University of California
Los Angeles

A dissertation submitted in partial satisfaction
of the requirements for the degree
Doctor of Philosophy in Mathematics

by

Andrew Soffer

2016

Abstract of the Dissertation

by

Andrew Soffer

Doctor of Philosophy in Mathematics

University of California, Los Angeles, 2016

Professor Igor Pak, Chair

A classical conjecture by Graham Higman states that the number of conjugacy classes in $U_n(q)$, the group of upper triangular $(n \times n)$-matrices over $\mathbb{F}_q$, is a polynomial function of q, for all n. This dissertation concerns itself with both enumerative and asymptotic results regarding the number of conjugacy classes in $U_n(q)$. We present both positive and negative evidence for Higman's conjecture, verifying the conjecture for $n \leq 16$, and suggesting that it probably fails for $n \geq 59$. The tools are both theoretical and computational. We introduce a new framework for testing Higman's conjecture, which involves recurrence relations for the number of conjugacy classes in pattern groups. These relations are proven by the orbit method for finite nilpotent groups.

We also improve the best known asymptotic upper bound on the number of conjugacy classes in $U_n(q)$, and introduce upper bounds on the number of conjugacy classes in groups in the lower central series for $U_n(q)$. To do so, we introduce a technique involving a combinatorial structure called a *gap array*. Gap arrays encode properties of centralizers of Jordan forms. By proving asymptotic results on the structure of gap arrays we can deduce asymptotic results about the number of conjugacy classes in $U_n(q)$.

The dissertation of Andrew Soffer is approved.

Raphäel Rouquier

Bruce Rothschild

Robert Guralnick

Igor Pak, Committee Chair

University of California, Los Angeles

2016

To my parents,

whose love and support have made

this both possible and probable.

Table of Contents

List of Figures

Acknowledgments

I must first thank my advisor, Igor Pak, for both his mathematical support and his relentless encouragement. His unique combination of humor and a never-give-up attitude have led me to take on mathematical problems harder than I could have attempted myself.

The results in this thesis could not have been discovered had it not been for the fruitful conversations with my colleagues. I would like to explicitly mention Adam Azzam, Stephen DeSalvo, Ted Dokos, Scott Garrabrant, Alejandro Morales, and Greta Panova. Each of these people have shared with me much of their valuable time and many of their valuable ideas.

More broadly, the environment of camaraderie and support in the UCLA mathematics department has been inspiring. My friends at UCLA made it easy to smile and laugh through the stresses of graduate school. Aside from providing the necessary emotional recharge, the social atmosphere in the UCLA mathematics department has explicitly led to fruitful mathematical ideas. The recurring joke[1] "did you try Jordan canonical forms?" was the basis for the ideas discussed in Chapter 9.

My family has been exceptionally supportive throughout my graduate studies. Leah, Michael, and Rachelli: I cannot overstate my appreciation for your patience, understanding, and encourangement.

For as long as I can remember, and certainly longer, my parents, Robin and Leonard Soffer, have instilled in me an eagerness to learn. I have always felt at home in an environment where I can learn and think critically, primarily because their home has always been such an environment.

[1]During our first year as graduate students, Dave Clyde managed to solve nearly every linear algebra problem using Jordan canonical forms, often in spite of the problem writer's intended solution. It became an ongoing joke that Dave could solve any problem with with Jordan canonical forms.

Vita

2011	B.A. in Mathematics, minor in Computer Science, Washington University in Saint Louis, Saint Louis, Missouri

CHAPTER 1

Introduction

This dissertation primarily concerns itself with the group $U_n(q)$ of upper-triangular matrices over the finite field $\mathbb{F}_q$. Little is known about the structure of this group or how the structure varies with the parameter q. Of particular interest will be questions regarding the number and structure of conjugacy classes in $U_n(q)$. The goal of understanding more about the group structure of such a ubiquitous linear-algebraic construction is certainly motivation enough to ask questions about their conjugacy classes, yet historically it did not happen this way. Rather, the question fell into existence by happenstance.

1.1 Historical motivation

The first mention of the number of conjugacy classes in the group of upper-triangular matrices was in Graham Higman's seminal 1960 paper "Enumerating p-groups I: Inequalities," [H1]. Upper-triangular matrices were the main content of neither this paper nor its successor [H2]. Rather, they were the development of tools to answer questions regarding the number of isomorphism classes of p-groups. For a prime p and integer n, let $f(p, n)$ denote the number of groups of order p^n. Higman's papers were primarily concerned with Questions 1.1 and 1.2.[2]

Question 1.1. For each fixed n, as a function of p, can $f(p, n)$ be described by a collection of polynomials $P_1, \ldots, P_m$ such that $f(p, n) = P_k(p)$ whenever $p \equiv k \pmod{m}$.

Stated more succinctly, Question 1.1, asks if for each fixed n, the function $f(p, n)$ is *polynomial on residue classes* (PORC). Higman's PORC conjecture, as it is now called, has

[2] Questions 1.1 and 1.2 are beyond the scope of this dissertation, and purely of historical and motivational interest.

received much attention since its inception (see, for example [H2, DV, VL]). While we will not explicitly discuss the PORC conjecture in this dissertation, we will devote much of our energy to a question of a similar flavor (cf. Conjecture 1.3). Higman's PORC conjecture has inspired many similar questions about "how nicely behaved" particular functions of primes or prime powers must [BB, HP, Vak]. The results of Halasi and Pálfy in [HP] will be relevant to our discusison in Chapter 5.

Higman also asked about the asymptotic behavior of the function $f(p, n)$.

Question 1.2. For each fixed p, as a function of n, how fast does $f(p, n)$ grow?

This question, or rather the partial answer to this question given in [H1], is the key connection between p-groups and upper-triangular matrices. In [H1] Higman proved an upper bound on the number of conjugacy classes in $U_n(q)$, and used this estimate to provide an upper bound for the number of groups of order p^n. In this way, he tied the asymptotic growth rate for p-groups to the asymptotic growth rate for the number of conjugacy classes in $U_n(q)$. Five years later, Charles Sims answered Question 1.2 using entirely different techniques [Si]. The question of determining the asymptotic behavior of $U_n(q)$ remained unresolved.

1.2 Problem statement and definitions

For a prime power q, let $\mathbb{F}_q$ denote the field with q elements, and let $\mathcal{U}_n(q)$ denote the nilpotent $\mathbb{F}_q$-algebra of strictly upper-triangular matrices. Concretely, we let $e_{i,j}$ denote the $(n \times n)$-matrix which has a 1 in cell (i, j) and has zeros everywhere else. Then

$$\mathcal{U}_n(q) = \bigoplus_{i<j} \mathbb{F}_q e_{i,j},$$

with multiplication defined by $e_{i,j} \cdot e_{k,\ell} = \delta_{j,k} e_{k,\ell}$ (where δ is the Kronecker δ-function).

We write $U_n(q)$ for the *unitriangular group*. That is, $U_n(q)$ denotes the group of upper-triangular matrices over $\mathbb{F}_q$ with ones on the diagonal, so

$$U_n(q) = 1 + \mathcal{U}_n(q).$$

The field $\mathbb{F}_q$ will almost always be clear from context. In such cases we will omit the parameter q, favoring the symbols U_n and $\mathcal{U}_n$.[3]

For a group G, let $\mathbf{k}(G)$ denote the number of conjugacy classes in G. We are now in a position to state precisely what is now known as *Higman's conjecture*.

Conjecture 1.3 (Higman, 1960)**.** *Let f_n denote the function on prime powers given by $f_n(q) = \mathbf{k}(U_n(q))$. Then $f_n \in \mathbb{Z}[q]$.*

Higman never stated this conjecture explicitly, nor did he claim his belief in the conjecture. Rather, the conjecture arose from a brief remark seemingly unrelated to the rest of [H1].

> "...it would also be interesting to know whether, as a consideration of small values of n suggests, for fixed n, the class number is a polynomial in q."

This question bears resemblance to those discussed in Higman second paper [H2] regarding counting problems whose solutions are PORC. In each case, the questions ask about how nicely behaved a family of functions must be. These questions are instances of a more general phenomenon known as *universality*, or sometimes *Murphy's law*. A *universality result* is one that says that a particular family of objects is "as bad as one desires." The precise meaning of this statement is dependent on the context; it is the general philosophy that ties the results together. In the case of $\mathbf{k}(U_n(q))$, a universality result would state that Higman's conjecture is badly false: That the functions $\mathbf{k}(U_n(q))$ could be perhaps as complicated as the number of $\mathbb{F}_q$-points on any algebraic variety.

Also of interest to us will be the question about the growth rate of $\mathbf{k}(U_n(q))$ as n tends to infinity.

Question 1.4. For each fixed prime power q, determine the asymptotic behavior of $f_n(q) = \mathbf{k}(U_n(q))$ as $n \to \infty$. Specifically, determine the constant c such that

$$\mathbf{k}(U_n(q)) = q^{cn^2(1+o(1))},$$

[3] More generally, throughout this dissertation when we define subalgebras of $\mathcal{U}_n$, we will use calligraphic letters such as $\mathcal{A}$, and their roman counterpart for the corresponding group, such as $1 + \mathcal{A} = A \le U_n$.

or prove that no such constant exists.

1.3 Results

Little is known about the exact functions values of $\mathbf{k}(U_n(q))$, even for n of moderate size. Gudivok et al. computed $\mathbf{k}(U_n(q))$ for all $n \leq 8$ [G$^+$]. The authors use a variation on the brute force algorithm. Later Arregi and Vera-López were able to use a different technique to compute $\mathbf{k}(U_n(q))$ for all $n \leq 13$. In Chapter 4 we develop combinatorial tools which extend the results of Gudivok et al. and Arregi and Vera-López, by proving the following:

Theorem 1.5. *Higman's conjecture holds for $n \leq 16$. Moreover, for all $n \leq 16$, we have $\mathbf{k}(U_n) \in \mathbb{N}[q-1]$.*

In Chapter 6 we describe our algorithm. The algorithm is based on counting co-adjoint orbits of pattern groups (see Chapter 4) using the combinatorial techniques developed in Section 4.2. These tools rely on a new structure which we call a *poset system*. When our algorithm cannot proceed due to the combinatorial complexity of the poset which must be we rely on a variant of the algorithm given by Arregi and Vera-López in [VA4].

In 2011, Halasi and Pálfy were able to show that a generalization of Higman's conjecture is false [HP]. In particular, they exhibitted the existence of pattern groups U_P for which $\mathbf{k}(U_P(q))$ is not a polynomial. We construct an explicit example of a 13-element poset for which $\mathbf{k}(U_P(q))$ is a polynomial function of q, and use our computational tools to prove the following result:

Theorem 1.6. *There is a pattern subgroup $U_{\mathfrak{P}}(q) \leq U_{13}(q)$ such that $\mathbf{k}(U_{\mathfrak{P}}(q))$ is* not *a polynomial function of q.*

In Chapter 5 we introduce a notion of pattern group embedding, and show that all pattern groups embed into U_n, if n is large enough. Taken with the results of Halasi and Palfy, we believe this to be good evidence that $\mathbf{k}(U_n(q))$ is not always a polynomial.

Regarding the asymptotic behavior of $\mathbf{k}(U_n(q))$, in [H1] Higman showed that

$$q^{\frac{1}{12}n^2(1+o(1))} \leq \mathbf{k}(U_n(q)) \leq q^{\frac{1}{4}n^2(1+o(1))}.$$

This lower bound with exponent $\frac{1}{12}n^2(1+o(1))$ has not been improved and is likely the correct asymptotic behavior of $\mathbf{k}(U_n(q))$ (cf. Section 10.4). The upper bound however has been improved. In 1992, Arregi and Vera-López showed $\mathbf{k}(U_n(q)) \leq q^{\frac{1}{6}n^2(1+o(1))}$ [VA1]. Their techniques focus on canonical matrices, a unique element in each conjugacy class in $U_n(q)$ defined in their papers [VA1, VA2, VA3, VA4]. Arregi and Vera-López were able to prove certain properties about the form of a canonical matrix and use this to obtain an upper bound for the number of such matrices.

In his undergraduate thesis [Mar1], Eric Marberg used the theory of supercharacters to obtain the same upper bound with exponent $\frac{1}{6}n^2(1+o(1))$. His proof highlighted the combinatorial nature of this problem.

In Chapter 8 we improve upon these bounds slightly. Specifically, we show

Theorem 1.7. *For every positive integer n and every prime power q, we have*

$$\mathbf{k}(U_n(q)) \leq p(n)^2 n!\; q^{\alpha n^2 + \frac{n}{2}},$$

where $p(n)$ denotes the number of integer partitions of n, and where

$$\alpha = \frac{40\sqrt{2}-41}{98} \approx 0.15886.$$

We also compute upper bounds on the number of conjugacy classes in each group in the lower central series of of $U_n(q)$. Let $U_{n,k}(q)$ denote the subgroup of $U_n(q)$ for which the first k diagonals above the main diagonal contain all zeros.

Theorem 1.8. *For each $m \in \mathbb{N}$, define*

$$\gamma_m := \frac{1}{6} - \frac{13}{24} \cdot 4^{-m} + 2^{-(m+1)} - 4^{-(m+1)}m.$$

Then for every q,

$$\mathbf{k}(U_{n,k}(q)) \leq q^{\gamma_m n^2(1+o_m(1))},$$

where $m = \left\lfloor \log_2 \binom{n}{k} \right\rfloor$, *and* $o_m(1)$ *denotes a function which, for each fixed* m, *tends to zero as* n *tends to infinity.*

1.4 Notation

For the most part, we introduce notation in the context in which it arises. However, there are certain conventions we use throughout this entire dissertation. We collect these conventions here for convenience. We use $\mathbb{N}$ to denote the natural numbers including zero. That is, $\mathbb{N} = \{n \in \mathbb{Z} : n \geq 0\}$. For the positive integers, we write $\mathbb{Z}^+ = \{n \in \mathbb{Z} : n > 0\}$.

