, occurrences of intervals of size 2 in a random permutation π can roughly be regarded as independent events, and as we know the expectation of X_2 is 2, the occurrence of any specific interval is relatively rare. Heuristically, this suggests that X_2 is asymptotically Poisson distributed with parameter 2. Using this heuristic, we have $\Pr(X_2 = 0) \to e^{-2}$ as $n \to \infty$, and so there are approximately $\frac{n!}{e^2}$ simple permutations of length n.

A formal argument for this is implicitly given in Uno and Yagiura [116], and was made explicit by Corteel, Louchard, and Pemantle [37]. The method, however, essentially dates back to the 1940s with Kaplansky [74] and Wolfowitz [121], who considered "runs" within permutations – a *run* is a set of points with contiguous positions whose values are $i, i+1, \ldots, i+r$ or $i+r, i+r-1, \ldots, i$, in that order.[4]

A non-probabilistic approach (but one still relying on the work of Kaplansky) producing more precise asymptotics is given by Albert, Atkinson, and Klazar [3]. They obtain the following theorem, and note that higher order terms are obtainable given sufficient computation:

Theorem 1.4 (Albert, Atkinson and Klazar [3])**.** *The number of simple permutations of length n is asymptotically given by*

$$\frac{n!}{e^2}\left(1 - \frac{4}{n} + \frac{2}{n(n-1)} + O(n^{-3})\right).$$

1.5 Inflations and the Substitution Decomposition

With the notion of simplicity established, we may now describe how all relational structures can be decomposed and written in terms of these simple objects. This is easier to establish by first defining the reverse process. Given an $\mathcal{L}$-structure $\mathcal{S}$, an *inflation* of $\mathcal{S}$ by the $\mathcal{L}$-structures $\mathcal{A}_s$ for each $s \in \mathrm{dom}(\mathcal{S})$ — denoted $\mathcal{S}[\mathcal{A}_s : s \in \mathrm{dom}(\mathcal{S})]$ — is the $\mathcal{L}$-structure obtained by replacing each element s of $\mathrm{dom}(\mathcal{S})$ with a set of elements $\mathrm{dom}(\mathcal{A}_s)$ that form an interval in the $\mathcal{L}$-structure $\mathcal{A} = \mathcal{S}[\mathcal{A}_s : s \in \mathrm{dom}(\mathcal{S})]$, i.e. for every $R \in \mathcal{L}$:

$$R^{\mathcal{A}}(a_1, \ldots, a_{n_R}) \iff \begin{cases} R^{\mathcal{A}_s}(a_1, \ldots, a_{n_R}) \text{ and } a_1, \ldots, a_{n_R} \in \mathrm{dom}(\mathcal{A}_s),\ s \in \mathrm{dom}(\mathcal{S}), \text{ or} \\ R^{\mathcal{S}}(s_1, \ldots, s_{n_R}) \text{ where each } s_i \in \mathrm{dom}(\mathcal{S}) \text{ and } a_i \in \mathcal{A}_{s_i}. \end{cases}$$

[4] Atkinson and Stitt [12] called permutations containing no runs *strongly irreducible*. Note that this is equivalent to a permutation containing no intervals of size two.

A *deflation* (or *decomposition*) of an $\mathcal{L}$-structure $\mathcal{A}$ is the reverse. We write $\mathcal{A} = \mathcal{S}[\mathcal{A}_s : s \in \mathrm{dom}(\mathcal{S})]$ to mean any deflation of $\mathcal{A}$ by disjoint intervals $\mathcal{A}_s$. We are primarily interested in the case where $\mathcal{S}$ is simple – the following theorem gives the uniqueness of such an $\mathcal{S}$, which will be called the *skeleton*.

Theorem 1.5 (The Substitution Decomposition). *Let $\mathcal{A}$ be an $\mathcal{L}$-structure for some language $\mathcal{L}$. Then there exists a unique simple $\mathcal{L}$-structure $\mathcal{S}$ such that $\mathcal{A} = \mathcal{S}[\mathcal{A}_s : s \in \mathrm{dom}(\mathcal{S})]$. Moreover, when $|\mathrm{dom}(\mathcal{S})| > 2$, every $\mathcal{A}_s$ is defined uniquely.*

Proof. Let M denote the set of all intervals, except $\mathrm{dom}(\mathcal{A})$, which are contained in no other proper intervals.

If two intervals $I, J \in M$ intersect, then Proposition 1.2.(b) shows that $I \cup J$ is also an interval, which, unless $I \cup J = \mathrm{dom}(\mathcal{A})$, contradicts the definition of M. If $I \cup J = \mathrm{dom}(\mathcal{A})$, then Proposition 1.2.(c) shows that $J \setminus I$ is an interval, so $\mathcal{A}$ can be written as the inflation of a two-element $\mathcal{L}$-structure, all of which are simple. If $\mathcal{A} = \mathcal{S}[\mathcal{A}_{s_1}, \mathcal{A}_{s_2}]$ and $\mathcal{A} = \mathcal{T}[\mathcal{A}_{t_1}, \mathcal{A}_{t_2}]$ are two different two-element decompositions, then we may assume that in $\mathcal{A}$ we have $\mathcal{A}_{s_1} \cap \mathcal{A}_{t_1} \neq \emptyset$ and $\mathcal{A}_{s_2} \cap \mathcal{A}_{t_2} \neq \emptyset$. Thus relations in $\mathcal{S}$ between s_1 and s_2 must agree with the relations in $\mathcal{A}$ between elements of the disjoint intervals $\mathcal{A}_{s_1}$ and $\mathcal{A}_{s_2}$. Since $\mathcal{A}_{s_1} \cap \mathcal{A}_{t_1} \subseteq \mathcal{A}_{s_1}$ and $\mathcal{A}_{s_2} \cap \mathcal{A}_{t_2} \subseteq \mathcal{A}_{s_2}$ are intervals, the relations between elements of $\mathcal{A}_{s_1}$ and $\mathcal{A}_{s_2}$ correspond to the relations between the elements of $\mathcal{A}_{s_1} \cap \mathcal{A}_{t_1}$ and $\mathcal{A}_{s_2} \cap \mathcal{A}_{t_2}$, which, by a similar argument must correspond to the relations between the elements of $\mathcal{A}_{t_1}$ and $\mathcal{A}_{t_2}$, and these are none other than the relations between t_1 and t_2 of $\mathrm{dom}(\mathcal{T})$. Similarly, relations involving just s_1 (respectively, s_2) correspond to relations involving just t_1 (t_2), and so $\mathcal{S}$ and $\mathcal{T}$ are isomorphic.

Otherwise, the sets in M partition $\mathrm{dom}(\mathcal{A})$. For each $I \in M$ choose a representative $x_I \in I$, and define the $\mathcal{L}$-structure $\mathcal{S}$ on $\{x_I\}$ by $\mathcal{A}|_{\{x_I\}} = \mathcal{S}$. Clearly $\mathcal{A}$ is the inflation of $\mathcal{S}$ by the structures $\mathcal{A}|_I$ for $I \in M$. The simplicity of $\mathcal{S}$ follows from the observation that if $\mathcal{S}$ contained a proper interval K, then $\bigcup_{x_I \in K} I$ would be a proper interval of $\mathcal{A}$ contradicting the definition of M. Furthermore, if $\mathcal{A} = \mathcal{T}[\mathcal{A}_t : t \in \mathrm{dom}(\mathcal{T})]$ for any other simple $\mathcal{L}$-structure $\mathcal{T}$, then each $\mathrm{dom}(\mathcal{A}_t)$ is an interval of $\mathcal{A}$ and so is contained in an interval in M. □

The non-unique cases which occur when $|\operatorname{dom}(\mathcal{S})| = 2$ may be dealt with in a number of ways, some of which are specific to particular types of structure, as we will see later. In the general setting, however, we can still find a unique substructure of $\mathcal{A}$ that is essentially from one of three groups.

Proposition 1.6 (Möhring and Radermacher [95, Theorems 3.4.3, 4.1.2 and 4.1.3]). *If $\mathcal{A}$ is an $\mathcal{L}$-structure whose skeleton $\mathcal{S}$ satisfies $|\operatorname{dom}(\mathcal{S})| = 2$, then there exists a unique maximal $\mathcal{L}$-structure $\mathcal{T}$ for which $\mathcal{A} = \mathcal{T}[\mathcal{A}_t : t \in \operatorname{dom}(\mathcal{T})]$ and, for every $R \in \mathcal{L}$, $R^{\mathcal{T}}$ is linear, complete or empty.*

Once we have established the substitution decomposition $\mathcal{A} = \mathcal{S}[\mathcal{A}_s : s \in \operatorname{dom}(\mathcal{S})]$, we may repeat the process on the substructures $\mathcal{A}_s$ for each $s \in \operatorname{dom}(\mathcal{S})$. Iterating this decomposition, we may continue until we are left only with substructures on singleton ground sets. We may represent this iterated substitution decomposition as a rooted tree – the *substitution decomposition tree*. Each node corresponds to a substructure of $\mathcal{A}$ whose ground set is an interval, with the root of the tree being $\mathcal{A}$ and the leaves being the singleton ground sets. For a given node with corresponding non-singleton structure $\mathcal{A}'$, the children of $\mathcal{A}'$ are the substructures $\mathcal{A}'_s$ in the decomposition $\mathcal{A}' = \mathcal{S}'[\mathcal{A}'_s : s \in \operatorname{dom}(\mathcal{S})]$. If $\mathcal{S}$ is a unique simple with $|\operatorname{dom}(\mathcal{S})| \geq 4$, label the node corresponding to $\mathcal{A}'$ with the symbol P (short for "proper"); if the (binary) relations in the language of $\mathcal{S}$ are linear and all other relations are complete or empty, label the node with the symbol L; if all the relations in the language of $\mathcal{S}$ are complete or empty, label the node D (short for "degenerate").

1.5.1 The Permutation Case

Restricting our attention to the permutation case, the substitution decomposition is somewhat easier to describe. Given a permutation σ of length m and nonempty permutations $\alpha_1, \ldots, \alpha_m$, the *inflation* of σ by $\alpha_1, \ldots, \alpha_m$ – denoted $\sigma[\alpha_1, \ldots, \alpha_m]$ – is the permutation obtained by replacing each entry $\sigma(i)$ by an interval that is order isomorphic to α_i. For example, $2413[1, 132, 321, 12] = 479832156$ (see Figure 1.5). Conversely, a *deflation* of π is

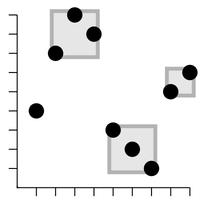

Figure 1.5: The plot of 479832156, an inflation of 2413.

any expression of π as an inflation $\pi = \sigma[\pi_1, \pi_2, \ldots, \pi_m]$, and we will call σ a *skeleton* of π. Theorem 1.5 then specialises to become:

Proposition 1.7 (Albert and Atkinson [2]). *Every permutation may be written as the inflation of a unique simple permutation. Moreover, if π can be written as $\sigma[\alpha_1, \ldots, \alpha_m]$ where σ is simple and $m \geq 4$, then the α_is are unique.*

The degenerate cases occur when a permutation can be written as an inflation of either 12 or 21, and we may choose a unique decomposition in these cases in a variety of ways. The principal decomposition that we will use for the substitution decomposition, however, is as described in Proposition 1.6.

The *direct sum* of two permutations α and β is the inflation $12[\alpha, \beta]$, and is usually denoted $\alpha \oplus \beta$. Similarly, the *skew sum* is the inflation $21[\alpha, \beta]$, and is denoted $\alpha \ominus \beta$. The direct sum operation acts as a dichotomy on the set of all permutations – dividing them into those that are *sum decomposable* (i.e. they can be represented as a direct sum), and those that are *sum indecomposable*. Similarly, the skew sum operation leads to the *skew decomposable* permutations, while those that cannot be represented as a skew sum are *skew indecomposable*.

With these definitions, if π can be written as a direct sum (i.e. an inflation of the simple permutation 12), then we may write $\pi = \iota_m[\alpha_1, \ldots, \alpha_m]$ uniquely where m is maximal, and each α_i is sum indecomposable. Similarly, if π is an inflation of 21, we may write $\pi = \delta_m[\alpha_1, \ldots, \alpha_m]$ where each α_i is skew indecomposable.

Alternatively, we may prefer to express π as the inflation of 12 or 21, in which case we will specify which deflation we want; the one that follows will be the decomposition we

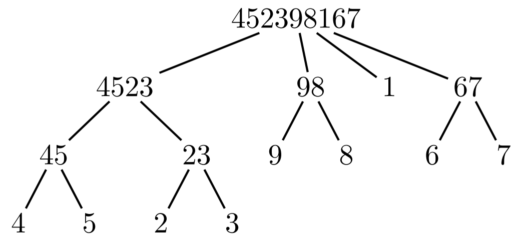

Figure 1.6: The substitution decomposition tree of $\pi = 452398167$.

mostly use.

Proposition 1.8 (Albert and Atkinson [2]). *If π is an inflation of 12, then there is a unique sum indecomposable α_1 such that $\pi = 12[\alpha_1, \alpha_2]$ for some α_2, which is itself unique. The same holds with 12 replaced by 21 and "sum" replaced by "skew".*

The substitution decomposition tree for a permutation then follows immediately. For example, consider the permutation $\pi = 452398167$. This is decomposed as

$$\begin{aligned} 452398167 &= 2413[3412, 21, 1, 12] \\ &= 2413[21[12, 12], 21[1, 1], 1, 12[1, 1]] \\ &= 2413[21[12[1, 1], 12[1, 1]], 21[1, 1], 1, 12[1, 1]] \end{aligned}$$

and its substitution decomposition tree is given in Figure 1.6.

CHAPTER 2

DECOMPOSITION

2.1 Background

SINCE simple permutations may be used to construct all other permutations via the substitution decomposition, it would be useful to know how simple permutations are themselves constructed. In particular, our aim is to find smaller "fundamental" simple permutations of some specified size within a given simple permutation. Some approaches to this question can be found in Schmerl and Trotter [107], in which the following is proved for all irreflexive binary relational structures.[1] Here, however, we will state only the permutation case, for which there is another proof by Murphy [97].

Theorem 2.1 (Schmerl and Trotter [107])**.** *Every simple permutation of length $n \geq 2$ contains a simple permutation of length $n-1$ or $n-2$.*

We will prove that long simple permutations must contain two long almost disjoint simple subsequences. Formally:

Theorem 2.2. *There is a function $f(k)$ such that every simple permutation of length at least $f(k)$ contains two simple subsequences, each of length at least k, sharing at most two entries.*

(The proof of Theorem 2.2 follows after establishing Theorem 2.14, found on Page 34.) The second "two" in the statement of Theorem 2.2 is best possible, as is demonstrated by

[1] A version of this theorem for *k-structures* – structures defined on a single k-ary relation in which every relation $(a_1, \ldots, a_k)$ has $a_i \neq a_j$ for some $i \neq j$ – can be found in Ehrenfeucht and McConnell [48].

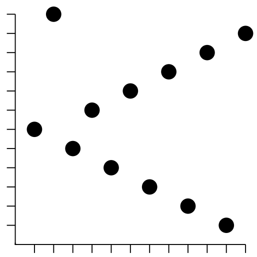

Figure 2.1: The plots of a wedge simple permutation. Note that every simple subsequence of length at least 4 must contain its first two entries.

the family of simple permutations of the form

$$m(2m)(m-1)(m+1)(m-2)(m+2)\cdots 1(2m-1);$$

the permutation in Figure 2.1 is of this form. On the other hand, no attempt has been made to optimise the function f; our proof gives an f of order about k^k.

This result alone, however, gives no real indication as to the underlying structure within the simple permutation; rather it is the method by which we arrive at Theorem 2.2. We give a Ramsey-type description of simple permutations in terms of some unavoidable substructures, similar to the Erdős-Szekeres Theorem as applied to arbitrary permutations:

Theorem 2.3 (Erdős and Szekeres [53]). *Every permutation of length n contains a monotone increasing or monotone decreasing subsequence of length at least $\sqrt{n}$.*

In particular, we will demonstrate how a sufficiently long simple permutation contains, in the first instance, a "parallel alternation" of length k, a "wedge alternation" of length k or a "pin sequence" of length k. By studying the decomposition of pin sequences, we can go further to provide a more straightforward result, namely every sufficiently long simple permutation contains either an "alternation" or an "oscillation".

A major motivation of this study is the enumeration of particular permutation classes. Although we will delay an in-depth discussion of this until Part II, it is worth noting that establishing a method of classifying the simple permutations brings us much closer to establishing what simple permutations lie in a given class.

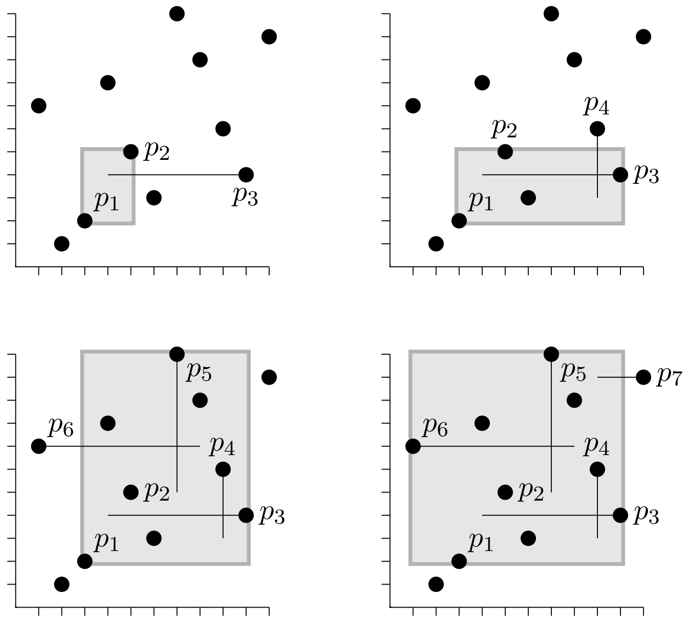

Figure 2.2: A pin sequence.

2.2 Pin Sequences

The core of the simple permutation decomposition is in understanding pin sequences. Empirically, they encapsulate precisely what it means to be simple: in the plot of a simple permutation, any set of points enclosed by an axis-parallel rectangle must be separated by at least one point lying outside the box above, below, to the left or to the right, and formalising the method of finding such a point is the motivation for defining pins, and subsequently sequences of pins.

While the viewpoint above will regard pins in their motivational setting as points *within* the plot of a permutation, when we come to discussing our final "unavoidable substructures" result, we are going to need to decompose these pin sequences. To do this, we will shift our viewpoint to building pin sequences from scratch by placing points in a plane, each of which will correspond to a pin. We will also need to consider subsequences of a given pin sequence, for which we will need to introduce "pin words".

Let us begin, however, with a more detailed motivational definition of pin sequences in our original setting. Recall the graphical representation of a permutation as described in Section 1.2. 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.

Choose two points p_1 and p_2 in the plot of a permutation π. If these two points do not form an interval then there is at least one point which lies outside $\mathrm{rect}(p_1, p_2)$ and slices $\mathrm{rect}(p_1, p_2)$ either horizontally or vertically. (This discussion is accompanied by the sequence of diagrams shown in Figure 2.2.) We call such a point a *pin*. Choose a pin and label it p_3. Now consider the larger rectangle $\mathrm{rect}(p_1, p_2, p_3)$. If this also does not form an interval in π then we can find another pin, p_4, which slices $\mathrm{rect}(p_1, p_2, p_3)$ either horizontally or vertically. Again, if $\mathrm{rect}(p_1, p_2, p_3, p_4)$ is not an interval then we can find another pin p_5. We refer to a sequence of pins constructed in this manner as a *pin sequence*.

Formally, a pin sequence is a sequence of points p_1, p_2, $\ldots$ in the plot of π such that for each $i \geq 3$,

- $p_i \notin \mathrm{rect}(p_1, \ldots, p_{i-1})$, and
- if $\mathrm{rect}(p_1, \ldots, p_{i-1}) = [a, b] \times [c, d]$ and $p_i = (x, y)$, we have either $a < x < b$ or $c < y < d$, or, in other words, p_i *slices* $\mathrm{rect}(p_1, \ldots, p_{i-1})$ either horizontally or vertically.

We describe pins as either *left*, *right*, *up*, or *down* based on their position relative to the rectangle that they slice. Thus in the pin sequence from Figure 2.2, p_3 and p_7 are right pins, p_4 and p_5 are up pins, p_6 is a left pin, and p_8 is a down pin (p_1 and p_2 lack direction).

A *proper pin sequence* is one that satisfies two additional conditions:

- *Maximality condition*: each pin must be maximal in its direction. For example, if $\mathrm{rect}(p_1, \ldots, p_{i-1}) = [a, b] \times [c, d]$ and $p_i = (x, y)$ is a right pin, then it is the right-most of all possible right pins for this rectangle, or, in other words, the region $(x, n] \times [c, d]$ is devoid of points.
- *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.

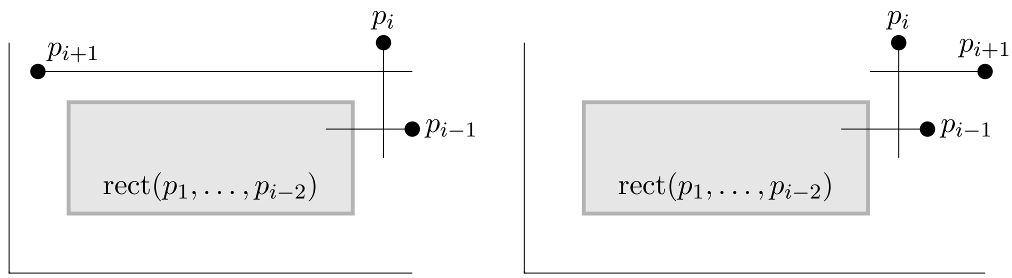

Figure 2.3: The two cases in the proof of Lemma 2.6.

For example, in the pin sequence shown in Figure 2.2, the choice of p_4 violates the maximality condition, while the choices of p_5, p_7, and p_8 violate the separation condition. The ultimate goal of the following succession of lemmas is to show (in Theorem 2.7) that all or all but one of the pins in a proper pin sequence themselves form a simple permutation. We begin by observing that proper pin sequences travel by 90° turns only.

Lemma 2.4. *In a proper pin sequence, p_{i+1} cannot lie in the same or opposite direction as p_i (for all $i \geq 3$).*

Proof. By the maximality condition, p_{i+1} cannot lie in the same direction as p_i. It cannot lie in the opposite direction by the separation condition. □

Lemma 2.5. *In a proper pin sequence, p_i does not separate any two members of $\{p_1, \ldots, p_{i-2}\}$.*

Proof. If p_i did separate $\text{rect}(p_1, \ldots, p_{i-2})$ into two parts then p_{i-1} would lie on one side of this divide, violating the separation condition. □

Lemma 2.6. *In a proper pin sequence, p_i and p_{i+1} are separated either by p_{i-1} or by each of $p_1, \ldots, p_{i-2}$.*

Proof. The lemma is vacuously true for $i = 1$ and $i = 2$, so let us assume that $i \geq 3$. Without loss we may assume that p_{i-1} is a right pin and p_i is an up pin. By Lemma 2.4, p_{i+1} must be either a right pin or a left pin. The remainder of the proof is evident from Figure 2.3. □

We are now ready to prove our main result about proper pin sequences.

Theorem 2.7. *If $p_1, \ldots, p_m$ is a proper pin sequence of length $m \geq 5$ 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.*

Proof. We are interested in the possible intervals in the subsequence given by the pins $p_1, \ldots, p_m$; we shall call these *intervals of pins*. The bulk of our proof is devoted to establishing the following claim: for any m, the only possible proper minimal nonsingleton intervals of pins in the proper pin sequence $\{p_1, \ldots, p_m\}$ are $\{p_1, p_m\}$, $\{p_2, p_m\}$, $\{p_1, p_3, \ldots, p_m\}$ or $\{p_2, \ldots, p_m\}$.

Take $M \subseteq \{p_1, \ldots, p_m\}$ to be a minimal non-singleton interval of pins. Note that M is therefore order isomorphic to a simple permutation. If M contains a pair of pins p_i and p_j with $i < j < m$ then by the separation condition $p_{j+1}, \ldots, p_m \in M$. Furthermore, because $j < m$, Lemma 2.6 shows that M contains either p_{j-1} or $p_1, p_2, \ldots, p_{j-2}$. In the latter case, if $j \geq 4$ then separation gives $p_{j-1} \in M$, as desired, while if $j \leq 3$, we have already found a minimal non-singleton interval of pins of the desired form. In the former case, the proof is completed by iterating this process. Only the case $M = \{p_i, p_m\}$ remains. If $3 \leq i \leq m-1$ then by the separation condition p_i separates $\{p_1, \ldots, p_{i-1}\}$, while Lemma 2.5 shows that p_m does not separate these points; thus at least one of them must lie in M, a contradiction which completes the proof of the claim.

Returning to the proof of the theorem, suppose that $\{p_1, \ldots, p_m\}$ is not itself order isomorphic to a simple permutation and that $m \geq 5$. Thus, by the claim, at least one of $\{p_1, p_m\}$, $\{p_2, p_m\}$, $\{p_1, p_3, \ldots, p_m\}$ or $\{p_2, \ldots, p_m\}$ forms a minimal nonsingleton interval of pins. The latter two cases give us a simple of the desired form, so now assume either $\{p_1, p_m\}$ or $\{p_2, p_m\}$ is an interval of pins. (Note that we cannot have both intervals since p_3 separates p_1 from p_2.) We assume the former as the latter is analogous. Consider the pin sequence $\{p_2, \ldots, p_m\}$. By the claim, the only possible minimal nonsingleton intervals of pins in this sequence are $\{p_2, p_m\}$, $\{p_3, p_m\}$, $\{p_2, p_4, \ldots, p_m\}$ or $\{p_3, \ldots, p_m\}$. The latter two cases may be ignored since the only interval of pins in the original sequence $\{p_1, \ldots, p_m\}$

was $\{p_1, p_m\}$, and hence the points that p_1 separated are the same as those separated by p_m. Thus it remains to eliminate the cases $\{p_2, p_m\}$ and $\{p_3, p_m\}$. Since p_3 separates p_1 from p_2 and $\{p_1, p_m\}$ is an interval, p_3 also separates p_2 from p_m, so $\{p_2, p_m\}$ cannot form an interval of pins for the sequence $\{p_2, \ldots, p_m\}$. Similarly, $\{p_3, p_m\}$ cannot be an interval of pins for $\{p_2, \ldots, p_m\}$ because p_4 separates p_3 from p_1 and thus also from p_m because we have assumed that $\{p_1, p_m\}$ forms an interval. Thus $\{p_2, \ldots, p_m\}$ contains no nontrivial intervals of pins and is therefore order isomorphic to a simple permutation, completing the proof. □

As a corollary of this theorem, we see that Theorem 2.2 (in fact, a stronger result) is true for simple permutations with long pin sequences.

Corollary 2.8. *If π contains a proper pin sequence of length at least $2k + 2$ (with $k \geq 4$) then π contains two disjoint simple subsequences, each of length at least k.*

Proof. Apply Theorem 2.7 to the two pin sequences $p_1, \ldots, p_{k+1}$ and $p_{k+2}, \ldots, p_{2k+2}$. □

We say that the pin sequence $p_1, \ldots, p_m$ for the permutation π of length n is *saturated* if $\mathrm{rect}(p_1, \ldots, p_m) = [n] \times [n]$. For example, the pin sequence in Figure 2.2 is saturated. Any two points $p_1 \neq p_2$ in the plot of a simple permutation can be extended to a saturated pin sequence, as we are forced to stop extending a pin sequence only upon finding an interval or when the rectangle contains every point in π.

It is important to note that two points in a simple permutation need not be extendable to a proper saturated pin sequence. For example, the permutation in Figure 2.2 does not have a proper saturated pin sequence beginning with p_1 and p_2. For this reason we work with a weaker requirement: the pin sequence $p_1, \ldots, p_m$ is said to be *right-reaching* if p_m is the right-most point of π.

Lemma 2.9. *For every simple permutation π and pair of points p_1 and p_2 (unless, trivially, p_1 is the right-most point of π), there is a proper right-reaching pin sequence beginning with p_1 and p_2.*

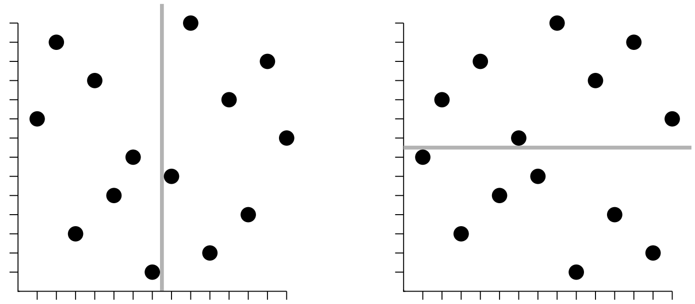

Figure 2.4: A horizontal alternation (left) and its inverse, a vertical alternation (right).

Proof. Clearly we can find a saturated pin sequence $p_1, p_2, \ldots$ in π that satisfies the maximality condition. Since this pin sequence is saturated, it includes the right-most point; label it p_{i_1}. Now take i_2 as small as possible so that $p_1, p_2, \ldots, p_{i_2}, p_{i_1}$ is a valid pin sequence. Note first that $i_2 < i_1$ because $p_1, \ldots, p_{i_1}$ is a valid pin sequence. Now observe that p_{i_1} separates p_{i_2} from $\mathrm{rect}(p_1, \ldots, p_{i_2-1})$, because $p_1, \ldots, p_{i_2-1}, p_{i_1}$ is not a valid pin sequence. Continuing in this manner, we find pins p_{i_3}, p_{i_4}, and so on, until we reach the stage where $p_{i_{m+1}} = p_2$. Then $p_1, p_2, p_{i_m}, p_{i_{m-1}}, \ldots, p_{i_1}$ is a proper right-reaching pin sequence. ☐

2.3 Simple Permutations without Long Proper Pin Sequences

It remains only to consider those simple permutations without long proper pin sequences. Lemma 2.9 shows that in such a permutation, any two points p_1, p_2 can be extended to a short proper right-reaching pin sequence. Our goal in this section is to use several of these short right-reaching sequences to prove that such permutations contain long "alternations".

We use the term *horizontal alternation* to refer to a permutation in which every odd entry lies to the left of every even entry, or the reverse of such a permutation. A *vertical alternation* is the group-theoretic inverse of a horizontal alternation. Examples are shown in Figure 2.4. Every sufficiently long vertical alternation contains either a long *parallel alternation* or a long *wedge alternation* (see Figure 2.5 for definitions):

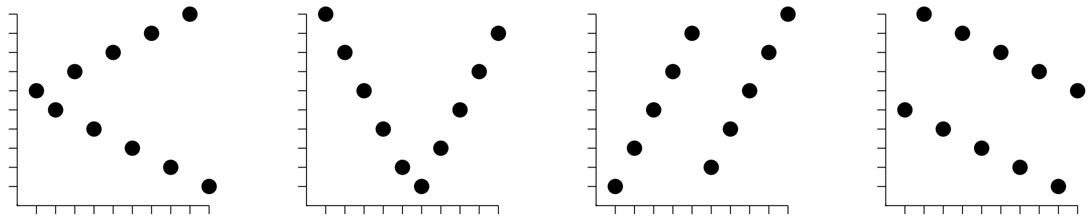

Figure 2.5: The two permutations on the left are wedge alternations, the two on the right are parallel alternations.

Proposition 2.10. *Every alternation of length at least $2k^4$ contains either a parallel or wedge alternation of length at least $2k$.*

Proof. Let π be a vertical alternation of length $2n \geq 2k^4$. By the Erdős-Szekeres Theorem 2.3, the sequence $\pi(1), \pi(3), \ldots, \pi(2n-1)$ contains a monotone subsequence of length at least k^2, say $\pi(i_1), \pi(i_2), \ldots, \pi(i_{k^2})$. Applying the Erdős-Szekeres Theorem to the subsequence $\pi(i_1+1), \pi(i_2+1), \ldots, \pi(i_{k^2}+1)$ completes the proof. □

Note that every parallel alternation of length $2k+2 \geq 10$ contains two disjoint simple permutations of length at least k. Thus Theorem 2.2 follows in the case where our simple permutation contains a long parallel alternation.

Returning to pin sequences, the pin sequences $p_1, p_2, \ldots$ and $q_1, q_2, \ldots$ are said to

- be *initially-nonoverlapping* if $\mathrm{rect}(p_1, p_2)$ and $\mathrm{rect}(q_1, q_2)$ are disjoint,
- *converge at the point x* if there exist i and j such that $p_i = q_j = x$ but $\{p_1, \ldots, p_{i-1}\}$ and $\{q_1, \ldots, q_{j-1}\}$ are disjoint.

A collection of pin sequences converges or is initially-nonoverlapping if they pairwise converge or are pairwise initially-nonoverlapping. Note that it is always possible to find a collection of $\lfloor n/2 \rfloor$ initially-nonoverlapping proper pin sequences in a permutation π of length n by taking proper pin sequences beginning with the first and second points, the third and fourth points, and so on, reading left to right.

Lemma 2.11. *If $16k$ initially-nonoverlapping proper pin sequences of π converge at the same point, then π contains an alternation of length at least $2k$.*

Proof. Suppose that $16k$ initially-nonoverlapping proper pin sequences converge at the point x. Note that x can be the first or second pin for at most one of these sequences because they are initially-nonoverlapping. Thus one of the following two possibilities must occur:

- at least $8k$ of the sequences have x as their third pin, or
- at least $8k$ of the sequences have x as their fourth or later pin.

Suppose that at least $8k$ of the sequences have x as their third pin. This point could be variously functioning as a left, right, down, or up pin for each of these $8k$ sequences, but x plays the same role for at least $2k$ sequences. Suppose, by symmetry, that x is a right pin for at least $2k$ sequences. Since x is the third pin for these sequences, one of their first two pins lies above x while the other lies below and because these sequences are initially-nonoverlapping, an alternation of length at least $2k$ can be obtained by choosing one point from each sequence.

Now suppose that at least $8k$ of the sequences have x as their fourth or later pin. Again we may assume without loss that x is a right pin for at least $2k$ of these sequences. Now consider the immediate predecessors to x in these sequences. These pins are either up pins or down pins (by Lemma 2.4). By symmetry we may assume that for at least k of these sequences the immediate predecessor to x is an up pin. Reading left to right, label these immediate predecessor pins $p^{(1)}$, $p^{(2)}, \ldots, p^{(k)}$ and let $R^{(i)}$ denote the rectangle for which $p^{(i)}$ is a pin. Note that each $R^{(i)}$ lies completely below x, as otherwise the separation condition would prevent x from following $p^{(i)}$ in the corresponding pin sequence. We now have the situation depicted in Figure 2.6.

It suffices to show, for each i, that π contains a point lying horizontally between $p^{(i)}$ and $p^{(i+1)}$ and below x since these points, together with the $p^{(i)}$'s and x, will give an alternation of length $2k$. However, if there is no such point then $p^{(i)}$ and $p^{(i+1)}$ could each function as up pins for both $R^{(i)}$ and $R^{(i+1)}$, and thus one of these choices would contradict the maximality condition, completing the proof. □

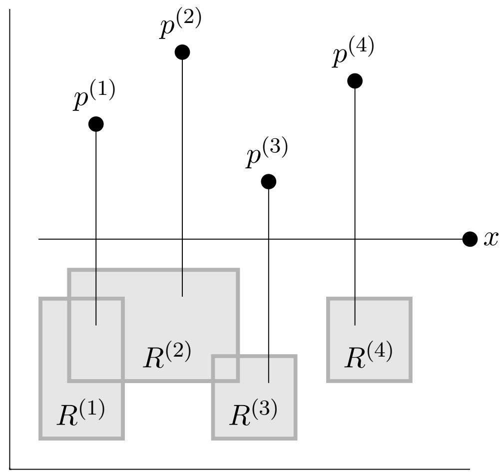

Figure 2.6: The situation that arises in the proof of Lemma 2.11.

Lemma 2.12. *Every simple permutation of length at least* $2(16k^4)^{2k}$ *contains either a proper pin sequence of length at least* $2k$ *or a parallel or wedge alternation of length at least* $2k$.

Proof. Suppose that a simple permutation π of length n contains neither a proper pin sequence of length at least $2k$ nor a parallel or wedge alternation of length at least $2k$. In particular, π does not contain a proper right-reaching pin sequence of length $2k$, and it follows from Proposition 2.10 that π has no alternations of length $2k^4$.

It follows from our earlier observations that π contains a collection of $\lfloor n/2 \rfloor$ initially-nonoverlapping proper right-reaching pin sequences. As these sequences are right-reaching, they all have the same final (right-most) pin which we denote by p. By Lemma 2.11, fewer than $16k^4$ of these pin sequences converge at p; equivalently, there are fewer than $16k^4$ distinct immediate predecessors to p, and we label these as $p^{(1)}, p^{(2)}, \ldots, p^{(m)}$. Again, fewer than $16k^4$ pin sequences converge at each of the $p^{(i)}$'s, so there are fewer than $(16k^4)^2$ immediate predecessors to these pins. Continue this process until we reach the sequences of length $2k$, of which we have assumed there are none. We have thus counted all $\lfloor n/2 \rfloor$ of our sequences, and have obtained the bound

$$\lfloor n/2 \rfloor < 1 + 16k^4 + (16k^4)^2 + (16k^4)^3 + \cdots + (16k^4)^{(2k-1)},$$

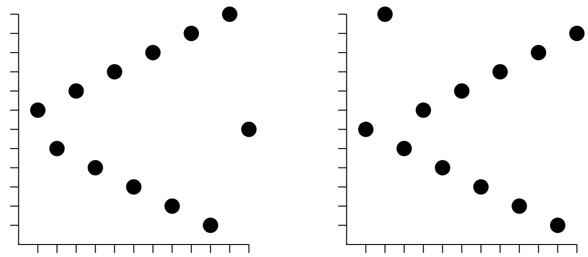

Figure 2.7: The two types of wedge simple permutations, type 1 (left) and type 2 (right).

so, simplifying,

$$n < 2(16k^4)^{2k}. \qquad \square$$

We are left to deal with simple permutations which do not have long proper pin sequences but do have long wedge alternations. We prove that these permutations contain long *wedge simple permutations*, of which there are two types (up to symmetry). Examples of these two types are shown in Figure 2.7.

Lemma 2.13. *If a simple permutation contains a wedge alternation of length $4k^2$ then it contains either a pin sequence of length at least $2k$ or a wedge simple permutation of length at least $2k$.*

Proof. Let π be a simple permutation containing a wedge alternation of length at least $4k^2$. By symmetry we may assume that this wedge alternation opens to the right (i.e. it is oriented as $<$). We call these the *wedge points* of π. Label the two left-most wedge points p_1 and p_2 and by Lemma 2.9 extend this into a proper right-reaching pin sequence $p_1, p_2, \ldots, p_m$.

Let R_i denote the smallest rectangle in the plot of π containing p_1, p_2, and p_i that is not sliced by a wedge point outside the rectangle. Define the *wedge sum* of the pin p_i, $\mathrm{ws}(p_i)$, to be the number of wedge points in R_i. For $i \geq 2$ define the *wedge contribution* of p_i by $\mathrm{wc}(p_i) = \mathrm{ws}(p_i) - \mathrm{ws}(p_{i-1})$ and set $\mathrm{wc}(p_1) = 1$. Regarding these quantities we make four observations:

(W1) the wedge sum of p_m is equal to the total number of wedge points and also to $\sum_{i=1}^{m} \mathrm{wc}(p_i)$,

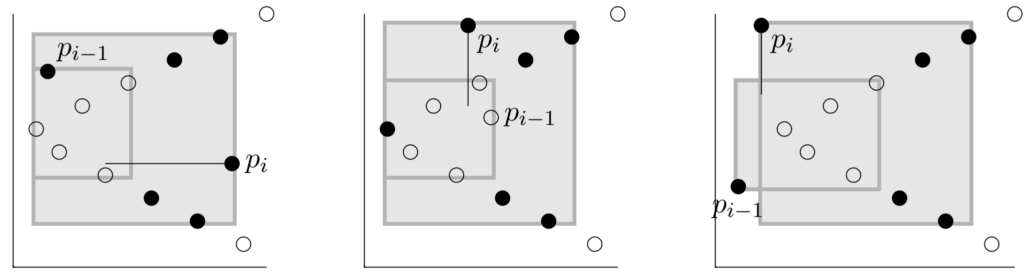

Figure 2.8: The three cases in the proof of Lemma 2.13; the solid points form simple permutations.

(W2) it is not hard to construct examples in which pins have negative wedge contributions; indeed,

(W3) left pins cannot have positive wedge contributions, and finally,

(W4) if p_i is an up pin, then the right-most wedge point in R_i is an upper wedge point.

We now claim that each p_i lies in a wedge simple permutation of length at least $\mathrm{wc}(p_i)+2$. This claim implies the theorem, because if no pin lies in a wedge simple permutation of length at least $2k$ then $\mathrm{wc}(p_i) \leq 2k-3$, so by (W1),

$$4k^2 \leq \sum_{i=1}^{m} \mathrm{wc}(p_i) \leq m(2k-3),$$

and thus $m \geq 2k$, giving the long pin sequence desired.

The claim is easily observed for $i = 1$ and, by (W3), vacuously true if p_i is a left pin. Thus by symmetry there are only three cases to consider: an up pin followed by a right pin, a right pin followed by an up pin, and a left pin followed by an up pin. These three cases are depicted in Figure 2.8.

Let us consider in detail the case of an up pin followed by a right pin. By (W4), the left-most wedge point in $R_i \setminus R_{i-1}$ lies below p_1. By separation, p_{i-1} lies above p_i, which is itself the right-most point in R_i. Therefore the wedge points in $R_i \setminus R_{i-1}$ together with p_i and p_{i-1} constitute a type 1 wedge simple permutation. The other cases follow by similar analysis; in the right-up case the wedge points in $R_i \setminus R_{i-1}$ together with p_1 and p_i give a wedge simple permutation of type 2, while in the left-up case a wedge simple permutation of type 2 can be formed from the wedge points in $R_i \setminus R_{i-1}$, p_{i-1}, and p_i. □

We have therefore established the following theorem.

Theorem 2.14. *Every simple permutation of length at least $2(256k^8)^{2k}$ contains a proper pin sequence of length $2k$, a parallel alternation of length $2k$, or a wedge simple permutation of length $2k$.*

The proof of Theorem 2.2 now follows by analysing each of these cases in turn. A parallel alternation of length $2k+2 \geq 10$ contains two disjoint simple permutations of length k. A type 1 wedge simple permutation of length $2k$ contains two type 1 wedge simple permutations of length k with only one entry in common, and a type 2 wedge simple permutation of length $2k$ contains two type 2 wedge simple permutations of length k which share two entries. Finally, Corollary 2.8 shows that a permutation with a proper pin sequence of length $2k+2$ contains two disjoint simple permutations of length k.

2.4 Pin Words

To explain how to expatiate Theorem 2.14 into a simpler "unavoidable substructures" result, we must first change our viewpoint so we can consider arbitrary proper pin sequences and their subsets, rather than pin sequences within a given simple permutation. This treatment will also be of use in Part II. To this end we extend the pin sequence definition to allow us to place points in the plane as they are required. While the precise coordinates of each pin will be far from unique, we do not encounter any difficulties as two sets of points in the plane constructed by the same pin sequence will be order isomorphic.

The changing viewpoint requires that we replace the maximality condition with the "externality" condition. Formally, a *proper pin sequence* is a sequence of points in the plane satisfying:

- *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.

- *Externality condition*: p_{i+1} must lie outside $\mathrm{rect}(p_1, \ldots, p_i)$.

Note that, as we are now building proper pin sequences from scratch, the externality and separation conditions together imply the maximality condition.

Proper pin sequences can essentially be described naturally by words over a four-letter alphabet consisting of the *directions* $\{L, R, U, D\}$ (standing for left, right, up and down). This does not, of course, precisely define how the pin sequence begins, a detail which we will deal with shortly.

A subsequence of a proper pin sequence, viewed in the same order as the original pin sequence, consists of some points which still satisfy the separation condition and some that do not. (Note that externality is always satisfied.) The points that do still separate can be described by one of the letters L, R, U or D as before, since they are still proper pins. Each point p not satisfying separation arose because its immediate predecessor pin in the proper pin sequence was not included in the subsequence. By externality, however, p must lie in one of the four *quadrants* as defined by the axis-parallel rectangle enclosing all points of the subsequence coming before p (see Figure 2.9). We may now represent p with a numeral corresponding to the *quadrant* in which it lies, and so to encode subsequences of proper pin sequences, we append to the alphabet $\{L, R, U, D\}$ the set of four *numerals* $\{1, 2, 3, 4\}$, indicating a point is to be placed in the appropriate quadrant.

Before our formal definition of a pin word, it remains to give an informal description of how to represent the start of a pin sequence. This may be done in a variety of ways, but the most effective method for our purposes will be to fix the placement of the origin, and regard it as a pin coming before the first pin of the original sequence. We can then represent the first pin with a numeral denoting its quadrant in relation to the origin, and thereafter proceed as already described.

Formally, 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

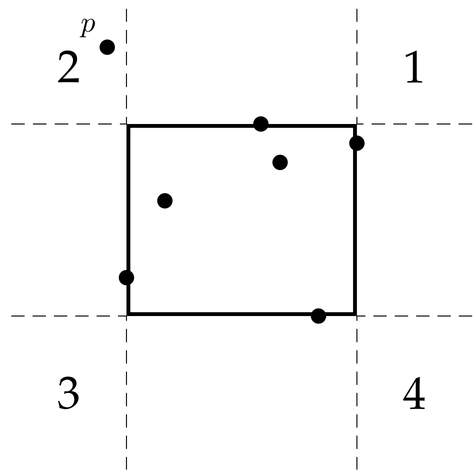

Figure 2.9: The point p lies in quadrant 2.

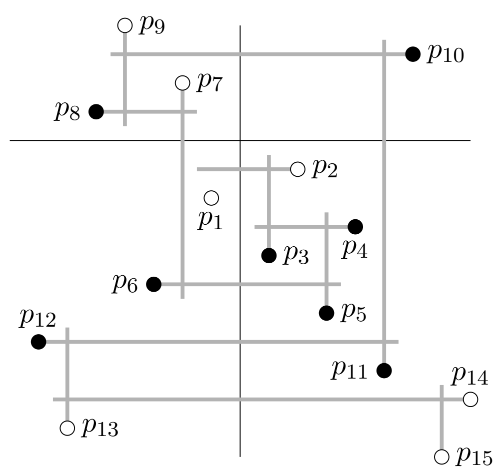

Figure 2.10: The proper pin sequence $p_1, \ldots, p_{15}$ shown corresponds to the strict pin word $w = 3RDRDLULURDLDRD$. The filled points correspond to the pin word $u = 4RDL21DL$, the permutation corresponding to this word, i.e., the permutation order isomorphic to the filled points, is 27453618.

(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 formalise 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 $0, p_1, \ldots, p_{i-1}$. 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 $p_1, \ldots, p_m$. See Figure 2.10 for an example. Conversely, we have the following result.

Lemma 2.15. *Every proper pin sequence corresponds to a 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. 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,

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

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 the *sequence of points corresponding to* w: $p_1, \ldots, p_m$. The *permutation corresponding to* w is again the permutation order isomorphic to this set of points. Again, Figure 2.10 gives an example of a nonstrict pin word.

We can now define an order, $\preceq$, on pin words. Let u and w be two pin words. 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 2.10, the division of u into strong numeral-led factors is $(4RDL)(2)(1DL)$, while w can be written $(3R)(DRDL)(U)(L)(U)(RDL)(DRD)$. 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 2.16. *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. ☐

2.5 Unavoidable Substructures in Simple Permutations

With the representation of pin sequences and their subsets in terms of pin words established, we may derive the promised unavoidable substructures result. 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 2.11; note that the sequence can be represented, for example, by the proper pin sequence $1RURU\cdots$.

We define an *increasing oscillation* to be any simple permutation that is contained in the increasing oscillating sequence, *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.

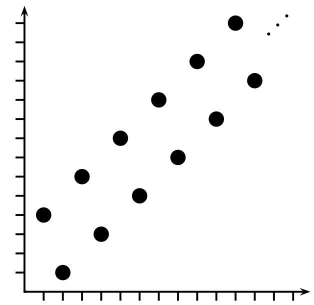

Figure 2.11: A plot of the increasing oscillating sequence.

Theorem 2.17. *Every sufficiently long simple permutation contains an alternation of length k or an oscillation of length k.*

Proof. By Theorem 2.14, 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 2.15, 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. $\square$

2.6 Other Contexts

Although our proof is highly permutation-centric, these is no reason why analogues of Theorem 2.2 cannot exist for other types of objects: we will shortly discuss the decomposition problem in the graph case. In the context of general relational structures, however, any analogue of Theorem 2.2 would need to allow for more intersection between the two simple substructures. For example, let $\mathcal{L}$ consist of a 2-ary relation $<$ and a k-ary relation R. Take $\mathcal{A}$ with $\mathrm{dom}(\mathcal{A}) = [2n]$ where $<$ is interpreted as the normal linear order on $[2n]$ and $R(1, 3, 5, \ldots, 2k-3, i)$ precisely for even $i \in [2k-2, 2n]$. This structure is simple, but all simple substructures (with at least two elements) of $\mathcal{A}$ must contain each of $1, 3, 5, \ldots, 2k-3$, and then to prevent these elements from containing a nontrivial interval, the simple substructure must also contain $2, 4, 6, \ldots, 2k-4$.

2.6.1 Pin Sequences in Graphs

Our approach for indecomposable graphs[3] follows the same principles as we used in the case of permutations. We want to define pin sequences and a set of "exceptional indecomposable graphs" (analogous to parallel and wedge simple alternations) in order to prove:

Conjecture 2.18. *Every sufficiently long indecomposable graph contains either a proper pin sequence of order k, or one of a finite number of families of exceptional indecomposable graphs with k vertices.*

We begin our discussion with some thoughts on pin sequences. Taking two vertices, p_1 and p_2 of an indecomposable graph G, $\{p_1, p_2\}$ cannot be an interval and so there must be a vertex p_3 which is adjacent to precisely one of p_1 or p_2, corresponding to a *pin*. Now since $\{p_1, p_2, p_3\}$ is not an interval, we may find a vertex p_4 adjacent to some but not all of p_1, p_2

[3]Recall that "simple" graphs are more usually called *indecomposable* graphs.

and p_3. We may continue in this manner to form a *pin sequence*, $p_1, \ldots, p_m$ for which p_i is adjacent to some but not all of $\{p_1, \ldots, p_{i-1}\}$. Any pin sequence within an indecomposable graph may be extended to form a *saturated* pin sequence, that is, one in which every vertex appears. Note that here our definition differs slightly from the permutation case; there we had defined saturated to mean that $\mathrm{rect}(p_1, \ldots, p_m)$ encloses all of our simple permutation π, while here we have no graphical representation where such an argument makes sense. As an immediate consequence of saturation, however, we may state our equivalent to "right reaching" pin sequences:

Lemma 2.19. *Given any three distinct vertices p_1, p_2 and w in an indecomposable graph G, there is a w-reaching pin sequence $p_1, p_2, \ldots, p_m = w$.*

It remains to define proper pin sequences for graphs. In the permutation case, we specified two conditions, namely separation and maximality (or externality in some viewpoints). Since maximality is essentially a feature arising from the pictorial representation of permutations, finding an equivalent for graphs is the first problem that arises. However, separation is easily converted into the *leaf condition*: for all $i \geq 3$, p_i is either a

- *Leaf:* p_i is adjacent to p_{i-1} and not to any of $p_1, \ldots, p_{i-2}$, or an
- *Antileaf:* p_i is adjacent to all of $p_1, \ldots, p_{i-2}$ and not to p_{i-1}.

It is worth noting that a similar construction called "reducing pseudopaths" can be found in the recent work of Zverovich [122]. Delaying the issue of maximality for the time being, we may proceed to derive results that look very similar to the permutation case. First, we have an analogue of Theorem 2.7:

Proposition 2.20. *If $p_1, \ldots, p_m$ is a proper pin sequence of length $m \geq 5$, then one of the sets of vertices $\{p_1, \ldots, p_m\}$, $\{p_1, \ldots, p_m\} \setminus \{p_1\}$ or $\{p_1, \ldots, p_m\} \setminus \{p_2\}$ induces an indecomposable graph.*[4]

[4]Note that we still require $m \geq 5$ as in the permutation case, as witnessed by the sequence $\{p_1, p_2, p_3, p_4\}$ with $p_1 \sim p_2$, p_3 a leaf and p_4 an antileaf, whence $\{p_1, p_4\}$ is an interval, but so is $\{p_1, p_3, p_4\}$.

We may also strengthen Lemma 2.19 in the desired way:

Lemma 2.21. *Given any three distinct vertices p_1, p_2 and w in an indecomposable graph G, there is a proper w-reaching pin sequence $p_1, p_2, \ldots, p_m = w$.*

Both proofs follow in the same (in fact, somewhat easier) way as the permutation versions (Theorem 2.7 and Lemma 2.9, respectively), noting that maximality may be removed without significant effect. However we now find that, without maximality, progress grinds to a halt. If an indecomposable graph contains a long proper pin sequence, then we can produce our sought-after substructure for Conjecture 2.18. On the other hand, if all the pin sequences are short, we must explore convergence of pin sequences and hence derive the set of "exceptional indecomposables", but it is in convergence that maximality plays its crucial rôle. We now present the current most promising definition of maximality, and approaches to the question of convergence.

Given an indecomposable graph G on n vertices, we may fix a labelling of $V(G)$ by $[n] = \{1, \ldots, n\}$. We are now concerned with a particular type of proper n-reaching pin sequence, starting from p_1, p_2: the pin sequence $p_1, p_2, \ldots, p_m = n$ is said to be a *proper quickly n-reaching pin sequence* if, for all $i \geq 3$, p_i has the greatest label of all vertices v such that $p_1, p_2, \ldots, p_{i-1}, v$ can be extended to a proper n-reaching pin sequence. We may now strengthen Lemma 2.19 yet further:

Lemma 2.22. *In an indecomposable graph on n vertices labelled by $[n]$, for any two vertices $p_1, p_2 \neq n$ there is a proper quickly n-reaching pin sequence $p_1, p_2, \ldots, p_m = n$.*

Two pin sequences $p_1, p_2, \ldots$ and $q_1, q_2, \ldots$ are said to *converge at the vertex x* if there exists i and j such that $p_i = q_j = x$, but $\{p_1, \ldots, p_{i-1}\}$ and $\{q_1, \ldots, q_{j-1}\}$ are disjoint. As we saw in the permutation case, however, convergence alone is not sufficient; we had to use initially-nonoverlapping pin sequences to see that those converging at their third pin still led to one of the exceptional simples. In the graph case, we may replace "initially-nonoverlapping" with *distinct third pins* – i.e. we must find pin sequences that do not converge until after their third pin. If this can be done, then together with the existing

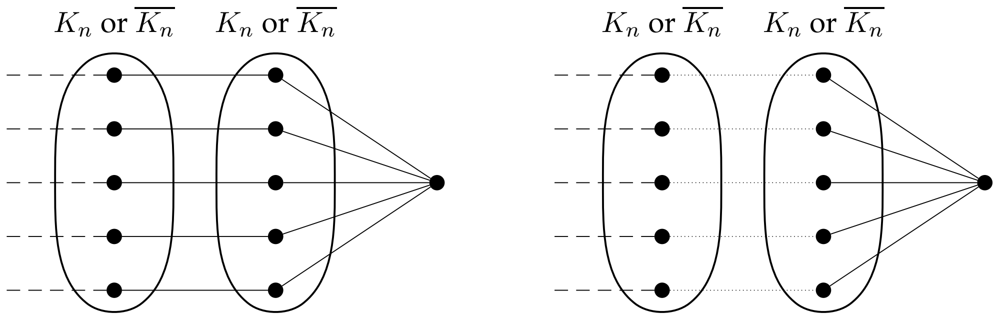

Figure 2.12: Forming exceptional indecomposable graphs from converging pin sequences.

maximality definition we should be able to rule out the type (iv) graphs we will encounter shortly in Figure 2.14, in which case Conjecture 2.18 would hopefully follow. Unfortunately, there remains the question of whether or not we can find sufficiently many pin sequences with distinct third pins:

Question 2.23. *In an indecomposable graph on n vertices, how many proper quickly n-reaching pin sequences with distinct third pins can be formed?*

The problem that distinct third pins is needed to solve is that convergence does not immediately lead us to exceptional indecomposables. In the permutation case we use the points in the pin sequences prior to convergence to construct an alternation, knowing by maximality that these sequences cannot "overlap". In the graph case, this ceases to be true, and even with our new notion of maximality we cannot rule out edges between vertices of different pin sequences. Thus either we need to adjust the definition of maximality, or introduce some further constraints on which pin sequences we select before any further progress can be made.

The Exceptional Indecomposables. Considering how the "well behaved" pin sequences converge, we may begin to describe the exceptional indecomposable graphs which contain only short pin sequences. Suppose a (large) set of pin sequences converges at the vertex x. By symmetry we may assume that for at least half of these sequences x is a leaf, so x

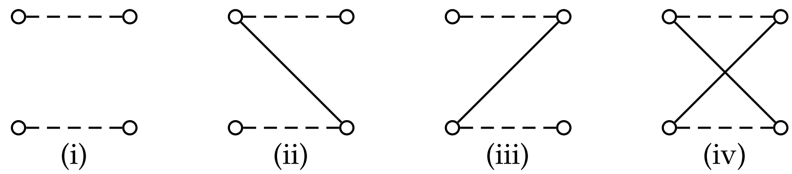

Figure 2.13: The four interactions between pin sequences.

is adjacent to the preceding pin of each of these pin sequences but to none of the earlier pins. Selecting these pin sequences, we now consider the set of immediate predecessor pins, each of which was either a leaf or an antileaf. We pick, using Ramsey's Theorem, the largest subset of these pins which forms a complete or independent subgraph, and which are all leaves or antileaves.

We now consider the pins occurring immediately before the predecessor pins in our chosen uniform subset. Again using Ramsey's Theorem, we may find a uniform subset of these vertices, and again we restrict our attention to the pin sequences corresponding to these vertices. Momentarily ignoring edge interactions between pin sequences at the predecessor and pre-predecessor levels, we now have one of the situations depicted in Figure 2.12.

We now consider the possible interactions between each pair of pin sequences, again with an aim to choosing a uniform subset. Listing these sequences in some order (in Figure 2.12 we view the order as going from top to bottom), there are essentially four different interactions between two pin sequences, types (i) — (iv) as shown in Figure 2.13.

A Ramsey-type argument may now be used to obtain a subset of these pin sequences whose pairwise interactions are uniform. The resulting graph needs to be either indecomposable or nearly so – as in the permutation case, we allow the removal of one or two points. In some cases the graph is immediately indecomposable (for example, the "double star" in Figure 2.14), while in others the removal of one or two points is sufficient (the "down and to the right" graph in Figure 2.14, the filled nodes form an indecomposable graph). However, in certain cases no exceptional indecomposable seems to be obtainable, and these structures are the ones that need to be ruled out by an appropriate definition

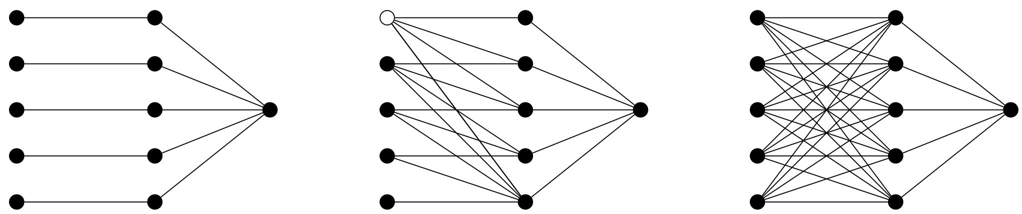

Figure 2.14: From left to right, the "double star", a "down and to the right" graph and a type (iv) highly decomposable graph.

of maximality (the type (iv) interaction graph in Figure 2.14). Note that, if we have also taken pin sequences with distinct third pins, we could, instead of looking at the penultimate pins before convergence, look at the antepenultimate pins and there perhaps rule out the existence of a large number of type (iv) interactions.

CHAPTER 3

SIMPLE EXTENSIONS

3.1 Introduction

OUR AIM in this chapter is to establish how we may embed any given $\mathcal{L}$-structure $\mathcal{A}$ into a simple $\mathcal{L}$-structure $\mathcal{B}$ containing as few extra elements as possible. Formally, we say that $\mathcal{B}$ is a *simple extension* of $\mathcal{A}$ if $\mathcal{B}$ is simple, $|\operatorname{dom}(\mathcal{A})| < |\operatorname{dom}(\mathcal{B})|$ and $\mathcal{B}|_{\operatorname{dom}(\mathcal{A})} = \mathcal{A}$. Our aim then is to minimise $|\operatorname{dom}(\mathcal{B})\backslash\operatorname{dom}(\mathcal{A})|$, writing it as a function of $n = |\operatorname{dom}(\mathcal{A})|$.

This work is partly motivated by the result for tournaments dating back to 1972, when Erdős, Fried, Hajnal and Milner [51] showed that every tournament may be extended to a simple tournament requiring at most two extra vertices (we will review this result in Section 3.4). Clearly, however, it will not be sufficient to consider just the two-point extensions for every relational structure. Nor do we need to look far to find an example: there is clearly no two-point simple extension of an arbitrary complete graph K_n. The permutation case is different again, while posets fall somewhere between the two. Thus asking for a solution for an arbitrary relational structure is somewhat meaningless – as we will see, even the well-known binary relational structures demonstrate a wide variety of results.

We may, however, follow a general approach by recalling the substitution decomposition (Theorem 1.5 on Page 16) of $\mathcal{A}$, and using induction. When the skeleton $\mathcal{S}$ of $\mathcal{A}$ defines a unique deflation $\mathcal{A} = \mathcal{S}[\mathcal{A}_s : s \in \operatorname{dom}(\mathcal{S})]$ into maximal intervals (i.e. when $|\operatorname{dom}(\mathcal{S})| \geq 3$), we can embed $\mathcal{A}$ into $\mathcal{B}$ inductively by embedding each $\mathcal{A}_s$ into $\mathcal{B}$ in a prescribed way. The degenerate and linear cases must in general be dealt with more carefully, although induction can still be used to produce the required result.

3.2 Permutations

We begin our study with the permutation case. Recall that, when viewing permutations graphically, an interval of a permutation π can be seen as a set of points enclosed by an axis-parallel rectangle with no other points above, below, to the left or to the right. To embed a given π in a simple permutation, therefore, we must ensure that every axis-parallel rectangle containing at least two points of π may be extended by a pin from the simple extension.

Lemma 3.1. *An increasing permutation of size n has a simple extension with $\lceil \frac{n+1}{2} \rceil$ additional points.*

Proof. For $n = 2$ the increasing permutation 12 is embeddable in the simple permutation 2413, so now suppose $n \geq 3$. Let $\pi = 12 \cdots n$. For $n = 2k$, we claim the permutation

$$k+1, 1, k+3, 2, \ldots, 3k-1, k, 3k+1, k+2, k+4, \ldots, 3k$$

is simple and contains $12 \cdots n$. For $n = 2k+1$, we claim

$$k+2, 1, k+4, 2, \ldots, 3k+2, k+1, k+3, k+5, \ldots, 3k+1$$

is simple. That both of these permutations are simple follows easily by checking Figure 3.1. $\square$

Note also that $m = \lceil \frac{n+1}{2} \rceil$ is the best possible bound. Every adjacent pair $i, i+1$ must be "separated" either horizontally or vertically by one of the additional points, and the points $\pi(1) = 1$ and $\pi(n) = n$ of π must not lie in the "corners" of the simple extension — a total of $n+1$ gaps to be filled. The bound on the number of additional points is then obtained by observing that each can fill at most two gaps (one horizontally, one vertically).

By symmetry, decreasing permutations may be extended in the same way:

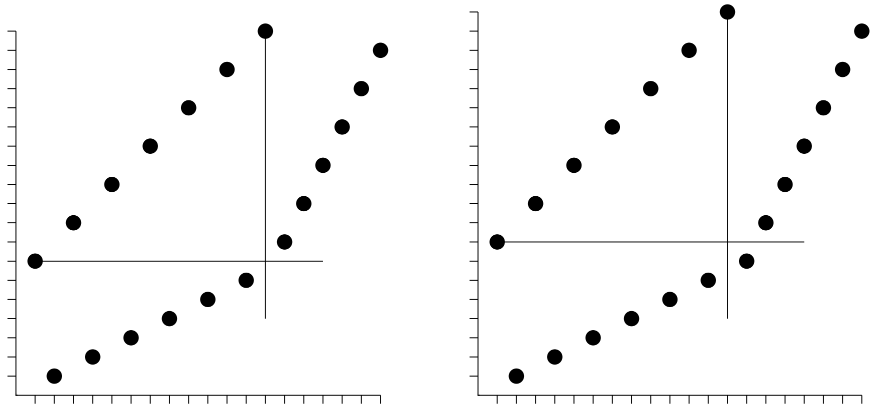

Figure 3.1: Simple permutations containing $12\cdots n$, for $n = 12$ (left) and $n = 13$ (right).

Lemma 3.2. *A decreasing permutation of size n has a simple extension with $\lceil \frac{n+1}{2} \rceil$ additional points.*

We are now ready to state the result in the general case.

Theorem 3.3. *Every permutation π on n symbols has a simple extension with at most $\lceil \frac{n+1}{2} \rceil$ additional points.*

Proof. We proceed by induction on $n \geq 2$, claiming that for each permutation π of length n we may construct two extensions, $\pi^{(M)}$ and $\pi^{(m)}$, satisfying:

- Viewed as extensions, both $\pi^{(M)}$ and $\pi^{(m)}$ have a new leftmost point which is neither a new maximum nor a new minimum, called the *entry point*.
- Both $\pi^{(M)}$ and $\pi^{(m)}$ have a new *exit point*; for $\pi^{(M)}$ this is a new maximum while for $\pi^{(m)}$ this is a new minimum, and in both cases it is neither a new leftmost point nor a new rightmost point.
- The only minimal non-singleton intervals of $\pi^{(M)}$ and $\pi^{(m)}$ contain the new exit point.
- At least one of $\pi^{(M)}$ and $\pi^{(m)}$ is simple.

In the base case $n = 2$, either $\pi = 12$ or $\pi = 21$. When $\pi = 12$, $\pi^{(M)} = 2413$ is simple, and the only minimal non-singleton interval of $\pi^{(m)} = 3124$ is 12, which contains the exit point. The case $\pi = 21$ is dealt with by symmetry.

So now suppose $n \geq 3$. If π is an increasing (respectively, decreasing) permutation, then Lemma 3.1 (resp., Lemma 3.2) proves the existence of a simple extension of the required size. Note further that the simple extension satisfies the requirements to act as $\pi^{(M)}$ (resp., $\pi^{(m)}$), using symmetry if required. When π is an increasing permutation, we obtain $\pi^{(m)}$ from $\pi^{(M)}$ by changing the new maximum for a new minimum using the mapping

$$\pi^{(m)}(i) = \pi^{(M)}(i) + 1 \mod |\pi^{(M)}|.$$

For decreasing permutations, $\pi^{(M)}$ is created similarly.

We may therefore assume that π is neither an increasing nor a decreasing permutation. Write π as the substitution decomposition, $\pi = \sigma[\pi_1, \pi_2, \ldots, \pi_m]$ where the simple skeleton, σ, is of length $m \geq 2$, and $\pi_1, \pi_2, \ldots, \pi_m$ are permutations of size $|\pi_i| = p_i$ for each i. First suppose $m > 2$ so that the substitution decomposition is unique. If $p_i = 1$ for all i, then $\pi = \sigma$ is already simple. We construct $\pi^{(M)}$ and $\pi^{(m)}$ by adding precisely two points. The first is a new leftmost point, which may be inserted vertically anywhere except as a new maximum, minimum, or adjacent to $\pi(1)$. The new maximum or minimum is inserted similarly, preserving simplicity.

So now suppose that at least one π_i contains at least two points. For every such π_i, the inductive hypothesis allows us to extend to either $\pi_i^{(M)}$ or $\pi_i^{(m)}$ by adding at most $\lceil \frac{p_i+1}{2} \rceil$ points. Our choice between $\pi_i^{(M)}$ or $\pi_i^{(m)}$ is made according to the location of the next leftmost non-singleton block, π_j say (i.e. $j > i$ and no k with $j > k > i$ and π_k non-singleton); if $\sigma(j) > \sigma(i)$, then we choose $\pi_i^{(M)}$, while if $\sigma(j) < \sigma(i)$, we choose $\pi_i^{(m)}$. In either case, the exit point of $\pi_i^{(M)}$ or $\pi_i^{(m)}$ is simultaneously used as the entry point for the extension of π_j to $\pi_j^{(M)}$ or $\pi_j^{(m)}$. In this way, we work left-to-right through σ connecting the non-singleton blocks π_i (see Figure 3.2). For the rightmost such block π_r, we use $\pi_r^{(M)}$ to form $\pi^{(M)}$, and $\pi_r^{(m)}$ to form $\pi^{(m)}$, the exit point being used as the new maximum for $\pi^{(M)}$ or the new minimum for $\pi^{(m)}$, respectively.

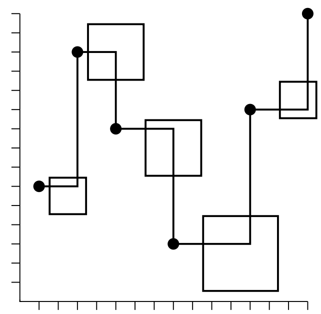

Figure 3.2: Connecting entry and exit points in the substitution decomposition.

These extensions will fail to be simple if the rightmost non-singleton block π_r is also the maximal or minimal block by value in the cases $\pi^{(M)}$ and $\pi^{(m)}$ respectively, and only then if $\pi_r^{(M)}$ or $\pi_r^{(m)}$ was not simple. Since π_r can only satisfy at most one of these, we may turn to the other for our simple extension. By symmetry, therefore, let us suppose that the rightmost non-singleton block π_r was not maximal in value.

Letting I be a non-singleton interval of $\pi^{(M)}$, first consider the case where I contains points from two distinct original (non-extended) blocks π_i and π_j. In this case the original simple skeleton σ of π forces us to include every such block, and subsequently all the extended points too. If on the other hand I contains two points in some extended block $\pi_i^{(M)}$ or $\pi_i^{(m)}$, then it must contain the exit point of that block and a point of the original π_i (else $\pi_i^{(M)}$ or $\pi_i^{(m)}$ did not satisfy the minimal proper interval property). Unless π_i was the rightmost non-singleton block, this exit point acts as the entry point of the extension of some other block π_j, which then requires us to include at least one other point of this extended block, and hence a point of the original block π_j, returning us to the previous case. Finally, if π_i was in fact the rightmost non-singleton block, then it was not the maximal block by value, and so the exit point of $\pi_i^{(M)}$ forces us to include the entirety of some other π_j block (note that such a π_j can be a singleton), again reducing to our first consideration.

In the case where $m = 2$ the substitution decomposition is not unique. Without loss we may assume that $\sigma = 12$, and so we may write $\pi = 12[\pi_1, \pi_2]$, where π_1 and π_2 may be chosen in a number of different ways. We begin by choosing π_1 to be as large as possible.

Unless π_2 is now a singleton, we will use this decomposition and proceed by extending π_1 to $\pi_1^{(M)}$ and π_2 to $\pi_2^{(M)}$ or $\pi_2^{(m)}$, and connecting the exit point of the first to the entry point of the second. If π_2 is a singleton and π_1 is not sum decomposable, we continue as above but with the exit point of $\pi_1^{(M)}$ placed above π_2. When π_1 is itself decomposable as $\pi_1 = 12[\pi_1', 1]$, we look at the decomposition $\pi = 12[\pi_1', 12]$. If again $\pi_1' = 12[\pi_1'', 1]$, then we repeat, so that $\pi = 12[\pi'', 123]$. Repeat this process, noting that it must terminate before we reach the end of π, as otherwise π is increasing, and at termination proceed as before. Simplicity follows in a similar manner to the unique decomposition case.

The number of points added in every one of the above cases is at most $\sum_{i=1}^{m} \left\lceil \frac{p_i+1}{2} \right\rceil - (m-1) \leq \lceil \frac{n+1}{2} \rceil$, noting that $\sum_{i=1}^{m} p_i = n$. □

3.3 Graphs

Recall that, in a graph G, an interval is a set of vertices $X \subseteq V(G)$ such that $N(v) \setminus X = N(w) \setminus X$ for every $v, w \in X$, and instead of "simple" we use the word *indecomposable* to describe a graph containing no proper intervals. We begin by specialising the Substitution Decomposition Theorem 1.5 for the context of graphs:

Proposition 3.4. *Let G be any graph. Then one of the following holds:*

(1) $G = H[J_v : v \in V(H)]$ where H is the simple skeleton of G, and this decomposition is unique.

(2) G is disconnected and can be written possibly non-uniquely as $G = \overline{K_2}[J_1, J_2]$.

(3) $\overline{G}$ is disconnected, and G can be written possibly non-uniquely as $G = K_2[J_1, J_2]$.

Our approach now follows the same pattern as the permutation case. We first consider simple extensions of the complete graph K_n, which is once more the "worst case" scenario. This result first appeared in Sumner's Ph.D. Thesis [115].

Lemma 3.5 (Sumner [115, Theorem 2.45])**.** *K_n has a simple extension with $\lceil \log_2(n+1) \rceil$ additional vertices.*

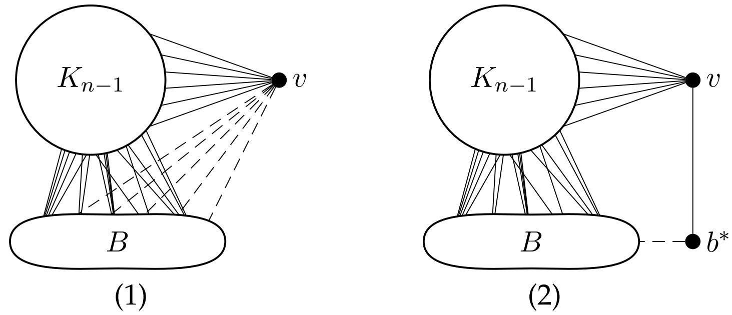

Figure 3.3: The two cases of Lemma 3.5.

Proof. We proceed by induction on n. The case $n = 1$ is trivial. For $n = 2$, we must add two new vertices. Regardless of whether the subgraph formed by the new vertices is connected or not, there is a way to add edges between the new and old vertices to form a graph isomorphic to P_4, the path of length four.

Now suppose $G \cong K_n$ for $n > 2$. There are two cases (these discussions are accompanied by Figure 3.3):

(1) $\lceil \log_2(n+1) \rceil = \lceil \log_2 n \rceil$. Choose a vertex $v \in V(G)$, and use induction to add a set of vertices B with edges to $G - v$ so that $(G - v) \cup B$ is simple. The remaining vertex v can be assigned a neighbourhood in B different to the neighbourhood of every other vertex in $G - v$, and so that $N(v) \cap B \neq B$. Since v has a different neighbourhood to every other vertex, it cannot lie in an interval with any other vertices, and so the graph is simple.

(2) $\lceil \log_2(n+1) \rceil = \lceil \log_2 n \rceil + 1$. Choose a vertex $v \in V(G)$, and use induction to add a set of vertices B with edges to $G - v$ so that $(G - v) \cup B$ is simple. For the remaining vertex v, we add a new vertex b^* and connect it to v.

Since $(G-v) \cup B$ is simple, any proper interval in the extended graph $G \cup B \cup \{b^*\}$ will need to involve either v or b^* (or both). We claim that any interval I in the extended graph of size ≥ 2 containing v also contains b^*. If I contains a vertex $x \in G - v$, then $b^* \notin N(x)$, so $b^* \in I$. The other case is where I contains a vertex $b \in B$, and then there is some $x \in G - v$ not connected to b, so $x \in I$, reducing to the previous case.

Now suppose we have an interval $I \supseteq \{x, b^*\}$ for $x \in (G-v) \cup B$. Since the only vertex in G connected to b^* is v, and x is connected to at least one other vertex $y \in G - v$, we have $y \in I$, and $x, y \in I$ implies $(G - v) \cup B \subset I$. The vertex v, if not already in I, must be included as $N(v) \cap B = \emptyset$.

□

Note that the above proof does not specify the internal edges of B, nor edges between any vertex in B and b^*, and so we may use any graph of size $\lceil \log_2(n+1) \rceil$ that we choose. Furthermore, by taking the complement, this immediately implies the following:

Lemma 3.6. *$\overline{K_n}$ has a simple extension with $\lceil \log_2(n+1) \rceil$ additional vertices.*

The bound $m = \lceil \log_2(n+1) \rceil$ is also the smallest possible, for if we were to add a set B of m vertices, with $n > 2^m - 1$, then either two vertices in G have the same neighbourhood in $G \cup B$, or one vertex of G is connected to every other vertex in $G \cup B$, both of which give an interval.

Theorem 3.7. *Every graph G has a simple extension with at most $m = \lceil \log_2(|V(G)| + 1) \rceil$ additional vertices.*

Proof. We proceed by induction on $n = |V(G)|$. The base cases $n = 1$ and $n = 2$ are covered by Lemmas 3.5 and 3.6, so now suppose $n \geq 3$. Write $G = H[J_v : v \in V(H)]$ where H is the simple skeleton of G. There are two cases when $|V(H)| = 2$; we will assume without loss in this case that $H = \overline{K_2}$, i.e. that G is disconnected. Further, we will choose the J_vs so that at least one of them is connected and has at least two vertices (having established the result for independent sets in Lemma 3.6).

If $H = G$ then the graph is already simple, but for the induction to work we must be able to extend to a larger simple graph. This we do by adding a single vertex, noting that the only intervals that need to be avoided in this case are either all of the old graph or intervals of size two involving the new vertex. The new vertex cannot therefore be adjacent to all or none of the old vertices, and it must also not have the same neighbourhood as any other vertex, but any other set of adjacencies is permitted (giving $2^n - 2n - 2$ possible one-point simple extensions).

Now assume that at least one interval J_v is non-trivial. Suppose first that $|V(H)| \geq 4$ so the substitution decomposition is unique. For each J_v we may add a set of vertices B_v which are connected to vertices in J_v so that $J_v \cup B_v$ is simple by induction. Fix an $x \in H$ for which B_x is of maximal size (note that $2 \leq |V(B_x)| \leq m$). For every other interval J_v, identify B_v with any subset of B_x, unless $|V(J_v)| = 1$, in which case we set $B_v = \emptyset$. Then we specify the edges between J_v and $B_x \setminus B_v$ such that:

(G1) Every pair of vertices $a \in J_v$ and $b \in B_x \setminus B_v$ disagree on at least one vertex of $J_x \cup J_v \cup B_x \setminus \{a, b\}$.

First consider the case where J_v is not a singleton. If there is a vertex in J_v that is adjacent to every other vertex in J_v, then we can satisfy (G1) by adding none of the edges between J_v and $B_x \setminus B_v$. Otherwise we can satisfy (G1) by adding all of the edges between J_v and $B_x \setminus B_v$.

If $J_v = \{a\}$ is a singleton, let us suppose $v \not\sim x$ in H by symmetry. Here we achieve (G1) by connecting a to no vertex of B_x; if $b \in B_x$ is connected to at least one vertex of J_x then a and b disagree on J_x, while if $b \in B_x$ is connected to no vertex of J_x then, to prevent $J_x \cup B_x \setminus \{b\}$ from being an interval of $J_x \cup B_x$, there must be a vertex of B_x to which b is adjacent and on which a and b will disagree.

We claim the resulting graph is simple. Consider an interval I with at least two vertices a and b. There are four cases:

- $a, b \in J_x \cup B_x$: simplicity implies that $J_x \cup B_x \subseteq I$. Then for any J_v such that $|V(J_v)| \geq 2$, there are at least two vertices of B_v in the interval, which forces $J_v \cup B_v \subseteq I$. When $|V(J_v)| = 1$, by (G1) the single vertex is adjacent to some but not all of $J_x \cup B_x$ and so must be included in I.

- $a, b \in J_v \cup B_v$, $v \neq x$: by the construction $|V(J_v)| \geq 2$, and by simplicity $J_v \cup B_v \subseteq$

I. There are now two vertices in I from $B_v \subseteq B_x$, a case which has already been considered.

- $a \in J_u$ and $b \in J_v$, $u \neq v$: first, if $|V(H)| \geq 4$ then the simplicity of H implies that $V(G) \subseteq I$, and in particular $J_x \subseteq I$, reducing to the first case above. Thus we have $H = \overline{K_2}$ and (say) J_u connected with at least two vertices, by our assumptions at the beginning of the proof. We then get that a has a neighbour in J_u while b does not, leading to the case above.

- $a \in J_v$, $v \neq x$, and $b \in B_x \setminus B_v$: by (G1) there must be at least one more vertex in I, and thus one of the other cases applies.

□

Although we know this bound is necessarily tight for complete or independent graphs, there does remain the question of whether or not we can do any better for an arbitrary graph G on n vertices, i.e. is there a smaller simple extension? Letting $\omega(G)$ denote the size of the largest clique (complete subgraph) of G, and $\alpha(G)$ the size of the largest independent set of G, we pose (without further discussion here) the following conjecture:

Conjecture 3.8. *Every graph G has a simple extension with at most $\lceil \log_2(m+1) \rceil$ additional vertices, where $m = \max[\omega(G), \alpha(G)]$ is the size of the largest clique or independent set in G.*

3.4 Tournaments

Recall that a tournament is a complete oriented graph, and an interval of a tournament T is a set $A \subseteq V(T)$ such that for all $v \notin A$, either $v \to A$ or $v \leftarrow A$. Given a tournament, we may define an abstract algebra (for a formal definition of abstract algebras, see Subsection 5.3.1) with two idempotent binary operations $A_T = \langle T, \vee, \wedge \rangle$, so that if $x \to y$, then $x \vee y = y \vee x = x$ and $x \wedge y = y \wedge x = y$. A tournament is simple if and only if its corresponding abstract algebra is also simple, i.e. the kernel of every homomorphism of an abstract algebra is either the whole structure or a single element. Simple extensions in tournaments have thus received some attention, and in particular it is known that at most two vertices are required in every case, while one vertex is sufficient in all but a certain family of cases.

Theorem 3.9 (Erdős, Fried, Hajnal and Milner [51])**.** *Every tournament has a simple extension with at most* 2 *additional vertices.*

Proposition 3.10 (Erdős, Hajnal and Milner [52])**.** *A tournament* T *has a one-vertex simple extension unless* $|T| = 3$ *or it has an odd number of vertices and is transitive.*

Note that these results hold for tournaments of arbitrary cardinality, though they had previously been proved for finite tournaments by Moon [96]. We give here a proof of the finite case using the substitution decomposition. Observe that the non-unique decompositions correspond precisely to transitive tournaments, i.e. tournaments for which $x \to y$ and $y \to z$ implies $x \to z$.