1.4.1 Algebraic notation

For a prime power q, let $\mathbb{F}_q$ denote the field with q elements, and let $\mathcal{U}_n(q)$ denote the $\mathbb{F}_q$-algebra of strictly upper-triangular matrices. Define the *unitriangular group* to be

$$U_n(q) := \{1 + X : X \in \mathcal{U}_n(q)\}.$$

The field $\mathbb{F}_q$ will almost always be clear from context. In such cases we will omit the parameter q.

For an associative algebra $\mathcal{A}$, define

$$\mathrm{Comm}(\mathcal{A}) = \{(X, Y) \in \mathcal{A} \times A : XY = YX\}.$$

For a group G, we use the same notation $\mathrm{Comm}(G)$ to denote the set of pairs of commuting elements in G. Note that $X, Y \in \mathcal{U}_n$ commute if and only if the elements $1 + X, 1 + Y \in U_n$ commute, so $|\mathrm{Comm}(\mathcal{U}_n)| = |\mathrm{Comm}\, U_n|$.

For a finite group G, let $\mathbf{k}(G)$ denote the number of conjugacy classes in G. From Burnside's lemma,

$$\mathbf{k}(G) = \frac{1}{|G|} \sum_{g \in G} |C_G(g)| = \frac{|\mathrm{Comm}(G)|}{|G|},$$

where $C_G(g)$ denotes the centralizer of $g \in G$. That is, $C_G(g) = \{h \in G : gh = hg\}$.

For positive integers a and b, let $\mathcal{M}_{a\times b}$ denote the vector space of $(a \times b)$-matrices over $\mathbb{F}_q$. We write $C_{\mathcal{M}}(X)$ for the centralizer of X in $\mathcal{M}_{n\times n}$. That is,

$$C_{\mathcal{M}}(X) := \{A \in \mathcal{M}_{n\times n} \mid AX = XA\}.$$

Similarly, define $C_{\mathcal{U}}(X) := \{A \in \mathcal{U}_n \mid AX = XA\}$.

1.4.2 Jordan canonical forms

Let $\lambda = (\lambda_1, \dots, \lambda_\ell)$ be a partition of n. By J_λ, we denote the Jordan canonical nilpotent matrix which has blocks of size $\lambda_1, \lambda_2, \dots, \lambda_\ell$. For example,

$$J_{(3,2)} = \left(\begin{array}{ccc|cc} 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ \hline 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 \end{array}\right).$$

Every upper-triangular[4] matrix $A \in \mathcal{U}_n$ is GL-conjugate to some matrix J_λ. We say that λ is the *shape* of A if A is GL-conjugate to J_λ, and write $\mathrm{sh}(A) = \lambda$. For $A \in \mathcal{U}_n$, and $1 \le k \le n$, let $A|_k$ denote the top-left $(k \times k)$-submatrix of A.

Remark 1.9. Though we do not use this fact explicitly, it is interesting to note that

$$\mathrm{sh}(A|_1) \subseteq \mathrm{sh}(A|_2) \subseteq \cdots \subseteq \mathrm{sh}(A|_n).$$

Thus, to each $A \in \mathcal{U}_n$, we can associate a Young tableau defined by the path in the Young lattice described above. For more information on standard Young tableaux, see [Sa].

1.4.3 Posets

For us, all posets will be finite and typically denoted by the letters P and Q. As a slight abuse of notation (though a common one), we identify the poset P with the ground set on

[4]Note that it is not the case that every matrix in $\mathrm{GL}_n(q)$ is conjugate to a Jordan canonical form. The fact that every $A \in \mathcal{U}_n$ is conjugate to some J_λ is a consequence of the fact that it is nilpotent (and thus has a characteristic polynomial which splits over any $\mathbb{F}_q$).

which partial order "$\prec$" is defined. We use $\mathbf{C}_n$ to denote an n-element chain, and $\mathbf{A}_n$ to denote an n-element anti-chain (a poset with n elements, none of which are comparable to any others). We use $\max(P)$ and $\min(P)$ to denote the set of maximal and minimal elements, respectively. The set of anti-chains in P is denoted by $\mathrm{ac}(P)$. The set of pairs of distinct related elements on a poset P is denoted

$$\mathrm{rel}(P) := \{(x,y) : x \prec_P y\}.$$

The *upper* and *lower bounds* of an element $x \in P$ are defined as

$$\mathrm{ub}_P(x) := \{y \in P : x \prec y\} \quad \text{and} \quad \mathrm{lb}_P(x) := \{y \in P : y \prec x\}.$$

For a subset $S \subseteq P$, let $P|_S$ denote the subposet of P induced on the set S. As a special case, for $x \in P$, we write $P - x$ for the subposet of P induced on $P \setminus \{x\}$.

For an element $x \in P$, let $P^{\langle x \rangle}$ denote the poset consisting exclusively of the relations where the larger element is x. That is, we have $\mathrm{rel}(P^{\langle x \rangle}) = \{(w,x) : w \in \mathrm{lb}_P(x)\}$.

We say that a poset P is Q-free if no induced subposet of P is isomorphic to Q. For example, a poset is $\mathbf{A}_2$-free if and only if it is a chain. Similarly, a poset is $\mathbf{C}_2$-free if and only if it is $\mathbf{A}_n$.

For a poset P, the *dual* poset P^* will be the one whose relations are reversed. That is, if $x \prec_P y$, then $y \prec_{P^*} x$. We also define two constructions of posets from smaller ones. First, for posets P and Q, their *disjoint union* $P \amalg Q$ is a poset whose elements are the elements of P and Q, and for which $x \prec y$ if either

1. $x, y \in P$, and $x \prec_P y$, or

2. $x, y \in Q$, and $x \prec_Q y$.

Up to isomorphism, the operation $\amalg$ is both commutative and associative. Second, the *lexicographic sum* $P + Q$ is the poset whose elements are the elements of P and Q, and for which $x \prec y$ if any of the following hold:

1. $x, y \in P$, and $x \prec_P y$,

2. $x, y \in Q$, and $x \prec_Q y$, or

3. $x \in P$ and $y \in Q$.

In terms of the Hasse diagrams (the usual graphical representation of a poset), the lexicographic sum is obtained by placing Q above P. The lexicographic sum is not commutative, but is associative (up to isomorphism).

1.4.4 Partitions

We adopt mostly standard notation regarding partitions (see, e.g. [Mac, St]). For a partition $\lambda \vdash n$, we let $m_i = m_i(\lambda) := \#\{j : \lambda_i = j\}$, the multiplicity of i in λ. We also use the notation

$$\mathbf{n}(\lambda) := \sum_i (i-1)\lambda_i = \sum_i \binom{\lambda_i'}{2}. \tag{1.1}$$

It is common to use the character "n" for both this function and this size of a partition. We choose to use the bold roman "**n**" for the function to avoid confusion.

It will be useful to treat partitions as infinite decreasing sequences of positive integers $(\lambda_1, \lambda_2, \dots)$ where $\lambda_i = 0$ for $i > \ell(\lambda)$. In this way, we can treat λ as an element of $\ell^1(\mathbb{Z}^+)$. Moreover, as $\ell^1(\mathbb{Z}^+) \subseteq \ell^2(\mathbb{Z}^+)$, it makes sense to talk about the inner product of two partitions. For partitions λ and μ (not necessarily of the same integer) define

$$\langle \lambda, \mu \rangle := \sum_i \lambda_i \mu_i, \quad \text{and} \quad \|\lambda\| := \sqrt{\langle \lambda, \lambda \rangle}.$$

These definitions agree with the definitions of $\langle \cdot, \cdot \rangle$ and $\| \cdot \|$ on $\ell^2(\mathbb{Z}^+)$. Note that combining (1.1) with this analytic notation, the function $\mathbf{n}$ can be expressed as

$$\mathbf{n}(\lambda) = \frac{\|\lambda'\|^2}{2} - \frac{n}{2} \tag{1.2}$$

Similarly, because the multiplicity m_i of the part i in λ can be computed as $m_i = \lambda_i' - \lambda_{i+1}'$, we have

$$(m_1, m_2, \dots) = \lambda' - L\lambda' \tag{1.3}$$

where L denotes the left-shift operator on $\ell^1(\mathbb{Z}^+)$.

Remark 1.10. We define $\|\cdot\|$ to be the ℓ^2-norm, rather than the ℓ^1-norm. Treating $\lambda \vdash n$ as an element of $\ell^1(\mathbb{Z}^+)$, its ℓ^1-norm is simply n, and therefore does not merit its own notation. When we need the ℓ^1-norm of a vector $v \in \ell^1(\mathbb{Z}^+)$, we will explicitly add a subscript to the norm, as in $\|v\|_1$.

CHAPTER 2

Pattern groups

2.1 Definitions

Pattern groups were first introduced in 1955 by A. J. Weir with the name *partition subgroup* [W]. More recently in [Is], Marty Isaacs rediscovered these objects and coined the term *pattern group*. We recall the definition of pattern algebras and pattern groups, but with slightly different notation in order to highlight their relationship with posets. This relationship between pattern groups and posets will be crucial to our discussion in Section 4.2.

Definition 2.1. Let P be a poset on $\{1, 2, \ldots, n\}$ which has the standard ordering as a linear extension. That is, whenever $i \preccurlyeq_P j$, then we also have $i \leq j$. Define the *pattern algebra* $\mathcal{U}_P(q)$ to be

$$\mathcal{U}_P(q) := \{X \in \mathcal{M}_{n\times n}(q) : X_{i,j} = 0 \text{ if } i \not\preccurlyeq_P j\}.$$

Every pattern algebra $\mathcal{U}_P(q)$ is a nilpotent $\mathbb{F}_q$-algebra. In fact, the pattern algebra $\mathcal{U}_P(q)$ is a subalgebra of the strictly upper-triangular matrices $\mathcal{U}_n(q)$.

Remark 2.2. Note that $\mathcal{U}_P$ depends not only on the isomorphism class of the poset P, but also on a specific linear extension of P. This definition is purely one of convenience. One could define $\mathcal{U}_P$ abstractly in terms of generators and relations in such a way as to make it clear that if P and Q are isomorphic posets, then $\mathcal{U}_P \cong \mathcal{U}_Q$. We use this isomorphism without further mention. For more details, the interested reader should read about *incidence algebras* (see, for example [Sp]).

Definition 2.3. For a poset P, define the *pattern group* $U_P(q)$ by

$$U_P(q) := 1 + \mathcal{U}_P(q) = \{1 + X : X \in \mathcal{U}_P(q)\}.$$

For general posets P, the group $U_P(q)$ is a subgroup of the unitriangular group $U_n(q)$. To simplify notation, we often omit the field and write U_P instead of $U_P(q)$. Similarly, we abbreviate $\mathcal{U}_P(q)$ by $\mathcal{U}_P$.

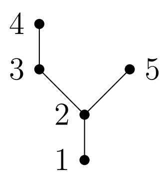
Example 2.4. If P is the n-element chain $\mathbf{C}_n$, then $\mathcal{U}_P = \mathcal{U}_n$, and $U_P = U_n$.

Example 2.5. If P is the n-element anti-chain $\mathbf{A}_n$, then $\mathcal{U}_P$ is the trivial algebra, and U_P is the trivial group.

Example 2.6. If P is the poset shown in Figure 2.1, then $\mathcal{U}_P$ consists of matrices of the form shown in the figure. We can see that as a vector space, $\mathcal{U}_P$ is generated by

$$\{e_{1,2},\ e_{1,3},\ e_{1,4},\ e_{1,5},\ e_{2,3},\ e_{2,4},\ e_{2,5},\ e_{3,4}\}.$$

These are precisely the elements $e_{i,j}$ where $i \prec_P j$. As an algebra, $\mathcal{U}_P$ can be generated by fewer elements. In particular, the pattern algebra $\mathcal{U}_P$ can be generated (as an algebra) by $\{e_{1,2}, e_{2,3}, e_{3,4}, e_{2,5}\}$. Note that $e_{i,j}$ is in this set precisely when i and j are connected by a line segment in the Hasse diagram (see Figure 2.1). The generators are the minimal relations (in the language of posets, the *cover relations*).



$$\begin{pmatrix} 1 & * & * & * & * \\ 0 & 1 & * & * & * \\ 0 & 0 & 1 & * & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix}$$

Figure 2.1: A poset P and the form of elements in the associated pattern group U_P.

Pattern groups have a particularly nice presentation which we will need in Section 4.2.

Proposition 2.7. *For every poset P, we have*

$$U_P(k) = \left\langle \mathcal{E}_{x,y}(\alpha) \middle| x \prec_P y, \alpha \in k^\times \right\rangle.$$

Moreover, for every $\alpha, \beta \in k^\times$, we have

$$[\mathcal{E}_{w,x}(\alpha), \mathcal{E}_{y,z}(\beta)] = \begin{cases} \mathcal{E}_{w,z}(\alpha\beta) & \text{if } x = y, \\ \mathcal{E}_{y,x}(-\alpha\beta) & \text{if } w = z, \\ 1 & \text{if } x \neq y. \end{cases}$$

2.2 Normal pattern groups

In the same 1955 paper where Weir introduced pattern groups, he also characterized those pattern groups which are normal and characteristic in U_n.

Theorem 2.8 (Theorem 2 in [W]). *A pattern group is normal in U_n if and only if whenever a cell can be non-zero, every cell above or to its right can be non-zero.*

Definition 2.9. Let P be a poset, and let Q be a subposet of P. We say Q is *normal* in P if both of the following conditions hold:

- Whenever $x \prec_P y$ and $y \prec_Q z$, then we have $x \prec_Q z$.
- Whenever $x \prec_Q y$ and $y \prec_P z$, then we have $x \prec_Q z$.