Proof. First observe that there are no simple tournaments on 4 vertices, and so a simple extension of a tournament on 3 vertices requires at least two vertices. There are, up to isomorphism, only two 3-vertex tournaments, and checking each case in turn shows that two vertices is sufficient.

Now suppose T is a finite transitive tournament, so we may label the vertices of T as $1, 2, \ldots$ so that $i \to j$ if and only if $i < j$. We add a single vertex x to the tournament satisfying $x \to i$ if i is odd and $i \to x$ if i is even. Unless T has an odd number of vertices, it is straightforward to check that the resulting tournament is simple. In the case where $|T| = 2n + 1$, we observe that the set of vertices with labels $\{1, 2, \ldots, 2n, x\}$ is an interval, as they all look at the vertex labelled $2n + 1$ in the same way. If alternatively we added a vertex y satisfying $y \to i$ if i is even and $i \to y$ if i is odd, then we find that the set $\{2, 3, \ldots, 2n + 1, y\}$ forms an interval. Note that for any other single vertex extension, z say, there must exist a label i for which $z \to i$ and $z \to i + 1$ or $i \to z$ and $i + 1 \to z$, and in either case $\{i, i + 1\}$ is an interval. Thus T has no single vertex simple extension. A 2-vertex simple extension is easily formed by, say, adjoining both the vertices x and y, as in Figure 3.4.

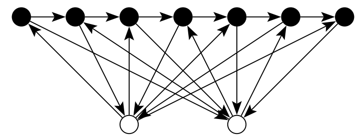

Figure 3.4: A 2-vertex simple extension of a transitive tournament on 7 vertices.

Having covered the transitive and 3-vertex cases, we claim that any other finite tournament T may be extended by a single vertex x to form a simple tournament. The substitution decomposition allows us to write $T = S[A_s : s \in S]$, where the skeleton S is either simple or transitive.

Where S is simple, if every A_s contains just one vertex then $T = S$. Unless $|T| = 3$ (a case that has already been covered), the addition of x will preserve simplicity providing it does not have the same connections as any existing vertex of T. (Note that if $|T| = n$, there are $2^n - n$ different ways of choosing x.) Where there is at least one non-singleton block A_s, we still attach x to every singleton block as before, ensuring x does not end up with the same adjacency as any of them.

This leaves just the non-singleton blocks, which we attach to x as follows. Any such A_s which is neither transitive of odd degree nor satisfies $|A_s| = 3$ may, by induction, be connected to x so that $A_s \cup \{x\}$ is simple. If, however, A_s is transitive and $|A_s| = 2n + 1$, then, labelling the vertices of A_s with $1, 2, \ldots, 2n + 1$ as before, set $x \to i$ if i is odd and $i \to x$ for i even. This makes the set $\{1, 2, \ldots, 2n, x\}$ a candidate to be an interval, but we may check that either (1) there is another non-singleton block $A_{s'}$ satisfying $A_s \to A_{s'}$ or $A_{s'} \to A_s$, but x looks at elements of $A_{s'}$ differently, or (2) all the other blocks of the substitution decomposition are singletons, but since x is already attached to all such blocks preserving simplicity there is a singleton block on which x and A_s disagree. A similar argument applies to the case where $|A_s| = 3$. The simplicity of the skeleton S now ensures this one-point extension is simple.

If the skeleton is transitive then we may take S maximally so that each A_s is uniquely defined. Moreover, at least one such A_s is not a singleton (as T is not transitive), and no non-singleton block can be transitive. The vertices of S may be labelled $1, 2, \ldots$ as before, but let us further identify the unique vertex s^* of S for which $s \to s^*$ for all $s \in S \setminus \{s^*\}$. We attach x to every non-singleton block in any way so that:

- If A_1 is a singleton, then $x \to A_1$.
- If A_{s^*} is a singleton, then $A_{s^*} \to x$.
- The vertex x looks at every pair A_i and A_{i+1} of adjacent singleton blocks differently.

This leaves the non-singleton blocks, which, by induction, are attached to x so that the resulting extension of each such block is simple. It is then easily checked that the resulting one-point extension of T is simple. □

3.5 Posets

Posets again give a different result, arising from the non-unique cases of the substitution decomposition – we encounter a "mix" of the results in the non-unique cases of permutations and graphs. For the former, recall that these cases correspond to the increasing and decreasing permutations, which (viewing them as relational structures) occur when the two linear orders agree – i.e. they correspond to a single linear order. For the latter, the non-uniqueness comes in the form of complete and independent graphs, arising from complete or empty edge sets – these are degenerate cases. Posets can be decomposed non-uniquely either through linearity or through degeneracy, and the simple extension in each case is significantly different.

We begin with the case where a poset $(P, <)$ is a linear order. This case is essentially identical to the increasing permutation case of Lemma 3.1. Indeed, there is a mapping between permutations and posets: letting π be a permutation on $[n]$, we may form the poset $(P, \prec)$ where $P = [n]$, and $i \prec j$ if and only if both $i < j$ and $\pi(i) < \pi(j)$. While poset intervals do not always correspond to permutation intervals, simple permutations do map to simple posets:

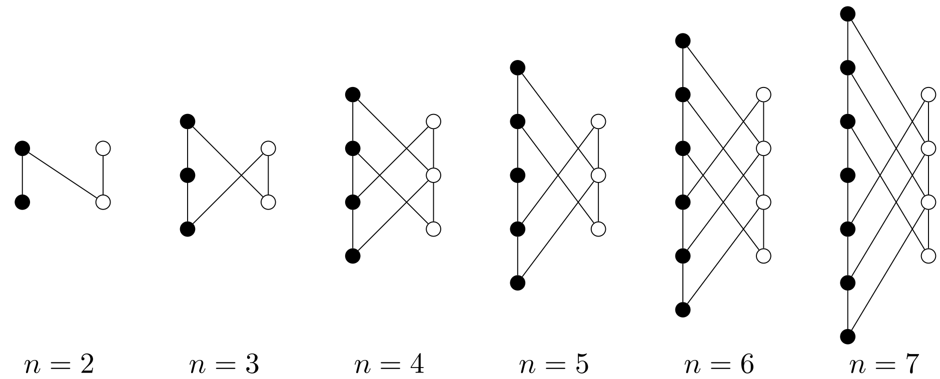

Figure 3.5: Simple extensions of short linear orders.

Lemma 3.11. *A permutation is simple if and only if its corresponding poset is simple.*

Proof. Suppose first that π is a simple permutation, that $(P, \prec)$ is its corresponding poset and that A is an interval in the poset. The corresponding set of points A_π of π cannot form an interval, so there must exist some point $(i, \pi(i))$ of π not in A_π which separates the points in $\mathrm{rect}(A_\pi)$ either horizontally or vertically. However, the element i of the poset corresponding to $(i, \pi(i))$ must then disagree on the elements of A, a contradiction since A was an interval.

Conversely, suppose $(P, \prec)$ is a simple poset corresponding to the permutation π, but that π contains some proper interval I. The set of elements I_P of P corresponding to I cannot form an interval, so there exists some element $p \in P \setminus I_P$ for which p is not related to every element of I_P in the same way. However, the point $(p, \pi(p))$ of π (which corresponds to $p \in P$) must then separate some points of I, a contradiction since I was an interval. $\square$

Observe that, although this mapping is not injective, increasing permutations map uniquely to linear orders, and thus:

Lemma 3.12. *A linear order $(P, <)$ on n elements has a simple extension containing at most $m = \left\lceil \frac{n+1}{2} \right\rceil$ additional elements.*

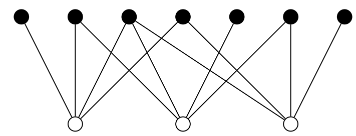

Figure 3.6: A 3-element simple extension of a 7-element antichain.

Proof. The linear order $(P, <)$ corresponds to an increasing permutation. By Lemma 3.1, an increasing permutation on n points has a simple extension with at most $m = \left\lceil \frac{n+1}{2} \right\rceil$ additional points. By Lemma 3.11 the corresponding poset is also simple, completing the proof. □

See Figure 3.5 for examples of the first few cases of this construction. Note that, as in Lemma 3.1, the case $n = 2$ must be handled separately, the resulting simple poset corresponding exactly to the permutation 2413.

Meanwhile, the degenerate case is an antichain, i.e. a poset containing no non-trivial relations. Recalling that every poset has a corresponding comparability graph, we may proceed in the same way as the graph case.

Lemma 3.13. *An n-element antichain has a simple extension requiring at most $\lceil \log_2(n+1) \rceil$ additional elements.*

Proof. The comparability graph of the poset $(P, <)$ is the independent graph $\overline{K_n}$, which, by Lemma 3.6, has a simple extension with $\lceil \log_2(n+1) \rceil$ additional vertices. Furthermore, the edges between these additional vertices are unspecified, so we may choose any set of edges that is transitively orientable. The extension for the graph was indecomposable, so by Lemma 1.1 (on Page 11) the corresponding poset will be simple. □

For example, Figure 3.6 shows a three-element simple extension of an antichain with seven elements, where the additional elements were taken to be incomparable. By the result for graphs, it follows that this is the best possible bound. Note also that the linear case of Lemma 3.12 is not easily solved by considering the corresponding comparability graph (equal to K_n) since any extension of the graph would need to be transitively orientable. Of course, the bound in Lemma 3.12 is also the best possible by its connection to the permutation case.

We now consider simple extensions of an arbitrary poset. Our approach takes much the same form as the permutation case, inductively "connecting" entry and exit points from the simple extensions of the intervals in the substitution decomposition.

Theorem 3.14. *A poset $(P,<)$ on n elements has a simple extension with at most $m = \left\lceil \frac{n+1}{2} \right\rceil$ additional elements.*

Sketch of proof. We proceed by induction on n, using the substitution decomposition. Our claim is that we may form three extensions $P^{(mm)}$, $P^{(MM)}$ and $P^{(Mm)}$ of a poset $(P,<)$, satisfying:

- Each of the three extensions has two new *distinguished elements*. For $P^{(mm)}$ these are both new minima, for $P^{(MM)}$ new maxima, and for $P^{(Mm)}$ there is one maximum and one minimum.
- The only minimal non-singleton intervals of $P^{(mm)}$, $P^{(Mm)}$ and $P^{(MM)}$ contain one of the distinguished elements.
- At least one of $P^{(mm)}$, $P^{(Mm)}$ and $P^{(MM)}$ is simple.

The base case is $n = 2$, in which case the poset is either linear or an antichain. Simple extensions have already been exhibited in Lemmas 3.12 and 3.13, and the extensions $P^{(mm)}$, $P^{(Mm)}$ and $P^{(MM)}$ are easily formed in each case.

So now suppose $n > 2$ and, by the Substitution Decomposition Theorem 1.5, our poset may be expressed as a deflation $P = S[A_s : s \in S]$ where $(S,<)$ is simple, linear or an antichain. When S is simple, we proceed in essentially the same way as the permutation case. If every A_s is a singleton, then $(P,<)$ is already simple, but for the purposes of the induction we can add two elements to form $P^{(mm)}$ and $P^{(MM)}$ in any way we choose, noting that any minimal non-singleton interval will necessarily involve at least one of the

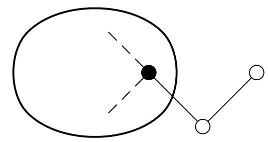

Figure 3.7: A 2-element simple extension of an arbitrary simple poset.

distinguished elements. Meanwhile, we may ensure that $P^{(Mm)}$ is a simple extension by adjoining two elements to any chosen element of P in the way shown in Figure 3.7.

When at least one A_s has more than one element, induction may be used on each such interval to form the three extensions $A_s^{(mm)}$, $A_s^{(Mm)}$ and $A_s^{(MM)}$. We choose the appropriate extension according to the following set of rules. Fix an order on the elements s of S for which the corresponding block A_s is not a singleton, labelling them as $1, 2, \ldots, k$. For $1 \leq i < k$, we pick the distinguished elements of the extension of A_i as follows:

- One of the distinguished elements is predetermined (for $i > 1$) by the extension of A_{i-1}. When $i = 1$, the distinguished element will act as one of the distinguished elements in the extension of P, and so must be chosen accordingly.
- If $A_i > A_{i+1}$, create a distinguished element that is both a new minimum for A_i and a new maximum for A_{i+1}.
- If $A_i < A_{i+1}$, create a distinguished element that is both a new maximum for A_i and a new minimum for A_{i+1}.
- If A_i and A_{i+1} are incomparable, create a distinguished element that is either a new maximum or a new minimum for both A_i and A_{i+1}.

The final distinguished element of A_k forms the other distinguished element in the extension of P, and so must be chosen accordingly. An argument similar to the permutation case proves that one of the extensions $P^{(mm)}$, $P^{(Mm)}$ or $P^{(MM)}$ is simple and of the required size. In the non-unique cases, pick S maximally so that S deflates P uniquely, and proceed as above. $\square$

In the way that simple extensions of posets seem to lie somewhere between the solution for permutations and graphs, we may be tempted to pose a result similar to Conjecture 3.8. Certainly the above bound can be improved when the skeleton turns out to be a linear order or an antichain by connecting more than two distinguished points together at a time, as dictated by Lemmas 3.12 and 3.13. Precisely *how* this improves the bound, though, is not clear. Even when the skeleton is not degenerate there are times when several distinguished points can be combined, but the rules for this seem difficult to establish. All we can do at this stage is to ask the following question:

Question 3.15. *How is the size of a minimal simple extension of a poset affected by the length of the largest chain or antichain in the poset?*

Chapter 4

Substitution Decomposition Algorithms

Much of the emphasis in the study of the substitution decomposition has been placed in its computation in optimal time. Finding an algorithm that is optimal for an arbitrary relational structure is possibly a worthy goal, though one that is likely to be difficult to achieve. For example, as we will shortly see the method used to derive an optimal algorithm to decompose permutations relies very heavily on their graphical presentation, which really is not extendable to more general structures. Although this doesn't rule out the discovery of an all-encompassing algorithm, it does indicate that such a method would be overly-complicated and most probably unenlightening.

We thus restrict our attention predominantly to the permutation case, though we will later discuss the same problem for graphs. The first algorithm which could compute the substitution decomposition of a permutation in linear time was given by Uno and Yagiura [116]. We will present a more recent and straightforward algorithm first published by Bergeron, Chauve, Montgolfier and Raffinot [17], and here rewritten to fit our treatment of permutations better.

In addition to the linear time substitution decomposition, Bergeron *et al.* [17] provide an optimal algorithm to compute the "common intervals" of a set of permutations on n elements, where a *common interval* is a set of (not necessarily contiguous) integers that, in each permutation π, is the image $\pi([i, j])$ of a contiguous set of positions. Our notion of interval is recovered from this definition by considering the common intervals of the set $\{\iota, \pi\}$, and our treatment here will be restricted just to this variety of interval. Note that, as there can be $N = n(n-1)/2$ intervals in a permutation π of length n (consider, for example, the intervals of an increasing permutation), we cannot expect to find an algorithm to compute all of these intervals in linear $O(n)$ time. Instead, the best-possible algorithm which we present works in $O(n+N)$ time. Despite it not being computable in linear time, this algorithm is interesting because of the importance intervals play in biomathematics, as mentioned in Chapter 1.

However, in order to compute the substitution decomposition of a permutation, we do not actually need to compute all the intervals; it is sufficient to compute the "strong intervals" (defined in Section 4.3, though essentially they may be viewed as the intervals occurring in the substitution decomposition tree), and there can be at most $2n-1$ of these. Thus we are able to hope for a linear time $O(n)$ algorithm, which is precisely what we obtain.

4.1 One- and Three-sided Intervals

We begin by considering an alternative way to view intervals; we may think of an interval of a permutation π as a set of points $\{p_1, \ldots, p_n\}$ which may be enclosed by the rectangle $\mathrm{rect}(p_1, \ldots, p_n)$ such that, in the plot of π,

- $\mathrm{rect}(p_1, \ldots, p_n)$ contains no points other than $p_1, \ldots, p_n$, and
- there are no pins separating any of $p_1, \ldots, p_n$ extending from $\mathrm{rect}(p_1, \ldots, p_n)$ in any direction (left, right, up or down).

If we weaken this second restriction by allowing pins to extend only in specified directions, we can obtain sets of points that are not intervals but look like intervals on the sides out of which pins are forbidden. For example, we may obtain a three-sided *right-open interval* by specifying that pins extending from $\mathrm{rect}(p_1, \ldots, p_n)$ can only be right pins. Our linear-time algorithm commences by first determining particular left-up-down- and right-

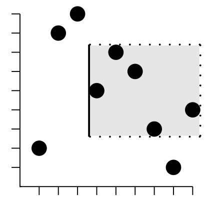

Figure 4.1: The shaded region denotes $\mathrm{rect}(I_{rud}(\pi; 4))$ of $\pi = 289576314$.

up-down-open intervals and using these to find the related right- and left-open intervals, which can then be used to "generate" the four-sided intervals.

Denote by $I_{rud}(\pi; i)$ the largest right-up-down-open interval of π for which i is the smallest (i.e. leftmost) position, i.e. $(i, \pi(i))$ defines the left edge of $\mathrm{rect}(I_{rud}(\pi; i))$. For example, if $\pi = 289576314$, then $I_{rud}(\pi; 4) = \{(4,5), (5,7), (6,6), (7,3), (9,4)\}$ (see Figure 4.1). Also, denote by $I_r(\pi; i)$ the largest right-open interval of π for which i is the smallest position. Returning to the previous example, $I_r(\pi; 4) = \{(4,5), (5,7), (6,6), (7,3)\}$. Similarly, $I_{\ell ud}(\pi; i)$ is the largest left-up-down-open interval, and $I_\ell(\pi; i)$ the largest left-open interval of π for which i is the *greatest* position. Since throughout this section we will be dealing only with a single permutation π, we will write $I_{rud}(\pi; i)$ more briefly as $I_{rud}(i)$, $I_r(\pi; i)$ as $I_r(i)$ and so on.

Our algorithm begins by computing $I_{rud}(i)$ and $I_{\ell ud}(i)$ for each i. Since the values of the points in each of $I_{rud}(i)$ and $I_{\ell ud}(i)$ form a contiguous set, it is sufficient to compute the points whose values are maximal and minimal for each. For a set of points P, denote by $\mathrm{maxval}(P)$ the *position* of the point in P whose value is maximal, and by $\mathrm{minval}(P)$ the position of the point whose value is minimal. Thus, our first step is to compute $\mathrm{minval}(I_{rud}(i))$, $\mathrm{maxval}(I_{rud}(i))$, $\mathrm{minval}(I_{\ell ud}(i))$ and $\mathrm{maxval}(I_{\ell ud}(i))$. The first of these is done using Algorithm 4.1, the others may be determined similarly.

Proposition 4.1 (Bergeron *et al.* [17, Proposition 4])**.** *Let π be a permutation of length n. Then Algorithm 4.1 computes* $\mathrm{minval}(I_{rud}(i))$ *for all $i \in [n]$ in $O(n)$ time.*

Proof. We assume that π^{-1} has been precomputed – a process which is easily done in $O(n)$

Algorithm 4.1 Computing $\mathrm{minval}(I_{rud}(i))$

```
S a stack recording point values, with topmost element s
π^{-1}(0) ← 0
push 0 on S
for i from 1 to n do
   while π^{-1}(i) < π^{-1}(s) do
      pop s from S
   end while
   minval(I_rud(π^{-1}(i))) ← π^{-1}(s + 1)
   push i on S
end for
```

time. At the beginning of the ith iteration of the **for** loop, the stack S contains, from top to bottom, a decreasing sequence of values whose sequence of corresponding positions as points in π is also decreasing. Among this sequence of values must be the largest value $j < i$ such that $\pi^{-1}(j) < \pi^{-1}(i)$, as the only way j could have been popped is if there were some j' with $j < j' < i$ and $\pi^{-1}(j') < \pi^{-1}(j)$, contradicting the definition of j. Furthermore, $\mathrm{minval}(I_{rud}(\pi^{-1}(i))) = \pi^{-1}(j+1)$, and so after popping all the values on top of j in the stack, the algorithm can return the position of the point whose value is $j + 1$. Since S stores every value $i \in [n]$ precisely once, it immediately follows that the algorithm has complexity $O(n)$. □

The next step is to find the three-sided intervals $I_r(i)$ and $I_\ell(i)$ for each $i \in [n]$. Note first that the set of positions in each $I_r(i)$ forms a contiguous set with smallest position equal to i, so for each i we only need to find the point in $I_r(i)$ whose position is greatest (i.e. the rightmost point). Similarly, the set of positions in $I_\ell(i)$ also forms a contiguous set, with maximum equal to $\pi(i)$, so here it is sufficient to find the point in $I_\ell(i)$ whose position is minimal.

Thus, for a set of points P let $\mathrm{minpos}(P)$ denote the position of the minimum (i.e. left-most) element, and $\mathrm{maxpos}(P)$ the position of the maximum (rightmost) element. Given the four bounds $\mathrm{minval}(I_{rud}(i))$, $\mathrm{maxval}(I_{rud}(i))$, $\mathrm{minval}(I_{\ell ud}(i))$ and $\mathrm{maxval}(I_{\ell ud}(i))$, we now seek $\mathrm{maxpos}(I_r(i))$ and $\mathrm{minpos}(I_\ell(i))$. The first of these is computed using Algorithm 4.2, while the second is done similarly.

Proposition 4.2 (Bergeron *et al.* [17, Proposition 3])**.** *Let π be a permutation of length n. Then, given* $\mathrm{minval}(I_{rud}(i))$ *and* $\mathrm{maxval}(I_{rud}(i))$*, Algorithm 4.2 computes* $\mathrm{maxpos}(I_r(i))$ *for all* $i \in [n]$ *in $O(n)$ time.*

Algorithm 4.2 Computing $\mathrm{maxpos}(I_r(i))$

```
for i from 1 to n do
  r_i ← i
end for
π(10) ← 10
for i from n to 1 do
  while π(minval(I_rud(i))) ≤ π(r_i + 1) ≤ π(maxval(I_rud(i))) do
    r_i ← r_{r_i+1}
  end while
  maxpos(I_r(i)) ← r_i
end for
```

Proof. Note first that $I_r(i)$ consists precisely of those points of $I_{rud}(i)$ whose positions form the longest contiguous sequence $[i, \mathrm{maxpos}(I_r(i))]$ for all $i \in [n]$. At the beginning of the ith iteration of the second **for** loop, we have found $\mathrm{maxpos}(I_r(i')) = r_{i'}$ for all $i' > i$, and r_i is still set to i. At all stages, r_i denotes the position of a point in $I_r(i)$, and hence $[i, r_i] \subseteq [i, \mathrm{maxpos}(I_r(i))]$. We next test whether the point with position immediately following r_i (i.e. $r_i + 1$) lies in $I_{rud}(i)$. If so, then $r_i + 1$ also lies in $I_r(i)$, as indeed does all of the right-open interval $I_r(r_i + 1)$. Thus we may now replace r_i with $\mathrm{maxpos}(I_r(r_i + 1)) = r_{r_i+1}$ and consider the new $r_i + 1$ at the start of the **while** loop. If, on the other hand, $r_i + 1 \notin I_{rud}(i)$, then r_i is the rightmost point of $I_r(i)$ and we have found $\mathrm{maxpos}(I_r(i))$ whence we may move on to consider the $(i-1)$th iteration. The complexity follows by observing that the contents of the **while** loop must be executed precisely $n - 1$ times in total. □

In the case of our ongoing example, $\pi = 289576314$, our list of bounds looks like:

i	1	2	3	4	5	6	7	8	9
$\mathrm{minpos}(I_\ell(i))$	1	2	2	4	5	2	7	7	1
$\mathrm{maxpos}(I_r(i))$	9	7	3	7	6	6	7	8	9

There remains one final prerequisite before we can show how to find intervals. For a permutation π of length n and position $i \in [n]$, define the *r-support* of i, denoted $\text{supp}_r(\pi; i)$, to be the largest position $j < i$ such that $I_r(i) \subset I_r(j)$. Similarly, define the *ℓ-support*, $\text{supp}_\ell(\pi; i)$, to be the smallest position $j > i$ such that $I_\ell(\pi; i) \subset I_\ell(\pi; j)$. Again we will use the more brief notation $\text{supp}_r(i)$ and $\text{supp}_\ell(i)$ since we are always working with the single permutation π. The r- and ℓ-supports will play a central role in finding the "strong intervals" of Section 4.3, and in Section 4.2 the r-support will reduce the number of candidate sets which we need to inspect in listing all the intervals. Given the bounds $\text{minpos}(I_\ell(i))$ and $\text{maxpos}(I_r(i))$, we may compute $\text{supp}_r(i)$ for all $i \in [n]$ using Algorithm 4.3, which clearly achieves this in $O(n)$ time.

Algorithm 4.3 Computing $\text{supp}_r(i)$

```
S a stack recording positions, with topmost element s
push 1 on S
supp_r(1) ← 1
for i from 2 to n do
   while maxpos(I_r(s)) < i do
      pop s from S
   end while
   supp_r(i) ← s
   push i on S
end for
```

The algorithm to find $\text{supp}_\ell(i)$ is analogous. For the example $\pi = 289576314$, we obtain:

i	1	2	3	4	5	6	7	8	9
$\text{supp}_\ell(i)$	7	3	6	6	6	9	8	9	9
$\text{supp}_r(i)$	1	1	2	2	4	5	4	1	4

There are now two avenues of exploration, each of which we will consider in turn. Section 4.2 computes all the intervals of a permutation π on $[n]$, which, if there are N such intervals, we show can be computed in $O(n + N)$ time. Section 4.3 shows how to search for the "strong intervals" of π (the intervals that define the substitution decomposition) showing that it can be done in $O(n)$ time, and from there compute the substitution decomposition of π.

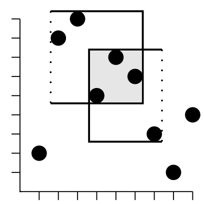

Figure 4.2: The intersection of $I_r(4)$ and $I_\ell(6)$ forms an interval of $\pi = 289576314$.

4.2 Generating Intervals

We have shown how to compute certain one- and three-sided intervals in linear time; it remains to show how these may be used to compute the (four-sided) intervals. Essentially, this is done by intersecting pairs of the three-sided intervals we computed in the previous subsection, and showing that what results is an interval (see Figure 4.2).

Proposition 4.3 (Bergeron *et al.* [17, Proposition 2])**.** *Let π be a permutation of length n, and let $i < j \in [n]$. Then the set of points with contiguous positions $[i, j]$ is an interval of π if and only if $i \geq \operatorname{minpos}(I_\ell(j))$ and $j \leq \operatorname{maxpos}(I_r(i))$.*

Proof. Suppose first that $[i, j]$ is a set of positions whose points in π form an interval P. Since P is an interval, we have both $[i, j] \subseteq [i, \operatorname{maxpos}(I_r(i))]$ and $[i, j] \subseteq [\operatorname{minpos}(I_\ell(j)), j]$, whence it follows that

$$[i, j] \subseteq [i, \operatorname{maxpos}(I_r(i))] \cap [\operatorname{minpos}(I_\ell(j)), j].$$

Conversely, suppose that for some $i < j \in [n]$ we have $i \geq \operatorname{minpos}(I_\ell(j))$ and $j \leq \operatorname{maxpos}(I_r(i))$. The set of points P with contiguous positions $[i, j]$ cannot be separated by a left pin since $(i, \pi(i))$ defines the left edge of $I_r(i)$, and it cannot be separated by a right pin since $(j, \pi(j))$ defines the right edge of $I_\ell(j)$. Finally, $(j, \pi(j)) \in I_r(i)$ and $(i, \pi(i)) \in I_\ell(j)$, and so, by the definitions of $I_r(i)$ and $I_\ell(j)$, P cannot be separated by up or down pins and hence forms a four-sided interval of π. $\square$

Proposition 4.3 alone will let us compute the intervals by examining the points with positions $[i, j]$ for every i, j with $1 \leq i \leq j \leq n$. We can reduce the number of these that need to be inspected, however, by making use of the r-support, a consideration which yields the sought-after $O(n + N)$ complexity.

Theorem 4.4 (Bergeron *et al.* [17, Theorem 2]). *The N intervals of a permutation π of length n can be computed in $O(n + N)$ time.*

Proof. For brevity, let us first set $\ell(i) = \text{minpos}(I_\ell(i))$, $r(i) = \text{maxpos}(I_r(i))$ and $s(i) = \text{supp}_r(i)$ for each $i \in [n]$. We must show that the output of Algorithm 4.4 is a complete list of the intervals of π. Suppose first that for some $i < j \in [n]$, the algorithm has printed $[i, j]$. This was output during the jth iteration of the **for** loop, and and within a **while** loop that ensures $i \geq \ell(j)$. Hence we only need to show that $j \leq r(i)$, as then Proposition 4.3 tells us that the points whose set of positions is $[i, j]$ will form an interval. This follows by studying how i evolved within the jth iteration before it was output; it was initiated by being set equal to j, and then was successively replaced by $s(j)$ finitely many times (possibly zero). Thus i is one of j, $s(j)$, $s(s(j)), \ldots$, and $j \leq r(i)$ then follows by considering the chain

$$j \leq r(j) \leq r(s(j)) \leq r(s(s(j))) \leq \ldots.$$

Conversely, for $i \leq j$, given the set of positions $[i, j]$ defines an interval of π, Proposition 4.3 implies that we have $i \geq \ell(j)$ and $j \leq r(i)$. Note that if $i = j$ then Algorithm 4.4 is guaranteed to return $[i, j]$ at the very start of the jth iteration, so we now assume $i < j$. Moreover, since $i \geq \ell(j)$, the algorithm will print $[i, j]$ providing we encounter the position i in the jth iteration of the **for** loop (as such an i will satisfy the **while** loop). Let i' be the smallest position such that $i < i' \leq j$ and $[i', j]$ was printed by the algorithm. By the minimality of i', we have $s(i') \leq i$. Now observe that $I_r(i') \subset I_r(i)$ as $i < i' \leq r(i)$, and so $r(i') > r(i)$ would contradict the maximality of $I_r(i)$. This implies that $s(i') \geq i$, and so $s(i') = i$, completing this part of the proof.

Finally, the complexity follows immediately since Algorithms 4.1, 4.2, and 4.3 have complexity $O(n)$, and the $O(n+N)$ complexity of Algorithm 4.4 follows by noting that the **while** loop will operate precisely N times. □

Algorithm 4.4 Computing the intervals of π

```
for j from n to 1 do
  i ← j
  while i ≥ minpos(I_ℓ(j)) do
    print [i, j]
    i ← supp_r(i)
  end while
end for
```

4.3 Strong Intervals and the Substitution Decomposition

Although we can now find all the intervals of π in optimal $O(n+N)$ time, we may prefer instead to find an $O(n)$ algorithm that is capable of telling us all that we really need to know, namely the substitution decomposition of π, and hence whether it is simple. To this end, define a *strong interval* of a permutation π to be an interval I of π for which every other interval J satisfies precisely one of $J \subseteq I$, $I \subseteq J$ or $J \cap I = \emptyset$ (i.e. I does not overlap with any other interval). The strong intervals of π are then precisely the intervals arising in the substitution decomposition, including both the whole of π and all the singleton intervals. Note that a permutation of length n has at most $2n-1$ strong intervals.

Up to now we have been working primarily with the three-sided intervals $I_\ell(i)$ and $I_r(i)$ for each $i \in [n]$ of a permutation π of length n. We have seen that they can be used to find all the intervals of π, but in order to restrict our attention to the strong intervals, we are going to want to replace our three-sided intervals with four-sided ones. Define, therefore, the *left-maximum* interval of a position $i \in [n]$ to be the largest interval of π whose rightmost point has position i, and write the leftmost position of this interval as $\mathrm{lmax}(\pi; i)$. Similarly, let $\mathrm{rmax}(\pi; i)$ denote the rightmost position of the largest interval of π whose leftmost point has position i (the *right-maximum* interval). Again we will abbreviate these to $\mathrm{lmax}(i)$ and $\mathrm{rmax}(i)$.

Trivially, we have $\mathrm{lmax}(i) \geq \mathrm{minpos}(I_\ell(i))$ and $\mathrm{rmax}(i) \leq \mathrm{maxpos}(I_r(i))$, and this suggests a starting point for finding the left-maximum and right-maximum intervals. However a direct search through the sets $I_\ell(i)$ and $I_r(i)$ cannot necessarily be performed in optimal time, so again we rely on the ℓ- and r-supports to reduce our search.

Proposition 4.5 (Bergeron *et al.* [17, Theorem 3])**.** *For a permutation π of length n, $\mathrm{rmax}(i)$ can be computed in $O(n)$ time.*

Proof. Note first that Algorithm 4.5 begins by setting $\mathrm{rmax}(i) = i$ for each i, with the exception of $\mathrm{rmax}(1)$ which is set to n, as expected. Note next that the **if** statement simply checks to see whether $[\mathrm{supp}_r(i), \mathrm{rmax}(i)]$ is a set of positions corresponding to an interval. If true, then $\mathrm{rmax}(j)$ for $j = \mathrm{supp}_r(i)$ is changed to $\mathrm{rmax}(i)$ if it is larger than the existing $\mathrm{rmax}(j)$. In either case, the set of points with positions $[j, \mathrm{rmax}(j)]$ will still correspond to an interval, so we need only check that the algorithm at some stage encounters the largest interval of π whose leftmost point is j.

Suppose for $j \in [n]$ that the set of points with positions $[j, j']$ correspond to the largest interval with leftmost point j, and that the algorithm has correctly found $\mathrm{rmax}(i)$ for all i such that $\mathrm{supp}(i) > j$. We may assume $j' > j$ as otherwise it is easy to see that Algorithm 4.5 correctly outputs $\mathrm{rmax}(j) = j$. By the maximality of j', we have $I_r(j') = \{(j', \pi(j'))\}$ and $\mathrm{rmax}(j') = j'$, so we are done if $\mathrm{supp}_r(j') = j$. (Note $\mathrm{supp}_r(j') < j$ is impossible since $[j, j']$ corresponds to an interval.) Let us therefore assume that $\mathrm{supp}_r(j') = j'' > j$, and note that the rightmost point in $I_r(j'')$ has position j', giving $\mathrm{rmax}(j'') = j'$ (since $I_r(j'')$ cannot be extended by a right pin). If $\mathrm{supp}_r(j'') = j$ then we are done, so instead suppose $\mathrm{supp}(j'') = j''' > j$, and observe that again we must have $\mathrm{rmax}(j''') = j'$. This process can only be repeated a limited number of times before we find some $i > j$ with $\mathrm{supp}(i) = j$ and $\mathrm{rmax}(i) = j'$. The complexity of Algorithm 4.5 follows immediately. $\square$

The computation for $\mathrm{lmax}(i)$ is similar, and for our running example $\pi = 289576314$ this gives:

i	1	2	3	4	5	6	7	8	9
$\mathrm{lmax}(i)$	1	2	2	4	5	2	7	8	1
$\mathrm{rmax}(i)$	9	6	3	6	6	6	7	8	9

Moving from the left-maximum and right-maximum intervals to the strong intervals is now a fairly straightforward process. We begin by listing the leftmost and rightmost

Algorithm 4.5 Computing $\mathrm{rmax}(i)$

$\mathrm{rmax}(1) \leftarrow n$
for i **from** 2 **to** n **do**
  $\mathrm{rmax}(i) \leftarrow i$
end for
for i **from** n **to** 2 **do**
  if $\mathrm{supp}_r(i) \geq \mathrm{minpos}(I_r(\mathrm{rmax}(i)))$ **and** $\mathrm{rmax}(i) \leq \mathrm{maxpos}(I_r(\mathrm{supp}_r(i)))$ **then**
    $\mathrm{rmax}(\mathrm{supp}_r(i)) \leftarrow \max(\mathrm{rmax}(i), \mathrm{rmax}(\mathrm{supp}_r(i)))$
  end if
end for

positions of the left-maximum and right-maximum intervals, marking right bounds with a bar, i.e. the set $\{i, \bar{i}, \mathrm{lmax}(i), \overline{\mathrm{rmax}(i)} : i \in [n]\}$ containing $4n$ bounds.

Next we sort this list into increasing order, $\{a_1, a_2, \ldots, a_{4n}\}$, listing left bounds before right bounds, noting that this can be done in linear time since there are only $2n$ possible values that the entries can take, each being either i or $\bar{i}$ for some $i \in [n]$. The sort can be further simplified by also noting that for each $i \in [n]$ we are guaranteed to see both i and $\bar{i}$ at least once. For our example ($\pi = 289576314$), this list is

$$\{1,1,1,\bar{1},2,2,2,2,\bar{2},3,\bar{3},\bar{3},4,4,\bar{4},5,5,\bar{5},6,\bar{6},\bar{6},\bar{6},\bar{6},\bar{6},7,7,\bar{7},\bar{7},8,8,\bar{8},\bar{8},9,\bar{9},\bar{9},\bar{9}\}.$$

We now work from left to right through this list, storing left bounds on a stack as they appear, and when we see a right bound r we take the top element s off the stack and return $[s, r]$ as a set of positions corresponding to a strong interval.

Theorem 4.6 (Bergeron *et al.* [17, Proposition 8])**.** *The strong intervals of a permutation π of length n can be computed in $O(n)$ time.*

Proof. If Algorithm 4.6 outputs an interval of the form $[i, \mathrm{rmax}(i)]$, then every interval whose positions are of the form $[\mathrm{lmax}(j), j]$ must have trivial intersection with $[i, \mathrm{rmax}(i)]$ (either $[\mathrm{lmax}(j), j] \subseteq [i, \mathrm{rmax}(i)]$ or $[\mathrm{lmax}(j), j] \cap [i, \mathrm{rmax}(i)]$ is empty). Subsequently, $[i, \mathrm{rmax}(i)]$ must intersect trivially with every interval of π since every interval is contained within a left-maximum or a right-maximum interval, and so $[i, \mathrm{rmax}(i)]$ is a strong inter-

Algorithm 4.6 Computing the strong intervals of π

S a stack recording positions, with topmost element s
for i **from** 1 **to** $4n$ **do**
  if a_i is a left bound **then**
    push a_i **on** S
  else
    print $[s, a_i]$
    pop s **from** S
  end if
end for

val. A similar argument can be applied if the algorithm outputs an interval of the form $[\mathrm{lmax}(j), j]$.

Now suppose the algorithm outputs the set of contiguous positions $[i, j]$ for which neither $\mathrm{lmax}(j) = i$ nor $\mathrm{rmax}(i) = j$. It follows that $[i, j] = [\mathrm{lmax}(j), j] \cap [i, \mathrm{rmax}(i)]$, and so $[i, j]$ corresponds to a set of points of π forming an interval. If $[i, j]$ does not correspond to a strong interval, then there exists a k for which either $i < k \leq j < \mathrm{rmax}(k)$ or $\mathrm{lmax}(k) < i \leq k < j$. In the former case, every interval $[k', \mathrm{rmax}(k')]$ with $i < \mathrm{rmax}(k') \leq k$ must satisfy $k' \geq k$, and so the algorithm would only permit the output of j as a right bound when paired with left bounds at least as big as k, a contradiction, proving that $[i, j]$ was strong.

Conversely, let $[i, j]$ correspond to a set of positions forming a strong interval of π, so there are no intervals of π whose positions have non-trivial intersection with $[i, j]$. To ensure the algorithm outputs $[i, j]$, we must find a left bound i and a right bound $\overline{j}$ in the ordered list of $4n$ bounds between which every left bound is matched by a right bound. Let x denote the number of positions k for which $\mathrm{lmax}(k) = i$ and $k < j$, and y the number of positions k for which $\mathrm{rmax}(k) = j$ and $i < k$. In the list of bounds $\{a_1, a_2, \ldots, a_{4n}\}$, there are $y - x$ more left bounds than right between the last occurrence of the left bound i and the first occurrence of the right bound $\overline{j}$. There are, however, at least $x + 1$ left bounds i and $y + 1$ right bounds $\overline{j}$ in this list, and so Algorithm 4.6 will output $[i, j]$. □

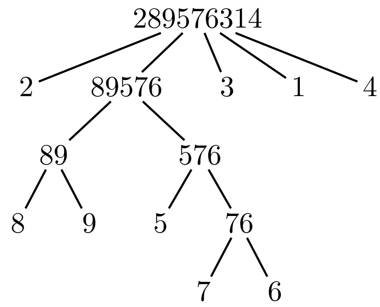

Figure 4.3: The substitution decomposition tree of $\pi = 289576314$.

For $\pi = 289576314$, after removing duplicates the output is

$$[1,1],[2,2],[3,3],[2,3],[4,4],[5,5],[6,6],[5,6],[4,6],[2,6],[7,7],[8,8],[9,9],[1,9].$$

We obtain the substitution decomposition tree by reading from right to left through our list of positions of strong intervals as output by Algorithm 4.6, noting that the strong intervals have been ordered as they would be output by a depth first search algorithm, working from right to left. Figure 4.3 shows the tree obtained for $\pi = 289576314$. Note that, by the definition of the strong intervals, in the cases where our permutation π is sum or skew decomposable, each sum or skew component will occupy a separate node. Where π is not sum or skew decomposable, the simple skeleton of π is easily obtained by taking the permutation order isomorphic to any chosen set of node representatives from the first level of the tree.

4.4 Graph Substitution Decomposition

The substitution decomposition has probably been studied most intensively in the context of graphs. It should come therefore as no great surprise that much time has been devoted to finding efficient algorithms to compute the substitution decomposition. Since 1972 algorithms that can compute the substitution decomposition tree for a graph with a variety of complexities ranging from $O(|V|^4)$ [73] to $O(|V| + |E| \log |V|)$ [38] have been found, while linear $O(|V| + |E|)$ complexity algorithms were found in 1994 by McConnell and Spinrad [88] and Cournier and Habib [39]. The former of these was later presented in more detail in [90]. A simpler divide-and-conquer algorithm was given by Dahlhaus, Gustedt and McConnell [41].

A related problem, and one that often appears alongside the substitution decomposition, is the transitive orientation of comparability graphs. The first $O(|V| + |E|)$ algorithm appears in McConnell and Spinrad [89], with a second algorithm by the same authors given in [90]. Armed with linear-time substitution decomposition and transitive orientation, one can solve many combinatorial problems in linear time. For example, the recognition of permutation graphs and two-dimensional posets (posets which are the intersection of two linear orders), and finding the maximum clique or minimum vertex colouring in comparability graphs. For further examples see [90].

Part II

Permutation Classes

Chapter 5

Containment as a Partial Order

As mentioned in Chapter 1, the pattern containment order is easily shown to be reflexive, transitive and antisymmetric, and hence forms a partial order on the set of all permutations (see Figure 5.1). Downsets of permutations under this order are called *permutation classes*. In other words, if $\mathcal{C}$ is a permutation class and $\pi \in \mathcal{C}$, then for any permutation σ with $\sigma \leq \pi$ we have $\sigma \in \mathcal{C}$. These sets have in the past also been labelled *closed classes* or *pattern classes*.

Permutation classes may be traced as far back as MacMahon [78], where $\mathrm{Av}(321)$ was enumerated by means of the study of "lattice permutations", though the more popular origin lies in Knuth [76]. It is not, however, until the last fifteen years that their study has become more intense, with a wide variety of questions being answered pertaining both to their structure and to their enumeration. These two varieties of question are not, of course, independent; greater knowledge of how permutation classes are constructed can often lead quickly to enumerative consequences, while the question of enumeration is frequently the motivation for the study of their structure. The structural work on simple permutations in Part I fits, to some extent, this mould; while their study was initially motivated by an enumeration problem, the consequences of the study extend well beyond the original question.

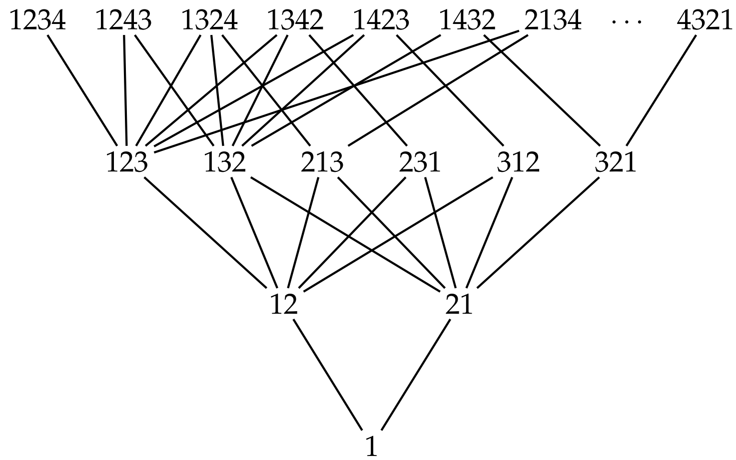

Figure 5.1: The start of the containment partial order.

5.1 Defining Permutation Classes

Permutation classes arise naturally in a variety of settings, ranging from sorting (see, for example, Bóna's survey [21]) to algebraic geometry (see, for example, Lakshmibai and Sandhya [77]). Typically, a permutation class is defined in one of the following ways:

- **Pattern avoidance.** A permutation class $\mathcal{C}$ can be regarded as a set of permutations which avoid certain patterns. The set B of minimal permutations not in $\mathcal{C}$ is known as the *basis* of $\mathcal{C}$. We write $\mathcal{C} = \mathrm{Av}(B)$ to mean the class $\mathcal{C} = \{\pi \mid \beta \not\leq \pi \text{ for all } \beta \in B\}$. Bases need not be finite – see the examples in Subsection 5.1.2 and the discussion on antichains in Section 5.3.

- **Permuting machines.** As already mentioned, permutation classes arise naturally as a result of machines which permute an input stream of symbols. Indeed, the set of stack-sortable permutations dates back to the major origin of permutation classes, Knuth [76]. Their study remains an area of active interest to this day – see the discussion at the end of Example 5.3.

- **Constructions.** New permutation classes can be formed using constructions involving one or more old classes (e.g. the union of two classes). See Subsection 5.1.2 for extensive examples.

- **Closures.** We may also define a class by taking the closure of some set of permutations, or even a set of functions that are order isomorphic to permutations. For two linearly ordered sets A and B and a bijection $f : A \to B$, we define the *closure* of f to be the permutation class $\mathcal{C} = \mathrm{Sub}(f : A \to B)$ as follows.[1] A permutation π of length n lies in $\mathcal{C}$ if there exists a sequence $a_1 < a_2 < \ldots < a_n$ of A for which $f(a_1), f(a_2), \ldots, f(a_n)$ is order isomorphic (under the linear order of B) to π. Similarly, we may define the closure of a set of bijections $\{f_i : A_i \to B_i, i \in I\}$ simply by taking the union,

$$\mathrm{Sub}(f_i : A_i \to B_i, i \in I) = \bigcup_{i \in I} \mathrm{Sub}(f_i : A_i \to B_i).$$

  Waton [118] introduced a geometrical approach to this notion of closure in his PhD thesis, whereby a permutation class is defined by the set of permutations which may be drawn by taking points that lie on a specified geometrical shape.

Once we have specified our chosen permutation class, we may wish to know answers to one or more of a wide variety of properties which the class may or may not possess. In all but the first case, our first problem is likely to be to find its basis, or at least whether the basis is finite or not, as this is arguably the most convenient way to represent a class. We will present many properties in the next two sections, but first, however, let us review some specific examples of permutation classes, the ways in which they may arise, and compute their bases.

5.1.1 Examples

Example 5.1 (Finite Classes)**.** By the Erdős-Szekeres Theorem 2.3, a class $\mathcal{C}$ is finite if and only if its basis B contains both an increasing permutation and a decreasing permutation. For example, the class $\mathcal{C} = \{1, 12, 21, 132, 213, 231, 312, 2143, 2413, 3142, 3412\}$ has basis $B = \{123, 321\}$.

[1] This is a special case of "ages" for classes of relational structures – see the discussion on atomicity in the general setting in Section 5.5.

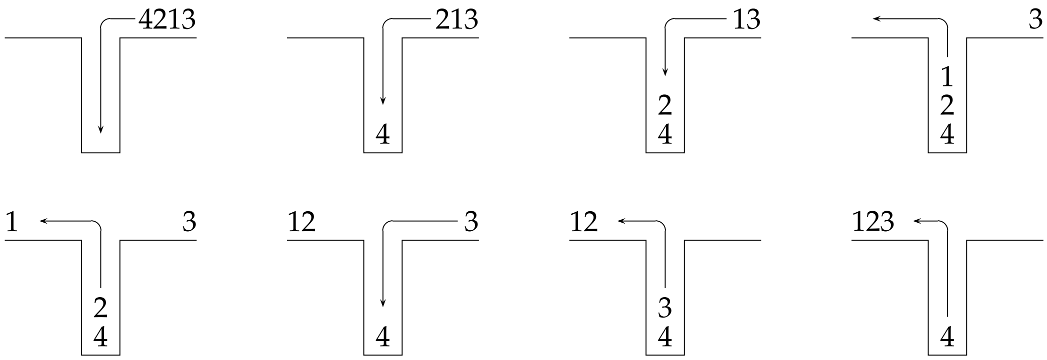

Figure 5.2: Sorting 4213 with a stack.

Example 5.2 (The set of Increasing Permutations)**.** The "smallest" infinite class is the set of increasing permutations $\mathcal{I} = \{1, 12, 123, 1234, 12345, \ldots\}$. It can easily be seen that every permutation in $\mathcal{I}$ avoids the permutation 21, and also that 21 is the only basis element, so that $\mathcal{I} = \mathrm{Av}(21)$.

Example 5.3 (The set of Stack-Sortable Permutations)**.** A *stack* is a one-dimensional array into which symbols may be "pushed", one on top of the other, with only the topmost symbol being available to be "popped" at each stage. A permutation of length n is *stack sortable* if it can be sorted into the increasing permutation $12 \cdots n$ by passing it through a stack, symbol by symbol (see, for example, Figure 5.2).

The set of stack sortable permutations clearly satisfies downward closure under the containment order, and so forms a permutation class. We next seek its basis, and first note that 231 is not stack sortable, since either the 2 must be popped before the 1 is pushed, or the 3 must be popped before the 2 can be popped. It is then fairly straightforward to show that every permutation that is not stack-sortable contains a copy of 231, and so $\mathrm{Av}(231)$ represents the set of stack sortable permutations.

There are many variants to this problem, several of which are discussed in Bóna's survey [21]. For example, we may connect two or more stacks in parallel or in series; we may restrict the depth of the stack by allowing it to contain at most m symbols at any one given time. The answers to some of these questions are immediate, while others remain open, and, indeed, some varieties do not form closed classes.

For example, in the case of connecting two stacks in series the general case is shown by Murphy [97] to be infinitely based with shortest basis elements of length 7, though a description of the complete basis is unknown. Atkinson, Murphy and Ruškuc [10] provide the complete but infinite basis for the subclass formed by imposing the condition that the stacks must be *ordered* – that is, from top to bottom the elements in each stack must form an increasing sequence. To achieve a finitely based class, we may restrict our attention to connecting a stack of depth 2 and an infinite stack in series, which has just 20 basis elements variously of lengths 5, 6, 7 and 8 [49].

Considerable study has been devoted to the *West-t-stack sortable permutations* [119], formed by adding a greedy algorithm to a sequence of ordered stacks: take the earliest available "push" onto a stack in the series if it exists, otherwise "pop" a new output symbol. However, the West-t-stack sortable permutations do not, in general, form a permutation class – for example, 35241 is West-2-sortable but 3241 is not.

Example 5.4 (The Separable Permutations)**.** We define the class $\mathcal{S}$ of separable permutations constructively. A permutation is *separable* if and only if it can be obtained by repeated application of direct and skew sums, starting with the permutation 1. For example,

$$\begin{aligned} 354621 &= 1324 \ominus 21 \\ &= (132 \oplus 1) \ominus 1 \ominus 1 \\ &= (1 \oplus 21 \oplus 1) \ominus 1 \ominus 1 \\ &= (1 \oplus (1 \ominus 1) \oplus 1) \ominus 1 \ominus 1. \end{aligned}$$

(Note the omission of certain brackets, which follows by the associativity of $\oplus$ and $\ominus$.)

It is then clear that the set of separable permutations is closed downwards under the containment order. It was shown by Bose, Buss and Lubiw [24] that the class of separable permutations is equal to $\mathrm{Av}(2413, 3142)$, and we may derive this result easily after considering Proposition 5.28 (see Page 106). Note that 2413 and 3142 are the two simple permutations of length 4, and that subsequently the only simple permutations in this class are 1, 12 and 21, which is precisely what we expect to see when we consider the substitution decomposition of a separable permutation.

The separable permutations seem to have made their first appearance as the permutations that can be sorted by pop-stacks in series, see Avis and Newborn [13]. Shapiro and Stephens [108] showed that the separable permutations are those that fill up under bootstrap percolation.[2] They are essentially the permutation analogue of series-parallel posets (see Stanley [113, Section 3.2]) and complement reducible graphs (see Corneil, Lerchs, and Burlingham [36]).

5.1.2 New Classes from Old

There is virtually an endless number of ways to define new sets of permutations from old, and only slightly fewer which construct permutation classes. Besides the obvious constructions given by the intersection and union of two classes, we can look at ways in which permutations themselves may be combined. For example, we may place permutations next to one another (horizontal juxtaposition) or one above the other (vertical juxtaposition); we may mix two permutations together (merge), or use inflations to place permutations inside one another (wreath product).

The Intersection of two Permutation Classes. Given two permutation classes defined by their bases $\mathcal{C} = \mathrm{Av}(A)$ and $\mathcal{D} = \mathrm{Av}(B)$, consider their intersection $\mathcal{C} \cap \mathcal{D}$. It is trivial to see that $\mathcal{C} \cap \mathcal{D}$ forms a permutation class, and also that its basis is given by the union $A \cup B$. If, therefore, $\mathcal{C}$ and $\mathcal{D}$ are finitely based, then so is $\mathcal{C} \cap \mathcal{D}$. Little more needs to be said – Murphy [97] "awaits questions about intersections that are worthy of attention!"

The Union of two Permutation Classes. Given two classes $\mathcal{C} = \mathrm{Av}(A)$ and $\mathcal{D} = \mathrm{Av}(B)$, the union $\mathcal{C} \cup \mathcal{D}$ is again a permutation class. Its basis is also easily determined; a per-

[2]Bootstrap percolation is a process defined on $n \times n$ 0-1 matrices, in which at each stage of the process every zero entry in the matrix becomes one if two or more of its neighbours are non-zero, while entries with value one remain the same. The process terminates when no more entries can be changed. Given an $n \times n$ permutation matrix, it will completely fill up with ones if and only if the permutation is separable.

mutation in the basis of $\mathcal{C} \cup \mathcal{D}$ must contain a copy of some $\alpha \in A$ and $\beta \in B$, and by its minimality it follows that such a basis element can contain no points other than these (such a permutation is known as a *minimal merge* of α and β). Thus, if $\mathcal{C}$ and $\mathcal{D}$ are finitely based, then so is $\mathcal{C} \cup \mathcal{D}$.

For example, letting $\mathcal{C} = \mathcal{I} = \mathrm{Av}(21)$ and $\mathcal{D} = \mathrm{Av}(12)$, then

$$\mathcal{C} \cup \mathcal{D} = \{1, 12, 21, 123, 321, 1234, 4321, \dots\},$$

and its basis consists of the minimal merges of 21 and 12, which are 132, 213, 231 and 312. Thus $\mathcal{C} \cup \mathcal{D} = \mathrm{Av}(132, 213, 231, 312)$.

Juxtaposition. Given two permutation classes $\mathcal{C}$ and $\mathcal{D}$, their *horizontal juxtaposition*, denoted $\left[\ \mathcal{C} \quad \mathcal{D}\ \right]$, 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 $\mathcal{D}$. In other words, the horizontal juxtaposition of $\mathcal{C}$ and $\mathcal{D}$ consists of those permutations π whose plot may be divided with a vertical line, so that the points on the left are order isomorphic to a permutation in $\mathcal{C}$ while those on the right are order isomorphic to a permutation in $\mathcal{D}$.

The question of finite basis is immediately answerable, and may be derived by following a similar argument to the one above for the union of two classes.

Proposition 5.5 (Atkinson [7])**.** *Let $\mathcal{C}$ and $\mathcal{D}$ be permutation classes. The basis elements of the class $\left[\ \mathcal{C} \quad \mathcal{D}\ \right]$ 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 $\mathcal{D}$ is denoted $\begin{bmatrix} \mathcal{C} \\ \mathcal{D} \end{bmatrix}$, and consists of those permutations π whose plot may be divided with a horizontal line, so that the points above the line are order isomorphic to a permutation in $\mathcal{C}$ while those below are order isomorphic to a permutation in $\mathcal{D}$.

Merge. A permutation π is a *merge* of the permutations α and β if π consists of two subsequences, one order isomorphic to α, the other to β. This may be written $\pi = \alpha \sqcup \beta$. Little is known about the basis of the merge $\mathcal{C} \sqcup \mathcal{D}$ of two classes – there are no counter examples to contradict the suggestion that $\mathcal{C} \sqcup \mathcal{D}$ is always finitely based if $\mathcal{C}$ and $\mathcal{D}$ are finitely based, but neither are there sufficient results to support such a conjecture. The merge of two permutations corresponds – somewhat roughly – to connecting permuting machines in parallel (see Atkinson and Beals [8]).

Grid Classes. An $m \times n$*-gridding* of a permutation π is a collection of $m-1$ distinct horizontal lines and $n-1$ vertical lines that divide the plot of π into mn cells.[3] Given an $m \times n$ matrix M of permutation classes, the *grid class of* M is the class $\mathcal{C}$ of all permutations π for which π is $m \times n$-griddable, with the points in each cell of the gridding being order isomorphic to a permutation from the class in the corresponding entry of the matrix. Grid classes may be considered to be a generalisation of the juxtaposition construction, though they are not merely compositions of juxtapositions. We may, however, ask the same questions. Pertinently:

Question 5.6. *If M is a matrix of permutation classes all of which are finitely based, when is the grid class of M finitely based?*

Obviously for matrices of dimensions $m \times 1$ or $1 \times n$, grid classes are equivalent to vertical and horizontal juxtapositions, respectively, and so the question of basis is known. In general it is not finitely based, consider, for example, the 2×2 matrix

$$M = \begin{pmatrix} \emptyset & \mathrm{Av}(321654) \\ \mathrm{Av}(321654) & \emptyset \end{pmatrix}.$$

[3] Most authors switch m and n to consider vertical lines first. Here, to avoid redefining the order in which the dimensions of a matrix are written for this brief review, we go against this convention.

The basis for the grid class of M is infinite – see Murphy [97]. There is more hope if we restrict M to contain only the monotone classes $\{1, 12, 123, \dots\}$ or $\{1, 21, 321, \dots\}$, but even here results can only be proved for a few specific 2×2 matrices. See Waton [118] for further discussion.

Conversely, we may ask when a given class may be gridded. Given two permutation classes $\mathcal{C}$ and $\mathcal{D}$, $\mathcal{C}$ is said to be *$\mathcal{D}$-griddable* if, for some m and n, $\mathcal{C}$ is a grid class of the $m \times n$ matrix M all of whose entries are $\mathcal{D}$. Huczynska and Vatter [70] characterise the $\mathcal{D}$-griddable classes where $\mathcal{D}$ is taken to consists precisely of the monotone permutations, while the following more general result appears in Vatter [117]:

Theorem 5.7 (Vatter [117])**.** *The permutation class $\mathcal{C}$ has a $\mathcal{D}$-gridding if and only if it does not contain arbitrarily long sums or skew sums of basis elements of $\mathcal{D}$, i.e. there exists a constant m so that $\mathcal{C}$ contains neither $\beta_1 \oplus \cdots \oplus \beta_m$ nor $\beta_1 \ominus \cdots \ominus \beta_m$ for basis elements β_i of $\mathcal{D}$.*

Direct and Skew Sums. There are several ways to use direct and skew sums to define new permutation classes. Naïvely, there is of course the set $\mathcal{C} \oplus \mathcal{D} = \{\alpha \oplus \beta : \alpha \in \mathcal{C}, \beta \in \mathcal{D}\}$, though this is only a permutation class if we force the empty permutation to be a member of both $\mathcal{C}$ and $\mathcal{D}$.

Of greater use is the "sum completion" of a class $\mathcal{C}$; a permutation class $\mathcal{C}$ is said to be *sum complete* if $\alpha, \beta \in \mathcal{C}$ implies $\alpha \oplus \beta \in \mathcal{C}$, and the *sum completion* of a class $\mathcal{C}$ is the smallest sum complete class containing $\mathcal{C}$. Similarly, we may define *skew complete* and the *skew completion* by replacing the operation $\oplus$ with $\ominus$. We may also mix these two operations; a class $\mathcal{C}$ is said to be *strongly complete* if $\mathcal{C}$ is both sum and skew complete. Accordingly, the *strong completion* of a permutation class $\mathcal{C}$ is the smallest strongly complete class containing $\mathcal{C}$.

We can tell if a class is sum, skew or strongly complete by looking at its basis.

Proposition 5.8. *A class $\mathcal{C}$ is sum (respectively, skew, strongly) complete if and only if every basis element is sum (respectivey skew, strongly) indecomposable.*

Proof. If we were to find a sum decomposable basis element π of the sum complete class $\mathcal{C}$, then we could write $\pi = \alpha \oplus \beta$ for some α and β, both of which necessarily lie in $\mathcal{C}$. But then, by its sum completion, $\mathcal{C}$ contains $\alpha \oplus \beta$, a contradiction. Conversely, if all the basis elements of $\mathcal{C}$ are sum indecomposable, then if for some α and β in $\mathcal{C}$ there is a copy of a basis element π in $\alpha \oplus \beta$, we would have either $\pi \leq \alpha$ or $\pi \leq \beta$, a contradiction.

The cases for skew complete and strongly complete classes are similar. □

Computing the basis of a sum, skew or strong completion of a class is not straightforward – in particular, if the class is finitely based then the sum, skew and strong completions need not be finitely based, examples of which we will see in Chapter 8.

The Wreath Product. The *wreath product* of two permutation classes $\mathcal{C}$ and $\mathcal{D}$ is the set $\mathcal{C} \wr \mathcal{D}$ of all permutations which can be expressed as an inflation of a permutation in $\mathcal{C}$ by permutations in $\mathcal{D}$, i.e. the set of permutations of the form $\pi[\alpha_1, \alpha_2, \ldots, \alpha_n]$ with $\pi \in \mathcal{C}$ and $\alpha_1, \alpha_2, \ldots, \alpha_n \in \mathcal{D}$.

It is easy to check that the wreath product of two permutation classes is again a permutation class. For example, the sum completion of a class $\mathcal{C}$ corresponds to the wreath product $\mathcal{I} \wr \mathcal{C}$, while the strong completion of $\mathcal{C}$ is the wreath product $\mathcal{S} \wr \mathcal{C}$ where $\mathcal{S} = \mathrm{Av}(2413, 3142)$ is the class of separable permutations.

The question of finite basis has been answered in only a few cases – if $\mathcal{C}$ and $\mathcal{D}$ are finitely based, when is $\mathcal{C} \wr \mathcal{D}$ finitely based? We take up this question in Chapter 8, establishing a more general finite basis result for wreath products.

Wreath Closure. A class $\mathcal{C}$ of permutations is *wreath-closed* if $\sigma[\alpha_1, \ldots, \alpha_m] \in \mathcal{C}$ for all $\sigma, \alpha_1, \ldots, \alpha_m \in \mathcal{C}$. The *wreath-closure* of a set X, $\mathcal{W}(X)$, is defined as the smallest wreath-closed class containing X. (This concept is well-defined because the intersection of wreath-closed classes is wreath-closed, and the set of all permutations is wreath-closed.)

Letting $\mathrm{Si}(\mathcal{C})$ denote the set of simple permutations in the class $\mathcal{C}$, we observe that $\mathrm{Si}(\mathcal{C}) = \mathrm{Si}(\mathcal{W}(\mathcal{C}))$ and indeed $\mathcal{W}(\mathcal{C})$ is the largest class with this property.[4] For example,

[4]While this claim may appear intuitively obvious, there are some technical subtleties. Every permutation

the wreath closure of $\mathrm{Av}(132)$ is the largest class whose only simple permutations are 1, 12, and 21, which is precisely the class of separable permutations of Example 5.4.

It is quite easy to decide if a permutation class given by a finite basis is wreath-closed:

Proposition 5.9 (Atkinson and Stitt [12])**.** *A permutation class is wreath-closed if and only if each of its basis elements is simple.*

One may also wish to compute the basis of $\mathcal{W}(\mathcal{C})$. This is routine for classes with finitely many simple permutations (see Proposition 5.28), but much less so in general. An example of a finitely based class whose wreath closure is infinitely based is $\mathrm{Av}(4321)$ – its wreath closure contains a variant of the increasing oscillating antichain, which we will define in Example 5.14.

The natural question is then:

Question 5.10. *Given a finite basis B, is it decidable whether $\mathcal{W}(\mathrm{Av}(B))$ is finitely based?*[5]

5.2 Enumeration

Probably the largest active area in the study of permutation classes is enumeration: given a class $\mathcal{C}$, how many permutations are there of length n, and is this sequence well-behaved? Once these questions are answered, we may be interested in finding out what other combinatorial structures are enumerated by this sequence, and whether bijections can be established between them. In the first instance, this may be done by looking at the Online Encyclopaedia of Integer Sequences [110].

For a permutation class $\mathcal{C}$, 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 $f(x) = \sum |\mathcal{C}_n| x^n$ as the *generating function for $\mathcal{C}$*. The generating function f is *algebraic* if it solves an equation of the form $p_n(x)f^n + p_{n-1}(x)f^{n-1} + \cdots + p_0(x)f^0 = 0$ for polynomials p_i. Similarly, a *rational generating function* is one that may

in $\mathcal{C}$ is an inflation of a member of $\mathrm{Si}(\mathcal{C})$ so it follows (e.g., inductively) that $\mathcal{C} \subseteq \mathcal{W}(\mathrm{Si}(\mathcal{C}))$. Thus $\mathcal{W}(\mathcal{C}) \subseteq \mathcal{W}(\mathrm{Si}(C))$, establishing that $\mathrm{Si}(\mathcal{C}) = \mathrm{Si}(\mathcal{W}(\mathcal{C}))$. As wreath closed classes are uniquely determined by their sets of simple permutations, $\mathcal{W}(\mathcal{C})$ is the largest class with this property.

[5] The analogous question for graphs was raised by Giakoumakis [60] and has received a sizable amount of attention, see for example Zverovich [122].

be written as a rational function, i.e. a function of the form $\frac{p(x)}{q(x)}$ where $p(x)$ and $q(x)$ are polynomials in x over the field of rational numbers.

As a trivial first example, consider the class $\mathcal{I} = \{1, 12, 123, \ldots\}$. There is precisely one permutation of each length, and so its generating function is $f(x) = \sum_{n=0}^{\infty} x^n = 1+x+x^2+\cdots$, or, in other words, $f = \frac{1}{1-x}$, a rational function. Note that here our sum begins at $n = 0$, implying that we are including the single permutation of length zero in the class. This is a convention that may or may not always be used – there are cases where including the empty permutation is convenient (particularly when considering recursive structures), while in other cases we may specifically not want it. It will be our convention to include the empty permutation unless required to do otherwise.

Our next example is somewhat more complicated, and the method employed to derive the enumeration is a classic recursive technique relying on knowledge of the structure of a permutation in the specified class. This is, of course, precisely where the rôle of simple permutations and the substitution decomposition will become invaluable.

Example 5.11 (The Stack Sortable Permutations)**.** As seen in Example 5.3, the set of stack sortable permutations is precisely the class $\mathrm{Av}(231)$. Within this class, the permutations of lengths $1, 2, 3, 4, 5 \ldots$ are enumerated by the sequence $1, 2, 5, 14, 42, \ldots$, which looks encouragingly like the sequence of Catalan Numbers (sequence A000108 of [110]), with general term $\frac{(2n)!}{n!(n+1)!}$.

We prove this fact by considering a permutation $\pi \in \mathrm{Av}(231)$ of length n. Since π must avoid 231, every point coming before the value n in π must lie below every point coming after the value n, i.e. $\pi = \alpha \oplus (1 \ominus \beta)$ for some α and β, which also of course must themselves avoid 231 (see Figure 5.3). Thus α and β must lie in $\mathrm{Av}(231)$, but there are no other restrictions on α and β save that we must of course have $|\alpha| + |\beta| + 1 = n$. Note also that this decomposition into α and β is unique, and hence can be used to decompose (or construct) every permutation in $\mathrm{Av}(231)$.

In terms of generating functions, if $f(x)$ is the generating function for $\mathcal{C} = \mathrm{Av}(231)$,

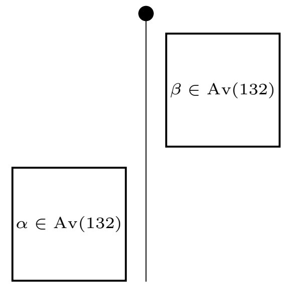

Figure 5.3: Generic structure of a 231-avoider.

then we can use the above consideration to derive the recursion

$$f = xf^2 + 1.$$

Note that here we have included the empty permutation, as we must allow α and/or β to be empty. Note further that the empty permutation cannot be decomposed as we did above because it has no maximum entry, hence the appearance of the "$+1$" term. Solving this algebraic equation is then straightforward, and gives

$$f = \frac{1-\sqrt{1-4x}}{2x} = 1 + x + 2x^2 + 5x^3 + 14x^4 + \ldots$$

as required.

Central to the enumeration problem is the classification of permutation classes with the same enumeration. We say that two permutations α and β are *Wilf equivalent* if $|\,\mathrm{Av}(\alpha)_n| = |\,\mathrm{Av}(\beta)_n|$ for all n, i.e. the classes $\mathrm{Av}(\alpha)$ and $\mathrm{Av}(\beta)$ are enumerated by the same generating function. We may also say that the permutations α and β belong to the same *Wilf class*. For example, the permutations 231 and 123 are Wilf equivalent, a fact which may be proved using several different bijections – see Richards [102], Rotem [104], Simion and Schmidt [109] or West [120] for various approaches to this problem. Since enumeration is then preserved under symmetry, this proves that all the permutations of length 3 belong to the same Wilf class. The computation of the Wilf classes up to length 7 were completed in 2001 by Stankova and West [112].

This term has since been extended in the natural way to sets of permutations – the permutation sets A and B are Wilf-equivalent if $|\operatorname{Av}(A)_n| = |\operatorname{Av}(B)_n|$. While this may open up an endless but for the most part uninteresting variety of problems, there are some very surprising results. Notably, Bóna [18] shows that the class $\operatorname{Av}(1342)$ has generating function $f = \dfrac{32x}{-8x^2+20x+1-(1-8x)^{3/2}}$. This is the same as the class of permutations which may be sorted with two ordered stacks in series, whose basis is infinite:

$$B = \{(2, 2m-1, 4, 1, 6, 3, 8, 5, \ldots, 2m, 2m-3) | m = 2, 3, 4, \ldots\}.$$

(This problem was previously discussed at the end of Example 5.3.)

Another approach to the problem of enumeration is that of asymptotics – how many permutations of length n are there in a given permutation class as n approaches infinity? In other words, we want to be able to say something about $\lim\limits_{n\to\infty} |\mathcal{C}_n|$, or, somewhat more usefully, $\lim\limits_{n\to\infty} \sqrt[n]{|\mathcal{C}_n|}$. As a first step, we have the "Stanley-Wilf conjecture", namely that for a given class $\mathcal{C}$ not containing every permutation, there exists a constant $\overline{K}$ such that

$$\limsup_{n\to\infty} \sqrt[n]{|\mathcal{C}_n|} = \overline{K}.$$

This result was proved in 2004 by Marcus and Tardos [87]. The constant $\overline{K}$ is known as the *upper growth rate* of the permutation class. We may similarly define the *lower growth rate*, $\liminf_{n\to\infty} \sqrt[n]{|\mathcal{C}_n|} = \underline{K}$. This naturally begs the question whether the upper and lower growth rates coincide, in which case $\lim_{n\to\infty} \sqrt[n]{|\mathcal{C}_n|} = K$ is called the *growth rate* of $\mathcal{C}$. It is conjectured that the growth rate always exists, a fact that has been shown in some cases. Arratia [6] proves this for sum or skew complete classes, among which are all of the permutation classes defined by a single basis element.

For example, the growth rate of the stack sortable permutations $\operatorname{Av}(231)$ is 4, a fact easily seen by recalling that $|\operatorname{Av}(231)_n| = \dfrac{(2n)!}{n!(n+1)!}$, and using Stirling's approximation $n! \approx \sqrt{2n\pi}\left(\dfrac{n}{e}\right)^n$.

5.3 Antichains, Partial Well Order and Atomicity

In any partial order, an *antichain* is a set of pairwise incomparable elements. Immediate from its definition, the basis of any permutation class is an antichain. As previously mentioned, there are infinitely based permutation classes, and hence there are infinite antichains. These have been widely studied – see for example Atkinson, Murphy and Ruškuc [9] and Murphy and Vatter [98].

An attempt at the classification of "fundamental" antichains was given in Murphy's PhD Thesis [97], though little progress has been made since. An infinite antichain A is said to be *fundamental* if its closure, $\mathrm{Sub}(A)$, contains no infinite antichains, except subsets of A itself. Other authors (see, for example, Gustedt [66]) refer to such antichains as minimal, because they are minimal under the following order on infinite antichains: $A \preceq B$ if A is contained in the closure of B. The need for identifying the fundamental antichains will become apparent when we introduce partial well order. Meanwhile, we offer the following conjecture:

Conjecture 5.12. *Every member of a fundamental infinite antichain contains at most two proper intervals.*

Example 5.13 (The Increasing Oscillating Antichain)**.** Let us consider the antichain based on the increasing oscillating sequence from Section 2.5. The first few elements of this antichain are $51234, 4127356, 412639578, \ldots$, with nth term $4126385 \cdots 2n+3, 2n-1, 2n+1, 2n+2$. The sixth term of this sequence is plotted in Figure 5.4); note the underlying pin sequence construction and the pair of points at either end of the sequence which form *anchors*, preventing its involvement in any other member of the antichain.

To prove that this is an antichain, we must show that no member is contained in any other. This may be done in a variety of ways, but a particularly neat method can be found in Klazar [75]. The *graph of a permutation* π of length n is the graph G_π whose vertex set is $V = [n]$, with $i \sim j$ if and only if $i < j$ and $\pi(i) > \pi(j)$ or vice versa ($j < i$ and $\pi(j) > \pi(i)$), i.e. if and only if there is a descent in π between i and j. For example, the

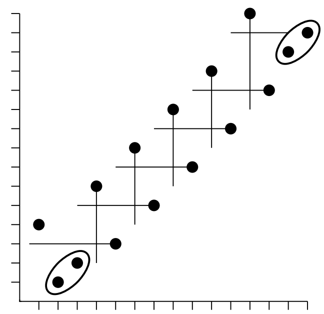

Figure 5.4: The sixth term of the increasing oscillating antichain.

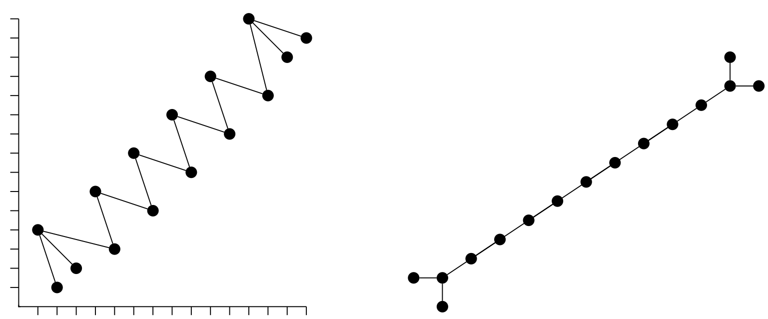

Figure 5.5: Forming the graph of the sixth term of the increasing oscillating antichain.

increasing permutation $12 \cdots n$ corresponds to the independent graph on n vertices, while the decreasing permutation $n \cdots 21$ corresponds to the complete graph K_n.

Although we lose uniqueness (for example, $G_{213} = G_{132}$), the pattern containment order translates to graph containment under taking induced subgraphs, that is, $\sigma \leq \pi$ implies $G_\sigma \leq G_\pi$. To show that two arbitrary members of the increasing oscillating antichain are not comparable under pattern containment, therefore, it is sufficient to show that their corresponding graphs are incomparable in the graph containment partial order. In some cases this may not make the containment problem any easier, but here the required result follows almost immediately.

The graph of the sixth term of the antichain is shown in Figure 5.5. Note that the nth member of the antichain will thus correspond to a graph consisting of a path of length $2n-1$ with a pair of leaves attached to each end. It is then clear that if we were to superimpose

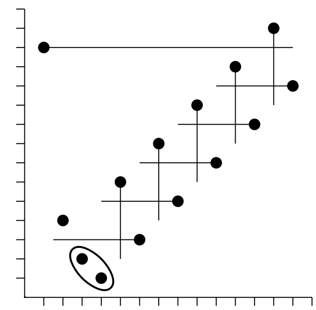

Figure 5.6: A basis element of the wreath closure of $\mathrm{Av}(4321)$.

the graph of a smaller member of the antichain onto the graph of a larger one, the end nodes of the smaller must correspond to the end nodes of the larger, leaving a path which cannot be superimposed onto the longer path without losing an edge. Thus the graphs are pairwise incomparable, and hence the permutations are pairwise incomparable.

Finally, we may observe that the antichain is fundamental since every subpermutation of an element of the antichain is either sum decomposable or lacks at least one anchor.

We may, of course, vary the anchors of the increasing oscillating sequence – and, indeed, most other antichains – to produce a complete variety of different antichains. We will use this fact in Chapter 8 to exhibit several antichains which lie in the basis of particular classes. Meanwhile, let us return to considering the basis of the wreath closure of $\mathrm{Av}(4321)$:

Example 5.14 (A Variant of the Increasing Oscillating Antichain)**.** We present here the variant of the increasing oscillating antichain, which, instead of having a pair of points at the top of the sequence to form an anchor, has a single point acting, essentially, as a left pin. The first two elements are 542163 and 74216385, and its nth term is $(2n+3)4216385\cdots(2n+4)(2n+1)$ (see Figure 5.6). A similar argument to Example 5.13 may be used to prove that it is an antichain.

5.3.1 Partial Well Order

While every basis forms an antichain, be it finite or infinite, we may also be interested in whether a class contains infinite antichains. A partial order is said to be a *partial well order* if it contains neither an infinite properly decreasing sequence nor an infinite antichain. In the case of permutation classes this first condition is always true (by the existence of a smallest element), and so a permutation class is *partially well ordered* if it contains no infinite antichain. For example, Knuth [76] shows that the set of stack sortable permutations, $\mathrm{Av}(231)$ is partially well ordered.

The decidability problem of whether a given permutation class is partially well ordered remains open:

Question 5.15. *Is it possible to decide if a permutation class given by a finite basis is partially well ordered?*[6]

Indeed there has been no recent major progress on the general problem. Alongside a variety of specific examples, Atkinson, Murphy and Ruškuc [9] showed that $\mathrm{Av}(\beta)$ is partially well ordered if and only if $\beta \in \{1, 12, 21, 132, 213, 231, 312\}$.

Showing that a class is not partially well ordered is simply a case of spotting an antichain inside it. For example, the class $\mathrm{Av}(321)$ contains the increasing oscillating antichain presented above. A non-partially well ordered class may contain many infinite antichains, but among them there must be at least one fundamental antichain.

Proposition 5.16 (Gustedt [66])**.** *Every non-partially well ordered permutation class contains an infinite fundamental antichain.*

Proof. With an eye toward applying Zorn's lemma, take an infinite descending chain $A^1 \succeq A^2 \succeq \cdots$ of infinite antichains and define

$$A^\infty = \{\alpha : \alpha \text{ is an element of all but finitely many } A^i\text{s}\}.$$

[6] This question is considered in more generality by Cherlin and Latka [34].

First observe that A^∞ is an antichain, and that $A^\infty \preceq A^i$ for all i. We claim that it is also infinite. Suppose to the contrary that A^∞ is finite. Thus A^∞ is a subset of all but finitely many of the A^is; without loss let us assume that it is contained in all the A^is. Now choose $\alpha_1 \in A^1 \setminus A^\infty$. For each $i \geq 2$, because $A^i \succeq A^{i-1}$, we may choose $\alpha_i \in A^i$ such that $\alpha_i \leq \alpha_{i-1}$. This gives a descending chain $\alpha_1 \geq \alpha_2 \geq \ldots$, so because permutation classes have no infinite strictly descending chains, there is some α_∞ and integer I such that $\alpha_i = \alpha_\infty$ for all $i \geq I$. However, this implies that $\alpha_1 \geq \alpha_I = \alpha_\infty \in A^\infty \subseteq A^1$, which requires (because A^1 is an antichain) $\alpha_1 = \alpha_\infty$, a contradiction to our choice of α_1. Thus Zorn's Lemma shows that the set of infinite antichains in a non-partially well ordered class has a minimal element under $\preceq$, as desired. □

Note that if A is a fundamental antichain then its *strict closure*, $\{\pi : \pi < \alpha \in A\}$, is partially well ordered.

On the other hand, showing that a class is partially well ordered is a considerably harder task. The primary tool here is a result of Higman [67], which we now state. We say that (A, M) is an *abstract algebra* if A is a set of elements and M a set of operations, for which each $\mu \in M$ is a k-ary operation, $\mu : A^k \to A$, for some positive integer k. Denote the set of k-ary operations by M_k, and suppose that M_k is empty for every $k > n$ for some n. (Note that we will allow 0-ary operations.) The abstract algebra (A, M) is said to be *minimal* if no subset B of A allows (B, M) to be an abstract algebra.

A partial order $\leq_A$ on the set of elements A is a *divisibility order* on (A, M) if every operation $\mu \in M_k$, $k = 0, 1, \ldots, n$, satisfies,

- $a \leq_A b$ implies $\mu(\mathbf{x}, a, \mathbf{y}) \leq_A \mu(\mathbf{x}, b, \mathbf{y})$,
- $a \leq_A \mu(\mathbf{x}, a, \mathbf{y})$,

where $\mathbf{x}$ and $\mathbf{y}$ are arbitrary sequences comprising elements of A whose lengths sum to $k - 1$. Furthermore, given partial orders $\leq_{M_k}$ on M_k, $k = 0, 1, \ldots, n$, we say that $\leq_A$ is *compatible* with these partial orders if, for $\lambda, \mu \in M_k$,

- $\lambda \leq_{M_k} \mu$ implies $\lambda(\mathbf{x}) \leq_A \mu(\mathbf{x})$ for all $\mathbf{x} \in A^k$.

Theorem 5.17 (Higman [67]). *Suppose that (A, M) is a minimal abstract algebra for which, for some n, the set M_k of k-ary operations in M is partially well ordered for each $k = 0, 1, \ldots, n$ and empty for $k > n$. Then (A, M) is partially well ordered under any divisibility ordering compatible with the orders of M_k.*

Higman's Theorem is applied to prove that a given permutation class is partially well ordered by showing how we may "build" the class from a smaller (very possibly finite) set.