Mimicking group theoretic notation, we write $Q \trianglelefteq P$.

It is worth considering explicitly the case where $P = \mathbf{C}_n$. In this case, the relation $\prec_P$ is just the stdandard ordering $<$ on $\{1, 2, \ldots, n\}$. The posets which are normal in $\mathbf{C}_n$ are those which satisfy the conditions

- $x < y \prec_Q z$ implies $x \prec_Q z$, and
- $x \prec_Q y < z$ implies $x \prec_Q z$.

These conditions can also be expressed in terms of the cells in U_Q which are allowed to be non-zero. The first condition says that if a cell is allowed to be non-zero, then so is any cell above it. The second conditions states that if a cell is allowed to be non-zero, then any cell to its right may be non-zero. In this way, Weir's result can be restated by saying that $U_Q \trianglelefteq U_n$ if and only if Q is normal in $\mathbf{C}_n$. With this example in hand, the following generalization of Weir's result should not be surprising.

Proposition 2.10. *Let P and Q denote posets on the base set $\{1, 2, \ldots, n\}$. The following are equivalent:*

1. *$U_Q(k) \trianglelefteq U_P(k)$ for every field k.*
2. *There exists a field k for which $U_Q(k) \trianglelefteq U_P(k)$.*
3. *$Q \trianglelefteq P$.*

Proof. It is evident that (1) implies (2). To see that (2) implies (3), suppose that $x \prec_Q y$ and $y \prec_P z$. Recall that $\mathcal{E}_{y,z}(t)^{-1} = \mathcal{E}_{y,z}(-t)$. For any $\alpha \in k$, conjugating $\mathcal{E}_{x,y}(\alpha)$ by $\mathcal{E}_{y,z}(-1)$ yields an element in $U_Q(k)$. Using Proposition 2.7,

$$\begin{aligned}\mathcal{E}_{y,z}(-1)\mathcal{E}_{x,y}(\alpha)\mathcal{E}_{y,z}(1) &= [\mathcal{E}_{y,z}(-1), \mathcal{E}_{x,y}(\alpha)]\mathcal{E}_{x,y}(\alpha) \\ &= \mathcal{E}_{x,z}(\alpha)\mathcal{E}_{x,y}(\alpha).\end{aligned}$$

As $\mathcal{E}_{x,y}(\alpha) \in U_Q(k)$, it follows that $\mathcal{E}_{x,z}(\alpha) \in U_Q(k)$ as well. Thus, $x \preccurlyeq_Q z$. The second condition for poset normality can be verified analogously, conjugating $\mathcal{E}_{y,z}(\alpha)$ by $\mathcal{E}_{x,y}(-1)$.

Lastly, to see that (3) implies (1), it suffices to show that $U_Q(k)$ is fixed by each $\mathcal{E}_{x,y}(\alpha)$ where $x \prec_P y$ and $\alpha \in k$. Moreover, it suffices to show that each element of a generating set for $U_Q(k)$, when conjugated by $\mathcal{E}_{x,y}(\alpha)$, is sent to another element of $U_Q(k)$.

We choose the generating set $\{\mathcal{E}_{x,y}(\beta) : x \prec_Q y, \beta \in k\}$. A calculation similar to the one above reveals

$$\mathcal{E}_{a,b}(\alpha)\mathcal{E}_{x,y}(\beta)\mathcal{E}_{a,b}(-\alpha) = \begin{cases} \mathcal{E}_{a,y}(\alpha\beta)\mathcal{E}_{x,y}(\beta) & \text{if } b = x \\ \mathcal{E}_{x,b}(-\alpha\beta)\mathcal{E}_{x,y}(\beta) & \text{if } a = y \\ \mathcal{E}_{x,y}(\beta) & \text{otherwise.} \end{cases}$$

By the normality of Q in P, if $a \prec_P b = x \prec_Q y$ then $a \prec_Q y$. This proves that the first of the three possibilities is an element of $U_Q(k)$. The second case is similar, relying on the fact that if $x \prec_Q y = a \prec_P b$, then $x \prec_Q b$. In the third case, $\mathcal{E}_{a,b}(\alpha)$ commutes with $\mathcal{E}_{x,y}(\beta)$, making the desired result trivial. Thus, each of the three possibilities consists of a product of elements all in U_Q, thereby completing the proof. □

2.3 Intersection of U_P with GL-classes

In Chapter 8 we stratify the elements $U_n(q)$ by their $\mathrm{GL}_n(q)$ conjugacy class. We will need to know how many upper-triangular matrices lie in each of these strata. Recall that, as $\mathbb{F}_q$ is not algebraically closed, the conjugacy classes in $\mathrm{GL}_n(q)$ cannot be given simply by Jordan canonical forms. However, because the characteristic polynomial of each $A \in U_n(q)$ splits (in fact, the characteristic polynomial of is always $(x-1)^n$), each $A \in U_n(q)$ is conjugate to a Jordan form. Rather than dealing with U_n, it is simpler to deal with $\mathcal{U}_n$. Let $F^\lambda(q)$ denote number of matrices in $\mathcal{U}_n$ which are GL-conjugate to a given Jordan form. That is, define

$$F^\lambda(q) := \#\{A \in \mathcal{U}_n(q) : \mathrm{sh}(A) = \lambda\}.$$

Martha Yip has calculated these numbers in [Y]. Specifically, she proved that $F^\lambda(q)$ is a polynomial in q with integer coefficients and degree $\binom{n}{2} - \mathbf{n}(\lambda)$. The leading coefficient of $F^\lambda(q)$ is f^λ, the number of standard Young tableaux of shape λ. It is clear from her proof that

$$F^\lambda(q) \leq f^\lambda q^{\binom{n}{2} - \mathbf{n}(\lambda)}. \tag{2.1}$$

This fact will be pertinent in Chapter 8.

Part I

Class enumeration

CHAPTER 3

Overview

In this part, we make a new push towards resolving Higman's conjecture, presenting both positive and negative evidence. Surprisingly, results of both types are united by the same underlying concept of embedding complicated pattern groups into simpler, yet larger ones.

Theorem 1.5. *Higman's conjecture holds for $n \leq 16$. Moreover, for all $n \leq 16$, we have $\mathbf{k}(U_n) \in \mathbb{N}[q-1]$.*

This theorem extends the results of Arregi and Vera-López and earlier computational results, and provides further evidence of Higman's conjecture. Our approach is based on computing the polynomials indirectly via a recursion over co-adjoint orbits arising in the finite field analogue of Kirillov's *orbit method* (see [K1, K3]). This approach is substantially different from, and turns out to be significantly more efficient than, the previous work which is based on direct enumeration of the conjugacy classes. We present the algorithm proving Theorem 1.5 in Chapter 6. We give a brief description of some earlier work in Chapter 10.

Our approach is based on a recursion over a large class of pattern groups (pattern groups were defined in Chapter 2. In a recent paper [HP], Halasi and Pálfy showed that there exist pattern groups for which the number of conjugacy classes is not given by a polynomial in the size of the field. In fact, they show that $\mathbf{k}(U_P(q))$ can be as bad as one desires. This work was the starting point of our investigation. Our next two result are also computational.

Theorem 3.1. *For every pattern subgroup $U_P(q) \leq U_9(q)$, we have $\mathbf{k}(U_P(q)) \in \mathbb{N}[q-1]$.*

While this shows that small pattern groups do exhibit polynomial behavior, this is false for larger n.

Theorem 1.6. *There is a pattern subgroup* $U_{\mathfrak{P}}(q) \leq U_{13}(q)$ *such that* $\mathbf{k}(U_{\mathfrak{P}}(q))$ *is* not *a polynomial function of* q.

While Halasi and Pálfy's approach is constructive, they do not give an explicit bound on the size of such a pattern group (cf. Section 10.5). We believe that the constant 13 in Theorem 1.6 is optimal, but this computation remains out of reach in part due to the excessively large number of pattern groups to consider (see Chapter 6).

The final result in this chapter offers an evidence against Higman's conjecture:

Theorem 3.2. *The pattern subgroup* $U_{\mathfrak{P}}(q)$ *from Theorem 1.6* embeds *into* $U_{59}(q)$.

Here the notion of *embedding* is somewhat technical and iterative. In Chapter 5, we prove that

$$\mathbf{k}(U_n(q)) = \sum_P A_P(q) \cdot \mathbf{k}(U_P(q)), \tag{3.1}$$

where $A_P(q) \in \mathbb{Z}[q]$ are polynomials and the sum is over pattern subgroups $U_P(q)$ which embed into $U_n(q)$, and are irreducible in a certain formal sense. Taking Theorem 1.6 into account, this strongly suggests that $\mathbf{k}(U_n(q))$ is not polynomial for sufficiently large n.

Conjecture 3.3. *The number of conjugacy classes* $\mathbf{k}(U_n(q))$ *is* not *polynomial for* $n \geq 59$.

This conjecture is hopelessly beyond the means of a computer experiment. We believe that Theorem 1.5 can in principle be extended to $n \leq 18$ by building upon our approach, and parallelizing the computation (see Section 10.6). It is unlikely however, that this would lead to a disproof of Higman's Conjecture 1.3 without a new approach.

Curiously, this brings the status of Higman's conjecture in line with that of Higman's related but more famous *PORC conjecture* (see Section 1.1 and [BNV]). Similarly to (3.1), it has been shown that the number $f(p, n)$ of groups of order p^n can be expressed as a large sum over certain *descendants*.

Recently Vaughan-Lee and du Sautoy showed that some of the terms counting the numbers of descendants are non-polynomial [DV]. Here is how Vaughan-Lee eloquently explains this in [VL]:

> *"The grand total might still be PORC, even though we know that one of the individual summands is not PORC. My own view is that this is extremely unlikely. But in any case I believe that Marcus's group provides a counterexample to what I hazard to call the* philosophy *behind Higman's conjecture."*

We hope the reader views our results in a similar vein (cf. Section 10.8).

The next several chapters are structured as follows. In Chapter 4, we prove some preliminary results on co-adjoint orbits of the pattern groups. We then proceed to develop combinatorial tools giving recursions for the number of co-adjoint orbits (Section 4.2). Chapter 5 is essentially poset theoretic, which allows us to prove Theorem 3.2. The experimental work which proves Theorems 1.5 and 1.6 is given in Chapter 6. Lastly, Chapter 7 sheds some light on a combinatrial coincidence appearing between the number of alternating permutations and the numbers $\mathbf{k}(U_n(2))$.

CHAPTER 4

Pattern groups and the co-adjoint action

4.1 Adjoint and co-adjoint actions for pattern groups

The *adjoint action* of U_P on $\mathcal{U}_P$ is defined by

$$\mathrm{Ad}_g : X \mapsto gXg^{-1}$$

for $g \in U_P$ and $X \in \mathcal{U}_P$. Enumerating conjugacy classes of a pattern group is equivalent to enumerating orbits of the adjoint action. Indeed, the action of U_P on itself by conjugation is equivariant with the adjoint action, as

$$1 + \mathrm{Ad}_g(X) = 1 + gXg^{-1} = g(1 + X)g^{-1}.$$

We consider the *co-adjoint action* of U_P on the dual of $\mathcal{U}_P$. For $f \in \mathcal{U}_P^*$ and $g \in U_P$, define $K_g(f) \in \mathcal{U}_P^*$ by

$$K_g(f) : X \mapsto f(g^{-1}Xg).$$

In other words, the co-adjoint action is given by $K_g(f) = f \circ \mathrm{Ad}_{g^{-1}}$.

Lemma 4.1. *The number of co-adjoint orbits for a pattern group $U_P(q)$ is equal to the number of adjoint orbits, and hence $\mathbf{k}(U_P(q))$.*

Note that several versions of Lemma 4.1 are known (cf. [K1, K2]). In particular, Kirillov proves the special case of U_n in [K3]. We present a full proof here for completeness.

Proof of Lemma 4.1. Extend both the adjoint and the co-adjoint actions by linearity from U_P to the entire group algebra $\mathbb{Z}[U_P]$. Then for $f \in \mathcal{U}_P^*$, the co-adjoint action K_{g-1} annihilates f

if and only if f vanishes on the image of $\mathrm{Ad}_{g^{-1}-1}$. Indeed,

$$K_{g-1}(f) = K_g(f) - f = f \circ \mathrm{Ad}_{g^{-1}} - f = f \circ \mathrm{Ad}_{g^{-1}-1}\,.$$

Let $I_g = \mathrm{Im}(\mathrm{Ad}_{g^{-1}-1})$. We apply Burnside's lemma to count the orbits of the co-adjoint action:

$$\begin{aligned}\left|\mathcal{U}_P^*/U_P\right| &= \frac{1}{|U_P|}\sum_{g\in U_P}\left|\ker(K_{g-1})\right| = \frac{1}{|U_P|}\sum_{g\in U_P}\#\{f\in\mathcal{U}_P^* | I_g \subseteq \ker f\} \\ &= \frac{1}{|U_P|}\sum_{g\in U_P} q^{\dim \mathcal{U}_P - \dim I_g} = \frac{1}{|U_P|}\sum_{g\in U_P}\left|\ker(\mathrm{Ad}_{g^{-1}-1})\right| = \left|\mathcal{U}_P/U_P\right|.\end{aligned}$$

This completes the proof. □

In place of functionals on pattern algebras, we identify $\mathcal{U}_P^*$ with a quotient space of $\mathcal{M}_{n\times n}$. Define

$$\mathcal{L}_P(q) := \mathcal{M}_{n\times n}(q) \Big/ \bigoplus_{i\not> j} \mathbb{F}_q e_{i,j}\,.$$

When $P = \mathbf{C}_n$ (the total order $\{1 < \cdots < n\}$), then $\mathcal{L}_P$ is the space of lower triangular matrices thought of as a quotient of all matrices by upper-triangular matrices (hence the notation "$\mathcal{L}$"). For general posets P, the space $\mathcal{L}_P$ is a further quotient of lower triangular matrices.