Example 5.18. By our definition in Example 5.4, the class $\mathrm{Av}(2413, 3142)$ of separable permutations is precisely the strong completion of the class $\{1\}$, i.e. the class formed from the permutation 1 using the binary operations $\oplus$ and $\ominus$. Higman's Theorem may now immediately be applied to show that $\mathrm{Av}(2413, 3142)$ is partially well ordered.

A permutation class $\mathcal{C}$ is *strongly finitely based* if it is finitely based and every closed subset of $\mathcal{C}$ is also finitely based.[7] Recalling that the basis of a class is an antichain, this definition immediately returns us to partial well order, and indeed we have a variety of equivalent conditions. A formal proof is provided by Atkinson, Murphy and Ruškuc [9].

Proposition 5.19. *Let $\mathcal{C}$ be a permutation class. Then the following are equivalent:*

(1) $\mathcal{C}$ is strongly finitely based.

(2) $\mathcal{C}$ has at most countably many closed subsets.

(3) $\mathcal{C}$ contains no infinite antichain.

(4) The subclasses of $\mathcal{C}$ satisfy the descending chain condition.

Partial well order also plays a rôle in some enumeration attempts. Klazar [75] shows that the smallest growth rate which admits uncountably many closed permutation classes lies between 2 and $2.33529\ldots$. This growth rate is determined by the smallest growth rate that a non-partially well ordered class can have – by Proposition 5.19, such a class will

[7] Higman [67] refers to this as the "finite basis property."

have uncountably many closed subsets, each of which cannot have a growth rate larger than the parent class. The lower bound arises by showing all classes with growth rate under 2 contain only finitely long alternations and oscillations, and these classes – via Higman – are partially well ordered. The upper bound arises by considering the class $\mathrm{Av}(321, 4123, 3412, 23451)$, and noting that it contains the increasing oscillating antichain (hence is not partially well ordered). This class has rational generating function

$$f(x) = \frac{x^5 + x^4 + x^3 + x^2 + x}{1 - x - 2x^2 - 2x^3 - x^4 - x^5}$$

and the growth rate $2.33529\ldots$ arises as the reciprocal of the smallest real root of the denominator (in fact, it is the only real root). Klazar mentions that Vatter and Murphy [private communication] can improve the upper bound to $2.20556\ldots$. The class which satisfies this is formed by appending the basis elements 134526, 134625, 314526 and 314625 to $\mathrm{Av}(321, 4123, 3412, 23451)$, and its growth rate is the dominant root of $x^3 - 2x^2 - 1$.

More recently, Vatter [117] proved that the bound is precisely $2.20556\ldots$ by computing the growth rates of all partially well ordered classes, a task relying on Proposition 5.22. He also makes the following conjecture:

Conjecture 5.20. *Every growth rate of permutation classes is also the growth rate of a partially well ordered permutation class.*

5.3.2 Atomicity

Recall in Subsection 5.1.2 how the union of two finitely based classes is again finitely based. It follows (by considering symmetries, if necessary) that the union of two strongly finitely based classes is again strongly finitely based, and subsequently we have the following.

Proposition 5.21 (Atkinson, Murphy and Ruškuc [9, Lemma 2.1])**.** *The union of a finite number of finitely based partially well ordered permutation classes is partially well ordered and finitely based.*

Conversely, how can we "break up" partially well ordered classes into a union of smaller "unbreakable" classes? This question motivates the study of atomic classes; a permutation class is *atomic* if it cannot be expressed as the union of two proper subclasses. This definition then allows us to provide a converse, as introduced in [9], though here we present an alternative proof based on the descending chain condition, first seen in Murphy's PhD thesis [97].

Proposition 5.22 (Atkinson, Murphy and Ruškuc [9, Theorem 2.2] and Murphy [97, Proposition 188]). *Every partially well ordered permutation class can be written as a finite union of atomic classes.*

Proof. Consider the binary tree whose root is the partially well ordered class $\mathcal{C}$, whose leaves are all atomic classes, and in which the children of the non-atomic class $\mathcal{D}$ are two proper subclasses $\mathcal{D}', \mathcal{D}'' \subsetneq \mathcal{D}$ such that $\mathcal{D}' \cup \mathcal{D}'' = \mathcal{D}$. Because $\mathcal{C}$ is partially well ordered its subclasses satisfy the descending chain condition by Proposition 5.19, so this tree contains no infinite paths and thus is finite. Its leaves give the desired atomic classes. □

In some sense, atomic classes can therefore be considered as the elemental classes from which all others are constructed by taking unions. In practice, however, outwith the comfortable realm of partial well order, atomicity does not behave as elegantly as we might hope – we can, for example, encounter atomic classes that are the union of infinitely many pairwise incomparable atomic classes (see Proposition 170 of Murphy [97]), while there are non-atomic finitely based classes which contain infinitely based maximal atomic subclasses (Proposition 186 of Murphy [97]). In its defence, however:

Proposition 5.23 (Murphy [97, Proposition 171]). *Every permutation class can be written as a union of maximal atomic classes.*

The question of uniqueness for this decomposition, however, falls short of what we would like. To ensure a union $\bigcup_{i\in I} \mathcal{C}^i$ of maximal atomic classes is unique, we must ensure that they are *independent*, that is, for every $i \in I$ we have

$$\bigcup_{j\neq i} \mathcal{C}^j \subset \bigcup_{j} \mathcal{C}^j,$$

and this is not always obtainable. Meanwhile, there remains the question of decidability:

Question 5.24. *Is it possible to decide whether a permutation class given by a finite basis is atomic?*

As with partial well-order, a general answer to this seems far off, though answers in specific cases are often obtainable. Cherlin, Shelah and Shi [33], however, suggest that the problem for general relational structures is not decidable.

Our toolbox for this question consists of a variety of equivalent definitions for atomicity. A class $\mathcal{C}$ is said to satisfy the *joint embedding property* if, for any two permutations α and β in $\mathcal{C}$, there exists π such that $\alpha \leq \pi$ and $\beta \leq \pi$.

Theorem 5.25 (Fraïssé [56]). *The following conditions on a permutation class $\mathcal{C}$ are equivalent:*

(1) $\mathcal{C} = \mathrm{Sub}(f : A \to B)$ for some linearly ordered sets A, B and bijection f.

(2) $\mathcal{C}$ cannot be expressed as a union of two proper closed subsets.

(3) $\mathcal{C}$ satisfies the joint embedding property.

(4) $\mathcal{C}$ contains permutations $\alpha_1 \leq \alpha_2 \leq \ldots$ such that for every $\pi \in \mathcal{C}$ we have $\pi \leq \alpha_n$ for some n.

Every sum, skew or strongly complete class is atomic. For example, given α and β in a sum complete class $\mathcal{C}$, we have $\alpha \oplus \beta \in \mathcal{C}$ and so $\mathcal{C}$ satisfies the joint embedding property. Since every permutation must be either sum or skew decomposable, it follows by Proposition 5.8 that every class having just one basis element is sum or skew complete, and hence atomic. Beyond that, however, decidability is not known – for example, we may write the class $\mathrm{Av}(321, 2143)$ as $\mathrm{Av}(321, 2143, 3142) \cup \mathrm{Av}(321, 2143, 2413)$.

Restricting our view to *natural classes* – that is, atomic classes defined via bijections of the natural numbers $f : \mathbb{N} \to \mathbb{N}$ – Atkinson, Murphy and Ruškuc [11] proved that it is decidable whether a finitely based permutation class is natural. It may also be decidable in other special cases; the author tried – and failed – to derive similar conditions for the "rational" case, namely $f : \mathbb{Q} \to \mathbb{Q}$.

5.4 Permutation Classes and Simple Permutations

By the central rôle which simple permutations take in forming the building blocks of permutations, it is not surprising that they also perform a similarly crucial job within permutation classes. Clearly every permutation of a class $\mathcal{C}$ may be broken down by means of its substitution decomposition, using only $\mathrm{Si}(\mathcal{C})$, the simple permutations from $\mathcal{C}$. For example, in the class $\mathcal{S} = \mathrm{Av}(2413, 3142)$ of separable permutations, we have $\mathrm{Si}(\mathcal{S}) = \{1, 12, 21\}$, and every permutation in $\mathcal{S}$ can be formed by repeated inflations of 12 and 21.

The converse, of course, is not true in general: we cannot reconstruct a class $\mathcal{C}$ by taking every possible inflation of the simple permutations $\mathrm{Si}(\mathcal{C})$ (for example, $\mathrm{Si}(\mathrm{Av}(231)) = \{1, 12, 21\}$, but $231 = 21[12, 1]$). This can only be done when a permutation class is wreath closed, as such a class then contains every inflation by its very definition.

When the set of simple permutations is infinite, there is not a great deal more that can be said. There is, however, a seemingly vast array of permutation classes that contain only finitely many simple permutations, and in this case there is much to say. In this section we will review a number of the known results, before contributing several more new results in Chapters 6 and 7.

Counting Simple Permutations. A first step towards determining whether a class contains only finitely many simple permutations is to use the Schmerl-Trotter Theorem 2.1 (found on page 21). By simply counting the simple permutations of size $n = 1, 2, \ldots$, if we encounter two consecutive lengths where there are no simple permutations, then the class can contain no longer simple permutations. For example, the number of simple permutations in $\mathrm{Av}(1324, 2143, 4231)$ of lengths 1 to 7 is $1, 2, 0, 2, 4, 0, 0$, and so the longest simple permutations in this class are of length 5. We will present a complete answer to this decidability problem in Chapter 7.

5.4.1 Finitely Many Simples

Classes with only finitely many simple permutations have nice properties. To name the three most significant: these classes have algebraic generating functions, are partially well ordered, and are finitely based. We will consider each of these topics in turn.

Algebraic Generating Functions. Albert and Atkinson [2] showed how every class containing only finitely many simple permutations is enumerated by an algebraic generating function, and this function is readily computable. This should come as no great surprise – expressing all permutations in such a class as the inflation of a simple skeleton gives us a recursive construction, in much the same way as when we enumerated the stack sortable permutations (Example 5.11), and such recursions immediately suggest that we should expect an algebraic generating function. We prove this fact, and a much more general result, in Chapter 6.

Partial Well Order. Since antichains (or, at least, fundamental antichains) rely heavily on the structure of simple permutations to maintain their incomparability (as witnessed by the statement of Conjecture 5.12), we can reasonably expect a permutation class containing only finitely many simple permutations to be partially well ordered. Before showing this, however, we exhibit an observation about partial well order that we will need.

Proposition 5.26. *The product $(P_1, \leq_1) \times \cdots \times (P_s, \leq_s)$ of a collection of partial orders is partially well ordered if and only if each of them is partially well ordered.*

Without further ado, we may now proceed to the desired result. Our proof follows Gustedt [66], although note that Albert and Atkinson [2] give a different proof, using Higman's Theorem 5.17.

Proposition 5.27 (Gustedt [66])**.** *Every permutation class with only finitely many simple permutations is partially well ordered.*

Proof. Suppose to the contrary that the class $\mathcal{C}$ contains an infinite antichain but only finitely many simple permutations. By Proposition 5.16, $\mathcal{C}$ contains an infinite fundamental antichain. Moreover, there is an infinite subset A of this antichain for which every element is an inflation of the same simple permutation, say σ. Let $\mathcal{D}$ denote the strict closure of A and note that A is also fundamental, so $\mathcal{D}$ is partially well ordered. It is easy to see that the permutation containment order, when restricted to inflations of σ, is isomorphic to a product order: $\sigma[\alpha_1, \ldots, \alpha_m] \le \sigma[\alpha'_1, \ldots, \alpha'_m]$ if and only if $\alpha_i \le \alpha'_i$ for all $i \in [m]$. However, this implies that A is an infinite antichain in a product $\mathcal{D} \times \cdots \times \mathcal{D}$ of partially well ordered posets, contradicting Proposition 5.26. □

Finitely Based. That a class containing only finitely many simple permutations is finitely based arises by first considering its wreath closure. Our first task is to compute the basis of a wreath closed class containing only finitely many simple permutations, which may be done using the Schmerl-Trotter Theorem 2.1 (Page 21):

Proposition 5.28. *If the longest simple permutations in $\mathcal{C}$ have length k then the basis elements of $\mathcal{W}(\mathcal{C})$ have length at most $k + 2$.*

Proof. The basis of $\mathcal{W}(\mathcal{C})$ is easily seen to consist of the minimal (under the pattern containment order) simple permutations not contained in $\mathcal{C}$ (cf. Proposition 5.9). Let π be such a permutation of length n. Theorem 2.1 shows that π contains a simple permutation σ of length $n - 1$ or $n - 2$. If $n \ge k + 3$, then $\sigma \notin \mathcal{C}$, so $\sigma \notin \mathcal{W}(\mathcal{C})$ and thus π cannot lie in the basis of $\mathcal{W}(\mathcal{C})$. □

For example, using this Proposition it can be computed that the wreath closure of 1, 12, 21, and 2413 is $\mathrm{Av}(3142, 25314, 246135, 362514)$ – we will encounter this class again in Example 6.10.

By Proposition 5.27, any permutation class – and in particular any wreath closed class – containing only finitely many simples is partially well ordered. Subsequently:

Theorem 5.29 (Albert and Atkinson [2]). *Every permutation class containing only finitely many simple permutations is finitely based.*

Proof. Let $\mathcal{C}$ be a class containing only finitely many simple permutations. By Proposition 5.28, $\mathcal{W}(\mathcal{C})$ is finitely based, and by Proposition 5.27 it is partially well ordered. The class $\mathcal{C}$ must therefore avoid all elements in the basis of $\mathcal{W}(\mathcal{C})$, together with the minimal elements of $\mathcal{W}(\mathcal{C})$ not belonging to $\mathcal{C}$, which form an antichain. By its partial well ordering any antichain in $\mathcal{W}(\mathcal{C})$ is finite, and so there can only be finitely many basis elements of $\mathcal{C}$. □

5.5 The Containment Partial Order in Other Structures

We may, of course, define the containment order on any relational structure and treat it as a partial order. Expanding upon the notion of extensions in Chapter 3, if $\mathcal{A}$ and $\mathcal{B}$ are relational structures over a common language $\mathcal{L}$ then an *embedding* of $\mathcal{A}$ into $\mathcal{B}$ is an injection $\varphi : \mathrm{dom}(\mathcal{A}) \to \mathrm{dom}(\mathcal{B})$ so that $\mathcal{B}|_{\varphi(\mathrm{dom}(\mathcal{A}))}$ is isomorphic to $\mathcal{A}$. If such an embedding exists, then we say $\mathcal{A} \leq \mathcal{B}$, a quasi order from which we may induce a partial order by considering the equivalence classes $\mathcal{A} \cong \mathcal{B}$, arising if and only if $\mathcal{A} \leq \mathcal{B}$ and $B \leq \mathcal{A}$.

In theory, one may then study any closed class of relational structures for a given language in the same way as one might study permutation classes. Formally, a set $\mathfrak{C}$ of relational structures over a common relational language $\mathcal{L}$ is an $\mathcal{L}$*-class* if $\mathcal{A} \in \mathfrak{C}$ and $\mathcal{B} \leq \mathcal{A}$ implies $\mathcal{B} \in \mathfrak{C}$. We might then if we wished define an $\mathcal{L}$-class in terms of structure avoidance and try to compute its generating function. We could consider intersections, unions and, by recalling the definition of inflation in this general setting, wreath products and wreath closures.

Antichains, partial well order and atomicity are notions taken from the theory of posets. Antichains are merely sets of pairwise incomparable elements; see Gustedt [66] for notions of minimality in antichains and some considerations on the existence of infinite antichains. Since every $\mathcal{L}$-class has a minimal element on one point, no $\mathcal{L}$-class can contain an infinite properly decreasing sequence. Thus an $\mathcal{L}$-class $\mathfrak{C}$ is *partially well ordered* if it contains no infinite antichains, and Higman's Theorem can be used in the general setting. Atomicity in the permutation class case is merely a special case of the "γ classes" of Fraïssé [56]; many of the results that are true for permutation classes are also true in the general case. For example, an atomic $\mathcal{L}$-class $\mathfrak{C}$ satisfies the joint embedding property, and is also expressible in a way analogous to the $\mathrm{Sub}(\pi : A \to B)$ notation. See Fraïssé [56], Hodges [68, Section 7.1], and, for a survey of more recent results, Pouzet [101].

Finitely many Simples. By means of the substitution decomposition, $\mathcal{L}$-classes which contain only finitely many simple $\mathcal{L}$-structures will have a recursive construction much as in the permutation class case. However in the general setting this does not correspond to an algebraic generating function, since structures in the partial order are defined only up to equivalence. In fact, it seems that having an algebraic generating function is special to the permutation case (for example, it is not true in the graph case).

All such $\mathcal{L}$-classes are, however, partially well ordered. As in the permutation case, antichains are instrinsically linked to simple permutations, and Proposition 5.27 is proved in the general case by Gustedt [66].

To answer the question of whether these classes are finitely based, we may obtain a partial answer by considering the most general setting of the Schmerl-Trotter Theorem 2.1 given in [107], namely that of binary, irreflexive relational structures, a set which includes graphs, tournaments and posets. Ehrenfeucht and McConnell [48] show that, for $k \geq 3$, a simple structure defined on a single k-ary relation must contain a simple substructure with k, $k-1$ or $k-2$ fewer points, and this was improved to just $k-1$ or $k-2$ fewer points by Bonizzoni and McConnell [23]. Further generalisations remain unknown.

The Graph Case. The "graph containment order" is in fact the order defined by induced subgraphs, and has been extensively studied. As with many other relational structures, classes of graphs closed under taking induced subgraphs are more often referred to as *hereditary properties*. A stronger condition is obtained by considering sets of graphs closed under taking subgraphs (rather than induced subgraphs), and these are referred to as *monotone* properties.

Properties need not be hereditary – consider, for example, the property consisting of all regular graphs. Examples of hereditary properties include the set of triangle-free graphs, all graphs of chromatic number at most k and the set of split graphs (graphs which may be partitioned into an independent set and a clique).

As with permutation classes, much of the study of hereditary graph properties is in their asymptotic enumeration. For a property $\mathcal{P}$, let $\mathcal{P}_n$ denote the set of graphs in $\mathcal{P}$ with n vertices, whence the function $|\mathcal{P}_n|$ defines the *speed* of the property. While little can be said about the speed of an arbitrary property, Scheinerman and Zito [106] prove that the speed of hereditary graph properties must, for sufficiently large n, be constant, polynomial, exponential, factorial or superfactorial. Subsequent study – in particular Balogh, Bollobás and Weinreich [15, 16] – has shown that there are many "jumps" within this already broken spectrum of speeds.

CHAPTER 6

ALGEBRAIC GENERATING FUNCTIONS

6.1 Introduction

When a class is enumerated by an algebraic generating function, we intuitively expect to find some recursive description of the permutations in the class. Such descriptions may arise in a variety of ways, but one of the most important is the substitution decomposition.

In a class which has only finitely many simple permutations, therefore, any long permutation must map nontrivial intervals onto intervals, and hence all the permutations of the class are constructed recursively via the substitution decomposition. With only finitely many simple permutations on which to "build", we expect the class to have an algebraic generating function:

Theorem 6.1 (Albert and Atkinson [2])**.** *A permutation class with only finitely many simple permutations has a readily computable algebraic generating function.*

Our aim in this chapter is to establish a generalisation of Theorem 6.1. We do this by observing that the recursive construction given by the substitution decomposition is not a feature merely of pattern avoidance in the containment order, but can be extended to enumerate a wide variety of other sets of permutations. In essence it can be extended to enumerate any set of permutations which can be built in the same way from a finite set of simple permutations, though we will still require that the set lies within a permutation class with only finitely many simple permutations.

Theorem 6.2. *Let $\mathcal{C}$ be a permutation class containing only finitely many simple permutations, $\mathcal{P}$ a finite query-complete set of properties, and $\mathcal{Q} \subseteq \mathcal{P}$. The generating function for the set of permutations in $\mathcal{C}$ satisfying every property in $\mathcal{Q}$ is algebraic over $\mathbb{Q}(x)$.*

The next section establishes the terminology required by Theorem 6.2, which we will then prove in Section 6.3. Section 6.4 shows how to describe some common families of permutations as query-complete sets of properties and hence demonstrates the scope of Theorem 6.2, with specific worked examples given in Section 6.5. In Sections 6.6 and 6.7 we adapt these techniques to enumerate two further families, namely involutions and cyclic closures, respectively. Some closing remarks are given in Section 6.8.

6.2 Properties and Query-completeness

As we saw at the end of Chapter 5, the term "property" has been used extensively in the study of other relational structures, and particularly in graph theory. It is natural, therefore, to use this term in the context of permutations in a similar way. To this end, define a *property*, P, to be any set of permutations, and say that a permutation π *satisfies* P if $\pi \in P$. Note that a permutation class is now simply an example of a property.

A set $\mathcal{P}$ of properties is *query-complete* if, for each simple permutation σ of length m and property $P \in \mathcal{P}$, there is a procedure to determine whether $\sigma[\alpha_1, \dots, \alpha_m]$ satisfies P based only on the knowledge of which properties of $\mathcal{P}$ each α_i satisfies. For example, the set of properties consisting of the 132-avoiding permutations, $\{\mathrm{Av}(132)\}$, is not query-complete, as witnessed by the fact that $12[1,1] \in \mathrm{Av}(132)$ but $12[1,21] \notin \mathrm{Av}(132)$, while both 1 and 12 avoid 132. However, $\{\mathrm{Av}(132), \mathrm{Av}(21)\}$ is query-complete:

$$\begin{aligned}
12[\alpha_1,\alpha_2] \in \mathrm{Av}(132) &\iff \alpha_1 \in \mathrm{Av}(132) \text{ and } \alpha_2 \in \mathrm{Av}(21),\\
21[\alpha_1,\alpha_2] \in \mathrm{Av}(132) &\iff \alpha_1 \in \mathrm{Av}(132) \text{ and } \alpha_2 \in \mathrm{Av}(132),\\
\sigma[\alpha_1,\dots,\alpha_m] \notin \mathrm{Av}(132) &\quad \text{if} \quad \sigma \notin \{1,12,21\} \text{ is simple},\\
12[\alpha_1,\alpha_2] \in \mathrm{Av}(21) &\iff \alpha_1 \in \mathrm{Av}(21) \text{ and } \alpha_2 \in \mathrm{Av}(21),\\
\sigma[\alpha_1,\dots,\alpha_m] \notin \mathrm{Av}(21) &\quad \text{if} \quad \sigma \notin \{1,12\} \text{ is simple}.
\end{aligned}$$

Note that since $\sigma[\alpha_1,\ldots,\alpha_m]$ is uniquely determined by σ and the α_is, every property P lies in some query-complete set, e.g., $\{P\} \cup \{\{\pi\} : \pi \text{ a permutation}\}$ is query-complete for every P. Thus the finiteness condition in Theorem 6.2 is essential. Another observation about query-complete sets, which will be liberally applied, is the following.

Proposition 6.3. *A union of query-complete sets of properties is itself query-complete.*

6.3 Proof of Main Result

We begin by recalling the substitution decomposition for permutations, which is encapsulated in two propositions from Chapter 1.

Proposition 1.7. *Every permutation may be written as the inflation of a unique simple permutation. Moreover, if π can be written as $\sigma[\alpha_1,\ldots,\alpha_m]$ where σ is simple and $m \geq 4$, then the α_is are unique.*

Proposition 1.8. *If π is an inflation of 12, then there is a unique sum indecomposable α_1 such that $\pi = 12[\alpha_1,\alpha_2]$ for some α_2, which is itself unique. The same holds with 12 replaced by 21 and "sum" replaced by "skew".*

Given a permutation class $\mathcal{C}$ and set $\mathcal{P}$ of properties, we write $\mathcal{C}_{\mathcal{P}}$ for the set of permutations in $\mathcal{C}$ that satisfy every property in $\mathcal{P}$, and write $f_{\mathcal{P}}$ for the generating function of $\mathcal{C}_{\mathcal{P}}$. Before beginning the proof of Theorem 6.2 we consider the case where $\mathcal{C}$ is wreath-closed and $\mathcal{P} = \emptyset$, which contains many of the main ideas of the proof in a more digestible form. (This presentation borrows heavily from Albert and Atkinson [2].)

We begin by introducing two properties,

$$\begin{aligned} \not\oplus &= \{\text{sum indecomposable permutations}\} \text{ and} \\ \not\ominus &= \{\text{skew indecomposable permutations}\}. \end{aligned}$$

Note that both $\{\not\oplus\}$ and $\{\not\ominus\}$ are query-complete, because for simple σ,

$$\begin{aligned} \sigma[\alpha_1,\ldots,\alpha_m] \in \not\oplus &\iff \sigma \neq 12 \text{ and} \\ \sigma[\alpha_1,\ldots,\alpha_m] \in \not\ominus &\iff \sigma \neq 21. \end{aligned}$$

We also introduce the notation

$$\sigma[\mathcal{C}^1, \dots, \mathcal{C}^m] = \{\sigma[\alpha_1, \dots, \alpha_m] : \alpha_i \in \mathcal{C}^i \text{ for all } i \in [m]\}.$$

By Propositions 1.7 and 1.8 and the assumption that $\mathcal{C}$ is wreath-closed, $\mathcal{C}$ can be written as

$$\mathcal{C} = \{1\} \uplus 12[\mathcal{C}_{\not\oplus}, \mathcal{C}] \uplus 21[\mathcal{C}_{\not\ominus}, \mathcal{C}] \uplus \biguplus_{\substack{\sigma \in \mathrm{Si}(\mathcal{C}) \\ |\sigma| \geq 4}} \sigma[\mathcal{C}, \dots, \mathcal{C}],$$

while $\mathcal{C}_{\not\oplus}$ and $\mathcal{C}_{\not\ominus}$ have the expressions

$$\begin{aligned} \mathcal{C}_{\not\oplus} &= \{1\} \uplus 21[\mathcal{C}_{\not\ominus}, \mathcal{C}] \uplus \biguplus_{\substack{\sigma \in \mathrm{Si}(\mathcal{C}) \\ |\sigma| \geq 4}} \sigma[\mathcal{C}, \dots, \mathcal{C}] &= \mathcal{C} \setminus 12[\mathcal{C}_{\not\oplus}, \mathcal{C}], \\ \mathcal{C}_{\not\ominus} &= \{1\} \uplus 12[\mathcal{C}_{\not\oplus}, \mathcal{C}] \uplus \biguplus_{\substack{\sigma \in \mathrm{Si}(\mathcal{C}) \\ |\sigma| \geq 4}} \sigma[\mathcal{C}, \dots, \mathcal{C}] &= \mathcal{C} \setminus 21[\mathcal{C}_{\not\ominus}, \mathcal{C}]. \end{aligned}$$

These give the system

$$\begin{cases} f = x + f_{\not\oplus} f + f_{\not\ominus} f + \displaystyle\sum_{\substack{\sigma \in \mathrm{Si}(\mathcal{C}) \\ |\sigma| \geq 4}} f^{|\sigma|}, \\ f_{\not\oplus} = x + f_{\not\ominus} f + \displaystyle\sum_{\substack{\sigma \in \mathrm{Si}(\mathcal{C}) \\ |\sigma| \geq 4}} f^{|\sigma|} = f - f_{\not\oplus} f = \dfrac{f}{1+f}, \\ f_{\not\ominus} = x + f_{\not\oplus} f + \displaystyle\sum_{\substack{\sigma \in \mathrm{Si}(\mathcal{C}) \\ |\sigma| \geq 4}} f^{|\sigma|} = f - f_{\not\ominus} f = \dfrac{f}{1+f}. \end{cases}$$

If we now let s denote the generating function for the simple permutations of length at least 4 in $\mathcal{C}$, we find that

$$f = x + \frac{2f^2}{1+f} + s(f),$$

so if s is algebraic, a fortiori if s is polynomial, f is algebraic. In particular, note that the separable permutations correspond to $s = 0$; substituting this value for s leaves $f = x + 2f^2/(1+f)$, and so we have proved that the generating function for the separables is

$$f = \frac{1 - x - \sqrt{1 - 6x + x^2}}{2} = x + 2x^2 + 6x^3 + 22x^4 + 90x^5 + \dots$$

giving the large Schröder numbers (sequence A006318 of [110]).

The following brief review of algebraic systems is a specialisation of the more general treatment in Stanley [114, Section 6.6]. Let $A = \{a_1, \ldots, a_n\}$ denote an alphabet. A *proper algebraic system* over $\mathbb{Q}[x_1, \ldots, x_m]$ is a set of equations $a_i = p_i(x_1, \ldots, x_m, a_1, \ldots, a_n)$ where each p_i is a polynomial with coefficients from $\mathbb{Q}$, has constant term 0, and contains no terms of the form ca_j where $c \in \mathbb{Q}$. The solution to such a system is a tuple $(f_1, \ldots, f_n)$ of formal power series from $\mathbb{Q}[[x_1, \ldots, x_m]]$ such that for all i, f_i is equal to $p_i(x_1, \ldots, x_m, a_1, \ldots, a_n)$ evaluated at $(a_1, \ldots, a_n) = (f_1, \ldots, f_n)$.

Theorem 6.4 (Stanley [114, Proposition 6.6.3 and Theorem 6.6.10])**.** *Every proper algebraic system $(p_1, \ldots, p_n)$ over $\mathbb{Q}[x_1, \ldots, x_m]$ has a unique solution $(f_1, \ldots, f_n)$. Moreover, each of these f_is is algebraic over $\mathbb{Q}(x_1, \ldots, x_m)$.*

The proof of Theorem 6.2 now follows, modulo the result of Lemma 6.5.

Theorem 6.2. *Let $\mathcal{C}$ be a permutation class containing only finitely many simple permutations, $\mathcal{P}$ a finite query-complete set of properties, and $\mathcal{Q} \subseteq \mathcal{P}$. The generating function for the set of permutations in $\mathcal{C}$ satisfying every property in $\mathcal{Q}$, i.e., $f_{\mathcal{Q}}$, is algebraic over $\mathbb{Q}(x)$.*

Proof. Let B denote the basis of $\mathcal{C}$, which is finite by Theorem 5.29 (on Page 106). Lemma 6.5 shows that for every $\beta \in B$, the property $\mathrm{Av}(\beta)$ lies in a finite query-complete set. Thus the set $\{\mathrm{Av}(\beta) : \beta \in B\}$ is contained in a finite query-complete set, and we have

$$\mathcal{C} = \mathcal{W}(\mathcal{C})_{\{\mathrm{Av}(\beta):\beta\in B\}}.$$

Therefore it suffices to prove the theorem for wreath-closed classes. Furthermore, if $\mathcal{P}$ is query-complete then $\mathcal{P} \cup \{\not\oplus, \not\ominus\}$ is also query-complete, so we may assume without loss that $\not\oplus, \not\ominus \in \mathcal{P}$.

Let $\mathcal{P}(\pi)$ denote the set of properties in $\mathcal{P}$ satisfied by π and, avoiding inclusion-exclusion, let $g_{\mathcal{R}}$ denote the generating function for the set of $\pi \in \mathcal{C}$ with $\mathcal{P}(\pi) = \mathcal{R}$, so

$$f_{\mathcal{Q}} = \sum_{\mathcal{Q}\subseteq\mathcal{R}\subseteq\mathcal{P}} g_{\mathcal{R}}.$$

As $\mathcal{P}$ is query-complete, for each simple σ, $\mathcal{P}(\sigma[\alpha_1, \ldots, \alpha_m])$ is completely determined by σ and $\mathcal{P}(\alpha_1), \ldots, \mathcal{P}(\alpha_m)$. Thus for each simple σ of length m, there is a finite collection of m-tuples of sets of properties such that $\mathcal{P}(\sigma[\alpha_1, \ldots, \alpha_m]) = \mathcal{R}$ precisely if $(\mathcal{P}(\alpha_1), \ldots, \mathcal{P}(\alpha_m))$ lies in this collection. If $m \geq 4$ then Proposition 1.7 implies that the generating function for all inflations π of σ with $\mathcal{P}(\pi) = \mathcal{R}$ can be expressed nontrivially as a polynomial in $\{g_{\mathcal{S}} : \mathcal{S} \subseteq \mathcal{P}\}$ of degree m. If $m = 2$, suppose $\sigma = 12$ without loss. By Proposition 1.8, all inflations of 12 have a unique decomposition as $12[\alpha_1, \alpha_2]$ where $\alpha_1 \in \not\oplus$. Thus the generating function for inflations π of 12 with $\mathcal{P}(\pi) = \mathcal{R}$ can be expressed as a sum of terms of the form $g_{\mathcal{S}} g_{\mathcal{T}}$ where $\not\oplus \in \mathcal{S}$.

Therefore $g_{\mathcal{R}}$ can be expressed as a polynomial in x (depending on whether $\mathcal{P}(1) = \mathcal{R}$) and $\{g_{\mathcal{S}} : \mathcal{S} \subseteq \mathcal{P}\}$. Moreover, these polynomials have no constant terms and no terms of the form $cg_{\mathcal{S}}$ for constant $c \neq 0$. Thus they form a proper algebraic system, so Theorem 6.4 implies that each $g_{\mathcal{S}}$ is algebraic. $\square$

6.4 Finite Query-Complete Sets

We exhibit several query-complete sets of properties in this section. The first of these is necessary for the proof of Theorem 6.2, the others for Corollary 6.21.

Lemma 6.5. *For every permutation β, the set $\{\mathrm{Av}(\delta) : \delta \leq \beta\}$ is query-complete.*

Proof. We prove the lemma by induction on the length of β. The base case $\beta = 1$ being trivial, let us suppose that β is of length at least 2. By induction, $\{\mathrm{Av}(\gamma) : \gamma \leq \delta\}$ is query-complete for all $\delta < \beta$, and thus by appealing to Proposition 6.3 it suffices to prove that whether $\pi = \sigma[\alpha_1, \ldots, \alpha_m]$ satisfies $\mathrm{Av}(\beta)$ can be decided entirely by knowing, for each i, which permutations δ satisfy $\delta \leq \alpha_i$ and $\delta \leq \beta$.

We define a *lenient inflation* to be an inflation $\sigma[\gamma_1, \ldots, \gamma_m]$ in which the γ_is are allowed to be empty. List all expressions of β as a lenient inflation of σ as

$$\begin{aligned} \beta &= \sigma[\gamma_1^{(1)}, \ldots, \gamma_m^{(1)}], \\ &\vdots \\ \beta &= \sigma[\gamma_1^{(t)}, \ldots, \gamma_m^{(t)}]. \end{aligned}$$

Clearly if we have, for some $s \in [t]$, $\alpha_i \geq \gamma_i^{(s)}$ for all $i \in [m]$, then $\pi \geq \beta$. Equivalently, to have $\pi \in \mathrm{Av}(\beta)$, for every $s \in [t]$ there must be at least one $i \in [m]$ for which $\alpha_i \ngeq \gamma_i^{(s)}$. Conversely, every embedding of β into π gives one of the lenient inflations in the list above, which completes the proof. □

In a *barred permutation*, one or more of the entries is barred; for π to avoid the barred permutation σ means that every set of entries of π order isomorphic to the nonbarred entries of σ can be extended to a set order isomorphic to σ itself. For example, 24315 avoids $21\overline{3}$ because every inversion (i.e., copy of 21) can be extended to a copy of 213 (append the 5), but 24315 contains $3\overline{1}2$ because the 3 and 1 of 24315 are order isomorphic to 32, but there is no way to extend this to a copy of 312. Barred permutations have arisen several times in the permutation pattern literature. For example, under West's notion of 2-stack sorting (see Example 5.3 on page 84) the permutations that can be sorted are those that avoid 2341 and $3\overline{5}241$, while Bousquet-Mélou and Butler [25] characterise the permutations corresponding to locally factorial Schubert varieties in terms of barred permutations.

A *blocked permutation* is a permutation containing dashes indicating the entries that need not occur consecutively (in the normal pattern-containment order, no entries need occur consecutively), or in the case of the beginning or trailing dashes, entries that need not occur at the beginning or end of the permutation, respectively. For example, 24135 contains only one copy of -1-23-, namely 235; the entries 245 do not form a copy of -1-23- because the 4 and 5 are not adjacent. Babson and Steingrímsson [14] introduced blocked permutations (although they called them generalised patterns, and implicitly assumed that their patterns had beginning and trailing dashes) and showed that they could be used to express most

Mahonian statistics. For example, the major index[1] of π is equal to the total number of copies of -1-32-, -2-31-, -3-21-, and -21- in π.

The proof of Lemma 6.5 extends in a straightforward manner to show that the property of avoiding a blocked or barred permutation (or, for that matter, a permutation combining these restrictions) also lies in a finite query-complete set, although the sets are not so easily described.[2]

The permutation $\pi \in S_n$ is said to be *alternating* if for all $i \in [2, n-1]$, $\pi(i)$ does not lie between $\pi(i-1)$ and $\pi(i+1)$.

Lemma 6.6. *The set of properties consisting of*

- $AL = \{$*alternating permutations*$\}$,
- $BR = \{$*permutations beginning with a rise, i.e., permutations with* $\pi(1) < \pi(2)\}$,
- $ER = \{$*permutations ending with a rise*$\}$, *and*
- $\{1\}$.

is query-complete.

Proof. Clearly $\{\{1\}, BR, ER\}$ is query-complete:

$$\begin{aligned} \sigma[\alpha_1, \ldots, \alpha_m] \in BR &\iff \alpha_1 \in BR \text{ or } (\alpha_1 = 1 \text{ and } \sigma \in BR)\,, \\ \sigma[\alpha_1, \ldots, \alpha_m] \in ER &\iff \alpha_m \in ER \text{ or } (\alpha_m = 1 \text{ and } \sigma \in ER)\,. \end{aligned}$$

For $\pi = \sigma[\alpha_1, \ldots, \alpha_m]$ to be an alternating permutation, we first need $\alpha_1, \ldots, \alpha_m \in AL$. Now suppose that the entries of π up to and including the $\sigma(i)$ interval are alternating (we have this for $i = 1$ from the above). If $\sigma(i) > \sigma(i+1)$ then π contains a descent between its $\sigma(i)$ interval and its $\sigma(i+1)$ interval. Thus α_i is allowed to be 1 (i.e., $\alpha_i \in \{1\}$) only if $i = 1$

[1] The major index is more commonly defined as the sum of the descents of π, $\displaystyle\sum_{\pi(i) > \pi(i+1)} i$.

[2] Consider, e.g., the problem of deciding whether $\pi = 3142[\alpha_1, \alpha_2, \alpha_3, \alpha_4]$ avoids -1-23-. First, each of the α_i's must avoid -1-23-. Then we also need α_3 and α_4 to not contain ascents (i.e., avoid -12-) since α_2 is nonempty, and α_2 to avoid -1-2, since otherwise the third element of the -1-23- could be chosen from α_3.

or $\sigma(i-1) < \sigma(i)$, while if $\alpha_i \neq 1$ then we must have $\alpha_i \in ER$, and whether or not α_i is 1 we must have $\alpha_{i+1} \in BR \cup \{1\}$. The case where $\sigma(i) < \sigma(i+1)$ is analogous, completing the proof. □

Recall that an *even permutation* is one that can be written as the product of an even number of transpositions, or (much more conveniently for our purposes) a permutation with an even number of inversions.

Lemma 6.7. *The set of properties consisting of*

- $EV = \{$*even permutations*$\}$ *and*
- $EL = \{$*permutations of even length*$\}$

is query-complete.

Proof. We have

$$\sigma[\alpha_1, \ldots, \alpha_m] \in EL \iff \text{an even number of } \alpha_i\text{'s fail to lie in } EL,$$

so $\{EL\}$ is query-complete. To see that $\{EV, EL\}$ is query-complete, we divide the inversions in $\sigma[\alpha_1, \ldots, \alpha_m]$ into two groups: inversions within a single $\sigma(i)$ interval and inversions between two intervals $\sigma(i)$ and $\sigma(j)$. We need to compute the parity of each of these numbers. The parity of the first type of inversions depends only on whether $\alpha_i \in EV$. For the second type, suppose $i < j$. If $\sigma(i) < \sigma(j)$ then there are an even number of inversions (more specifically, 0) between the intervals $\sigma(i)$ and $\sigma(j)$ while if $\sigma(i) > \sigma(j)$ then the number of inversions between these intervals is $|\alpha_i||\alpha_j|$, which is even if α_i or α_j lie in EL and odd otherwise. □

We say that the entry $\pi(i)$ *begins a descent* if $\pi(i) > \pi(i+1)$ and *begins an ascent* if $\pi(i) < \pi(i+1)$. A permutation is *Dumont of the first kind* if each even entry begins a descent and each odd entry either begins an ascent or occurs last (this dates back to Dumont [42]). For example, 5642137 is a Dumont permutation of the first kind. We further say that a permutation is *almost Dumont* if every non-terminal even entry begins a descent and every non-terminal odd entry begins an ascent, or *anti-almost Dumont* if every non-terminal odd entry begins a descent and every non-terminal even entry begins an ascent.

Lemma 6.8. *The set of properties consisting of*

- $DU = \{$*Dumont permutations of the first kind*$\}$,
- $AD = \{$*almost Dumont permutations*$\}$,
- $AAD = \{$*anti-almost Dumont permutations*$\}$,
- $EO = \{$*permutations which end with an odd entry*$\}$ *and*
- $EL = \{$*permutations of even length*$\}$

is query-complete.