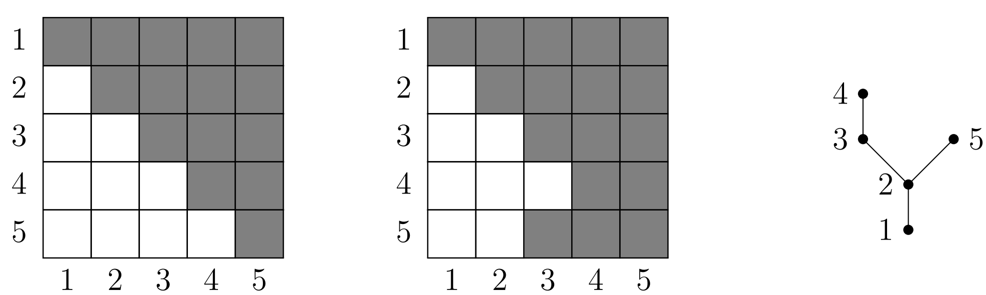


Figure 4.1: The spaces $\mathcal{L}_{\mathbf{C}_5}$ (left) and $\mathcal{L}_P$ (right) where P is the poset shown to the right. The shaded cells represent those components of $\mathcal{M}_{n\times n}$ which are quotiented away in $\mathcal{L}_P$.

The space $\mathcal{L}_P$ is isomorphic to $\mathcal{U}_P^*(q)$. Specifically, for each $A \in \mathcal{L}_P$, define the functional $f_X \in \mathcal{U}_P^*$ by

$$f_X(A) := \mathrm{tr}(X \cdot A).$$

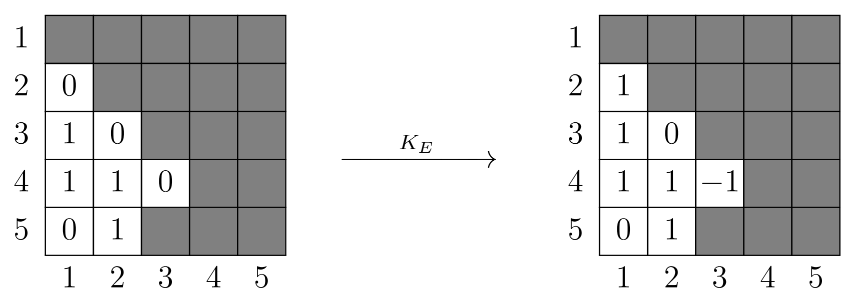
Every functional in $\mathcal{U}_P^*$ can be expressed as f_X for some $X \in \mathcal{L}_P$. This identification of X with f_X is well-defined, as the quotiented cells in $\mathcal{L}_P$ (those $e_{i,j}$ with $i \not\succ j$) will precisely align with the cells that are forced to be zero in $\mathcal{U}_P$. That is, if $i \not\succ j$, then for $A \in U_P$, we have $A_{j,i} = 0$. Thus, their contribution to the trace will be zero.

Pushing the co-adjoint action through this identification yields an action of U_P on $\mathcal{L}_P$, which we also call the co-adjoint action (and also write K_g). For $g \in U_P$ and $L \in \mathcal{L}_P$, the action becomes

$$K_g(L) = gLg^{-1}.$$

To be precise, let $\rho : \mathcal{M}_{n\times n} \to \mathcal{L}_P$ denote the canonical projection map. For $X \in \mathcal{L}_P$, pick a representative $X' \in \mathcal{M}_{n\times n}$ so that $\rho(X') = X$. Then $K_g(X) = \rho(gX'g^{-1})$. It is evident that the choice of X' is irrelevant.

Example 4.2. Let P denote the poset shown in Figure 4.1, and let $X \in \mathcal{L}_P(q)$ denote the element shown below on the left. We consider the co-adjoint action of the elementary matrix $E = \mathcal{E}_{2,3}(1)$ on X.



Consider the left multiplication $X \mapsto EX$. This action adds the contents of row 3 to row 2. Thus, for $Y = K_E(X)$, we have $Y_{3,1} = X_{3,1} + X_{3,2}$. All other cells in row 2 are trivial in $\mathcal{L}_P$. For the right multiplication $EX \mapsto EXE^{-1}$, we take the contents of column 2 and subtract them from the contents of column 3. We see that $Y_{4,3} = X_{4,3} - X_{4,2}$. All other cells in column 3 are trivial in $\mathcal{L}_P$, so this is the only relevant data.

Definition 4.3. Observe that conjugation, the adjoint action, and the co-adjoint actions on $\mathcal{U}_P^*$ and $\mathcal{L}_P$ all have the same number of orbits. Therefore, we define the quantity $\mathbf{k}(P)$ to be

$$\mathbf{k}(P) := \mathbf{k}(U_P) = |\mathcal{U}_P/U_P| = |\mathcal{U}_P^*/U_P| = |\mathcal{L}_P/U_P|\,.$$

The field $\mathbb{F}_q$ will nearly always be evident from the context and will be omitted for the sake of brevity. In the few cases where we want to specify the field, we will write $\mathbf{k}(P;q)$ to denote $\mathbf{k}(U_P(q))$.

4.2 Combinatorial tools for computing $\mathbf{k}(P)$

4.2.1 Elementary operations

We begin with a the following result which can be seen easily in the language of the co-adjoint action on $\mathcal{L}_P$. However, we prove the result via elementary group theory.

Proposition 4.4. *For posets P and Q, we have*

1. $\mathbf{k}(P) = \mathbf{k}(P^*)$

2. $\mathbf{k}(P) = \mathbf{k}(P_1) \cdot \mathbf{k}(P_2)$ *where $P_i = P|_{S_i}$ for $i = 1, 2$, and $S_1, S_2 \subseteq P$ such that $S_1 \cup S_2 = P$ and $P|_{S_1 \cap S_2}$ contains no relations.*

3. $\mathbf{k}(P \amalg Q) = \mathbf{k}(P) \cdot \mathbf{k}(Q)$

Proof. For (1), we must label the elements P^* appropriately so that $i \leq j$ whenever $i \preccurlyeq_{P^*} j$ (as required by the definition of pattern groups). Let $n = |P|$, and for each $i \in P$, relabel the element i with the label $n + 1 - i$. This will reverse the total ordering on P so that it agrees with the partial ordering on P^*. In terms of matrices, we have expressed U_{P^*} as the elements of U_P "transposed" about the anti-diagonal. Let ϕ denote this anti-diagonal transposition. Then the map $g \mapsto \phi(g^{-1})$ is an isomorphism between the groups U_P and U_{P^*}, proving $\mathbf{k}(P) = \mathbf{k}(P^*)$.

For (2), let $P_i = P|_{S_i}$ for $i = 1, 2$. We claim that the following map $\psi : U_{P_1} \times U_{P_2} \to U_P$ defined by $\psi(g_1, g_2) = g_1 g_2$ is the isomorphism. First, note that for $g_i \in U_{P_i}$, the elements g_1 and g_2 commute. To this end, it suffices to see that generators commute, which follows from the fact that $P|_{S_1 \cap S_2}$ has no relations, and Proposition 2.7. Then ψ is a homomorphism, as

$$\psi(g_1, g_2)\psi(h_1, h_2) = g_1 g_2 h_1 h_2 = g_1 h_1 g_2 h_2 = \psi(g_1 h_1, g_2 h_2), \quad \text{and}$$

$$\psi(g_1, g_2)^{-1} = g_2^{-1} g_1^{-1} = g_1^{-1} g_2^{-1} = \psi(g_1^{-1}, g_2^{-1}).$$

Whenever $x \prec_P y$, then either $x \prec_{P_1} y$ or $x \prec_{P_2} y$. Therefore, every generator $\mathcal{E}_{x,y}(\alpha)$ of U_P is either a generator of U_{P_1} or U_{P_2}, so ψ is surjective. It follows from the fact that $|U_{P_1} \times U_{P_2}| = |U_P|$ that ψ is an isomorphism, proving that $\mathbf{k}(P) = \mathbf{k}(P_1) \cdot \mathbf{k}(P_2)$. Finally, to see (3), apply (2) to the poset $P \amalg Q$ with $S_1 = P$ and $S_2 = Q$. □

4.2.2 Poset systems

Let Q be a subposet of P. Then the algebra $\mathcal{U}_Q$ canonically injects into $\mathcal{U}_P$, and so we obtain a canonical projection

$$\Pi_{P,Q} : \mathcal{L}_P \to \mathcal{L}_Q.$$

This projection sends $e_{i,j}$ to zero whenever $i \prec_P j$, but $i \not\prec_Q j$. For specific choices of Q and P, this map can be used effectively to enumerate $\mathbf{k}(P)$.

Fix a maximal element $m \in \max(P)$. Of particular interest will be the poset $P^{\langle m \rangle}$, defined by

$$\operatorname{rel}\left(P^{\langle m \rangle}\right) = \{(x, m) : x \prec_P m\}.$$

This is the poset whose only relations are those taken from P which involve the element m. To simplify notation, for the remainder of this subsection, let $Q = P^{\langle m \rangle}$, and let $\Pi = \Pi_{P,Q}$. That is, the projection Π annihilates all $e_{i,j} \in \mathcal{L}_P$ which are not of the form $e_{m,x}$ for $x \prec_P m$ (see Figure 4.2).

The map Π induces an action of U_P on $\mathcal{L}_Q$, the orbits of which are easy to analyze. Define the *support* of an element $X \in \mathcal{L}_Q$ to be

$$\operatorname{supp}(X) := \{x \in Q : X_{m,x} \neq 0\}.$$

Each U_P-orbit of $\mathcal{L}_Q$ contains precisely one element whose support is an anti-chain in $\operatorname{lb}(m)$. (Note that it does not matter if we mean $\operatorname{lb}_P(m)$ or $\operatorname{lb}_Q(m)$, as these sets are equal.) We can stratify the U_P-orbits of $\mathcal{L}_P$ by their image in $\mathcal{L}_Q$ under them map Π. That is,

$$\mathbf{k}(P) = \sum_X \left| \Pi^{-1}(X) \middle/ \operatorname{stab}_{U_P}(X) \right| \tag{4.1}$$

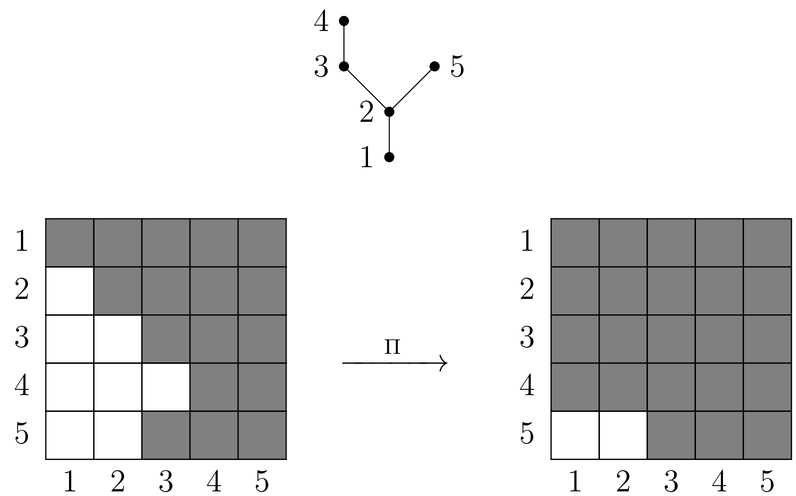


Figure 4.2: A poset P, and the projection map $\Pi : \mathcal{L}_P \to \mathcal{L}_{P\langle m\rangle}$ for $m = 5$.

where the sum is over all elements in $\mathcal{L}_Q$ whose support is an anti-chain in P.

Moreover, if $X, Y \in \mathcal{L}_Q$ have the same support $A \in \mathrm{ac}(\mathrm{lb}(m))$, then the corresponding summands for X and for Y in (4.1) are equal. This can be seen by allowing the diagonal matrices to act on $\mathcal{L}_P$ by conjugation, and noting that for an appropriate choice of diagonal matrix δ, we have

$$\delta \cdot \Pi^{-1}(X) \cdot \delta^{-1} = \Pi^{-1}(Y).$$

Furthermore, for the same diagonal matrix δ, we have

$$\delta \cdot \mathrm{stab}_{U_P}(X) \cdot \delta^{-1} = \mathrm{stab}_{U_P}(Y).$$

Therefore we can sum over a single representative for each anti-chain, and take each summand with multiplicity $(q-1)^{|A|}$. That is,

$$\mathbf{k}(P) = \sum_{A \in \mathrm{ac}(\mathrm{lb}(m))} (q-1)^{|A|} \left| \Pi^{-1}(1_A) \big/ \mathrm{stab}_{U_P}(1_A) \right| \tag{4.2}$$

where $1_A = \sum_{a \in A} e_{m,x}$, the indicator function on A. Pictorially, we are stratifying the U_P-orbits of $\mathcal{L}_P$ by the bottom row in their associated diagram (see Figure 4.3).

The notation in (4.2) is quite cumbersome, even after suppressing some of the subscripts. We make the following definition which keeps track of the essential data.

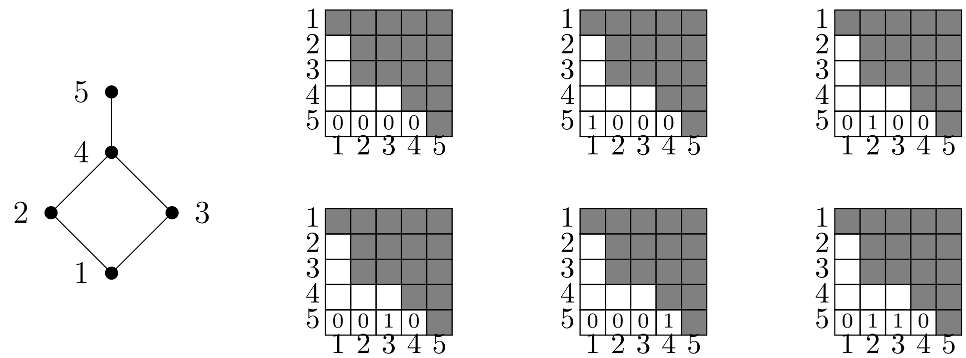


Figure 4.3: The stratification of $\mathcal{L}_P$, where P is the displayed poset. The maximal element is 5, and $\mathrm{lb}_P(5) = \{1, 2, 3, 4\}$. Each diagram shown represents the form of an element in $\Pi^{-1}(1_A)$ for a different anti-chain A. The possible anti-chains consist of the empty set, the singletons, and $\{2, 3\}$ (displayed in order). We take the first with multiplicity 1, the next four with multiplicity $(q-1)$, and the last one with multiplicity $(q-1)^2$.

Definition 4.5. A *poset system* is a triple (P, m, A) consisting of a poset P, a maximal element $m \in \max(P)$, and an anti-chain $A \in \mathrm{ac}(\mathrm{lb}_P(m))$.

Let $S = (P, m, A)$ be a poset system. By a slight abuse of notation, we define $\mathbf{k}(S) = \mathbf{k}(S; q)$ as follows:

$$\mathbf{k}(S) := \left| \Pi^{-1}(1_A) \big/ \mathrm{stab}_{U_P}(1_A) \right|,$$

where $\Pi = \Pi_{P,Q}$ and $Q = P^{\langle m \rangle}$ as above. For any poset P, and any $m \in \max(P)$, we may rewrite (4.2) in this more condensed notation,

$$\mathbf{k}(P) = \sum_{A \in \mathrm{ac}(\mathrm{lb}(m))} (q-1)^{|A|} \mathbf{k}(P, m, A). \tag{4.3}$$

This relation is our main tool for computing $\mathbf{k}(U_n)$. We show that under certain conditions on poset systems S, there exists a poset Q for which $\mathbf{k}(S) = \mathbf{k}(Q)$. When such a poset exists, we then recursively apply (4.3).

Formally, whenever $\mathbf{k}(S) = \mathbf{k}(P)$ for a poset P and poset system S, we say that S *reduces to* P, and that S is *reducible*.

Remark 4.6. If every poset system was reducible, an inductive argument would imply that $\mathbf{k}(P)$ was a polynomial for every poset P. This is certainly not the case, as Halasi and

Pálfy have constructed posets for which $\mathbf{k}(P)$ is not a polynomial [HP]. However, by adding suitable combinatorial constraints to S, we can guarantee reducibility.

Lemma 4.7. *Let $S = (P, m, A)$ be a poset system such that there exists no pair of elements $(a, x) \in A \times P$ for which $a \prec x \prec m$. Then S reduces to $P - m$.*

Proof. We begin by showing that the entire group U_P stabilizes 1_A. Let $\alpha \in \mathbb{F}_q^\times$, and let $E = \mathcal{E}_{x,y}(\alpha)$ be a generator of U_P. If $x \notin A$, then it is easy to see that $K_E(1_A) = 1_A$. On the other hand, if $x \in A$, then by assumption $y \not\prec m$. Therefore, we have $K_E(1_A) = 1_A - e_{m,y} = 1_A$, as $e_{m,y}$ is trivial in $\mathcal{L}_Q$.

From Proposition 2.7, we know that each of the generators $\mathcal{E}_{x,y}(\alpha)$ of U_P is either a generator of U_{P-m}, or of the form $\mathcal{E}_{x,m}(\alpha)$ for $\alpha \in \mathbb{F}_q^\times$. Because each generator of the form $\mathcal{E}_{x,m}(\alpha)$ acts trivially on $\mathcal{L}_P$, we have

$$\mathbf{k}(S) = \left|\Pi^{-1}(1_A)/\operatorname{stab}_{U_P}(1_A)\right| = \left|\Pi^{-1}(1_A)/U_{P-m}\right|.$$

Now every element of U_{P-m} acts trivially on row m (the $\mathbb{F}_q$-linear span of $e_{m,x}$). Simply removing this row yields the co-adjoint action of U_{P-m} on $\mathcal{L}_{P-m}$, so we obtain

$$\mathbf{k}(S) = \left|\Pi^{-1}(1_A)/U_{P-m}\right| = |\mathcal{L}_{P-m}/U_{P-m}| = \mathbf{k}(P - m)$$

as desired. ☐

Lemma 4.8. *Let (P, m, A) be a poset system, and suppose that $a, b \in A$ such that*

$$\operatorname{ub}_P(a) \supseteq \operatorname{ub}_P(b) \text{ and } \operatorname{lb}_P(a) \subseteq \operatorname{lb}_P(b).$$

Then $\mathbf{k}(P, m, A) = \mathbf{k}(P, m, A - \{b\})$.

Proof. Let $\Phi : \mathcal{L}_P \to \mathcal{L}_P$ denote conjugation by $E = \mathcal{E}_{a,b}(1)$. Note that $E \notin U_P$, since a and b are incomparable. However, E normalizes U_P, and so the map Φ is well-defined. As a slight abuse of notation, we also use Φ to denote the conjugation map $\Phi : U_P \to U_P$ given by $\Phi(g) = EgE^{-1}$. It is routine to verify that for $X \in \mathcal{L}_P$ and $g \in U_P$, we have

$$\Phi(K_g(X)) = K_{\Phi(g)}(\Phi(X)).$$

Now let $Q = P^{\langle m \rangle}$ and $\Pi = \Pi_{P,Q}$. Pushing Φ through Π to an action of $\mathcal{L}_Q$, we have

$$\Phi(1_A) = E(1_A)E^{-1} = 1_A - e_{m,b} = 1_{A-\{b\}}.$$

Moreover, as Φ commutes with Π, we have $\Phi(\Pi^{-1}(1_A)) = \Pi^{-1}(1_{A-\{b\}})$. Lastly, note that $\Phi(\mathrm{stab}_{U_P}(1_A)) = \mathrm{stab}_{U_P}(1_{A-\{b\}})$. Thus, we have

$$\begin{aligned}\mathbf{k}(P, m, A) &= \left|\Pi^{-1}(1_A)\big/\mathrm{stab}_{U_P}(1_A)\right| \\ &= \left|\Pi^{-1}(1_A)\big/\mathrm{stab}_{U_P}(1_{A-\{b\}})\right| = \mathbf{k}(P, m, A - \{b\}),\end{aligned}$$

which is the desired result. □

4.2.3 The operator $\mathcal{D}$

Let $S = (P, m, A)$ be a poset system. Define $\mathcal{D}(S)$ to be a poset obtained from P by removing relations $a \prec x$ whenever the following two criteria hold:

1. $a \in A$, and $a \prec x \prec m$.
2. If $a' \in A$ and $a' \prec x$, then $a' = a$.

Stated more concisely, the set of pairs of related elements in $\mathcal{D}(S)$ is given by

$$\mathrm{rel}(\mathcal{D}(S)) = \mathrm{rel}(P) \setminus \{(a, x) : a \prec x \prec m,\ |A \cap \mathrm{lb}(x)| = 1\}.$$

In Figures 4.4 and 4.5, we provide examples of poset systems S and the application of the operator $\mathcal{D}(S)$. Poset systems are shown graphically as the Hasse diagram of their underlying poset with special marked elements. Generic elements of P will be denoted by "•" as they normally are a Hasse diagram. The elements of the anti-chain A will be denoted by "◦." The maximal element m will be denoted by "▫."

Lemma 4.9. *For any poset system $S = (P, m, A)$, we have $\mathbf{k}(S) = \mathbf{k}(\mathcal{D}(S), m, A)$.*

Figure 4.4: The poset system $S = (\mathbf{C}_5, 5, \{3\})$ shown on the left and the poset $\mathcal{D}(S)$ shown on the right.

Figure 4.5: The poset system $S = (P, 6, \{3\})$ shown on the left and the poset $\mathcal{D}(S)$ shown on the right.

Proof. Let $Q = P^{\langle m \rangle}$. Not only is Q a subposet of P, but it is also a subposet of $\mathcal{D}(S)$. Therefore every element of A is less than m in $\mathcal{D}(S)$ as well as in P, so the poset system $(\mathcal{D}(S), m, A)$ is well-defined.

We first show that $\mathrm{stab}_{U_P}(1_A) = \mathrm{stab}_{U_{\mathcal{D}(S)}}(1_A)$. As $\mathcal{D}(S)$ is a subposet of P, we have $U_{\mathcal{D}(S)} \leq U_P$, and so clearly $\mathrm{stab}_{U_{\mathcal{D}(S)}}(1_A) \leq \mathrm{stab}_{U_P}(1_A)$. To show equality, it suffices to show that the two stabilizers have the same cardinality. Moreover, by the orbit-stabilizer theorem, it suffices to show

$$\frac{|U_P|}{\left|U_{\mathcal{D}(S)}\right|} = \frac{|\Omega_P|}{\left|\Omega_{\mathcal{D}(S)}\right|},$$

where Ω_P denotes the U_P-orbit of $\mathcal{L}_Q$ containing 1_A, and $\Omega_{\mathcal{D}(S)}$ denotes the $U_{\mathcal{D}(S)}$-orbit of $\mathcal{L}_Q$ containing 1_A.

It is immediate from the definition of pattern groups that $|U_P| = q^{|\mathrm{rel}(P)|}$. Thus, we have $|U_P| \,/\, \left|U_{\mathcal{D}(S)}\right| = q^{|R|}$, where $R = \mathrm{rel}(P) \setminus \mathrm{rel}(\mathcal{D}(S))$. We may characterize R in a different way:

$$R = \{(a,x) \in A \times \mathrm{lb}(m) : a \text{ is the unique element of } A \text{ below } x\}.$$

For pairs $(a,x) \in R$, the element $a \in A$ is uniquely defined by x, and so R is in bijection with the set

$$R' = \{x \prec_P m : |\mathrm{lb}(x) \cap A| = 1\}.$$

We now turn to the orbits Ω_P and $\Omega_{\mathcal{D}(S)}$ in $\mathcal{L}_Q$. Certainly $X_{m,a} = 1$ for each $a \in A$, and $X_{m,x} = 0$ if $x \not\succ_P a$ for all $a \in A$. If, on the other hand, there does exist some $a \in A$ for which $a \prec_P x$, then by conjugation one can obtain any value at $X_{m,x}$. Specifically, note that for $E = \mathcal{E}_{a,x}(\alpha)$, we have $K_E(X) = X - \alpha e_{m,x}$. It follows that

$$\frac{|\Omega_P|}{\left|\Omega_{\mathcal{D}(S)}\right|} = q^{|R_1|-|R_2|}, \quad \text{where}$$

$$R_1 = \{x \prec_P m \ : \ a \prec_P x \text{ for some } a \in A\} \quad \text{and}$$

$$R_2 = \{x \prec_P m \ : \ a \prec_{\mathcal{D}(S)} x \text{ for some } a \in A\}.$$

From the definition of $\mathcal{D}(S)$, we have $a \prec_{\mathcal{D}(S)} x$ if and only if there is more than one element of A which is less than x in P. Hence,

$$\frac{|\Omega_P|}{\left|\Omega_{\mathcal{D}(S)}\right|} = q^{\#\{x \ : \ |\mathrm{lb}(x) \cap A| = 1\}} = q^{|R'|} = q^{|R|}.$$

This proves that $\mathrm{stab}_{U_P}(1_A) = \mathrm{stab}_{U_{\mathcal{D}(S)}}(1_A)$. For the remainder of the proof, we let G denote both of these groups. We are now left to show that $\Pi^{-1}_{P,Q}(1_A)$ and $\Pi^{-1}_{\mathcal{D}(S),Q}(1_A)$ are isomorphic G-sets. There is a natural choice for such a G-equivariant bijection: The canonical projection $\Pi_{P,\mathcal{D}(S)} : \mathcal{L}_P \to \mathcal{L}_{\mathcal{D}(S)}$ restricts to

$$\rho : \Pi^{-1}_{P,Q}(1_A) \longrightarrow \Pi^{-1}_{\mathcal{D}(S),Q}(1_A).$$

We now argue that ρ preserves G-orbits. More precisely, we claim that for all $X, Y \in \Pi^{-1}_{P,Q}(1_A)$, the elements X and Y belong to the same G-orbit if and only if $\rho(X)$ and $\rho(Y)$ belong to the same G-orbit.

Because ρ respects the co-adjoint action, it is clear that $\rho(X)$ and $\rho(Y)$ belong to the same G-orbit whenever X and Y belong to the same G-orbit. In the other direction, suppose $\rho(X) = K_g(\rho(Y))$ for some $g \in G$. Then $X - K_g(Y) \in \ker \rho$. It is easy to see that

$$\ker \rho = \bigoplus_{(a,x)\in R} \mathbb{F}_q e_{x,a} \, .$$

Indeed, the pairs $(a, x) \in R$ are precisely the pairs of elements for which $a \prec_P x$ but $a \not\prec_{\mathcal{D}(S)} x$, so linear combinations of the $e_{x,a}$ are exactly the elements which are projected away by ρ. Now let $(a, x) \in R$, and let $E = \mathcal{E}_{x,m}(\alpha)$. For $Z \in \Pi^{-1}_{P,Q}(1_A)$, we have

$$K_E(Z) = Z + \alpha e_{a,x}.$$

Thus, if two elements of $\Pi^{-1}_{P,Q}(1_A)$ differ by an element of $\ker \rho$, they must belong to the same G-orbit. In particular, X and $K_g(Y)$ belong to the same G-orbit. This proves

$$\mathbf{k}(S) = \left|\Pi^{-1}_{P,Q}(1_A)/G\right| = \left|\Pi^{-1}_{\mathcal{D}(S),Q}(1_A)/G\right| = \mathbf{k}(\mathcal{D}(S), m, A),$$

which completes the proof. □

Lemma 4.10. *Let $S = (P, m, A)$ be a poset system with $A = \{a_1, \ldots, a_k\}$ such that*

$$\mathrm{lb}_P(a_1) \subseteq \mathrm{lb}_P(a_2) \subseteq \cdots \subseteq \mathrm{lb}_P(a_k) \text{ and}$$
$$\mathrm{ub}_P(a_1) \subseteq \mathrm{ub}_P(a_2) \subseteq \cdots \subseteq \mathrm{ub}_P(a_k).$$

Further suppose that m is the unique maximum above a_k. Then S is reducible.