Proof. First note that $DU = AD \cap EO$, so it suffices to show that $\{AD, AAD, EO, EL\}$ is query-complete. By the proof of Proposition 6.7 we have that $\{EL\}$ is query-complete. Using the EL property, we can determine the parity of the number of entries of lesser value than any given interval; there are an even number of entries below the $\sigma(i)$ interval if and only if an even number of the permutations $\alpha_{\sigma^{-1}(1)}, \alpha_{\sigma^{-1}(2)}, \ldots, \alpha_{\sigma^{-1}(\sigma(i)-1)}$ fail to lie in EL. From this, it follows readily that the set $\{EO, EL\}$ is query-complete: $\sigma[\alpha_1, \ldots, \alpha_m] \in EO$ if $\alpha_m \in EO$ and an even number of entries lie below the $\sigma(m)$ interval, or if $\alpha_m \notin EO$ and an odd number of entries lie below the $\sigma(m)$ interval.

We are reduced to the problem of determining membership in AD and AAD. As the cases are analogous, we consider only the former. Consider the permutation $\pi = \sigma[\alpha_1, \ldots, \alpha_m]$. We divide our task into two parts: first, we check that the entries corresponding to each $\sigma(i)$ interval satisfy the desired properties, and second, we check that the "transitions" between successive intervals satisfy these properties. To resolve the first, for π to lie in AD, we must have that each α_i lies in AD (resp., AAD) if and only if there are an even (resp., odd) number of entries below the $\sigma(i)$ interval. For the second, if $\sigma(i) < \sigma(i+1)$ then the $\sigma(i)$ interval must end in an odd entry. This requires that $\alpha_i \in EO$ if there are an even number of entries below the $\sigma(i)$ interval, and $\alpha_i \notin EO$ otherwise. The $\sigma(i) > \sigma(i+1)$ case follows similarly, completing the proof. $\square$

The imaginative reader should at this point have no trouble constructing many other properties that lie in finite query-complete sets. Examples include the property of beginning with a 1, or more generally of mapping any fixed i to any fixed j, or of having major index congruent to $1 \bmod 3$, or of having an odd number of left-to-right minima.

6.5 Examples

While we have already shown how to enumerate the separable permutations in Section 6.3, here we use the approach of Theorem 6.2.

Example 6.9 (Separable permutations)**.** With the notation from the proof of Theorem 6.2, we have that for the separable permutations:

$$\begin{cases} g_{\not\oplus,\not\ominus} & = & x, \\ g_{\not\oplus} & = & (g_{\not\oplus,\not\ominus} + g_{\not\ominus})(g_{\not\oplus,\not\ominus} + g_{\not\oplus} + g_{\not\ominus}), \\ g_{\not\ominus} & = & (g_{\not\oplus,\not\ominus} + g_{\not\oplus})(g_{\not\oplus,\not\ominus} + g_{\not\oplus} + g_{\not\ominus}), \end{cases}$$

where our universe of properties $\mathcal{P}$ is $\{\not\oplus, \not\ominus\}$. We are interested in $f = g_{\not\oplus,\not\ominus} + g_{\not\oplus} + g_{\not\ominus}$. By summing the three equalities above and simplifying one obtains $f = x + (x + f)f$, which leads, reassuringly, to the generating function for the large Schröder numbers,

$$f = \frac{1 - x - \sqrt{1 - 6x + x^2}}{2}.$$

This system does not change dramatically when another simple permutation is introduced, as shown by the next example.

Example 6.10 (The wreath closure of $1, 12, 21$, and 2413)**.** Here we again take $\mathcal{P} = \{\not\oplus, \not\ominus\}$ and the system is

$$\begin{cases} g_{\not\oplus,\not\ominus} & = & x + (g_{\not\oplus,\not\ominus} + g_{\not\oplus} + g_{\not\ominus})^4, \\ g_{\not\oplus} & = & (g_{\not\oplus,\not\ominus} + g_{\not\ominus})(g_{\not\oplus,\not\ominus} + g_{\not\oplus} + g_{\not\ominus}), \\ g_{\not\ominus} & = & (g_{\not\oplus,\not\ominus} + g_{\not\oplus})(g_{\not\oplus,\not\ominus} + g_{\not\oplus} + g_{\not\ominus}). \end{cases}$$

The generating function for this class, $f = g_{\not\oplus,\not\ominus} + g_{\not\oplus} + g_{\not\ominus}$, satisfies

$$f^5 + f^4 + f^2 + (x-1)f + x = 0,$$

and the first terms of the sequence are $1, 2, 6, 23, 102, 492, \ldots$ (sequence A120346 of [110]).

Example 6.11 ($\mathrm{Av}(132)$)**.** The wreath closure of $\mathrm{Av}(132)$ is the class of separable permutations, so to enumerate $\mathrm{Av}(132)$ we need to refine Example 6.9. While Proposition 6.5 shows that $\{\mathrm{Av}(1), \mathrm{Av}(12), \mathrm{Av}(21), \mathrm{Av}(132)\}$ is query-complete, it is sufficient to set $\mathcal{P} = \{\not\oplus, \not\ominus, \mathrm{Av}(21), \mathrm{Av}(132)\}$ by our remarks in Section 6.2. Our system is then

$$\begin{cases} g_{\not\oplus,\not\ominus,\mathrm{Av}(21)} & = & x, \\ g_{\not\ominus,\mathrm{Av}(21)} & = & g_{\not\oplus,\not\ominus,\mathrm{Av}(21)}(g_{\not\oplus,\not\ominus,\mathrm{Av}(21)} + g_{\not\ominus,\mathrm{Av}(21)}), \\ g_{\not\oplus} & = & (g_{\not\oplus,\not\ominus,\mathrm{Av}(21)} + g_{\not\ominus,\mathrm{Av}(21)} + g_{\not\ominus})(g_{\not\oplus,\not\ominus,\mathrm{Av}(21)} + g_{\not\ominus,\mathrm{Av}(21)} + g_{\not\oplus} + g_{\not\ominus}), \\ g_{\not\ominus} & = & g_{\not\oplus}(g_{\not\oplus,\not\ominus,\mathrm{Av}(21)} + g_{\not\ominus,\mathrm{Av}(21)}). \end{cases}$$

(As we are only interested in 132-avoiding permutations we have suppressed the subscript $\mathrm{Av}(132)$, which would otherwise be present in all these terms.) Setting

$$f = g_{\not\oplus,\not\ominus,\mathrm{Av}(21)} + g_{\not\ominus,\mathrm{Av}(21)} + g_{\not\oplus} + g_{\not\ominus}$$

and solving yields

$$f = \frac{1 - 2x - \sqrt{1-4x}}{2x},$$

the generating function for the Catalan numbers, as expected.

Example 6.12 ($\mathrm{Av}(2413, 3142, 2143)$)**.** Here we take $\mathcal{P} = \{\not\oplus, \not\ominus, \mathrm{Av}(21), \mathrm{Av}(2143)\}$ and our system is

$$\begin{cases} g_{\not\oplus,\not\ominus,\mathrm{Av}(21)} & = & x, \\ g_{\not\ominus,\mathrm{Av}(21)} & = & g_{\not\oplus,\not\ominus,\mathrm{Av}(21)}(g_{\not\oplus,\not\ominus,\mathrm{Av}(21)} + g_{\not\ominus,\mathrm{Av}(21)}), \\ g_{\not\oplus} & = & (g_{\not\oplus,\not\ominus,\mathrm{Av}(21)} + g_{\not\ominus,\mathrm{Av}(21)} + g_{\not\ominus})(g_{\not\oplus,\not\ominus,\mathrm{Av}(21)} + g_{\not\ominus,\mathrm{Av}(21)} + g_{\not\oplus} + g_{\not\ominus}), \\ g_{\not\ominus} & = & g_{\not\oplus,\not\ominus,\mathrm{Av}(21)}(g_{\not\oplus} + g_{\not\ominus}) + g_{\not\oplus}(g_{\not\oplus,\not\ominus,\mathrm{Av}(21)} + g_{\not\ominus,\mathrm{Av}(21)}), \end{cases}$$

where here we have suppressed the $\mathrm{Av}(2143)$ subscript. This gives the generating function

$$\frac{1 - 3x + 2x^2 - \sqrt{1 - 6x + 5x^2}}{2x(2-x)},$$

and thus the number of permutations of length n in this class is $\sum \binom{n}{k} F_{n-k}$ (sequence A033321 of [110]), where F_n denotes the nth term in Fine's sequence.[3]

Example 6.13 (Alternating separable permutations)**.** Lemma 6.6 shows that we need to introduce the properties AL (alternating permutations), BR (permutations beginning with a rise), ER (permutations ending with a rise), and $\{1\}$. In the separable case $\{1\} = \not\varnothing \cap \varnothing$ so we take $\mathcal{P} = \{\not\varnothing, \varnothing, BR, ER, AL\}$, and as AL occurs in each of the terms of our system we suppress it. We then have

$$\begin{cases}
g_{\not\varnothing,\varnothing} & = & x, \\
g_{\not\varnothing} & = & (g_{\not\varnothing,\varnothing} + g_{\varnothing,ER})(g_{\not\varnothing,\varnothing} + g_{\not\varnothing,BR} + g_{\varnothing,BR}), \\
g_{\not\varnothing,BR} & = & g_{\varnothing,BR,ER}(g_{\not\varnothing,\varnothing} + g_{\not\varnothing,BR} + g_{\varnothing,BR}), \\
g_{\not\varnothing,ER} & = & (g_{\not\varnothing,\varnothing} + g_{\varnothing,ER})(g_{\not\varnothing,BR,ER} + g_{\varnothing,BR,ER}), \\
g_{\not\varnothing,BR,ER} & = & g_{\varnothing,BR,ER}(g_{\not\varnothing,BR,ER} + g_{\varnothing,BR,ER}), \\
g_{\varnothing} & = & g_{\not\varnothing}(g_{\not\varnothing} + g_{\varnothing}), \\
g_{\varnothing,BR} & = & (g_{\not\varnothing,\varnothing} + g_{\not\varnothing,BR})(g_{\not\varnothing} + g_{\varnothing}), \\
g_{\varnothing,ER} & = & g_{\not\varnothing}(g_{\not\varnothing,\varnothing} + g_{\not\varnothing,ER} + g_{\varnothing,ER}), \\
g_{\varnothing,BR,ER} & = & (g_{\not\varnothing,\varnothing} + g_{\not\varnothing,BR})(g_{\not\varnothing,\varnothing} + g_{\not\varnothing,ER} + g_{\varnothing,ER}).
\end{cases}$$

The generating function for these permutations satisfies

$$f^3 - (2x^2 - 5x + 4)f^2 - (4x^3 + x^2 - 8x)f - (2x^4 + 5x^3 + 4x^2) = 0,$$

and the first few terms of the sequence are $1, 2, 4, 8, 20, 48, \ldots$ (sequence A121703 of [110]).

6.6 Involutions

Unfortunately, involutionhood lies just outside the scope of our query-complete-property machinery: letting I denote the set of involutions we have that $12[\alpha_1, \alpha_2] \in I \iff \alpha_1, \alpha_2 \in I$, but when is $21[\alpha_1, \alpha_2] \in I$?

We begin by considering the effect of inversion on the substitution decomposition. First observe that

$$(\sigma[\alpha_1, \ldots, \alpha_m])^{-1} = \sigma^{-1}[\alpha^{-1}_{\sigma^{-1}(1)}, \ldots, \alpha^{-1}_{\sigma^{-1}(m)}].$$

[3] *Fine's sequence* is defined by $2F_n + F_{n-1} = C_n$ for $n \geq 1$, where C_n denotes the nth Catalan number (sequence A000957 of [110]).

Recalling the first part of Proposition 1.7 ("every permutation is the inflation of a unique simple permutation"), we have that if π is an involution then it must be the inflation of a simple involution. By the second part of Proposition 1.7 we then obtain the following:

Proposition 6.14. *If $\pi = \sigma[\alpha_1, \ldots, \alpha_m]$ is an involution and $\sigma \neq 21$ is a simple permutation then σ is an involution and $\alpha_i = \alpha^{-1}_{\sigma^{-1}(i)} = \alpha^{-1}_{\sigma(i)}$ for all $i \in [m]$.*

The case $\sigma = 21$ must be handled separately but is not any more difficult.

Proposition 6.15. *The involutions that are inflations of 21 are precisely those of the form*

- $21[\alpha_1, \alpha_2]$ *for skew indecomposable α_1 and α_2 with $\alpha_1 = \alpha_2^{-1}$, and*
- $321[\alpha_1, \alpha_2, \alpha_3]$, *where α_1 and α_3 are skew indecomposable, $\alpha_1 = \alpha_3^{-1}$, and α_2 is an involution.*

Define the *inverse* of the property P by $P^{-1} = \{\pi^{-1} : \pi \in P\}$, and for a set of properties $\mathcal{P}$, $\mathcal{P}^{-1} = \{P^{-1} : P \in \mathcal{P}\}$.

Theorem 6.16. *Let $\mathcal{C}$ be a permutation class containing only finitely many simple permutations, $\mathcal{P}$ a finite query-complete set of properties, and $\mathcal{Q} \subseteq \mathcal{P}$. The generating function for the set of involutions in $\mathcal{C}$ satisfying every property in $\mathcal{Q}$ is algebraic over $\mathbb{Q}(x)$.*

Proof. We assume (without loss) both that $\not\geq, \not\leq \in \mathcal{P}$ and that $\mathcal{P} = \mathcal{P}^{-1}$. As in the proof of Theorem 6.2, let $\mathcal{P}(\pi)$ denote the set of properties in $\mathcal{P}$ satisfied by π and $g_{\mathcal{R}}$ denote the generating function for the set of $\pi \in \mathcal{C}$ with $\mathcal{P}(\pi) = \mathcal{R}$. Also let $h_{\mathcal{R}}$ denote the generating function for the set of involutions $\pi \in \mathcal{C}$ with $\mathcal{P}(\pi) = \mathcal{R}$. It suffices to show that each $h_{\mathcal{R}}$ is algebraic over $\mathbb{Q}(x)$.

As Propositions 6.14 and 6.15 indicate, we need to count pairs (α, α^{-1}) where α and α^{-1} satisfy certain sets of properties. To this end define

$$p_{\mathcal{R}} = \sum_{\substack{\alpha \in \mathcal{C} \\ \mathcal{P}(\alpha) = \mathcal{R}}} x^{|\alpha| + |\alpha^{-1}|}.$$

Note that $p_{\mathcal{R}}$ is nothing other than $g_{\mathcal{R}}(x^2)$.

Now take σ to be a simple permutation. We need to compute the contribution to $h_{\mathcal{R}}$ of inflations of σ. If σ is not an involution, Proposition 6.14 shows that this contribution is 0. Otherwise since $\mathcal{P}$ is query-complete, $\mathcal{P}(\sigma[\alpha_1,\ldots,\alpha_m]) = \mathcal{R}$ if and only if $(\mathcal{P}(\alpha_1),\ldots,\mathcal{P}(\alpha_m))$ lies in a certain collection of m-tuples of sets of properties. Choose one of these m-tuples, say $(\mathcal{R}_1,\ldots,\mathcal{R}_m)$, and suppose first that $m = |\sigma| \geq 4$. It suffices to calculate the contribution of involutions of the form $\sigma[\alpha_1,\ldots,\alpha_m]$ with $\mathcal{P}(\alpha_i) = \mathcal{R}_i$ for all $i \in [m]$. If there is some $j \in [m]$ for which $\mathcal{R}_j \neq \mathcal{R}^{-1}_{\sigma(j)}$ then this contribution is 0 by Proposition 6.14. Otherwise the contribution is a single term in which each fixed point j corresponds to an $h_{\mathcal{R}_j}$ factor and each non-fixed-point pair $(j,\sigma(j))$ corresponds to a $p_{\mathcal{R}_j}$ factor. A similar analysis of inflations of 12 and 21 — in the latter case appealing to Proposition 6.15 — allows us to compute their contributions.

Therefore each $h_{\mathcal{R}}$ can be expressed nontrivially as a polynomial in x, $\{h_{\mathcal{S}} : \mathcal{S} \subseteq \mathcal{P}\}$, and $\{p_{\mathcal{S}} : \mathcal{S} \subseteq \mathcal{P}\}$. Viewing x and $\{p_{\mathcal{S}} : \mathcal{S} \subseteq \mathcal{P}\}$ as variables, Theorem 6.4 implies that each $h_{\mathcal{R}}$ is algebraic over $\mathbb{Q}(x, \{p_{\mathcal{S}} : \mathcal{S} \subseteq \mathcal{P}\})$. Furthermore, $p_{\mathcal{S}} = g_{\mathcal{S}}(x^2)$, so $\mathbb{Q}(x, \{p_{\mathcal{S}} : \mathcal{S} \subseteq \mathcal{P}\})$ is an algebraic extension of $\mathbb{Q}(x)$ by Theorem 6.2, proving the theorem. $\square$

One could adapt the proof of Theorem 6.16 to count the permutations in $\mathcal{C}$ that are invariant under other symmetries. For example, the permutations invariant under the composition of reverse and complement studied by Guibert and Pergola [64]. Egge [43] considers the enumeration of restricted permutations invariant under other symmetries.

Example 6.17 (Separable involutions)**.** We take $\mathcal{P} = \{\not\oplus, \not\ominus\}$. Using the notation from the proof of Theorem 6.16, we wish to find $f = h_{\not\oplus,\not\ominus} + h_{\not\oplus} + h_{\not\ominus}$. These generating functions are related to each other and to the p generating functions by

$$\begin{cases} h_{\not\oplus,\not\ominus} & = & x, \\ h_{\not\oplus} & = & (p_{\not\oplus,\not\ominus} + p_{\not\ominus}) + (p_{\not\oplus,\not\ominus} + p_{\not\ominus})(h_{\not\oplus,\not\ominus} + h_{\not\oplus} + h_{\not\ominus}), \\ h_{\not\ominus} & = & (h_{\not\oplus,\not\ominus} + h_{\not\oplus})(h_{\not\oplus,\not\ominus} + h_{\not\oplus} + h_{\not\ominus}). \end{cases}$$

From Example 6.9 it can be computed that

$$\begin{aligned}
p_{\not\oplus,\not\ominus} - x^2 &= 0, \\
2p_{\not\oplus}^2 + (3x^2-1)p_{\not\oplus} + x^4 &= 0, \\
2p_{\not\ominus}^2 + (3x^2-1)p_{\not\ominus} + x^4 &= 0.
\end{aligned}$$

Combining these with the system above and solving as usual shows that

$$x^2f^4 + (x^3+3x^2+x-1)f^3 + (3x^3+6x^2-x)f^2 + (3x^3+7x^2-x-1)f + x^3+3x^2+x = 0,$$

and the first few terms of the sequence are $1, 2, 4, 10, 24, 64, \ldots$ (sequence A121704 of [110]).

6.7 Cyclic Closures

In order to demonstrate that the framework developed here can be applied in less obvious situations, we present an application which differs in flavour from our previous examples. The permutation τ is said to be a *cyclic rotation* (or simply, *rotation*) of the permutation π, both of length n, if there is an $i \in [n]$ for which $\tau = \pi(i+1)\ldots\pi(n)\pi(1)\ldots\pi(i)$. Given a permutation class $\mathcal{C}$, its *cyclic closure*, $\mathrm{cc}(\mathcal{C})$, consists of all rotations of members of $\mathcal{C}$. This operation has been studied by the Otago group [1], who proved several basis and enumeration results. The main result of this section, Theorem 6.19, shows that the cyclic closure of a class with finitely many simple permutations has an algebraic generating function.

The cyclic closure of the class $\mathcal{C}$ can be partitioned into orbits of permutations under rotation. As the orbit of a permutation of length n has precisely n elements, to enumerate a cyclic closure it suffices to count orbits. We do this by distinguishing one permutation per orbit and then counting these permutations. For us, a *distinguished* member of $\mathrm{cc}(\mathcal{C})$ is a permutation π that satisfies:

(1) $\pi \in \mathcal{C}$ (this can clearly be achieved, because every orbit in $\mathrm{cc}(\mathcal{C})$ contains at least one element of $\mathcal{C}$) and

(2) among all permutations in its orbit satisfying (1), π is the one in which the entry 1 lies furthest to the left.

For example, one orbit in $\mathrm{cc}(\mathrm{Av}(132))$ is

$$12534, 41253, 34125, 53412, 25341.$$

Only two of these permutations avoid 132, 34125 and 53412. Since the entry 1 lies further to the left in 34125, this is the distinguished permutation of its orbit.

Our goal is to show that the property of distinction lies in a finite query-complete set of properties. We begin by offering a different viewpoint in which instead of rotating permutations we divide them into two parts. A *divided permutation* is a permutation equipped with a divider $|$, i.e., $\pi_1|\pi_2$, and we refer to $\pi_1|\pi_2$ as a *division* of the concatenation $\pi_1\pi_2$. We say that the divided permutation $\sigma_1|\sigma_2$ is contained in the divided permutation $\pi_1|\pi_2$ if $\pi_1\pi_2$ contains a subsequence order isomorphic to $\sigma_1\sigma_2$ in which the entries corresponding to σ_1 come from π_1 and the entries corresponding to σ_2 come from π_2. For example, $513|42$ contains $32|1$ because of the subsequence 532, but $32|1$ is not contained in $51|342$.

Suppose now that we are given a permutation $\pi \in \mathcal{C} = \mathrm{Av}(B)$ and we wish to decide if π is a distinguished member of $\mathrm{cc}(\mathcal{C})$. According to (2) above, we need to check all rotations of π in which the 1 lies further to the left. Instead, let us consider all divisions $\pi_1|\pi_2$ of π in which π_1 is nonempty and π_2 contains the entry 1, thinking of such a division as corresponding to the rotation $\pi_2\pi_1$. For π to be distinguished, each of these divisions must contain $\beta_2|\beta_1$ for some $\beta_1\beta_2 \in B$, because that will imply that the corresponding rotation contains $\beta_1\beta_2$ and thus fails to lie in $\mathcal{C}$.

For a set of divided permutations Δ, let us therefore define the property $DP_1(\Delta)$ to consist of all permutations π for which every division $\pi_1|\pi_2$ where π_1 is nonempty and the 1 lies in π_2 contains at least one of the divided permutations in Δ. Our set of distinguished permutations for $\mathrm{cc}(\mathcal{C})$ will then consist of those permutations from $\mathcal{C}$ which satisfy

$$DP_1(\{\beta_2|\beta_1 : \beta_1\beta_2 \in B\}).$$

We also need a similar family: $DP(\Delta)$ consists of all permutations π for which every division $\pi_1|\pi_2$ of π in which π_1 is nonempty contains at least one of the divided permutations in Δ. (Note that we allow π_2 to be empty.)

Lemma 6.18. *For any finite set B of permutations, the property $DP_1(\{\beta_2|\beta_1 : \beta_1\beta_2 \in B\})$ lies in a finite query-complete set of properties.*

Proof. The finite query-complete set we take consists of

$$\{\mathrm{Av}(\delta) : \delta \leq \beta \text{ for some } \beta \in B\}$$

and the properties $DP(\Delta)$ and $DP_1(\Delta)$ for all $\Delta \subset \{\delta_2|\delta_1 : \delta_1\delta_2 \leq \beta \text{ for some } \beta \in B\}$.

Let $\pi = \sigma[\alpha_1, \ldots, \alpha_m]$. Propositions 6.3 and 6.5 show that the Av properties form a query-complete set, so it suffices to prove that membership in the DP and DP_1 can be decided based on σ and which of these properties each α_i satisfies. Since these properties are very similar, we consider only the $DP_1(\Delta)$ case.

Suppose that $\sigma(\ell) = 1$, so that the entry 1 in π occurs in its $\sigma(\ell)$ interval. First, for each $k < \ell$, we need to consider divisions of π which slice its $\sigma(k)$ interval (or slice between this interval and the next). As in the proof of Proposition 6.5 we consider lenient inflations (inflations in which intervals are allowed to be empty), although we now insist that the divider occur in the kth interval of the lenient inflations (we allow that interval to contain the divider alone). List all such lenient inflations of all divided permutations in Δ as

$$\sigma[\gamma_1^{(1)}, \ldots, \gamma_m^{(1)}], \ldots, \sigma[\gamma_1^{(t)}, \ldots, \gamma_m^{(t)}].$$

We need to determine whether every division of π which slices its $\sigma(k)$ interval contains one of these lenient inflations. If for some $s \in [t]$ and $j \neq k$, α_j does not contain $\gamma_j^{(s)}$ (which can be determined from the Av properties), then none of these divisions of π can contain that lenient inflation. Remove these infeasible inflations from the list, leaving

$$\sigma[\gamma_1^{(u_1)}, \ldots, \gamma_m^{(u_1)}], \ldots, \sigma[\gamma_1^{(u_v)}, \ldots, \gamma_m^{(u_v)}].$$

Now a division of π slicing its $\sigma(k)$ interval contains the ith lenient inflation in this list if and only if $\gamma_k^{(u_i)}$ is either a lone divider or is contained (as a divided permutation) in the resulting, divided α_k. Thus every division of π which slices its $\sigma(k)$ interval contains a divided permutation from Δ if and only if

$$\alpha_k \in DP(\{\gamma_k^{(u_1)}, \ldots, \gamma_k^{(u_v)}\}),$$

and this property is in our set of properties. The analysis for divisions of π which slice the $\sigma(\ell)$ interval (the block containing the entry 1) is identical, except that DP is replaced by DP_1. □

Theorem 6.19. *If a permutation class $\mathcal{C}$ contains only finitely many simple permutations then its cyclic closure $\mathrm{cc}(\mathcal{C})$ has an algebraic generating function over $\mathbb{Q}(x)$.*

Proof. Let $\mathcal{C} = \mathrm{Av}(B)$ contain only finitely many simple permutations, so by Theorem 5.29, B is finite. Lemma 6.18 the shows that the property $DP_1(\{\beta_2|\beta_1 : \beta_1\beta_2 \in B\})$ lies in a finite query-complete set. Thus the distinguished permutations, which are the permutations in $\mathcal{C}$ that satisfy this property, have an algebraic generating function by Theorem 6.2. Call this generating function f. Since every orbit of length n permutations in $\mathrm{cc}(C)$ contains n elements, precisely one of which is distinguished, the generating function for $\mathrm{cc}(\mathcal{C})$ is $xf'(x)$, which is also algebraic. □

We conclude the section with an abridged example.

Example 6.20 (The cyclic closure of $\mathrm{Av}(132)$)**.** The distinguished elements for $\mathrm{cc}(\mathrm{Av}(132))$ are those that lie in $\mathrm{Av}(132)$ and satisfy

$$DP_1(\{\beta_2|\beta_1 : \beta_1\beta_2 = 132\}) = DP_1(132|, 32|1, 2|13, |132).$$

If any division of a permutation contains $132|$ or $|132$ then the permutation itself contains 132; since we are only counting 132-avoiding permutations, we may write the generating function for the distinguished elements as $f_{DP_1(32|1,2|13)}$, where $f_{\mathcal{Q}}$ denotes the generating function for the permutations in $\mathrm{Av}(132)$ which satisfy every property in $\mathcal{Q}$ but may satisfy additional properties. In the other examples we have given the complete system of g generating functions. Owing to the number of properties involved and the labour necessary for their specification, here we only describe how to compute two of the f generating functions.

Let us begin with the $f_{\not\emptyset,DP_1(32|1,2|13)}$ term. Since our only simple permutations are 1, 12, 21, the $\oplus$-indecomposable permutations are 1 and those that can be expressed uniquely as $21[\alpha_1, \alpha_2]$ where $\alpha_1 \in \not\varnothing$. First consider divisions of $21[\alpha_1, \alpha_2]$ which slice α_1; for these to contain either $32|1$ or $2|13$, the divided α_1 must contain either $21|$, which can be extended to $32|1$ by including an entry of α_2, or $2|13$. All such permutations must contain 21, so they are counted by $f_{\not\varnothing,DP(21|,2|13)} - f_{\not\varnothing,\mathrm{Av}(21),DP(21|,2|13)}$. Now observe that the divisions which slice α_2 before its entry 1 necessarily contain a copy of $32|1$ where the '3' comes from α_1 and the '2' comes from an entry of α_2 preceding 1 (if there is no such entry, then none of these divisions need checking), and so every 132-avoiding permutation may serve as α_2. Thus we have

$$f_{\not\varnothing,DP_1(32|1,2|13)} = x + \left(f_{\not\varnothing,DP(21|,2|13)} - f_{\not\varnothing,\mathrm{Av}(21),DP(21|,2|13)}\right) f.$$

This leaves us to determine $f_{\not\varnothing,DP(21|,2|13)}$. These permutations (except for 1) can be written uniquely as $\pi = 12[\alpha_1, \alpha_2]$ where $\alpha_1 \in \not\varnothing$ and as they avoid 132 we have $\alpha_2 \in \mathrm{Av}(21)$. The divisions slicing α_1 must create $21|$ or $2|13$ patterns in π, which will occur if and only if $\alpha_1 \in DP(21|, 2|1)$. This rules out $\alpha_1 = 1$, so these permutations are counted by $f_{\not\varnothing,DP(21|,2|1)} - x$. Because $\alpha \in DP(21|, 2|1)$, α_1 must contain 21, and thus all divisions which slice α_2 will contain $21|$. Therefore the only restriction on α_2 is that it must avoid 21, giving the equation

$$f_{\not\varnothing,DP(21|,2|13)} = x + \left(f_{\not\varnothing,DP(21|,2|1)} - x\right) f_{\mathrm{Av}(21)}.$$

Similar reasoning allows one to compute the entire system, which leads to the solution

$$f_{DP_1(32|1,2|13)} = \frac{(1-2x)(1-2x-\sqrt{1-4x})}{2x(1-x)}.$$

From this we find that the generating function for $\mathrm{cc}(\mathrm{Av}(132))$ is

$$xf'_{DP_1(32|1,2|13)} = \frac{1-4x+4x^2-4x^3-(1-2x)\sqrt{1-4x}}{2x(1-x)^2\sqrt{1-4x}},$$

which agrees with the results of Albert *et al.* [1]. The first few terms of the sequence are $1, 2, 6, 24, 100, \ldots$.

6.8 Applicability and Application

With the results of the paper now established, we conclude by discussing their use. First, let us summarise the finite query-complete sets that we have covered in this chapter as a corollary of Theorem 6.2.

Corollary 6.21. *In a permutation class $\mathcal{C}$ with only finitely many simple permutations, the generating functions for the following sequences are algebraic over $\mathbb{Q}(x)$:*

- *the number of permutations in $\mathcal{C}_n$ (this is the result of Albert and Atkinson [2]),*
- *the number of alternating permutations in $\mathcal{C}_n$,*
- *the number of even permutations in $\mathcal{C}_n$,*
- *the number of Dumont permutations of the first kind in $\mathcal{C}_n$,*
- *the number of permutations in $\mathcal{C}_n$ avoiding any finite set of blocked or barred permutations, and*
- *the number of involutions in $\mathcal{C}_n$.*

Moreover, these conditions can be combined in any finite manner desired.

As mentioned previously, $\mathrm{Av}(132)$ contains only three simple permutations, so Corollary 6.21 explains, e.g., why the even permutations in $\mathrm{Av}(132, \beta)$ have an algebraic generating function for every β, first proved in Mansour [83]. Other results in the literature to which Corollary 6.21 applies appear in [44, 45, 46, 50, 62, 63, 79, 80, 82, 84].

Other reasons for algebraicity. Having finitely many simple permutations is a sufficient condition for a class to possess an algebraic generating function, but it is by no means necessary. Consider $\mathrm{Av}(123)$, which, like $\mathrm{Av}(132)$, is enumerated by the Catalan numbers. However, $\mathrm{Av}(123)$ contains the infinite sequence of simple permutations $2n-1, 2n-3, \ldots, 3, 1, 2n, 2n-2, \ldots, 4, 2$ (one such permutation is plotted in Figure 1.3 on page 10). Indeed, every class of the form $\mathrm{Av}(\beta)$ where $|\beta| \geq 4$ contains either this infinite family or a symmetry of it. Thus our approach cannot be used to derive Bóna's result [18] that $\mathrm{Av}(1342)$ has an algebraic generating function. Nor can it be used to prove the fact that, for a surprising number of length 4 permutations β, the β-avoiding involutions are counted by the Motzkin numbers, as has been established by numerous researchers including Guibert [61], Guibert, Pergola and Pinzani [65], Jaggard [72] and Bousquet-Mélou and Steingrímsson [26]. The method also cannot be used to enumerate West-two-stack-sortable permutations [119].

Derangements. Notably absent from our list of finite query-complete sets in Section 6.4 are derangements, despite the fact that the 132-avoiding derangements are counted by Fine's sequence (Robertson, Saracino, and Zeilberger [103]), which has an algebraic generating function. To see that the set of derangements does not lie in a finite query-complete set of properties, for $\alpha \in S_n$ define $D(\alpha) = \{\alpha(i) - i : i \in [n]\}$. Then $21[12\cdots j, \alpha]$ is a derangement if and only if $j \notin D(\alpha)$. This shows that α_1 and α_2 must lie in different sets of properties whenever $D(\alpha_1) \cap \mathbb{N} \neq D(\alpha_2) \cap \mathbb{N}$, implying that the set of derangements can only lie in an infinite query-complete set of properties.

6.8.1 Simple Decomposition Revisited

We have not yet discussed the consequences of the decomposition of simple permutations for our knowledge of permutation classes. In the next chapter we will cover the problem of decidability for simple permutations, but this is by no means the only use of the decomposition. Indeed, our initial motivation was to derive the following theorem, whose importance has so far been left unspoken:

Theorem 2.2. *There is a function $f(k)$ such that every simple permutation of length at least $f(k)$ contains two simple subsequences, each of length at least k, sharing at most two entries.*

This result helps us in the enumeration of certain permutation classes, which we will introduce by means of a motivational example. As we have seen, the simple permutations of the class $\mathrm{Av}(132)$ are precisely 1, 12 and 21. Theorems 2.2 and 6.1 (on Page 111) combine to give a short proof of the following result.

Theorem 6.22 (Bóna [19]; Mansour and Vainshtein [85])**.** *For every r, the class of all permutations containing at most r copies of 132 has an algebraic generating function.*

For example, the generating function in the $r = 1$ case is

$$\frac{1-\sqrt{1-4x}}{2x} + \frac{8x^3}{\sqrt{1-4x}\left(1+\sqrt{1-4x}\right)^3},$$

due, originally, to Bóna [20].

Proof of Theorem 6.22 via Theorems 2.2 and 6.1. We wish to show that only finitely many simple permutations contain at most r copies of 132, or in other words, that there is a function $g(r)$ so that every simple permutation of length at least $g(r)$ contains more than r copies of 132. Since the only simple permutations in $\mathrm{Av}(132)$ are 1, 12 and 21, we may take $g(0) = 3$. We now proceed by induction, setting $g(r) = f(g(\lfloor r/2 \rfloor))$, where f is the function from Theorem 2.2. By that theorem, every simple permutation π of length at least $g(r)$ contains two simple subsequences of length at least $g(\lfloor r/2 \rfloor)$. By induction each of these simple subsequences contains more than $\lfloor r/2 \rfloor$ copies of 132. Moreover, because these simple subsequences share at most two entries, their copies of 132 are distinct, and thus π contains more than r copies of 132, as desired. □

Indeed, the proof above shows that every permutation class whose members contain a bounded number of copies of 132 has an algebraic generating function, whereas Theorem 6.22 is concerned only with the entire class of permutations with at most r copies of 132. There is of course nothing special about 132. Denote by $\mathrm{Av}(\beta_1^{\le r_1}, \beta_2^{\le r_2}, \dots, \beta_k^{\le r_k})$ the class of permutations that have at most r_1 copies of β_1, at most r_2 copies of β_2, and so on.[4] The proof just given can be adapted to prove the following result.

Corollary 6.23. *If the class $\mathrm{Av}(\beta_1, \beta_2, \dots, \beta_k)$ contains only finitely many simple permutations then for all choices of nonnegative integers r_1, r_2, …, r_k, the class $\mathrm{Av}(\beta_1^{\le r_1}, \beta_2^{\le r_2}, \dots, \beta_k^{\le r_k})$ also contains only finitely many simple permutations.*

The largest permutation class whose only simple permutations are 1, 12, and 21 is of course the class of separable permutations, $\mathrm{Av}(2413, 3142)$. Thus as another instance of

[4] That this is a permutation class is clear, although finding its basis may be less obvious. An easy argument shows that the basis elements of this class have length at most $\max\{(r_i+1)|\beta_i| : i \in [k]\}$; see Atkinson [7] for the details. One such computation: $\mathrm{Av}(132^{\le 1}) = \mathrm{Av}(1243, 1342, 1423, 1432, 2143, 35142, 354162, 461325, 465132)$.

Corollary 6.23, we have the following.

Corollary 6.24. *For all r and s, every subclass of* $\mathrm{Av}(2413^{\leq r}, 3142^{\leq s})$ *contains only finitely many simple permutations and thus has an algebraic generating function.*

This chapter has extended the scope of Theorem 6.1 to finite query-complete sets of properties, and we may combine Corollary 6.21 with Theorem 2.2 to give easy proofs of several results in the literature. For example, the even permutations in $\mathrm{Av}(132^{\leq r})$ are enumerated by an algebraic generating function, due originally to Mansour [81]. (Note that, when counting even permutations, unlike when counting all permutations, symmetry considerations reduce us to three cases of length three permutations – 123, 132, and 231 – not two, and thus there is another result we can state at this point: the even permutations in $\mathrm{Av}(231^{\leq r})$ have an algebraic generating function for all r, although this result seems to have escaped print.[5])

Other results to which Theorem 2.2 and Corollary 6.21 may be applied can be found in [35, 80, 86].

6.8.2 Linear Time Membership

Out of some of the machinery developed in this chapter comes an indication that, given a permutation class $\mathcal{C}$ containing only finitely many simple permutations, it may be decided in linear time whether an arbitrary permutation π of length n lies in $\mathcal{C}$. The approach relies first and foremost on the fact that we may compute the substitution decomposition of any permutation in linear time, as per Chapter 4. We begin by first performing some precomputations specific to the class $\mathcal{C}$, all of which may be done essentially in constant time:

- Compute $\mathrm{Si}(\mathcal{C})$, the number of simple permutations in $\mathcal{C}$.

[5]We cannot say anything about the other case, $\mathrm{Av}(123)$, since it contains infinitely many simple permutations, and hence so does $\mathrm{Av}(123^{\leq r})$. The class $\mathrm{Av}(123^{\leq 1})$ was, however, counted by Noonan [99], while $\mathrm{Av}(123^{\leq 2})$ was counted by Fulmek [57], proving a conjecture of Noonan and Zeilberger [100]. No results for larger values are known, although Fulmek conjectures formulas for $r = 3$ and $r = 4$, and that $\mathrm{Av}(123^{\leq r})$ has an algebraic generating function for all r.

- Compute the basis B of $\mathcal{C}$, noting that permutations in B can be no longer than $\max_{\sigma \in \mathrm{Si}(\mathcal{C})} |\sigma| + 2$ by the Schmerl-Trotter Theorem 2.1.
- For every β either lying in B or contained in a permutation lying in B, list all expressions of β as a lenient inflation of each $\sigma \in \mathrm{Si}(\mathcal{C})$.

(Recall that a lenient inflation is an inflation $\sigma[\gamma_1, \ldots, \gamma_m]$ in which the γ_is are allowed to be empty.)

With these precomputations performed, we now take our candidate permutation π of length n and compute its substitution decomposition, $\pi = \sigma[\alpha_1, \ldots, \alpha_m]$. Now, after first trivially checking that the skeleton σ lies in $\mathcal{C}$, we look at all the expressions of each $\beta \in B$ as lenient inflations of σ. Note that if $\beta \leq \pi$, there must exist an expression of β as a lenient inflation $\beta = \sigma[\gamma_1, \ldots, \gamma_m]$ so that $\gamma_i \leq \alpha_i$ for every $i = 1, \ldots, m$.

Thus, taking each lenient inflation $\beta = \sigma[\gamma_1, \ldots, \gamma_m]$ in turn, we look recursively at each block, testing to see if $\gamma_i \leq \alpha_i$ is true. Though this recursion makes the linear-time complexity non-obvious, note that the number of levels of recursion that are required cannot be more than the maximum depth of the substitution decomposition tree, which itself cannot have more than $2n$ nodes. The recursion will eventually reduce the problem to making only trivial comparisons, each of which is immediately answerable in constant time. The author would be keen to see a more rigourous treatment of this problem, and indeed an implementation of any subsequent algorithm.

CHAPTER 7

DECIDABILITY AND UNAVOIDABLE SUBSTRUCTURES

7.1 Introduction

HAVING DEFINED permutation classes and observed in Section 5.4 and Chapter 6 how simple permutations control many of their properties, it seems essential now to ask which finitely based classes contain only finitely many simple permutations. Our decomposition of simple permutations and identification of their unavoidable substructures in Chapter 2 puts us in a strong position to establish whether this question is decidable. Our main result establishes that this can be done algorithmically:

Theorem 7.1. *It is possible to decide if a permutation class given by a finite basis contains infinitely many simple permutations.*

We first begin by reminding the reader of pin sequences, as defined in Chapter 2. In particular, here we will be constructing pin sequences from scratch, before studying their possible subsequences. As we saw in Section 2.4, this treatment requires us to consider a slight variant of the original definition of pin sequences, namely that a proper pin sequence $p_1, \ldots, p_m$ must satisfy the following two conditions:

- *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.
- *Externality condition*: p_{i+1} must lie outside $\mathrm{rect}(p_1, \ldots, p_i)$.

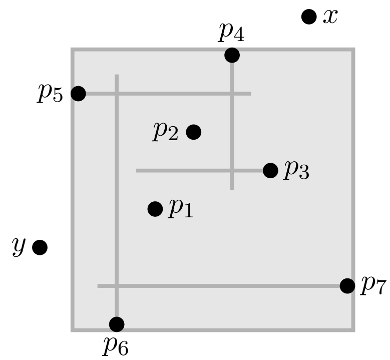

Figure 7.1: The points $p_1, \ldots, p_7$ form a proper pin sequence, and $\mathrm{rect}(p_1, \ldots, p_7)$ is denoted by the grey box. The point x satisfies the externality and separation conditions for this pin sequence and thus could be chosen as p_8; y, however, fails the separation condition.