Proof. We proceed by induction on $|A|$. If $A = \varnothing$, then $\mathbf{k}(S) = \mathbf{k}(P - m)$ by Lemma 4.7. If $|A| = 1$, then $\mathbf{k}(S) = \mathbf{k}(\mathcal{D}(S) - m)$ by Lemmas 4.9 and 4.7 applied in succession.

Now suppose the result holds whenever the anti-chain has fewer than k elements, and let $|A| = k$. Applying Lemma 4.9, we have $\mathbf{k}(S) = \mathbf{k}(\mathcal{D}(S), m, A)$. Let

$$R := \mathrm{rel}(P) \setminus \mathrm{rel}(\mathcal{D}(S)),$$

and note that because m is the unique maximum above a_k, we have

$$R = \{(a, x) \in A \times P : \mathrm{lb}(x) \cap A = \{a\}\} \, .$$

If $(a_i, x) \in R$, then $a_i \prec_P x$, and for all $j \neq i$, it must be that $a_j \not\prec_P x$. However if there exists some j satisfying $i < j \leq k$, then because $\mathrm{ub}(a_i) \subseteq \mathrm{ub}(a_j)$, it follows that $a_j \prec_P x$. As a_i is the unique element of A below x, it must be that there is no j satisfying $i < j \leq k$, so it must be that $i = k$. Hence,

$$R = \{(a_k, x) : x \in \mathrm{ub}_P(a_k) \setminus \mathrm{ub}_P(a_{k-1})\}.$$

Therefore $\mathrm{ub}_{\mathcal{D}(S)}(a_k) = \mathrm{ub}_{\mathcal{D}(S)}(a_{k-1})$, and so $(\mathcal{D}(S), m, A)$ satisfies the hypotheses of Lemma 4.8. This tells us that $\mathbf{k}(\mathcal{D}(S), m, A) = \mathbf{k}(\mathcal{D}(S), m, A - \{a_k\})$. By inductive hypothesis, there exists a poset Q for which $\mathbf{k}(\mathcal{D}(S), m, A - \{a_k\}) = \mathbf{k}(Q)$. Stringing these equalities together yields $\mathbf{k}(S) = \mathbf{k}(Q)$, as desired. □

4.2.4 Reduction of wishbone-free posets

With suitable constraints on the poset, we may obtain a recurrence relation for the number of conjugacy classes in its pattern group. One such constraint is as follows. Define the poset $\curlywedge$ (pronounced "wishbone") as in Figure 4.6. We say a poset is $\curlywedge$-*free* if it does not have the wishbone poset as an induced subposet.

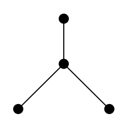


Figure 4.6: The wishbone-poset $\curlywedge = \mathbf{A}_2 + \mathbf{C}_2$.

Theorem 4.11. *Let P be a $\curlywedge$-free poset, and let $m \in \max(P)$. Then*

$$\mathbf{k}(P) = \sum_{S=(P,m,A)} (q-1)^{|A|}\mathbf{k}(\mathcal{D}(S) - m).$$

Proof. Let $S = (P, m, A)$ be a poset system. In light of (4.3), it suffices to show that

$$\mathbf{k}(S) = \mathbf{k}(\mathcal{D}(S) - m).$$

By Lemma 4.9, we see that $\mathbf{k}(S) = \mathbf{k}(\mathcal{D}(S), m, A)$. We claim that if P is $\curlywedge$-free, then $\mathcal{D}(S)$ has no element x for which $a \prec_{\mathcal{D}(S)} x \prec_{\mathcal{D}(S)} m$. Once this claim is established, Lemma 4.7

proves that $\mathbf{k}(S) = \mathbf{k}(\mathcal{D}(S) - m)$. Suppose for the sake of contradiction that $x \in \mathcal{D}(S)$, and $a \in A$ such that

$$a \prec_{\mathcal{D}(S)} x \prec_{\mathcal{D}(S)} m.$$

Because $\mathcal{D}(S)$ is obtained from P by removing relations, certainly $a \prec_P x \prec_P m$. Moreover, because $a \prec_P x$ was not removed, we know that $|A \cap \mathrm{lb}(x)| > 1$. Thus, there must be some other $b \in A$ with $b \prec_P x$. Now $\{a, b, x, m\}$ induces a copy of $\curlywedge$ in P, which is a contradiction. □

Remark 4.12. Theorem 4.11 does not use the full strength of the $\curlywedge$-freeness condition. It is only necessary that P be $\curlywedge$-free below a single maximal element. Hence, we have the following strengthening of the theorem.

Theorem 4.13. *Let P be a poset, and suppose that there exists some $m \in \max(P)$ such that the poset induced on $\{x : x \preccurlyeq_P m\}$ is $\curlywedge$-free. Then*

$$\mathbf{k}(P) = \sum_{S=(P,m,A)} (q-1)^{|A|} \mathbf{k}(\mathcal{D}(S) - m).$$

4.2.5 Interval posets

In a different direction, we consider interval posets. Given a collection of closed intervals $I_k = [\ell_k, r_k]$ in $\mathbb{R}$, one can define a partial order called the *interval order* on $\{I_k\}$ by declaring $I_j \preccurlyeq I_k$ whenever $r_j \leq \ell_k$. An *interval poset* is a poset which is the interval order of some family of intervals on a line. The class of interval posets has been well-studied (see e.g. [Tr]), and has several equivalent characterizations. For our purposes, the important properties of interval poset will be items (3) and (4) in the following theorem.

Theorem 4.14 (Theorem 3.2 in [Tr])**.** *For a poset P, the following are equivalent:*

1. *P is an interval poset,*

2. *P is $(\mathbf{C}_2 \amalg \mathbf{C}_2)$-free,*

3. *the collection of sets $\mathrm{ub}(x)$ for $x \in P$ is totally ordered by inclusion, and*

4. the collection of sets lb(x) *for* $x \in P$ *is totally ordered by inclusion.*

From here we have the following positive result.

Theorem 4.15. *Every interval poset with a unique maximal element is reducible.*

Proof. From (4.3), it suffices to show that every poset system $S = (P, m, A)$ is reducible. We do so by induction on $|A|$. If $|A| = 0$ then the result follows from Lemma 4.7. If $|A| = 1$, the result follows from Lemmas 4.9 and 4.7 applied in succession. Otherwise, suppose that $|A| \geq 2$. Because P is an interval poset, we may assume without loss of generality, that the elements of $A = \{a_1, \ldots, a_k\}$ satisfy

$$\mathrm{lb}(a_1) \subseteq \mathrm{lb}(a_2) \subseteq \cdots \subseteq \mathrm{lb}(a_k).$$

If there exist $a, b \in A$ which satisfy the conditions of Lemma 4.8, then $\mathbf{k}(S) = \mathbf{k}(P, m, A - \{b\})$, and the inductive hypothesis proves the claim. We may therefore assume that for every $i < j$, we have $\mathrm{ub}(a_j) \nsubseteq \mathrm{ub}(a_i)$. However, in an interval poset, the sets $\mathrm{ub}(x)$ are also totally ordered by inclusion. We conclude that

$$\mathrm{ub}(a_1) \subseteq \mathrm{ub}(a_2) \subseteq \cdots \subseteq \mathrm{ub}(a_k),$$

and the result follows from Lemma 4.10. □

CHAPTER 5

Embedding

5.1 Embedding sequences

Consider an attempt to compute $\mathbf{k}(U_n) = \mathbf{k}(\mathbf{C}_n)$ by recursively applying (4.3) along with the other tools developed in Section 4.2. If a poset system S appears in a computation and is reducible to a poset P, we can replace $\mathbf{k}(S)$ with $\mathbf{k}(P)$, and compute $\mathbf{k}(P)$, applying (4.3) again. We show that for every poset P, one can take n sufficiently large so that $\mathbf{k}(P)$ appears in the recursive expansion of $\mathbf{k}(U_n)$. With the following definition, we make this statement precise in Theorem 5.5.

Definition 5.1. We say that a poset P *strongly embeds*[5] into a poset Q if there exists a sequence of poset systems $S_1, \ldots, S_n$ with $S_i = (P_i, m_i, \{a_i\})$, such that

1. $P_0 = P$,

2. $P_n = Q$,

3. for $0 \le i < n$, we have $P_i \cong \mathcal{D}(S_{i+1}) - m_{i+1}$.

When P strongly embeds into Q, we write $P \rightsquigarrow Q$. The sequence

$$P = P_0 \rightsquigarrow P_1 \rightsquigarrow \cdots \rightsquigarrow P_{n-1} \rightsquigarrow P_n = Q$$

is called a *strong embedding sequence*. When we wish to signify that the strong embedding sequence has length n, we write $P \overset{n}{\rightsquigarrow} Q$.

[5] We define a weaker notion which we call *embedding* later on in Definition 5.7.

Note that the anti-chains in each poset system are required to have exactly one element. Thus, Lemmas 4.7 and 4.9 can be applied, and $\mathbf{k}(S_i) = \mathbf{k}(P_{i+1})$. The following observations regarding strong embedding are easy.

Proposition 5.2. *Let P, Q, and R be posets such that $P \overset{k}{\rightsquigarrow} Q$. Then*

$$R + P \overset{k}{\rightsquigarrow} R + Q, \quad \text{and} \quad R \amalg P \overset{k}{\rightsquigarrow} R \amalg Q.$$

The next few lemmas are technical, so we provide an outline of our methods for showing that every poset strongly embeds into a chain. First, Lemma 5.3 tells us that if we have a poset P sitting inside a larger poset $P + \mathbf{C}_k$, it is safe to focus just on P. That is, any strong embedding of P into a chain can be transformed into a strong embedding of $P + \mathbf{C}_k$ into an even larger chain. With this in mind, we may safely assume that P does not have a unique maximum.

Next, Lemma 5.4 proves that we can take a maximal element m of P and connect it to each of the other elements in P. The result will be a poset which has a chain sitting atop it which can safely be ignored.

Finally, the content of Theorem 5.5 applies Lemma 5.4 inductively, proving that each poset strongly embeds into a chain.

Lemma 5.3. *Let P, Q, and R denote posets, and suppose $P \overset{k}{\rightsquigarrow} Q$. Then we have*

$$P + R \overset{2k}{\rightsquigarrow} Q + R + \mathbf{C}_k.$$

Proof. We proceed by induction on k. We first show the result for $P \overset{1}{\rightsquigarrow} Q$. Let $(Q, m, \{a\})$ be a poset system for which $P \cong \mathcal{D}(Q, m, \{a\}) - m$. Then

$$\operatorname{rel}(P) = \operatorname{rel}(Q - m) \setminus \{(a, x) : a \prec_Q x \prec_Q m\}.$$

We define poset systems S_1 and S_2 to yield a strong embedding sequence for $P + R \rightsquigarrow Q + R + \mathbf{C}_1$. We work backwards from $Q + R + \mathbf{C}_1$, first defining S_2, then defining S_1 in terms of S_2.

Let m' denote the unique maximal element in $Q + R + \mathbf{C}_1$, and define

$$S_2 = (Q + R + \mathbf{C}_1, m', \{m\}) \text{ and}$$
$$S_1 = (\mathcal{D}(S_2) - m', m, \{a\}).$$

We aim to show that $\mathcal{D}(S_1) - m \cong P + R$. To this end, we begin with $Q + R + \mathbf{C}_1$ and follow backwards through the strong embedding sequence to determine which relations were removed. First, for $\mathcal{D}(S_2) - m'$, the relations removed were all the relations of the form (m, r) for $r \in R$. It follows that m is maximal in $\mathcal{D}(S_2) - m'$.

Next $\mathcal{D}(S_1) - m$ removes all of the relations (a, x) where $a \prec_Q x \prec_Q m$. The result is that $P + R$ and $\mathcal{D}(S_1) - m$ have precisely the same relations and are therefore isomorphic posets. This proves that $P + R \overset{2}{\rightsquigarrow} Q + R + \mathbf{C}_1$, which concludes the base case.

Assume that for all posets P, Q, and R, whenever $P \overset{k}{\rightsquigarrow} Q$ we have $P + R \overset{2k}{\rightsquigarrow} Q + R + \mathbf{C}_k$. Suppose we have posets P and Q for which $P \overset{k+1}{\rightsquigarrow} Q$. Write the strong embedding sequence

$$P = P_0 \overset{k}{\rightsquigarrow} P_k \overset{1}{\rightsquigarrow} Q.$$

By the inductive hypothesis $P + R \overset{2k}{\rightsquigarrow} P_k + R + \mathbf{C}_k$. Furthermore, because $P_k \overset{1}{\rightsquigarrow} Q$, the base case shows us that

$$P_k + (R + \mathbf{C}_k) \overset{2}{\rightsquigarrow} Q + (R + \mathbf{C}_k) + \mathbf{C}_1 = Q + R + \mathbf{C}_{k+1}.$$

Together, we have $P + R \overset{2k+2}{\rightsquigarrow} Q + R + \mathbf{C}_{k+1}$, which completes the induction. □

Lemma 5.4. *Let P be a poset, and let $m \in \max(P)$. Then*

$$P \overset{k}{\rightsquigarrow} (P - m) + \mathbf{C}_{k+1},$$

where $k = |P| - |\mathrm{lb}_P(m)| - 1$.