(See Figure 7.1 for an illustration.) To consider subsequences of a given pin sequence, as we must, we refer the reader to the discussion on pin words given in Section 2.4.

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

Theorem 2.7. *If $p_1, \ldots, p_m$ is a proper pin sequence of length $m \geq 5$ 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, we also saw that there were other "fundamental" types of simple permutation – in particular, we recall the definitions of parallel and wedge alternations. 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 simple permutations (called *wedge simple permutations of types* 1 *and* 2). These three families are plotted in Figure 7.2.

We recall that these families of permutations capture, in a sense, the diversity of simple permutations:

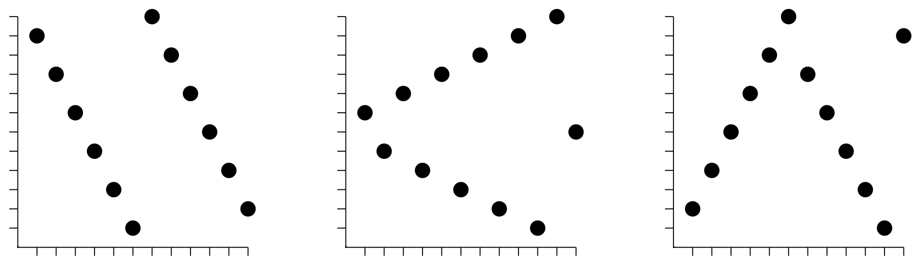

Figure 7.2: From left to right: a parallel alternation, a wedge simple permutation of type 1, and a wedge simple permutation of type 2.

Theorem 2.14. *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 2.7 and 2.14 show that Theorem 7.1 will follow if we can decide when a class has arbitrarily long parallel alternations, wedge simple permutations and proper pin sequences. The first two of these considerations are straightforward, and form the subject of the next section, while the question for proper pin sequences requires a little more work. Essentially, the problem of deciding whether a permutation class contains arbitrarily long pin sequences is equivalent to the problem of determining whether a permutation class admits arbitrarily long pin words. Thus converting the problem to one of languages, we will review in Section 7.3 the required results from formal language theory before going on to prove in Section 7.4 that the language of pins is regular, and hence the problem is decidable.

7.2 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 $\backslash\backslash$, as in Figure 7.2. 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 points 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. This, however, is done simply by using the juxtaposition, as defined in Subsection 5.1.2. The basis of the juxtaposition of two classes is decidable by Proposition 5.5 (Page 87), and this is all we need to solve the parallel alternation decision problem.

Proposition 7.2. *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, all but finitely many) parallel alternation(s) oriented $\backslash\backslash$ is

$$\left[\ \mathrm{Av}(12) \quad \mathrm{Av}(12)\ \right] = \mathrm{Av}(123, 2413, 3412),$$

as desired. □

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

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

Proof. The wedge simple permutations of type 1 that are oriented $<$, as in Figure 7.2, are contained in

$$\begin{aligned}\left[\ \left[\begin{array}{c}\mathrm{Av}(21)\\ \mathrm{Av}(12)\end{array}\right] \quad \{1\}\ \right] &= \left[\ \mathrm{Av}(132, 312) \quad \mathrm{Av}(12, 21)\ \right]\\ &= \mathrm{Av}(1324, 1423, 1432, 2431, 3124, 4123, 4132, 4231).\end{aligned}$$

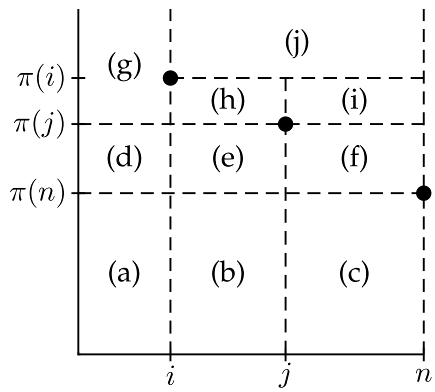

Figure 7.3: The situation in the proof of Proposition 7.3.

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 $\pi \in \begin{bmatrix} \mathrm{Av}(21) \\ \mathrm{Av}(12) \end{bmatrix}$ 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 complement of π. 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 7.3. 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. □

Proposition 7.4. *The permutation class* $\mathrm{Av}(B)$ *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 permutations of type 2 that are oriented Λ, as in Figure 7.2, 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

$$\begin{aligned} \begin{bmatrix} \mathrm{Av}(21) & \mathrm{Av}(12) & \{1\} \end{bmatrix} &= \begin{bmatrix} \mathrm{Av}(213, 312) & \mathrm{Av}(12, 21) \end{bmatrix} \\ &= \mathrm{Av}(2134, 2143, 3124, 3142, 3241, 4123, 4132, 4231), \end{aligned}$$

and so in particular, the permutation obtained by removing the rightmost element of π, say $\pi(n)$, is contained in $\begin{bmatrix} \mathrm{Av}(21) & \mathrm{Av}(12) \end{bmatrix}$. 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 7.4.

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 $\begin{bmatrix} \mathrm{Av}(21) & \mathrm{Av}(12) \end{bmatrix}$ shows that π itself lies in $\begin{bmatrix} \mathrm{Av}(21) & \mathrm{Av}(12) \end{bmatrix}$,

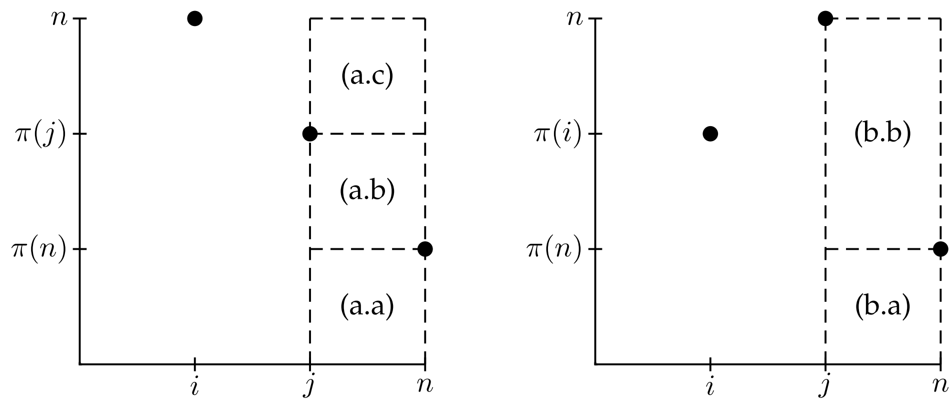

Figure 7.4: The two situations in the proof of Proposition 7.4.

and so π is contained in one of the desired wedge simple permutations, completing the proof. □

7.3 Review of Regular Languages and Automata

The classic results mentioned here are covered more comprehensively in many texts, for example, Hopcroft, Motwani, and Ullman [69], 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 *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 7.5 shows the transition diagram for the automaton that accepts strict pin words (in this automaton, all states are accept states).

A language that is accepted by a finite automaton is called *recognisable*. By Kleene's

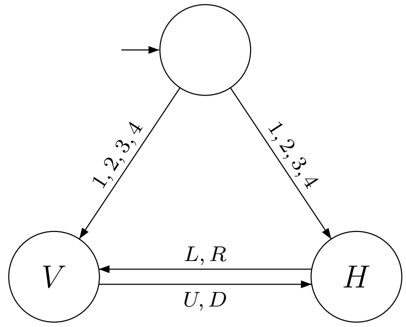

Figure 7.5: The automaton that accepts the language of strict pin words (V and H are accept states).

theorem, the recognisable languages are precisely the *regular languages*,[1] 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 7.5. *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

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

[1]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_is and w_is are allowed to be ε). We denote the set of words that the transducer T produces from the set of input words $\mathcal{L}$ by $T(\mathcal{L})$.

Proposition 7.6. *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. $\square$

7.4 Decidability

We are now in a position to prove our main result. We wish to decide whether the finitely based class $\mathrm{Av}(B)$ contains only finitely many simple permutations. Propositions 7.2–7.4 show how to decide if $\mathrm{Av}(B)$ contains arbitrarily long parallel alternations or wedge simple permutations, so by Theorem 2.14 (repeated in this chapter on page 138) it suffices to decide whether $\mathrm{Av}(B)$ contains arbitrarily long proper pin sequences.

We first recall two lemmas concerning pin words that we will require. The first shows that we may convert every proper pin sequence to a strict pin word. The proof is given on Page 37.

Lemma 2.15. *Every proper pin sequence corresponds to a strict pin word.*

The other lemma we must recall shows us how to relate subsequences of proper pin sequences with pin words, and vice versa. The proof may be found on Page 38.

Lemma 2.16. *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.*

Now, consider a permutation π that is order isomorphic to a proper pin sequence and thus, by Lemma 2.15, 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 2.16, β corresponds to a pin word $u \preceq w$. Conversely, if $w \succeq u$ for some u corresponding to $\beta \in B$, then Lemma 2.16 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 7.5.

Lemma 7.7. *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.*

Proof. Let T denote the transducer in Figure 7.6. 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.

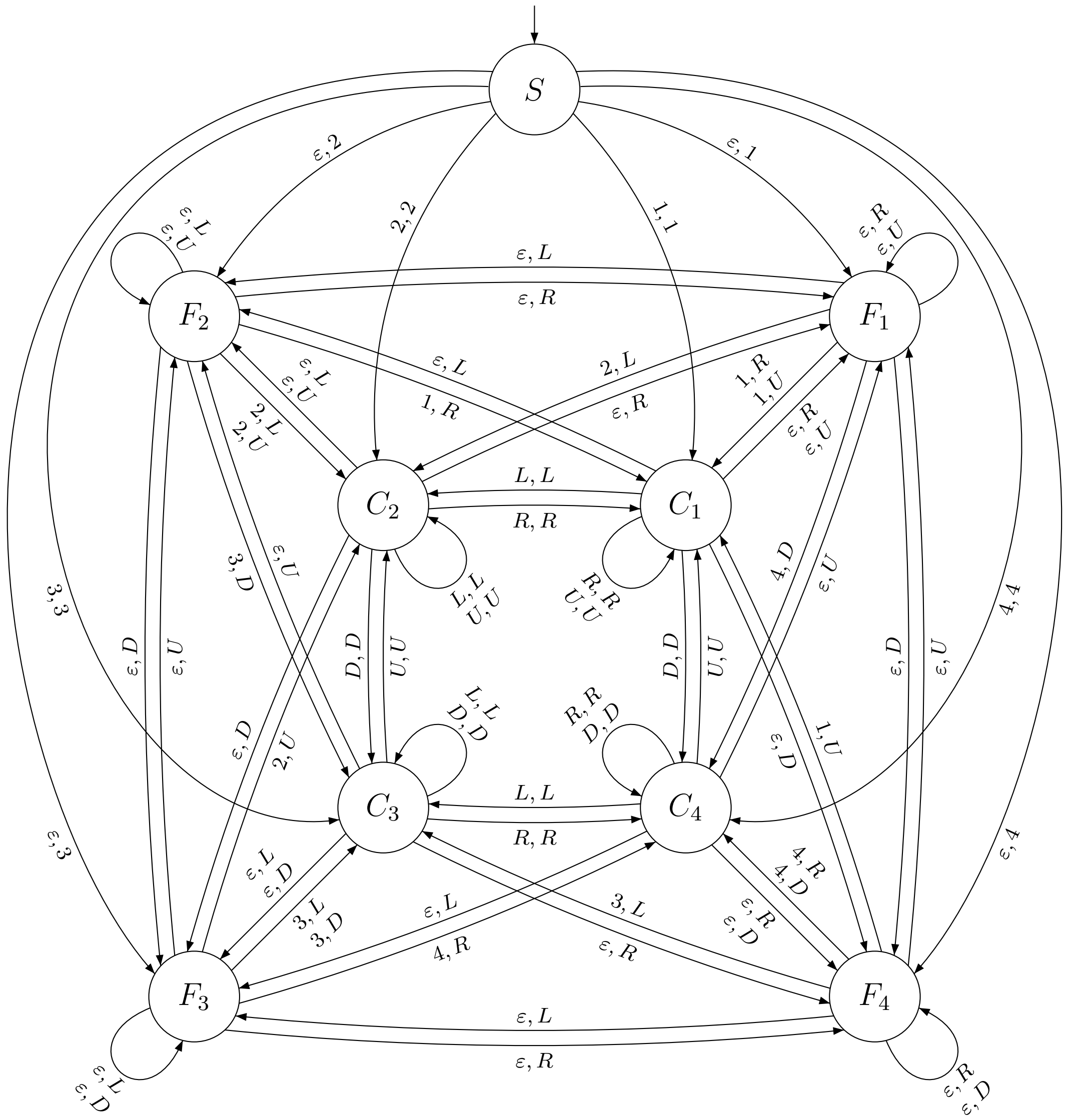

Figure 7.6: The transducer that produces all strict pin words containing the input pin word.

(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_k or F_k, then p_n lies in quadrant k. 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_k, then q_m lies in quadrant k.

(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. □

The proof of Theorem 7.1 now follows from the discussion at the beginning of the section.

7.5 An Easier Sufficient Condition

Though we have now seen a complete answer to the decidability problem, putting this method into practical use may, in some cases, be more work than is actually required. We can in fact derive a much easier-to-check set of conditions by recalling the unavoidable substructures result of Chapter 2:

Theorem 2.17. *Every sufficiently long simple permutation contains an alternation of length k or an oscillation of length k.*

Thus a permutation 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 $\begin{bmatrix} \mathrm{Av}(21) & \mathrm{Av}(12) \end{bmatrix}$ 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 [5] also attests to the strength of these conditions; they prove that classes without long alternations have rational generating functions.

As we have already shown how to decide if $\mathrm{Av}(B)$ contains arbitrarily long alternations, to convert Theorem 2.17 from a theorem about unavoidable substructures to an easily checked sufficient condition for containing only finitely many simple permutations we need to decide if $\mathrm{Av}(B)$ contains arbitrarily long oscillations. As with the parallel and wedge alternations from Section 7.2, 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, that are order isomorphic to a subset of the increasing oscillating sequence. This computation is given without proof in Murphy's thesis [97]. Here we provide the proof.

Proposition 7.8. *The class of all permutations contained in all but finitely many increasing oscillations is* $\mathrm{Av}(321, 2341, 3412, 4123)$.

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*[2] 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 $\pi \in \mathrm{Av}(321, 2341, 3412, 4123)$. 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 Figure 7.7 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 of our embedding of $\pi(i)$ is d_i. This follows from the rules above. □

7.6 Other Contexts.

To the best of our knowledge, no analogue of Theorem 7.1 is known for other relational structures. If we were to follow the pattern laid down in this thesis, our approach would be to decompose the simple structures and then establish an algorithmic method to avoid these structures. We discussed in Section 2.6 some possibilities to generalise the decomposition methods of Chapter 2, and saw in particular the problems encountered in the

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

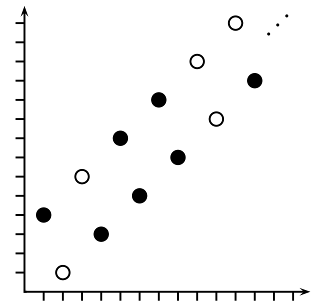

Figure 7.7: The filled points show the embedding of 2153647, with rank encoding 1020100, given by the proof of Proposition 7.8.

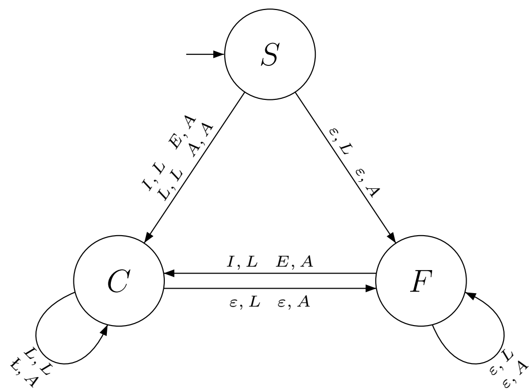

Figure 7.8: The prototype transducer for graphs.

graph case. On the assumption that these difficulties may be overcome (particularly in the graph case, but perhaps more generally) it seems likely that decidability would most likely follow. Our approach, therefore, remains furtive.

Determining the Language of Pins in Graphs. Assuming the existing definition from Section 2.6 for pin sequences in the graph case is nearly correct, it will actually turn out to be somewhat easier to construct an analogue to Lemma 7.7. To begin with, recall that the alphabet for the language of pins in graphs consists of only four letters, namely $\{L, A, I, E\}$, where L corresponds to adding a leaf, A an antileaf (connected to all but the last pin), I

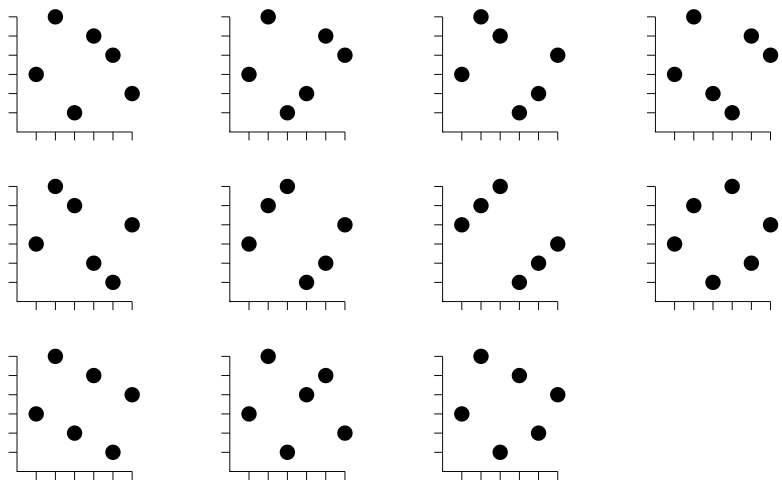

Figure 7.9: The basis elements of length 6 for the pin class (up to symmetry).

an independent point (i.e. connected to nothing) and E a point connected to everything. The transducer producing all strict pin words for graphs is thus much smaller than the permutation case of Figure 7.6, and a prototype is given in Figure 7.8. Note that since we do not have the issue of quadrants in graphs, there is only one fabrication state F and one copy state C.

7.7 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 2.16. As this class arises from words, it has a distinctly "regular" feel, and thus we offer:

Conjecture 7.9. *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 7.9). The class also has 220 basis elements of length 7. The class of course contains arbitrarily long simple permutations, and it is trivially not partially well ordered – the members of the variant of the increasing oscillating antichain from Example 5.14 (Page 97) may be encoded by words of the form $122RURU \cdots RUL$.

CHAPTER 8

THE WREATH PRODUCT

WE NOW CONSIDER a somewhat different problem, following the classic problem of determining the basis of a permutation class defined in one of the ways described in Section 5.1. As we mentioned there, the question for the wreath product of two permutation classes is known in only a few specific cases. Atkinson [12] shows that for any finitely based class $\mathcal{C}$, the wreath product $\mathcal{C} \wr \mathrm{Av}(21)$ is finitely based, but that $\mathrm{Av}(21) \wr \mathrm{Av}(321654)$ is not finitely based. There remains to be seen precisely what distinguishes these two cases. Our aim in this chapter is to find an answer to that question. In particular, we establish the following:

Theorem 8.1. *For any finitely based class $\mathcal{D}$ not admitting arbitrarily long pin sequences, the wreath product $\mathcal{C} \wr \mathcal{D}$ is finitely based for all finitely based classes $\mathcal{C}$.*

The approach is constructive, and will rely on our knowledge of the substitution decomposition learnt from Chapter 1, and our results concerning pin sequences from Chapter 2. We first introduce $\mathcal{D}$-profiles, which give us the ability to decompose permutations arising in wreath products into components belonging to the two original classes. For a permutation not arising in such a wreath product, we prove the existence of a subsequence order isomorphic to a basis element of the class $\mathcal{C}$. Moreover, there is a basis element of $\mathcal{D}$ lying within the "minimal block" defined by any two points of this subsequence. It is then a matter of using these considerations to show that, when the class $\mathcal{D}$ admits only finite pin sequences, the minimal elements not in the wreath product have bounded size.

Our secondary aim, arising as a result of the above considerations, is to exhibit a number of classes of the form $\mathcal{D} = \mathrm{Av}(\alpha)$ for $|\alpha| \leq 3$, or $\mathcal{D} = \mathrm{Av}(\alpha, \beta)$ with $|\alpha| \leq 4$, $|\beta| \leq 4$ which do not satisfy Theorem 8.1, and to demonstrate how an infinitely based wreath product $\mathcal{C} \wr \mathcal{D}$ can be found in each case.

8.1 $\mathcal{D}$-Profiles

We need to be able to know when a given permutation lies in the wreath product of two permutation classes. This could be done by inspecting all possible decompositions and checking for membership of the original classes, but this is liable to be computationally intensive. Instead, we would prefer only to check a single decomposition, from which membership or otherwise of the wreath product is immediately obvious.

The *profile* of a permutation π is the unique permutation obtained by contracting every maximal consecutive increasing sequence in π into a single point [7]. For example, the profile of 3415672 is 3142 because of the segments 34, 1, 567 and 2.

The notion of a "$\mathcal{D}$-profile" connects this idea with the definition of the substitution decomposition $\pi = \sigma[\pi_1, \ldots, \pi_m]$. We want the $\mathcal{D}$-profile of π to be the shortest possible deflation of π, given that we may only deflate by elements from the class $\mathcal{D}$. However, this is not clearly well-defined, so before we can proceed, we must first introduce $\mathcal{D}$-deflations.

Formally, let $\mathcal{D}$ be a permutation class, and π any permutation. Then a *$\mathcal{D}$-deflation* of π is a permutation π' for which π can be expressed as $\pi'[\alpha_1, \alpha_2, \ldots, \alpha_k]$ with $\alpha_1, \alpha_2, \ldots, \alpha_k \in \mathcal{D}$. For an arbitrary permutation π, there are many different $\mathcal{D}$-deflations. However, the shortest one is unique, and it is this one that gives rise to the $\mathcal{D}$-profile.

Lemma 8.2. *For every closed class $\mathcal{D}$ and permutation π, the shortest $\mathcal{D}$-deflation of π is unique.*

Proof. We proceed by induction on $n = |\pi|$. The case $n = 1$ is trivial, so now suppose $n > 1$. Fix a shortest $\mathcal{D}$-deflation of the permutation π, and label this permutation $\pi^{\mathcal{D}}$. If $\pi \in \mathcal{D}$ then $\pi^{\mathcal{D}} = 1$ is unique, so we will assume $\pi \notin \mathcal{D}$.

Let σ, of length $m \geq 2$, be the skeleton of π, and first consider the case where $m \geq 4$, whereby we have the unique substitution decomposition $\pi = \sigma[\pi_1, \pi_2, \ldots, \pi_m]$. By the inductive hypothesis, the shortest $\mathcal{D}$-deflations of $\pi_1, \pi_2, \ldots, \pi_m$ are unique, and we will label them $\pi_1^{\mathcal{D}}, \pi_2^{\mathcal{D}}, \ldots, \pi_m^{\mathcal{D}}$. We claim that $\pi^{\mathcal{D}} = \sigma[\pi_1^{\mathcal{D}}, \pi_2^{\mathcal{D}}, \ldots, \pi_m^{\mathcal{D}}]$. Consider any other $\mathcal{D}$-deflation of π, $\pi = \pi'[\alpha_1, \alpha_2, \ldots, \alpha_k]$. Since $\pi \notin \mathcal{D}$, π' cannot be trivial, and so $\sigma \leq \pi'$, and indeed σ is the skeleton of π', giving a unique deflation $\pi' = \sigma[\pi'_1, \ldots, \pi'_m]$. Moreover, π'_i is a $\mathcal{D}$-deflation of π_i for all i. Since $\pi_i^{\mathcal{D}}$ is the unique shortest $\mathcal{D}$-deflation, we must have $\pi_i^{\mathcal{D}} \leq \pi'_i$, which implies $\pi^{\mathcal{D}} \leq \pi'$.

When $m = 2$, more care is required. In this case π is either sum or skew decomposable, and without loss of generality we may assume the former. Write $\pi = 12\cdots t[\pi_1, \pi_2, \ldots, \pi_t]$ where each π_i is sum indecomposable. If every $\pi_i \in \mathcal{D}$, then any shortest $\mathcal{D}$-deflation of π will be an increasing permutation of length at most t, and as there is only one increasing permutation of each length, $\pi^{\mathcal{D}}$ will be unique. So now suppose that there exists at least one i such that $\pi_i \notin \mathcal{D}$, so that $|\pi_i^{\mathcal{D}}| \geq 2$. Since π_i is sum indecomposable, $\pi_i^{\mathcal{D}}$ is also sum indecomposable. We claim the shortest $\mathcal{D}$-deflation of π will be

$$\pi^{\mathcal{D}} = (\pi_1 \oplus \cdots \oplus \pi_{i-1})^{\mathcal{D}} \oplus \pi_i^{\mathcal{D}} \oplus (\pi_{i+1} \oplus \cdots \oplus \pi_t)^{\mathcal{D}}.$$

Any other $\mathcal{D}$-deflation will also have to be written as a direct sum of three permutations in this way, and by induction each of these will involve the respective shortest $\mathcal{D}$-deflation.

□

Thus, for any class $\mathcal{D}$ and permutation π, the *$\mathcal{D}$-profile* of π is the unique shortest $\mathcal{D}$-deflation of π, and is denoted $\pi^{\mathcal{D}}$. Note that setting $\mathcal{D} = \mathrm{Av}(21)$, the set of increasing permutations, returns the original definition of the profile, but if we set $\mathcal{D} = \mathcal{S}$, the set of all permutations, we do not get the substitution decomposition back, as $\pi^{\mathcal{S}} = 1$ for any permutation. However, an easy consequence of the above proof is that if $\pi \notin \mathcal{D}$, and σ is the skeleton of π, then $\sigma \leq \pi^{\mathcal{D}}$.

As mentioned at the beginning of this section, our aim with $\mathcal{D}$-profiles is to be able to to move from the permutations of the wreath product $\mathcal{C} \wr \mathcal{D}$ down to the permutations in the two classes $\mathcal{C}$ and $\mathcal{D}$ in a single step. Thus although initially we may know very little about the structure of a permutation in the basis of $\mathcal{C} \wr \mathcal{D}$, by taking its $\mathcal{D}$-profile we should be left with a permutation involving a (known) basis element of $\mathcal{C}$. Conversely, we want to be able to construct basis elements of $\mathcal{C} \wr \mathcal{D}$ given only the bases of $\mathcal{C}$ and $\mathcal{D}$. These ideas are encapsulated in the following theorem.

Theorem 8.3. *Let $\mathcal{C}$ and $\mathcal{D}$ be two arbitrary permutation classes. Then $\pi \in \mathcal{C} \wr \mathcal{D}$ if and only if $\pi^{\mathcal{D}} \in \mathcal{C}$.*

Proof. One direction is immediate. For the converse, since $\pi \in \mathcal{C} \wr \mathcal{D}$, there exists $\pi' \in \mathcal{C}$ which is a deflation of π by permutations in $\mathcal{D}$. The proof of Lemma 8.2 then tells us that $\pi^{\mathcal{D}} \leq \pi'$, completing the proof. □

Any expression of the form $\pi = \pi^{\mathcal{D}}[\alpha_1, \ldots, \alpha_k]$ is called a *$\mathcal{D}$-profile decomposition* of π, and the blocks α_i are called the *$\mathcal{D}$-profile blocks*. These blocks are not typically uniquely defined. For example, the $\mathrm{Av}(123)$-profile of 234615 is 23514, but it can be decomposed either as $23514[12, 1, 1, 1, 1]$ or $23514[1, 12, 1, 1, 1]$. Thus it will be useful to fix a particular $\mathcal{D}$-profile decomposition, especially as later we are going to need to know about the structure of each of the $\mathcal{D}$-profile blocks.

The *left-greedy $\mathcal{D}$-profile* of π is the decomposition $\pi = \pi^{\mathcal{D}}_{\lambda}[\lambda_1, \lambda_2, \ldots, \lambda_\ell]$ with $\lambda_i \in \mathcal{D}$ for all i, in which λ_1 is first chosen maximally, then λ_2, and so on. Each λ_i is called a *left-greedy $\mathcal{D}$-profile block* of π. This yields the usual, unique, $\mathcal{D}$-profile:

Lemma 8.4. *For any class $\mathcal{D}$ and permutation π, $\pi^{\mathcal{D}} = \pi^{\mathcal{D}}_{\lambda}$.*

Proof. Again, we use induction on $n = |\pi|$. The base case $n = 1$ is trivial, so now suppose $n > 1$. Assume further that $\pi \notin \mathcal{D}$, as otherwise $\pi^{\mathcal{D}} = \pi^{\mathcal{D}}_{\lambda} = 1$ follows immediately. Let $\pi = \pi^{\mathcal{D}}_{\lambda}[\lambda_1, \lambda_2, \ldots, \lambda_\ell]$ be the left-greedy $\mathcal{D}$-profile of π, let $\pi^{\mathcal{D}}[\alpha_1, \alpha_2, \ldots, \alpha_k]$ be any other $\mathcal{D}$-profile decomposition of π, and let $\sigma[\pi_1, \pi_2, \ldots, \pi_m]$ be the substitution decomposition.

Consider first the case where $m = |\sigma| \geq 4$. By the proof of Lemma 8.2, we have $\pi^{\mathcal{D}} = \sigma[\pi_1^{\mathcal{D}}, \pi_2^{\mathcal{D}}, \ldots, \pi_m^{\mathcal{D}}]$. A similar argument shows that $\pi^{\mathcal{D}}_{\lambda} = \sigma[(\pi_1)^{\mathcal{D}}_{\lambda}, (\pi_2)^{\mathcal{D}}_{\lambda}, \ldots, (\pi_m)^{\mathcal{D}}_{\lambda}]$, and by induction $\pi_i^{\mathcal{D}} = (\pi_i)^{\mathcal{D}}_{\lambda}$ for all i, giving the required result.

When $m = 2$, π is either sum or skew decomposable, and we may assume the former. Write $\pi = 12 \cdots t[\pi_1, \pi_2, \ldots, \pi_t]$ where each π_i is sum indecomposable. In the case where every $\pi_i \in \mathcal{D}$, both $\pi^{\mathcal{D}}$ and $\pi_\lambda^{\mathcal{D}}$ will be increasing permutations with $k \le \ell \le t$. When using the left-greedy $\mathcal{D}$-profile decomposition, the block λ_1 was chosen maximally, and so $\alpha_1 \le \lambda_1$. Then the block λ_2 was taken maximally, so the $\mathcal{D}$-profile block α_2 cannot extend further right than the end of λ_2, hence $\alpha_2 \le \lambda_1 \oplus \lambda_2$. Continuing in this manner, we see that, for all i, $\alpha_i \le \lambda_1 \oplus \lambda_2 \oplus \cdots \oplus \lambda_i$, and in particular $\alpha_k \le \lambda_1 \oplus \lambda_2 \oplus \cdots \oplus \lambda_k$. But we must have $k \le \ell$, and so $k = \ell$. The remaining case is where at least one $\pi_i \notin \mathcal{D}$. Pick i to be minimal with this property, and then by the proof of Lemma 8.2,the $\mathcal{D}$-profile breaks into three pieces,

$$\pi^{\mathcal{D}} = (\pi_1 \oplus \cdots \oplus \pi_{i-1})^{\mathcal{D}} \oplus \pi_i^{\mathcal{D}} \oplus (\pi_{i+1} \oplus \cdots \oplus \pi_t)^{\mathcal{D}}.$$

A similar argument holds for the left-greedy $\mathcal{D}$-profile, and then by induction each of the three pieces in the left-greedy $\mathcal{D}$-profile is equal to the corresponding piece in the $\mathcal{D}$-profile. □

There is, of course, nothing special about the left-greedy $\mathcal{D}$-profile; it can be seen that any algorithm to compute a $\mathcal{D}$-profile-like decomposition in which at each stage the blocks are chosen maximally will yield a $\mathcal{D}$-profile deflation. For our purposes, however, when required we will always use the left-greedy algorithm.

8.2 The Minimal Block

The primary aim of this section is to be able to tell if any two points in a permutation belong to the same left-greedy $\mathcal{D}$-profile block, and also a partial converse: given the $\mathcal{D}$-profile deflation, what can we say about the points "between" two specified points? To this end, we define a new concept as follows. Let π be any permutation of length n. For all $1 \le i < j \le n$, the *minimal block of π that contains $\pi(i)$ and $\pi(j)$*, denoted $\mathrm{mb}(\pi; i, j)$, is the segment of π which forms the shortest interval containing both $\pi(i)$ and $\pi(j)$. In other words, there exists $k \le i$ and $\ell \ge j - k$ such that $\mathrm{mb}(\pi; i, j) = \pi(k) \cdots \pi(k + \ell)$ forms an interval but no subsegment of this contains both $\pi(i)$ and $\pi(j)$ and forms an interval.

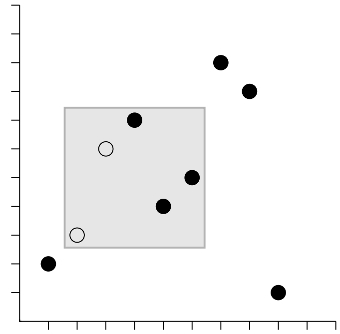

Figure 8.1: The minimal block $\mathrm{mb}(\pi; 2, 3)$ in $\pi = 236745981$.

For example, if $\pi = 236745981$, then the minimal block on $\pi(2) = 3$ and $\pi(3) = 6$ is $\mathrm{mb}(\pi; 2, 3) = 36745$ (See Figure 8.1).

It follows from the observation that the intersection of two intervals itself forms an interval (see Proposition 1.2 (a)) that the minimal block is always uniquely defined. Before we can proceed to the main result of this section, we make one further observation.

Lemma 8.5. *Let π be any permutation and let $i \neq j$ be any pair of positions in π. Then if $k, \ell \in \mathrm{mb}(\pi; i, j)$ with $k \neq \ell$ we have*

$$\mathrm{mb}(\pi; k, l) \subseteq \mathrm{mb}(\pi; i, j).$$

Moreover, if both i and j separate k from ℓ by position, then $\mathrm{mb}(\pi; k, \ell) = \mathrm{mb}(\pi; i, j)$.

Proof. That $\mathrm{mb}(\pi; k, \ell)$ is contained in $\mathrm{mb}(\pi; i, j)$ is obvious. Now suppose i and j separate k from ℓ by position, i.e. $k \leq i < j \leq \ell$. Then $\mathrm{mb}(\pi; k, \ell)$ is an interval of π containing both $\pi(i)$ and $\pi(j)$. As $\mathrm{mb}(\pi; i, j)$ is minimal with this property, we have $\mathrm{mb}(\pi; i, j) \subseteq \mathrm{mb}(\pi; k, \ell)$ and so $\mathrm{mb}(\pi; i, j) = \mathrm{mb}(\pi; k, \ell)$. □

We are now ready to prove our main technical result of this section.

Lemma 8.6. *Let $\mathcal{D}$ be a permutation class, and let $\pi \in S_n$ be any permutation. Then for any pair i, j with $1 \leq i < j \leq n$:*

(i) If the permutation order isomorphic to $\mathrm{mb}(\pi; i, j)$ does not lie in $\mathcal{D}$, then $\pi(i)$ and $\pi(j)$ lie in different $\mathcal{D}$-profile blocks.

(ii) *Conversely, if $\pi(a_i)$ and $\pi(a_j)$ are the first symbols of two distinct left greedy $\mathcal{D}$-profile blocks α_i and α_j respectively, then the permutation order isomorphic to $\mathrm{mb}(\pi; i, j)$ does not lie in $\mathcal{D}$.*

Proof. (i) By minimality and uniqueness of the minimal block, every block in π containing both $\pi(i)$ and $\pi(j)$ must contain the minimal block $\mathrm{mb}(\pi; i, j)$. Hence every such block does not lie in $\mathcal{D}$, so cannot be a $\mathcal{D}$-profile block.

(ii) Write $\pi = \pi^{\mathcal{D}}[\alpha_1, \alpha_2, \ldots, \alpha_k]$, and let the sequence $\pi(a_1), \pi(a_2), \ldots, \pi(a_k)$ represent the leading points in π of the left-greedy $\mathcal{D}$-profile blocks $\alpha_1, \alpha_2, \ldots, \alpha_k$. Let α_i and α_j, $i < j$, be a pair of $\mathcal{D}$-profile blocks. We prove the statement by induction on i.

When $i = 1$, the block α_1 was picked maximally subject to $\alpha_1 \in \mathcal{D}$. For any $j > 1$, the minimal block $\mathrm{mb}(\pi; a_1, a_j)$ strictly contains α_1 and then the maximality of α_1 is contradicted unless $\mathrm{mb}(\pi; a_1, a_j) \notin \mathcal{D}$.

Suppose now that $i > 1$, and that $\mathrm{mb}(\pi; a_\ell, a_j) \notin \mathcal{D}$ for any $\ell < i$ and $j > \ell$. The $\mathcal{D}$-profile block α_i was picked maximally to avoid basis elements of $\mathcal{D}$, subject to starting at symbol $\pi(a_i)$. Consider, for some $j > i$, the minimal block $\mathrm{mb}(\pi; a_i, a_j)$, necessarily containing all of α_i. If the leftmost point of $\mathrm{mb}(\pi; a_i, a_j)$ is $\pi(a_i)$, then since α_i is the maximal block lying in $\mathcal{D}$ which starts at $\pi(a_i)$, we must have $\mathrm{mb}(\pi; a_i, a_j) \notin \mathcal{D}$. So now suppose that $\mathrm{mb}(\pi; a_i, a_j)$ contains at least one symbol $\pi(h)$ from π with $h < a_i$. Let the $\mathcal{D}$-profile block containing $\pi(h)$ be α_ℓ; we claim that α_ℓ is completely contained in $\mathrm{mb}(\pi; a_i, a_j)$. If not, then part of α_ℓ lies outside $\mathrm{mb}(\pi; a_i, a_j)$ in both position and value, and so the part lying inside $\mathrm{mb}(\pi; a_i, a_j)$ itself forms an interval in either the top-left or bottom-left corner of the minimal block, but yet it contains neither $\pi(a_i)$ nor $\pi(a_j)$, contradicting the minimality of $\mathrm{mb}(\pi; a_i, a_j)$. In particular, the first symbol $\pi(a_\ell)$ of α_ℓ is in $\mathrm{mb}(\pi; a_i, a_j)$, and by Lemma 8.5, we have $\mathrm{mb}(\pi; a_\ell, a_j) = \mathrm{mb}(\pi; a_i, a_j)$. By the inductive hypothesis $\mathrm{mb}(\pi; a_\ell, a_j) \notin \mathcal{D}$, and so $\mathrm{mb}(\pi; a_i, a_j) \notin \mathcal{D}$. □

Using this result, we now know when two points of a permutation will lie in the same $\mathcal{D}$-profile block, and, more importantly for what follows, we know that a basis element of $\mathcal{D}$ exists in the minimal block of the first symbols of any two $\mathcal{D}$-profile blocks. What we do not yet know is how to find it; given such a minimal block, we need a method to search through the block systematically and locate the points that form this basis element within a bounded number of steps. Once again it is pin sequences that will provide the solution.

8.3 Pin Sequences and the Wreath Product

For the pin sequences in this chapter, we will revert to considering those that occur within a given permutation, or, indeed, part of a permutation. Recall that for this purpose a proper pin sequence uses the separation condition instead of the externality condition, together with maximality:

- *Maximality*: each pin must be taken maximally in its direction. For example, a proper left pin out of $\mathrm{rect}(p_1, p_2, \ldots, p_{i-1})$ must be the left pin slicing $\mathrm{rect}(p_1, p_2, \ldots, p_{i-1})$ with smallest position.
- *Separation*: in slicing $\mathrm{rect}(p_1, p_2, \ldots, p_i)$, the pin p_{i+1} must lie either horizontally or vertically between p_i and $\mathrm{rect}(p_1, p_2, \ldots, p_{i-1})$.

Also, while we have thus far used pin sequences solely with simple permutations, here we will need to use them in a more general setting. We cannot, of course, expect the same results to hold, but we may prove some that are similar for minimal blocks. Recall that, in a permutation π, a pin sequence $p_1, p_2, \ldots, p_m$ is said to be saturated if $\mathrm{rect}(p_1, p_2, \ldots, p_m)$ encloses all of π. Whereas in simple permutations any pin sequence may be extended to one that is saturated, this is not true for arbitrary permutations, but a weaker condition does hold – we may saturate the minimal block defined on $(i, \pi(i))$ and $(j, \pi(j))$ if these points form the first two points of our pin sequence.

To convert a saturated pin sequence to a proper pin sequence, we first had to restrict our attention towards attaining just one of the boundaries of the permutation. We said that a pin sequence $p_1, p_2, \ldots, p_m$ of π is right-reaching if p_m is the rightmost position of π:

Lemma 2.9. *For every simple permutation π and pair of points p_1 and p_2 (unless, trivially, p_1 is the right-most point of π), there is a proper right-reaching pin sequence beginning with p_1 and p_2.*

We want the same lemma to hold within a minimal block, defined as usual by two points, which will also form the first two points of our proper pin sequence. In the minimal block case, right-reaching means that the last pin is the right-most point of the minimal block, rather than of the whole permutation. Hence:

Lemma 8.7. *Let $\pi \in S_n$ be any permutation, and let $1 \le i < j \le n$. Then there exists a proper pin sequence with starting points $p_1 = (i, \pi(i))$ and $p_2 = (j, \pi(j))$ which is right-reaching in* $\mathrm{mb}(\pi; i, j)$.