Proof. Let $X = \{x : x \not\preccurlyeq_P m\}$, and note that $|X| = k$. Order the elements of $X = \{x_1, x_2, \ldots, x_k\}$ according to some reverse linear extension of P, so that if $x_i \preccurlyeq_P x_j$, then $i \geq j$.

Let $Q_0 = (P - m) + \mathbf{C}_{|X|+1}$, and label the elements in

$$\mathbf{C}_{|X|+1} = \{m < p_k < p_{k-1} < \cdots < p_1\}.$$

For $1 \leq i \leq k$, define Q_i recursively as $Q_i = \mathcal{D}(Q_{i-1}, p_i, \{x_i\}) - p_i$.

The relations removed from Q_i are simple to describe:

$$\mathrm{rel}(Q_{i-1}) \setminus \mathrm{rel}(Q_i) = \{(x_i, p_j) : i+1 \leq j \leq k\} \cup \{(x_i, m)\}.$$

Note that Q_k is a poset which has $p_1, \ldots, p_k$ removed. Thus, the fact that we removed the relations $\{(x_i, p_j) : i+1 \leq j \leq k\}$ from Q_{i-1} to obtain Q_i is not relevant. However, we did remove (x_i, m) for each i. By the definition of X, we have the equality $Q_k = P$. Thus, we have constructed a strong embedding sequence

$$P = Q_k \rightsquigarrow Q_{k-1} \rightsquigarrow \cdots \rightsquigarrow Q_0 = (P - m) + \mathbf{C}_{|X|+1},$$

which proves the result. □

Theorem 5.5. *Every poset strongly embeds into a chain. Specifically,*

$$P \rightsquigarrow \mathbf{C}_{|P|^2 - 2|\mathrm{rel}(P)|}.$$

Proof. Let $F(P)$ denote the set of elements which are not comparable to every element in P. We proceed by induction on $|F(P)|$. If $F(P) = \varnothing$, then P is a chain and the result is trivial.

Otherwise, let $m \in F(P)$ be maximal amongst elements of $F(P)$. As every element of $\mathrm{ub}(m)$ is comparable to every element in P, the elements in $\mathrm{ub}(m)$ are totally ordered. Thus, we may dissect P into

$$P = P_0 + \mathbf{C}_\ell,$$

where $\ell = |\mathrm{ub}(m)|$, and where $m \in \max(P_0)$.

By Lemma 5.4, we know that

$$P_0 \overset{k}{\rightsquigarrow} (P_0 - m) + \mathbf{C}_{k+1},$$

where $k = |P_0| - |\mathrm{lb}_P(m)| - 1$. Applying Lemma 5.3, we see that

$$P = P_0 + \mathbf{C}_\ell \overset{2k}{\rightsquigarrow} (P_0 - m) + \mathbf{C}_{2k+\ell+1}.$$

Let $Q = (P_0 - m) + \mathbf{C}_{2k+1+\ell}$. Note that $F(Q) = F(P) \setminus \{m\}$, and so by inductive hypothesis,

$$P \rightsquigarrow Q \rightsquigarrow \mathbf{C}_{|Q|^2 - 2|\mathrm{rel}\, Q|}.$$

It now suffices to show that $|P|^2 - 2\,|\mathrm{rel}(P)| = |Q|^2 - 2\,|\mathrm{rel}(P)|$. To this end, note that $|Q| = |P| + 2k$, and so $|Q|^2 - |P|^2 = 4k(k + |P|)$.

We now express both $|\mathrm{rel}(Q)|$ and $|\mathrm{rel}(P)|$ in terms of $|\mathrm{rel}(P_0 - m)|$ by conditioning each pair of related elements on whether or not each element of the pair is contained in $P_0 - m$. We have

$$\begin{aligned} 2\,|\mathrm{rel}(P)| &= 2\,|\mathrm{rel}(P_0 - m)| + 2\,|\mathrm{lb}_P(m)| + 2\ell\,|P_0| + l(l-1), \text{ and} \\ 2\,|\mathrm{rel}(Q)| &= 2\,|\mathrm{rel}(P_0 - m)| + 2(2k + \ell + 1)(|P_0| - 1) + (2k + \ell + 1)(2k + \ell). \end{aligned}$$

Recalling that $|\mathrm{lb}_P(m)| = |P_0| - k - 1$, we have

$$\begin{aligned} 2\,|\mathrm{rel}(Q)| - 2\,|\mathrm{rel}(P)| &= 4k(|P_0| + \ell + k) \\ &= 4k(k + |P|) \\ &= 2k(|P| + |P| + 2k) \\ &= (|P| - |Q|)(|P| + |Q|) = |P|^2 - |Q|^2\,, \end{aligned}$$

which completes the proof. □

5.2 Consequences for U_n

Recall that Halasi and Pálfy proved the existence of a poset P for which $\mathbf{k}(P)$ is not a polynomial [HP]. Modifying their construction, we obtained a 13-element poset $\mathfrak{P}$ shown in Figure 5.1, such that $\mathbf{k}(\mathfrak{P})$ is not a polynomial in q (c.f. Section 10.5). Using Lemma 3.1 of [HP], we have computed $\mathbf{k}(\mathfrak{P})$.

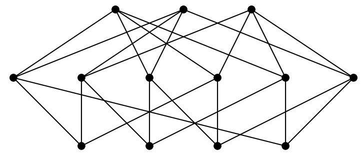


Figure 5.1: A 13-element poset $\mathfrak{P}$, for which $\mathbf{k}(\mathfrak{P}) \notin \mathbb{Z}[q]$.

Proposition 5.6. *Let $\mathfrak{P}$ denote the poset shown in Figure 5.1. Then*

$$\begin{aligned}\mathbf{k}(\mathfrak{P}) = 1 &+ 36t + 582t^2 + 5628t^3 + 36601t^4 + 170712t^5 + 594892t^6 \\ &+ 1593937t^7 + 3355488t^8 + 5646608t^9 + 7705410t^{10} \\ &+ 8631900t^{11} + 8023776t^{12} + 6248381t^{13} + 4111322t^{14} \\ &+ 2302222t^{15} + 1102490t^{16} + 451836t^{17} + 157555t^{18} \\ &+ 46042t^{19} + 10971t^{20} + 2040t^{21} + 276t^{22} + 24t^{23} + t^{24} \\ &+ \delta(q) \cdot t^{12}(t+2)^6,\end{aligned}$$

where $t = q - 1$ and

$$\delta(q) = \begin{cases} 2 & \text{if } q \text{ is odd,} \\ 1 & \text{otherwise.} \end{cases}$$

This proposition proves Theorem 1.6. It follows from Theorem 5.5 that $\mathfrak{P} \twoheadrightarrow \mathbf{C}_{97}$. However, the strong embedding sequence can be made more efficient by weakening our definition.

Definition 5.7. A poset P *embeds* into a poset Q if there exists a sequence of poset systems $S_1, \ldots, S_n$ with $S_i = (P_i, m_i, A_i)$, such that

1. $P_0 = P$,
2. $P_n = Q$, and
3. for $0 \leq i < n$, we have $\mathbf{k}(P_i) = \mathbf{k}(S_{i+1})$.

When P embeds into Q, we write $P \rightsquigarrow Q$. The sequence

$$P = P_0 \rightsquigarrow P_1 \rightsquigarrow \cdots \rightsquigarrow P_{n-1} \rightsquigarrow P_n = Q$$

is called an *embedding sequence*. When we wish to signify that the embedding sequence has length n, we write $P \overset{n}{\rightsquigarrow} Q$.

Note that if $P \twoheadrightsquigarrow Q$, then $P \rightsquigarrow Q$ as well. One of the tools we are able to use with embeddings which was not possible with the stricter notion of strong embeddings is the fact that $\mathbf{k}(P^*) = \mathbf{k}(P)$. This is the fact that we will exploit to show that $\mathfrak{P} \rightsquigarrow \mathbf{C}_{59}$ in Proposition 5.9.

Lemma 5.8. *For nonnegative integers a and b, we have $\mathbf{C}_a \amalg \mathbf{C}_b \twoheadrightsquigarrow \mathbf{C}_{2a+b}$*

Proof. We proceed by induction on a. When $a = 0$, the result is trivial. Otherwise, let $P = \mathbf{C}_1 + (\mathbf{C}_{a-1} \amalg \mathbf{C}_{b+1})$, let m be the maximal element in $\mathbf{C}_{b+1}$, and let $\hat{0}$ be the unique minimal element of P. By inductive hypothesis, we know that $\mathbf{C}_{a-1} \amalg \mathbf{C}_{b+1} \twoheadrightsquigarrow \mathbf{C}_{2a+b-1}$, and so $P \twoheadrightsquigarrow \mathbf{C}_{2a+b}$. Note that $\mathcal{D}(P, m, \{\hat{0}\}) - m$ is isomorphic to $\mathbf{C}_a \amalg \mathbf{C}_b$, so

$$\mathbf{C}_a \amalg \mathbf{C}_b \twoheadrightsquigarrow P \twoheadrightsquigarrow \mathbf{C}_{2a+b},$$

proving the result. □

Proposition 5.9. *Let $\mathfrak{P}$ denote the poset shown in Figure 5.1. Then $\mathfrak{P} \rightsquigarrow \mathbf{C}_{59}$.*

Proof. We use the techniques in the proof of Lemma 5.4 to attach each of the maximal elements to each of the non-maximal elements. Most of these relations are already present. For each maximal element, we must only add two relations for each maximal element. Next, we dualize and apply the same process to the newly maximal elements (the elements which were minimal in $\mathfrak{P}$). For each of these, we must only add three relations per maximal element. The resulting poset, shown in Figure 5.2. Symbolically, this poset can be described as

$$P' = (\mathbf{C}_3 \amalg \mathbf{C}_3 \amalg \mathbf{C}_3) + \mathbf{A}_6 + (\mathbf{C}_4 \amalg \mathbf{C}_4 \amalg \mathbf{C}_4 \amalg \mathbf{C}_4).$$

Using Lemma 5.8 and dualizing, we obtain that

$$P' \rightsquigarrow \mathbf{C}_{28} + \mathbf{A}_6 + \mathbf{C}_{15}.$$

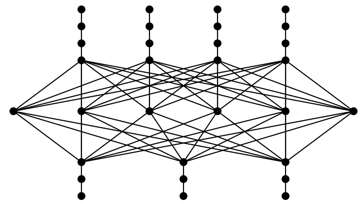


Figure 5.2: A poset P' used as an intermediate step in an embedding sequence for the poset $\mathfrak{P}$ shown in Figure 5.1.

Finally, because $\mathbf{A}_6 \overset{5}{\rightsquigarrow} \mathbf{C}_{11}$, we know that $\mathbf{C}_{28} + \mathbf{A}_6 \overset{5}{\rightsquigarrow} \mathbf{C}_{39}$. Applying Lemma 5.3 yields

$$\mathfrak{P} \rightsquigarrow P' \rightsquigarrow \mathbf{C}_{28} + \mathbf{A}_6 + \mathbf{C}_{15} \overset{10}{\rightsquigarrow} \mathbf{C}_{59},$$

which completes the proof. □

Remark 5.10. As a consequence of the preceding result, one can express $\mathbf{k}(U_{59})$ as a $\mathbb{Z}[q]$-linear combination of terms of the form $\mathbf{k}(P)$ and $\mathbf{k}(S)$ for posets P and poset systems S such that one of these terms is $\mathbf{k}(\mathfrak{P})$. It seems implausible that the remaining terms would contribute in such a way as to cancel out the contribution of $\mathbf{k}(\mathfrak{P})$, and render $\mathbf{k}(U_{59})$ a polynomial, thus leading to our Conjecture 3.3. Unfortunately explicit computation of $\mathbf{k}(U_{59})$ is well beyond the capabilities of any modern computer, and likely to remain so in the foreseeable future. (cf. Section 10.6).

CHAPTER 6

Algorithm and experimental results

6.1 Algorithmic Details

Given a poset P, to test whether or not $\mathbf{k}(P)$ is a polynomial in q, we apply the following recursive algorithm. Pick a maximal element m of P and iterate through all poset systems of the form $S = (P, m, A)$. If we can apply the equivalences given in Lemmas 4.7, 4.8, and 4.9 to obtain a reduction of S to some poset Q, then recursively compute $\mathbf{k}(Q)$. If there is even one poset system S which cannot be reduced via these methods, we try another maximal element. If we exhaust all maximal elements in this way, we try the same procedure on the dual poset P^*. If this also fails, we fall back on a slower approach to compute the values $\mathbf{k}(S)$ which the algorithm otherwise failed to compute. This slower approach is a modification of the algorithm discussed in [VA1, VA2, VA3]. We call this modification the *VA-algorithm*, and give a brief description of the necessary adaptations.

6.1.1 VA-algorithm

Order the cells in $\mathcal{L}_P$ from bottom to top, and reading each row left to right. This is the ordering induced from

$$(n, 1) < (n, 2) < \cdots < (n, n-1) < (n-1, 1) < \cdots < (3, 1) < (3, 2) < (2, 1).$$

The computation starts at the least cell in the ordering, and iterates through the cells recursively, branching when necessary. When the algorithm reaches a cell, it attempts to conjugate the cell to zero while fixing all previously seen cells. If this is possible, the cell is called *inert*. The algorithm sets the cell to zero, and continues on to the next cell in the ordering. If the cell is not inert, it is called a *ramification cell*. The algorithm will branch into two cases: one where the ramification cell contains a zero, and one it does not.