Proof. In the minimal block $\mathrm{mb}(\pi; i, j)$, there exists a saturated (non-proper) pin sequence $p_1, p_2, \ldots$ starting from the pins $p_1 = (i, \pi(i))$ and $p_2 = (j, \pi(j))$. If there were no such sequence, then some corner of the minimal block, not including either $\pi(i)$ or $\pi(j)$, would form an interval by itself, contradicting the minimality of $\mathrm{mb}(\pi; i, j)$. Moreover, we may assume, by removing unnecessary pins and relabelling, that every pin is maximal in its direction.

The proof then follows the proof in Chapter 2 of Lemma 2.9. Since the pin sequence is saturated, it includes the rightmost point of π. Label this point p_{i_1}. Next, take the smallest $i_2 < i_1$ such that $p_1, p_2, \ldots, p_{i_2}, p_{i_1}$ is a valid pin sequence, and observe that p_{i_1} separates p_{i_2} from $\mathrm{rect}(p_1, p_2, \ldots, p_{i_2 - 1})$, as $p_1, p_2, \ldots, p_{i_2-1}, p_{i_1}$ is not a valid pin sequence. Continue in this manner, finding pins $p_{i_3}, p_{i_4}, \ldots$ until we reach $p_{i_{m+1}} = p_2$, and then $p_1, p_2, p_{i_m}, p_{i_{m-1}} \ldots, p_{i_1}$ is a proper right-reaching pin sequence. $\square$

Lemma 2.9 is easily recovered from Lemma 8.7 by setting π to be a simple permutation, and observing that all minimal blocks in a simple permutation are the whole permutation. This is, therefore, a true generalisation of that lemma.

We are now ready to prove our main result.

Theorem 8.8. *Let $\mathcal{D} = \mathrm{Av}(B)$ be a finitely based permutation class not admitting arbitrarily long pin sequences. Then $\mathcal{C} \wr \mathcal{D}$ is finitely based for all finitely based classes $\mathcal{C} = \mathrm{Av}(D)$.*

Proof. Let $b = \max_{\beta \in B}(|\beta|)$, $d = \max_{\delta \in D}(|\delta|)$, and π be any permutation in the basis of $\mathcal{C} \wr \mathcal{D}$. By Theorem 8.3, we have $\pi^{\mathcal{D}} \notin \mathcal{C}$, and so there exists some $\delta \in D$ such that $\delta \le \pi^{\mathcal{D}}$. We will be done if we can identify a bounded subsequence of π order isomorphic to a permutation ω, say, for which $\delta \leq \omega^{\mathcal{D}}$, as then $\omega^{\mathcal{D}} \notin \mathcal{C}$ implies $\omega \notin \mathcal{C} \wr \mathcal{D}$, and hence $\omega = \pi$.

First include in our subsequence of π the set of points order isomorphic to δ with positions $d_1, d_2, \ldots, d_k$ ($k = |\delta|$), chosen so that each $\pi(d_i)$ is the leftmost point of a distinct left greedy $\mathcal{D}$-profile block, and the choice of blocks is also leftmost. For every pair d_i, d_{i+1}, Lemma 8.6 tells us that the minimal block $\mathrm{mb}(\pi; d_i, d_{i+1})$ involves some $\beta \in B$, and we include one such occurrence of this β in our subsequence. Our aim now is to add a bounded number of points so that β still lies in the minimal block of the permutation ω on the points corresponding to $\pi(d_i)$ and $\pi(d_{i+1})$, as then these two points are preserved distinctly in $\omega^{\mathcal{D}}$. We do this by taking a proper right-reaching and a proper left-reaching pin sequence of $\mathrm{mb}(\pi; d_i, d_{i+1})$ (which exist by Lemma 8.7), and including them in the subsequence. These pin sequences are only guaranteed to be bounded when $\mathcal{D}$ does not admit arbitrarily long pin sequences, as then there exists a number N so that every pin sequence of length $N+2$ involves some basis element of $\mathcal{D}$.

Thus $\omega^{\mathcal{D}}$ still involves a subsequence order isomorphic to δ, and $|\omega| \leq d + (d-1)(2(N-1) + b)$. $\square$

We saw in Chapter 7 that it is decidable whether a finitely based class admits arbitrarily long pin sequences or not, and therefore given any pattern class we can tell whether Theorem 8.8 applies.

8.4 Infinitely Based Examples

For a class $\mathcal{D}$ which admits infinite pin sequences, Theorem 8.8 gives us no information on whether the basis of $\mathcal{C} \wr \mathcal{D}$ (here for a specified class $\mathcal{C}$) is finite. However, the proof does tell us what some of the basis elements look like. A basis element β of a wreath product $\mathcal{C} \wr \mathcal{D}$ is built around a *core* of points order isomorphic to a basis element of $\mathcal{C}$. To preserve all the points of this core when taking the $\mathcal{D}$-profile of β (as required by Theorem 8.3), every minimal block between any two points of the core must involve a basis element of $\mathcal{D}$. If we can embed arbitrarily long pin sequences in these minimal blocks, β may itself be

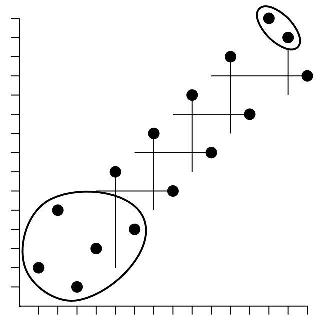

Figure 8.2: The element β_5 in the basis of $\mathrm{Av}(25134) \wr \mathrm{Av}(321)$.

made arbitrarily long. For example, the class $\mathrm{Av}(321)$ admits the increasing oscillating pin sequence encoded $RU\,RU\,RU\cdots$, and so we have:

Theorem 8.9. $\mathrm{Av}(25134) \wr \mathrm{Av}(321)$ *is not finitely based.*

Proof. We exhibit an infinite antichain generated by repeatedly taking up and right pins lying in the basis of $\mathrm{Av}(25134) \wr \mathrm{Av}(321)$. The first few elements of the antichain are

$$\begin{aligned}
\beta_1 &= 2,5,1,3,7,6,4 \\
\beta_2 &= 2,5,1,3,7,4,9,8,6 \\
\beta_k &= 2,5,1,3,7,4 \mid 9,6,11,8,\ldots,2k+3,2k \mid 2k+5,2k+4,2k+2 \quad (k \geq 3).
\end{aligned}$$

Here, as in [9], the $|$ symbol is used only to clarify the structure of the permutation. See Figure 8.2 for an illustration of a typical member of this antichain. We observe:

(i) The set $\{\beta_k \mid k \geq 1\}$ is an antichain.

(ii) The only occurrence of 321 in each β_k is $2k+5, 2k+4, 2k+2$.

(iii) The only occurrence of 25134 in each β_k is $2,5,1,3,\cdot,4$, and hence this forms the core.

(iv) Each β_k is neither sum nor skew decomposable.

(v) The $\mathrm{Av}(321)$-profile of β_k is $2,5,1,3,7,4,\ldots,2k+3,2k,2k+4,2k+2$ (the only non-trivial deflation occurs between $2k+5$ and $2k+4$). In particular, $25134 \prec \beta_k^{\mathrm{Av}(321)}$ for all k, hence by Theorem 8.3 $\beta_k \notin \mathrm{Av}(25134) \wr \mathrm{Av}(321)$.

It only remains to show that β_k is minimally not in $\mathrm{Av}(25134) \wr \mathrm{Av}(321)$. Consider the effect of removing any symbol j. If $j = 2k+5, 2k+4$ or $2k+2$ then by (ii) this no longer involves 321 so $\beta_k - j \in \mathrm{Av}(321) \subset \mathrm{Av}(25134) \wr \mathrm{Av}(321)$. Similarly, if $j = 2, 5, 1, 3$ or 4 then by (iii) $\beta_k - j$ no longer has a core, so $\beta_k - j \in \mathrm{Av}(25134) \subset \mathrm{Av}(25134) \wr \mathrm{Av}(321)$.

For any other j, $\beta_k - j$ is sum decomposable. Under the $\mathrm{Av}(321)$-profile, the first (lower) component deflates to a single point, and hence $(\beta_k - j)^{\mathrm{Av}(321)} \in \mathrm{Av}(25134)$. Thus $\beta_k - j \in \mathrm{Av}(25134) \wr \mathrm{Av}(321)$, completing the proof. □

Note that in the above example, the class $\mathcal{C} = \mathrm{Av}(25134)$ was specifically chosen so that the basis element 25134 is not contained in the repeated pin sequence used to build the antichain, but it does lie in the class $\mathcal{D}$. This ensures that the core, 25134, acts as an anchor at the base of the antichain, but yet the only instance of the basis element 321 is in the upper anchor.

As a result, for any class $\mathcal{D}$ which contains both the infinite pin sequence formed by alternating between up and right pins, and the permutation 25134, the wreath product $\mathrm{Av}(25134) \wr \mathcal{D}$ will always contain an infinite antichain similar to the one above.

Example 8.10. (i) The classes $\mathcal{D} = \mathrm{Av}(321, 2341)$ and $\mathcal{D} = \mathrm{Av}(321, 3412)$ both avoid the permutation 321 and so the antichain in the proof of Theorem 8.9 lies in the basis of $\mathrm{Av}(25134) \wr \mathcal{D}$ in both cases.

(ii) All of the classes $\mathcal{D} = \mathrm{Av}(\alpha, \beta)$ where the pair (α, β) is one of

$$(4321, 4312), (4321, 4231), (4321, 4213), (4321, 3412) \text{ and } (4321, 3214)$$

avoid 4321, and so the antichain with terms

$$\begin{aligned}
\beta_1 &= 2, 5, 1, 3, 8, 7, 6, 4 \\
\beta_2 &= 2, 5, 1, 3, 7, 4, 10, 9, 8, 6 \\
\beta_k &= 2, 5, 1, 3, 7, 4 \mid 9, 6, 11, 8, \ldots, 2k+3, 2k \mid 2k+6, 2k+5, 2k+4, 2k+2 \;\; (k \geq 3)
\end{aligned}$$

lies in the basis of $\mathrm{Av}(25134) \wr \mathcal{D}$ in each case.

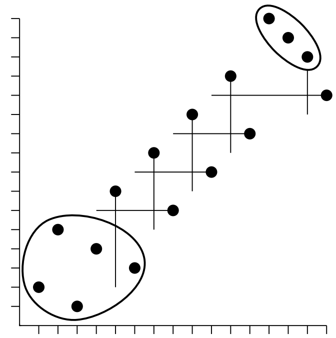

Figure 8.3: The element β_5 in the basis of $\mathrm{Av}(25143) \wr \mathrm{Av}(4321, 4123)$.

(iii) The classes $\mathcal{D} = \mathrm{Av}(4312, 4231)$, $\mathcal{D} = \mathrm{Av}(4312, 4213)$ and $\mathcal{D} = \mathrm{Av}(4312, 3421)$ all avoid 4312, so swapping the order of the final two points of each β_k in case (ii) gives the required antichain.

Example 8.11. The two classes $\mathcal{D} = \mathrm{Av}(4321, 4123)$ and $\mathcal{D} = \mathrm{Av}(4312, 4123)$ both admit the pin sequence formed by repeatedly taking up and right pins, but do not contain the permutation 25134, because of the basis element 4123. However, the class $\mathcal{C} = \mathrm{Av}(25143)$ may be used instead. In the first case, the antichain is (see Figure 8.3 for an illustration):

$$\begin{aligned}
\beta_1 &= 2, 5, 1, 4, 8, 7, 6, 3 \\
\beta_2 &= 2, 5, 1, 4, 7, 3, 10, 9, 8, 6 \\
\beta_k &= 2, 5, 1, 4, 7, 3 \mid 9, 6, 11, 8, \ldots, 2k+3, 2k \mid 2k+6, 2k+5, 2k+4, 2k+2 \quad (k \geq 3).
\end{aligned}$$

All the examples so far have admitted the same "up-right" pin sequence, corresponding to variants of the increasing oscillating antichain. Another commonly found infinite pin sequence is formed by repeating the pattern left, down, right, up,[1] and there are (to within symmetry) two classes of the form $\mathcal{D} = \mathrm{Av}(\alpha, \beta)$ with $|\alpha| = |\beta| = 4$ which admit this sequence: $\mathcal{D} = \mathrm{Av}(3412, 2413)$ and $\mathcal{D} = \mathrm{Av}(3412, 2143)$. Each one must be handled separately.

Example 8.12. (i) $\mathcal{D} = \mathrm{Av}(3412, 2413)$ may be paired with $\mathcal{C} = \mathrm{Av}(31542)$ to produce

[1]This repeating pattern is the foundation for the "Widdershins" antichain of [97].

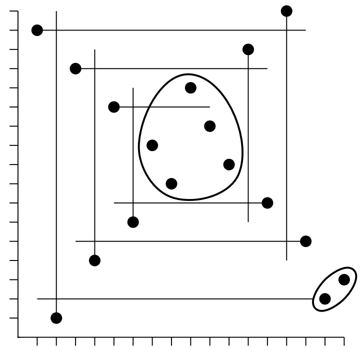

Figure 8.4: The basis element β_3 in $\mathrm{Av}(31542) \wr \mathrm{Av}(3412, 2413)$.

the antichain with terms

$$\begin{aligned}
\beta_1 &= 8, 1, 6, 4, 9, 7, 5, 2, 3 \\
\beta_k &= 4k+4, 1, 4k+2, 4, 4k, 6, \ldots 2k+6, 2k \mid \\
& \quad 2k+4, 2k+2, 2k+7, 2k+5, 2k+3 \mid \\
& \quad 2k+9, 2k+1, \ldots, 4k+5, 5 \mid 2, 3 \quad (k \geq 2).
\end{aligned}$$

See Figure 8.4 for an illustration. Note that the occurrence of 3412 in any β_k is not unique, but every occurrence requires the final two symbols $2, 3$ of β_k, and so these points still behave in the same way as in previous examples.

(ii) $\mathcal{D} = \mathrm{Av}(3412, 2143)$ may be paired with $\mathcal{C} = \mathrm{Av}(412563)$ to produce the antichain with terms:

$$\begin{aligned}
\beta_1 &= 10, 1, 8, 4, 6, 9, 11, 7, 5, 2, 3 \\
\beta_k &= 4k+6, 1, 4k+4, 4, 4k+2, 6, \ldots, 2k+8, 2k \mid \\
& \quad 2k+6, 2k+2, 2k+4, 2k+7, 2k+9, 2k+5, 2k+3 \mid \\
& \quad 2k+11, 2k+1, \ldots, 4k+7, 5 \mid 2, 3 \quad (k \geq 2).
\end{aligned}$$

8.5 Concluding Remarks and Conjectures

The above examples suggest, to some extent, a general method for finding infinite bases. However, these examples rely on just one method for constructing antichains, and there is no reason why this method should always work. For example, a somewhat different construction was used by Atkinson and Stitt [12] to demonstrate an infinite antichain in the basis of $\mathrm{Av}(21) \wr \mathrm{Av}(321654)$, relying on the sum decomposability of the basis element 321654. The other difficulty in finding infinite bases is that, for each given class $\mathcal{D}$, the search for a suitable class $\mathcal{C}$ is very specific, and rarely seems to be applicable to more than a handful of other classes.

In fact, it is unlikely that we can always find such a class $\mathcal{C}$. For example, we saw in Proposition 7.8 that the closure of the increasing oscillating sequence $416385\cdots$ is given by $\mathrm{Av}(321, 2341, 3412, 4123)$. This class, of course, admits the infinite proper pin sequence alternating between up and right pins, but, there are no other permutations in this class which can be used to anchor an infinite antichain based around this pin sequence, so the method described hitherto does not work here. We therefore pose the following question.

Question 8.13. *Is there a finitely based class $\mathcal{C}$ for which $\mathcal{C} \wr \mathrm{Av}(321, 2341, 3412, 4123)$ is not finitely based?*

The Other Direction. Given a finitely based class $\mathcal{C}$, can we tell if $\mathcal{C} \wr \mathcal{D}$ is finitely based for all finitely based permutations classes $\mathcal{D}$? Noting that even $\mathcal{C} = \mathrm{Av}(21)$ does not satisfy this (as witnessed by the infinite basis within $\mathrm{Av}(21) \wr \mathrm{Av}(321654)$), it might be that there are no classes which satisfy this. However, we must not be deceived into thinking that the more well-behaved a class $\mathcal{C}$ is, the more likely $\mathcal{C} \wr \mathcal{D}$ is to be finitely based, as there is no real evidence to support this. We will, however, offer the following conjecture anyway.

Conjecture 8.14. *For any finitely based class $\mathcal{C}$, there exists a finitely based class $\mathcal{D}$ such that $\mathcal{C} \wr \mathcal{D}$ is not finitely based.*

Wreath Basis Decidability. The ultimate aim, of course, is to be able to answer the following question: given two finitely based classes $\mathcal{C}$ and $\mathcal{D}$, what is the basis of $\mathcal{C} \wr \mathcal{D}$? Trivially, if $\mathcal{C}$ and $\mathcal{D}$ both contain finitely many simple permutations, then so does $\mathcal{C} \wr \mathcal{D}$ and so the basis is finite, but this result follows as a special case of Theorem 8.8. A general decision procedure is not likely to be straightforward, and remains somewhat remote. A first step towards such a result would be a better understanding of the structure of infinite antichains.

BIBLIOGRAPHY

[1] ALBERT, M. H., ALDRED, R. E. L., ATKINSON, M. D., VAN DITMARSCH, H. P., HANDLEY, C., HOLTON, D. A., MCCAUGHAN, D. J., AND MONTEITH, C. Cyclically closed pattern classes of permutations. *Australasian J. Combinatorics* , to appear. Cited on pages 126 and 130.

[2] ALBERT, M. H., AND ATKINSON, M. D. Simple permutations and pattern restricted permutations. *Discrete Math. 300*, 1-3 (2005), 1–15. Cited on pages 18, 19, 105, 106, 111, 113, and 131.

[3] ALBERT, M. H., ATKINSON, M. D., AND KLAZAR, M. The enumeration of simple permutations. *J. Integer Seq. 6*, 4 (2003), Article 03.4.4, 18 pp. (electronic). Cited on page 15.

[4] ALBERT, M. H., ATKINSON, M. D., AND RUŠKUC, N. Regular closed sets of permutations. *Theoret. Comput. Sci. 306*, 1-3 (2003), 85–100. Cited on page 151.

[5] ALBERT, M. H., LINTON, S., AND RUŠKUC, N. The insertion encoding of permutations. *Electron. J. Combin. 12*, 1 (2005), Research paper 47, 31 pp. (electronic). Cited on page 150.

[6] ARRATIA, R. On the Stanley-Wilf conjecture for the number of permutations avoiding a given pattern. *Electron. J. Combin. 6* (1999), Note, N1, 4 pp. (electronic). Cited on page 94.

[7] ATKINSON, M. D. Restricted permutations. *Discrete Math. 195*, 1-3 (1999), 27–38. Cited on pages 87, 133, and 156.

[8] ATKINSON, M. D., AND BEALS, R. Permuting mechanisms and closed classes of permutations. In *Combinatorics, computation & logic '99 (Auckland)*, vol. 21 of *Aust. Comput. Sci. Commun.* Springer, Singapore, 1999, pp. 117–127. Cited on page 88.

[9] ATKINSON, M. D., MURPHY, M. M., AND RUŠKUC, N. Partially well-ordered closed sets of permutations. *Order 19*, 2 (2002), 101–113. Cited on pages 95, 98, 100, 101, 102, and 165.

[10] ATKINSON, M. D., MURPHY, M. M., AND RUŠKUC, N. Sorting with two ordered stacks in series. *Theoret. Comput. Sci. 289*, 1 (2002), 205–223. Cited on page 85.

[11] ATKINSON, M. D., MURPHY, M. M., AND RUŠKUC, N. Pattern avoidance classes and subpermutations. *Electron. J. Combin. 12*, 1 (2005), Research paper 60, 18 pp. (electronic). Cited on page 103.

[12] ATKINSON, M. D., AND STITT, T. Restricted permutations and the wreath product. *Discrete Math. 259*, 1-3 (2002), 19–36. Cited on pages 15, 91, 155, and 169.

[13] AVIS, D., AND NEWBORN, M. On pop-stacks in series. *Utilitas Math. 19* (1981), 129–140. Cited on page 86.

[14] BABSON, E., AND STEINGRÍMSSON, E. Generalized permutation patterns and a classification of the Mahonian statistics. *Sém. Lothar. Combin. 44* (2000), Art. B44b, 18 pp. (electronic). Cited on page 117.

[15] BALOGH, J., BOLLOBÁS, B., AND WEINREICH, D. The speed of hereditary properties of graphs. *J. Combin. Theory Ser. B 79*, 2 (2000), 131–156. Cited on page 109.

[16] BALOGH, J., BOLLOBÁS, B., AND WEINREICH, D. A jump to the Bell number for hereditary graph properties. *J. Combin. Theory Ser. B 95*, 1 (2005), 29–48. Cited on page 109.

[17] BERGERON, A., CHAUVE, C., DE MONTGOLFIER, F., AND RAFFINOT, M. Computing common intervals of K permutations, with applications to modular decompo-

sition of graphs. In *Algorithms—ESA 2005*, vol. 3669 of *Lecture Notes in Comput. Sci.* Springer, Berlin, 2005, pp. 779–790. Cited on pages 65, 67, 69, 71, 72, 74, and 75.

[18] BÓNA, M. Exact enumeration of 1342-avoiding permutations: a close link with labeled trees and planar maps. *J. Combin. Theory Ser. A 80*, 2 (1997), 257–272. Cited on pages 94 and 131.

[19] BÓNA, M. The number of permutations with exactly r 132-subsequences is P-recursive in the size! *Adv. in Appl. Math. 18*, 4 (1997), 510–522. Cited on page 132.

[20] BÓNA, M. Permutations with one or two 132-subsequences. *Discrete Math. 181*, 1-3 (1998), 267–274. Cited on page 133.

[21] BÓNA, M. A survey of stack-sorting disciplines. *Electron. J. Combin. 9*, 2 (2003), Article 1, 16 pp. (electronic). Cited on pages 82 and 84.

[22] BÓNA, M. *Combinatorics of permutations*. Discrete Mathematics and its Applications (Boca Raton). Chapman & Hall/CRC, Boca Raton, FL, 2004. With a foreword by Richard Stanley. Cited on page 5.

[23] BONIZZONI, P., AND MCCONNELL, R. M. Nesting of prime substructures in k-ary relations. *Theoret. Comput. Sci. 259*, 1-2 (2001), 341–357. Cited on page 108.

[24] BOSE, P., BUSS, J. F., AND LUBIW, A. Pattern matching for permutations. *Inform. Process. Lett. 65*, 5 (1998), 277–283. Cited on page 85.

[25] BOUSQUET-MÉLOU, M., AND BUTLER, S. Forest-like permutations. arxiv:math.CO/0603617. Cited on page 117.

[26] BOUSQUET-MÉLOU, M., AND STEINGRÍMSSON, E. Decreasing subsequences in permutations and Wilf equivalence for involutions. *J. Algebraic Combin. 22*, 4 (2005), 383–409. Cited on page 132.

[27] 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. Cited on page 10.

[28] BRIGNALL, R. Wreath products of permutation classes. *Electron. J. Combin. 14* (2007), Research paper 46, 15 pp. (electronic). Cited on page xi.

[29] BRIGNALL, R., HUCZYNSKA, S., AND VATTER, V. Decomposing simple permutations, with enumerative consequences. *Combinatorica*, accepted. Cited on page xi.

[30] BRIGNALL, R., HUCZYNSKA, S., AND VATTER, V. Simple permutations and algebraic generating functions. *J. Combin. Theory Ser. A*, accepted. Cited on page xi.

[31] BRIGNALL, R., RUŠKUC, N., AND VATTER, V. Simple permutations: decidability and unavoidable substructures. *Theoret. Comput. Sci.*, accepted. Cited on page xi.

[32] BUER, H., AND MÖHRING, R. H. A fast algorithm for the decomposition of graphs and posets. *Math. Oper. Res. 8*, 2 (1983), 170–184. Cited on page 11.

[33] CHERLIN, G., SHELAH, S., AND SHI, N. Universal graphs with forbidden subgraphs and algebraic closure. *Adv. in Appl. Math. 22*, 4 (1999), 454–491. Cited on page 103.

[34] 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. Cited on page 98.

[35] CLAESSON, A., AND MANSOUR, T. Counting occurrences of a pattern of type $(1, 2)$ or $(2, 1)$ in permutations. *Adv. in Appl. Math. 29*, 2 (2002), 293–310. Cited on page 134.

[36] CORNEIL, D. G., LERCHS, H., AND BURLINGHAM, L. S. Complement reducible graphs. *Discrete Appl. Math. 3*, 3 (1981), 163–174. Cited on page 86.

[37] CORTEEL, S., LOUCHARD, G., AND PEMANTLE, R. Common intervals of permutations. *Discrete Math. Theor. Comput. Sci. 8*, 1 (2006), 189–214 (electronic). Cited on pages 9 and 15.

[38] COURNIER, A., AND HABIB, M. An efficient algorithm to recognize prime undirected graphs. In *Graph-theoretic concepts in computer science (Wiesbaden-Naurod, 1992)*, vol. 657 of *Lecture Notes in Comput. Sci.* Springer, Berlin, 1993, pp. 212–224. Cited on page 77.

[39] COURNIER, A., AND HABIB, M. A new linear algorithm for modular decomposition. In *Trees in algebra and programming—CAAP '94 (Edinburgh, 1994)*, vol. 787 of *Lecture Notes in Comput. Sci.* Springer, Berlin, 1994, pp. 68–84. Cited on page 77.

[40] CRVENKOVIĆ, S., DOLINKA, I., AND MARKOVIĆ, P. A survey of algebra of tournaments. *Novi Sad J. Math. 29* (1999), 95–130. Cited on page 10.

[41] DAHLHAUS, E., GUSTEDT, J., AND MCCONNELL, R. M. Efficient and practical modular decomposition. In *Proceedings of the Eighth Annual ACM-SIAM Symposium on Discrete Algorithms (New Orleans, LA, 1997)* (New York, 1997), ACM, pp. 26–35. Cited on page 78.

[42] DUMONT, D. Interprétations combinatoires des nombres de Genocchi. *Duke Math. J. 41* (1974), 305–318. Cited on page 119.

[43] EGGE, E. S. Restricted symmetric permutations. Preprint. Cited on page 125.

[44] EGGE, E. S. Restricted 3412-avoiding involutions, continued fractions, and Chebyshev polynomials. *Adv. in Appl. Math. 33*, 3 (2004), 451–475. Cited on page 131.

[45] EGGE, E. S., AND MANSOUR, T. 132-avoiding two-stack sortable permutations, fibonacci numbers, and pell numbers. *Discrete Appl. Math. 143*, 1-3 (2004), 72–83. Cited on page 131.

[46] EGGE, E. S., AND MANSOUR, T. 231-avoiding involutions and Fibonacci numbers. *Australas. J. Combin. 30* (2004), 75–84. Cited on page 131.

[47] EHRENFEUCHT, A., HARJU, T., AND ROZENBERG, G. *The theory of 2-structures*. World Scientific Publishing Co. Inc., River Edge, NJ, 1999. Cited on page 10.

[48] EHRENFEUCHT, A., AND MCCONNELL, R. A k-structure generalization of the theory of 2-structures. *Theoret. Comput. Sci. 132*, 1-2 (1994), 209–227. Cited on pages 21 and 108.

[49] ELDER, M. Permutations generated by a stack of depth 2 and an infinite stack in series. *Electron. J. Combin. 13*, 1 (2006), Research Paper 68, 12 pp. (electronic). Cited on page 85.

[50] ELIZALDE, S., AND MANSOUR, T. Restricted Motzkin permutations, Motzkin paths, continued fractions and Chebyshev polynomials. *Discrete Math. 305*, 1-3 (2005), 170–189. Cited on page 131.

[51] ERDŐS, P., FRIED, E., HAJNAL, A., AND MILNER, E. C. Some remarks on simple tournaments. *Algebra Universalis 2* (1972), 238–245. Cited on pages 13, 47, and 57.

[52] ERDŐS, P., HAJNAL, A., AND MILNER, E. C. Simple one-point extensions of tournaments. *Mathematika 19* (1972), 57–62. Cited on page 57.

[53] ERDŐS, P., AND SZEKERES, G. A combinatorial problem in geometry. *Compos. Math. 2* (1935), 463–470. Cited on page 22.

[54] FÖLDES, S. On intervals in relational structures. *Z. Math. Logik Grundlag. Math. 26*, 2 (1980), 97–101. Cited on pages 9 and 11.

[55] FRAÏSSÉ, R. On a decomposition of relations which generalizes the sum of ordering relations. *Bull. Amer. Math. Soc. 59* (1953), 389. Cited on page 3.

[56] FRAÏSSÉ, R. Sur l'extension aux relations de quelques propriétés des ordres. *Ann. Sci. Ecole Norm. Sup. (3) 71* (1954), 363–388. Cited on pages 103, 107, and 108.

[57] FULMEK, M. Enumeration of permutations containing a prescribed number of occurrences of a pattern of length three. *Adv. in Appl. Math. 30*, 4 (2003), 607–632. Cited on page 134.

[58] GALLAI, T. Transitiv orientierbare Graphen. *Acta Math. Acad. Sci. Hungar 18* (1967), 25–66. Cited on page 3.

[59] GALLAI, T. *A translation of T. Gallai's paper: "Transitiv orientierbare Graphen"*. Wiley-Intersci. Ser. Discrete Math. Optim. Wiley, Chichester, 2001, pp. 25–66. Cited on page 3.

[60] GIAKOUMAKIS, V. On the closure of graphs under substitution. *Discrete Math. 177*, 1-3 (1997), 83–97. Cited on page 91.

[61] GUIBERT, O. *Combinatoire des permutations à motifs exclus, en liaison avec mots, cartes planaires et tableaux de Young*. PhD thesis, LaBRI, Université Bordeaux 1, 1995. Cited on page 132.

[62] GUIBERT, O., AND MANSOUR, T. Restricted 132-involutions. *Sém. Lothar. Combin. 48* (2002), Art. B48a, 23 pp. (electronic). Cited on page 131.

[63] GUIBERT, O., AND MANSOUR, T. Some statistics on restricted 132 involutions. *Ann. Comb. 6*, 3-4 (2002), 349–374. Cited on page 131.

[64] GUIBERT, O., AND PERGOLA, E. Enumeration of vexillary involutions which are equal to their mirror/complement. *Discrete Math. 224*, 1-3 (2000), 281–287. Cited on page 125.

[65] GUIBERT, O., PERGOLA, E., AND PINZANI, R. Vexillary involutions are enumerated by motzkin numbers. *Ann. Comb. 5*, 2 (2001), 153–147. Cited on page 132.

[66] GUSTEDT, J. Finiteness theorems for graphs and posets obtained by compositions. *Order 15* (1999), 203–220. Cited on pages 95, 98, 105, 107, and 108.

[67] HIGMAN, G. Ordering by divisibility in abstract algebras. *Proc. London Math. Soc. (3) 2* (1952), 326–336. Cited on pages 99 and 100.

[68] HODGES, W. *Model theory*, vol. 42 of *Encyclopedia of Mathematics and its Applications*. Cambridge University Press, Cambridge, 1993. Cited on page 108.

[69] 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. Cited on page 143.

[70] HUCZYNSKA, S., AND VATTER, V. Grid classes and the Fibonacci dichotomy for restricted permutations. *Electron. J. Combin. 13* (2006), Research paper 54, 14 pp. (electronic). Cited on page 89.

[71] ILLE, P. Indecomposable graphs. *Discrete Math. 173*, 1-3 (1997), 71–78. Cited on page 10.

[72] JAGGARD, A. D. Prefix exchanging and pattern avoidance by involutions. *Electron. J. Combin. 9*, 2 (2002/03), Research paper 16, 24 pp. (electronic). Permutation patterns (Otago, 2003). Cited on page 132.

[73] JAMES, L. O., STANTON, R. G., AND COWAN, D. D. Graph decomposition for undirected graphs. In *Proceedings of the Third Southeastern Conference on Combinatorics, Graph Theory, and Computing (Florida Atlantic Univ., Boca Raton, Fla., 1972)* (Boca Raton, Fla., 1972), Florida Atlantic Univ., pp. 281–290. Cited on page 77.

[74] KAPLANSKY, I. The asymptotic distribution of runs of consecutive elements. *Ann. Math. Statistics 16* (1945), 200–203. Cited on page 15.

[75] KLAZAR, M. On the least exponential growth admitting uncountably many closed permutation classes. *Theoret. Comput. Sci. 321*, 2-3 (2004), 271–281. Cited on pages 95 and 100.

[76] KNUTH, D. E. *The art of computer programming. Vol. 1: Fundamental algorithms.* Addison-Wesley Publishing Co., Reading, Mass., 1969. Cited on pages 81, 82, and 98.

[77] 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. Cited on page 82.

[78] MACMAHON, P. A. *Combinatory Analysis*. Cambridge University Press, London, 1915/16. Cited on page 81.

[79] MANSOUR, T. Restricted 1-3-2 permutations and generalized patterns. *Ann. Comb. 6*, 1 (2002), 65–76. Cited on page 131.

[80] MANSOUR, T. Restricted 132-alternating permutations and Chebyshev polynomials. *Ann. Comb. 7*, 2 (2003), 201–227. Cited on pages 131 and 134.

[81] MANSOUR, T. Counting occurrences of 132 in an even permutation. *Int. J. Math. Math. Sci.*, 25-28 (2004), 1329–1341. Cited on page 134.

[82] MANSOUR, T. Restricted 132-Dumont permutations. *Australas. J. Combin. 29* (2004), 103–117. Cited on page 131.

[83] MANSOUR, T. Restricted even permutations and Chebyshev polynomials. *Discrete Math. 306*, 12 (2006), 1161–1176. Cited on page 131.

[84] MANSOUR, T., AND VAINSHTEIN, A. Restricted 132-avoiding permutations. *Adv. in Appl. Math. 26*, 3 (2001), 258–269. Cited on page 131.

[85] MANSOUR, T., AND VAINSHTEIN, A. Counting occurrences of 132 in a permutation. *Adv. in Appl. Math. 28*, 2 (2002), 185–195. Cited on page 132.

[86] MANSOUR, T., YAN, S. H. F., AND YANG, L. L. M. Counting occurrences of 231 in an involution. *Discrete Math. 306*, 6 (2006), 564–572. Cited on page 134.

[87] MARCUS, A., AND TARDOS, G. Excluded permutation matrices and the Stanley-Wilf conjecture. *J. Combin. Theory Ser. A 107*, 1 (2004), 153–160. Cited on page 94.

[88] McConnell, R. M., and Spinrad, J. P. Linear-time modular decomposition and efficient transitive orientation of comparability graphs. In *Proceedings of the Fifth Annual ACM-SIAM Symposium on Discrete Algorithms (Arlington, VA, 1994)* (New York, 1994), ACM, pp. 536–545. Cited on page 77.

[89] McConnell, R. M., and Spinrad, J. P. Linear-time transitive orientation. In *Proceedings of the Eighth Annual ACM-SIAM Symposium on Discrete Algorithms (New Orleans, LA, 1997)* (New York, 1997), ACM, pp. 19–25. Cited on page 78.

[90] McConnell, R. M., and Spinrad, J. P. Modular decomposition and transitive orientation. *Discrete Math. 201*, 1-3 (1999), 189–241. Cited on page 78.

[91] Möhring, R. H. On the distribution of locally undecomposable relations and independence systems. In *Colloquium on Mathematical Methods of Operations Research (Aachen, 1980)*, vol. 42 of *Methods Oper. Res.* Athenäum/Hain/Hanstein, Königstein/Ts., 1981, pp. 33–48. Cited on page 13.

[92] Möhring, R. H. An algebraic decomposition theory for discrete structures. In *Universal algebra and its links with logic, algebra, combinatorics and computer science (Darmstadt, 1983)*, vol. 4 of *Res. Exp. Math.* Heldermann, Berlin, 1984, pp. 191–203. Cited on page 4.

[93] Möhring, R. H. Algorithmic aspects of comparability graphs and interval graphs. In *Graphs and order (Banff, Alta., 1984)*, vol. 147 of *NATO Adv. Sci. Inst. Ser. C Math. Phys. Sci.* Reidel, Dordrecht, 1985, pp. 41–101. Cited on page 8.

[94] Möhring, R. H. Algorithmic aspects of the substitution decomposition in optimization over relations, sets systems and Boolean functions. *Ann. Oper. Res. 4*, 1-4 (1985), 195–225. Cited on page 4.

[95] Möhring, R. H., and Radermacher, F. J. Substitution decomposition for discrete structures and connections with combinatorial optimization. In *Algebraic and*

combinatorial methods in operations research, vol. 95 of *North-Holland Math. Stud.* North-Holland, Amsterdam, 1984, pp. 257–355. Cited on pages 4, 7, 12, and 17.

[96] MOON, J. W. Embedding tournaments in simple tournaments. *Discrete Math. 2*, 4 (1972), 389–395. Cited on page 57.

[97] MURPHY, M. M. *Restricted permutations, antichains, atomic classes, and stack sorting*. PhD thesis, Univ. of St Andrews, 2002. Cited on pages 21, 85, 86, 89, 95, 102, 150, and 167.

[98] MURPHY, M. M., AND VATTER, V. Profile classes and partial well-order for permutations. *Electron. J. Combin. 9*, 2 (2003), Research paper 17, 30 pp. (electronic). Cited on page 95.

[99] NOONAN, J. The number of permutations containing exactly one increasing subsequence of length three. *Discrete Math. 152*, 1-3 (1996), 307–313. Cited on page 134.

[100] NOONAN, J., AND ZEILBERGER, D. The enumeration of permutations with a prescribed number of "forbidden" patterns. *Adv. in Appl. Math. 17*, 4 (1996), 381–407. Cited on page 134.

[101] POUZET, M. The profile of relations. arXiv:math.CO/0703211. Cited on page 108.

[102] RICHARDS, D. Ballot sequences and restricted permutations. *Ars Combin. 25* (1988), 83–86. Cited on page 93.

[103] ROBERTSON, A., SARACINO, D., AND ZEILBERGER, D. Refined restricted permutations. *Ann. Comb. 6*, 3-4 (2002), 427–444. Cited on page 132.

[104] ROTEM, D. On a correspondence between binary trees and a certain type of permutation. *Information Processing Lett. 4*, 3 (1975/76), 58–61. Cited on page 93.

[105] SABIDUSSI, G. Graph derivatives. *Math. Z. 76* (1961), 385–401. Cited on page 10.

[106] SCHEINERMAN, E. R., AND ZITO, J. On the size of hereditary classes of graphs. *J. Combin. Theory Ser. B 61*, 1 (1994), 16–39. Cited on page 109.

[107] 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. Cited on pages 9, 21, and 108.

[108] SHAPIRO, L., AND STEPHENS, A. B. Bootstrap percolation, the Schröder numbers, and the n-kings problem. *SIAM J. Discrete Math. 4*, 2 (1991), 275–280. Cited on page 86.

[109] SIMION, R., AND SCHMIDT, F. W. Restricted permutations. *European J. Combin. 6*, 4 (1985), 383–406. Cited on page 93.

[110] SLOANE, N. J. A. The on-line encyclopedia of integer sequences. Available online at http://www.research.att.com/˜njas/sequences/. Cited on pages 9, 91, 92, 114, 122, 123, and 126.

[111] SMITH, R. Permutation reconstruction. *Electron. J. Combin. 13* (2006), Note 11, 8 pp. (electronic). Cited on page 6.

[112] STANKOVA, Z., AND WEST, J. A new class of Wilf-equivalent permutations. *J. Algebraic Combin. 15*, 3 (2002), 271–290. Cited on page 93.

[113] STANLEY, R. P. *Enumerative combinatorics. Vol. 1*, vol. 49 of *Cambridge Studies in Advanced Mathematics*. Cambridge University Press, Cambridge, 1997. Cited on page 86.

[114] STANLEY, R. P. *Enumerative combinatorics. Vol. 2*, vol. 62 of *Cambridge Studies in Advanced Mathematics*. Cambridge University Press, Cambridge, 1999. Cited on page 115.

[115] SUMNER, D. P. *Indecomposable graphs*. PhD thesis, Univ. of Massachusetts, 1971. Cited on page 52.

[116] UNO, T., AND YAGIURA, M. Fast algorithms to enumerate all common intervals of two permutations. *Algorithmica 26*, 2 (2000), 290–309. Cited on pages 15 and 65.

[117] VATTER, V. Small permutation classes. In preparation. Cited on pages 89 and 101.

[118] WATON, S. *On Permutation Classes Generated by Token Passing Networks, Gridding Matrices and Pictures: Three Flavours of Involvement*. PhD thesis, Univ. of St Andrews, 2007. Cited on pages 6, 83, and 89.

[119] WEST, J. Sorting twice through a stack. *Theoret. Comput. Sci. 117*, 1-2 (1993), 303–313. Cited on pages 85 and 132.

[120] WEST, J. Generating trees and the Catalan and Schröder numbers. *Discrete Math. 146*, 1-3 (1995), 247–262. Cited on page 93.

[121] WOLFOWITZ, J. Note on runs of consecutive elements. *Ann. Math. Statistics 15* (1944), 97–98. Cited on page 15.

[122] ZVEROVICH, I. A finiteness theorem for primal extensions. *Discrete Math. 296*, 1 (2005), 103–116. Cited on pages 42 and 91.

INDEX