It often happens that some cells will be inert or ramification cells depending on some algebraic conditions on the previously visited cells. For example, it may be the case that cell $(5, 2)$ will be a ramification cell if $X_{5,1} = X_{6,2}$, and inert otherwise, where $X_{i,j}$ denotes the value in cell (i, j). In such instances, the algorithm will branch into three different cases:

1. The condition to be inert holds, and the cell is set to zero.

2. The condition to be inert fails (so the cell is a ramification cell), but the cell happens to be zero anyway.

3. The condition to be inert fails (so the cell is a ramification cell), and the cell is non-zero.

Determining how often the algebraic conditions hold requires counting $\mathbb{F}_q$-points on algebraic varieties. In general this is an extraordinarily difficult task, we employ several techniques which take care of the vast majority of the conditions that show up in practice.[6]

To apply the VA-algorithm to a poset system $S = (P, m, A)$, rather than starting at the beginning of the ordering, we start with some seeded data. Specifically, we start with the value 1 in each cell (m, a), where $a \in A$, and the value 0 in each cell (m, x) for $x \notin A$.

6.1.2 Pseudocode

In the two subsequent boxed figures, we provide pseudocode for our algorithm (excluding the VA-algorithm). For further details, we refer the reader to our `C++` source code, which is available at `http://www.math.ucla.edu/~asoffer/content/pgcc.zip`).

[6] Our precise techniques for point-counting are somewhat involved. We use techniques to repeatedly reduce the complexity of the varieties. For example, if a variety is linear in any variable, we can reduce it. If a variety can be factored in a relatively simple way, we can reduce its complexity as well.

6.2 Small posets

Gann and Proctor maintain a list of posets with 9 or fewer elements on their website [GP]. We use their lists of connected posets in our verification. Without using the VA-algorithm, our program verifies that $\mathbf{k}(P) \in \mathbb{Z}[q]$ for every poset P with 7 or fewer elements. Furthermore, using the VA-algorithm when necessary as described above, our program verifies that $\mathbf{k}(P) \in \mathbb{Z}[q]$ for every poset P with 9 or fewer elements. Moreover, for each such poset P, we have $\mathbf{k}(P) \in \mathbb{N}[q-1]$. This proves Theorem 3.1. A text file containing all posets on 9 or fewer elements along with their associated polynomials is available at `http://www.math.ucla.edu/~asoffer/kunq/posets.txt`.

6.3 Chains

Our program computes $\mathbf{k}(U_n)$ for every $n \leq 11$ without needing to employ the VA-algorithm. For $12 \leq n \leq 16$, our program verifies the polynomiality modulo the computation of several "exceptional poset systems" which are tackled with the VA-algorithm. This verifies the results of Arregi and Vera-López in [VA3], and extends their results to the computation to all $n \leq 16$.

As n grows, the number of exceptional poset systems which require the use of the VA-algorithm grows quickly, as shown in Figure 6.1.[7] The polynomials $\mathbf{k}(U_n)$, for $n \leq 16$ are given in the Appendix A and prove Theorem 1.5.

[7]Computations made with an Intel®Xeon® CPU X5650 2.67GHz and 50Gb of RAM.

Input: A poset P.

Output: The function $\mathbf{k}(P)$.

```
function compute_poset( P ):
  output = 0
  let m in max(P)

  for each A in antichains(P) below m:
    Q = compute_poset_system(P, m, A)
    if Q is not "FAILURE":
      output = output + (q-1)^size(A) * compute_poset(Q)
    else:
      if have not tried some max m':
        restart with m'
      else if have not tried P*:
        compute_poset( P* )
      else:
        output = output + VA_algorithm(P, m, A)
  return output
```

Input: A poset system (P, m, A).

Output: A poset Q with $\mathbf{k}(Q) = \mathbf{k}(P, m, A)$, or "FAILURE" if none can be found.

```
function compute_poset_system( P, m, A ):
  while P is changing:
    P = D(P, m, A)
    for a,b in A:
      if above(a) contains above(b) and
         below(b) contains below(a):
        A = A - b

  if no element below m and above member of A:
    return P
  else:
    return "FAILURE"
```

n	Exceptional poset systems	Computation time (sec.)	
≤ 11	0	≤ 0.2	
12	1	0.5	
13	8	4.4	
14	64	120.7	($\sim$ 2 minutes)
15	485	4456	($\sim$ 1.2 hours)
16	3550	164557	($\sim$ 46 hours)

Figure 6.1: Computation time and number of exceptional poset systems.

CHAPTER 7

A combinatorial coincidence

7.1 Alternating permutations and chains

Recall the sequence $\{\mathfrak{A}_n\}_{n=1}^{\infty}$ of alternating permutations [St], defined by

$$\mathfrak{A}_n = \#\{\sigma \in \mathfrak{S}_n : \sigma(1) < \sigma(2) > \sigma(3) < \cdots\}.$$

Comparing the first few terms of $\mathfrak{A}_n$ with $\mathbf{k}(U_n(2))$ reveals a remarkable coincidence (see Figure 7.1). The first five terms in the sequences are identical, and only differ slightly in the next few terms.

n	1	2	3	4	5	6	7	
$\mathfrak{A}_{n+1}$	1	2	5	16	61	272	1385	(OEIS A000111)
$\mathbf{k}(U_n(2))$	1	2	5	16	61	275	1430	(OEIS A007976)

Figure 7.1: The sequence counting the number of alternating permutations and the number of conjugacy classes in $U_n(2)$.

This data led Kirillov to conjecture $\mathfrak{A}_{n+1} \leq \mathbf{k}(U_n(2))$ in [K3]. The asymptotic behavior of these sequences reveal that the conjecture is true once n is large enough: $\mathfrak{A}_{n+1} = e^{O(n \log n)}$, whereas Higman's lower bound for $\mathbf{k}(U_n(2))$ ensures us that $\mathbf{k}(U_n(2)) \geq e^{O(n^2)}$. Using the explicit bound $\mathbf{k}(U_n(2)) \geq 2^{n^2/12}$ derivable from [VA4, Theorem 12], and the recurrence relation for $\mathfrak{A}_n$ reveals that Kirillov's conjecture holds for $n \geq 43$.

However the intent of Kirillov's conjecture was not so much to establish the explicit inequality, but rather to shed light on the numeric coincidence. Perhaps there exists an injective map assigning to each alternating permutation a conjugacy class or co-adjoint orbit of $U_n(2)$ (cf. Lemma 4.1). If this were the case, it would imply $\mathfrak{A}_{n+1} \leq \mathbf{k}(U_n(2))$, and yield more insight into the combinatorial nature of $\mathbf{k}(U_n(q))$.

7.2 Entringer numbers and Y-posets

In [K3] Kirillov noticed that if one stratified the co-adjoint orbits of $U_n(2)$ by the first non-zero cell in the bottom row of the lower-triangular model, further similarities to alternating permutations appeared. In light of Subsection 4.2.3, we know that the number of co-adjoint orbits in each stratum is equal to the number of co-adjoint orbits in a pattern group. Specifically, Let $\mathbf{Y}_a^b$ denote the $(a+b+1)$-element poset $\mathbf{Y}_a^b := \mathbf{C}_a + (\mathbf{C}_1 \amalg \mathbf{C}_{b-1})$.

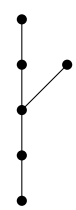


Figure 7.2: The poset $\mathbf{Y}_3^3$.

For the sake of simplifying notation, define

$$a_{n,k} := \begin{cases} 0 & \text{if } k = 0 \\ \mathbf{k}\Big(\mathbf{Y}_{k-1}^{n-k}; 2\Big) & \text{if } 1 < k < n \\ \mathbf{k}(\mathbf{C}_{n-1}) & \text{if } k = n. \end{cases}$$

Let $\mathfrak{A}_{n,k}$ denote the number of alternating permutations $\sigma \in \mathfrak{S}_{n+1}$ such that $\sigma(1) = k$. Thus, $\mathfrak{A}_{n,n} = \mathfrak{A}_n$. The numbers $\mathfrak{A}_{n,k}$ are known as the Entringer numbers. Small values of $\mathfrak{A}_{n,k}$ are provided in Figure 7.3. Note that the nth row holds the values $\mathfrak{A}_{n,k}$ for $k = 0, 1, \ldots, n$, but the values are not always read left-to-right. Rather, the values are read in alternating directions, according to the arrows shown.

In the language of $\mathfrak{A}_{n,k}$ and $a_{n,k}$, Kirillov's observation was the conjectural relationship that $\mathfrak{A}_{n,k} \leq a_{n,k}$ (see Figures 7.3 and 7.4).

$$\begin{array}{c}
1 \\
0 \rightarrow 1 \\
1 \leftarrow 1 \leftarrow 0 \\
0 \rightarrow 1 \rightarrow 2 \rightarrow 2 \\
5 \leftarrow 5 \leftarrow 4 \leftarrow 2 \leftarrow 0 \\
0 \rightarrow 5 \rightarrow 10 \rightarrow 14 \rightarrow 16 \rightarrow 16 \\
61 \leftarrow 61 \leftarrow 56 \leftarrow 46 \leftarrow 32 \leftarrow 16 \leftarrow 0 \\
0 \rightarrow 61 \rightarrow 122 \rightarrow 178 \rightarrow 224 \rightarrow 256 \rightarrow 272 \rightarrow 272
\end{array}$$

Figure 7.3: The triangle of Entringer numbers

$$\begin{array}{c}
1 \\
0 \rightarrow 1 \\
1 \leftarrow 1 \leftarrow 0 \\
0 \rightarrow 1 \rightarrow 2 \rightarrow 2 \\
5 \leftarrow 5 \leftarrow 4 \leftarrow 2 \leftarrow 0 \\
0 \rightarrow 5 \rightarrow 10 \rightarrow 14 \rightarrow 16 \rightarrow 16 \\
61 \leftarrow 61 \leftarrow \mathbf{59} \leftarrow 46 \leftarrow 32 \leftarrow 16 \leftarrow 0 \\
0 \rightarrow 61 \rightarrow 122 \rightarrow \mathbf{181} \rightarrow \mathbf{236} \rightarrow \mathbf{280} \rightarrow \mathbf{275} \rightarrow \mathbf{275}
\end{array}$$

Figure 7.4: The **Y**-triangle, consisting of values $a_{n,k}$. The values in boldface are those which differ from the corresponding entries in the Entringer triangle.

Entringer numbers can be computed with the recursive formula

$$\mathfrak{A}_{n,k} = \begin{cases} 1 & \text{if } n = k = 0 \\ 0 & \text{if } n \geq 1, k = 0 \\ \mathfrak{A}_{n,k-1} + \mathfrak{A}_{n-1,n-k} & n \geq k > 0 \end{cases} \tag{7.1}$$

There is no hope that an analogous inequality holds for $a_{n,k}$, as such an inequality would guarantee that each row is monotone, which is not the case for the last row shown in Figure 7.4. However, there is hope for a slightly modified variant of (7.1). Partially unravelling the recursion yields

$$\mathfrak{A}_{n,k} = \sum_{i=1}^{k} \mathfrak{A}_{n-1,n-i}. \tag{7.2}$$

Stated another way, an entry in the Entringer triangle can be computed by summing all the numbers in the previous row to one side (which side to use alternates according to the parity of the row number). The available evidence supports the following conjecture:

Conjecture 7.1. *For every* $n \in \mathbb{N}$ *and every* $0 \leq k \leq n$*,*

$$a_{n,k} \geq \sum_{i=1}^{k} a_{n-1,n-i}.$$

Proposition 7.2. *For every* n*, when* $k = 0, 1, 2,$ *or* n*,*

$$a_{n,k} = \sum_{i=1}^{k} a_{n-1,n-i}.$$

That is for these values of k *and* n*, Conjecture 7.1 holds* with equality.

Proof. For $k = 0$, the claim is that $a_{n,0} = 0$, which is true by definition. For $k = 1$, the claim is that $a_{n,1} = a_{n-1,n-1}$. Note that

$$\begin{aligned} a_{n,1} = \mathbf{k}(\mathbf{Y}_0^{n-1}; 2) &= \mathbf{k}(\mathbf{C}_1 \amalg \mathbf{C}_{n-2}; 2) \\ &= \mathbf{k}(\mathbf{C}_1; 2)\mathbf{k}(\mathbf{C}_{n-2}; 2) \\ &= \mathbf{k}(\mathbf{C}_{n-2}; 2) \\ &= a_{n-1,n-1}. \end{aligned}$$

For $k = 2$, the claim states that $a_{n,2} = a_{n-1,n-2} + a_{n-1,n-1}$. To see that this inequality holds, we expand the poset $\mathbf{k}(\mathbf{Y}_1^{n-1})$ using part 2 of Proposition 4.4. Specifically, we see that $\mathbf{k}(\mathbf{Y}_1^{n-1}; 2) = \mathbf{k}(\mathbf{C}_2; 2)\mathbf{k}(\mathbf{C}_n; 2) = 2a_{n-1,n-1}$. However,

$$a_{n-1,n-2} = \mathbf{k}(\mathbf{Y}_{n-3}^1; 2) = \mathbf{k}(\mathbf{C}_{n-2}; 2) = a_{n-1,n-1},$$

so once again, equality holds.

When $k = n$, the claim is that $a_{n,n} = a_{n-1,0} + \cdots + a_{n-1,n-1}$. In this case, $a_{n,n} = \mathbf{k}(\mathbf{C}_{n-1}; 2)$, and the sum on the right-hand side is exactly the result of applying Theorem 4.11 to $\mathbf{C}_{n-1}$. $\square$

An inequality such as the one in Conjecture 7.1 would imply Kirillov's conjecture via an inductive argument. Moreover, $a_{n,k}$ naturally decomposes into a sum of k terms, each of which could be compared to one of the terms $a_{n-1,j}$ on the right-hand side of the inequality in the conjecture.

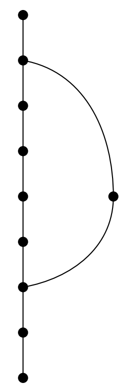


Figure 7.5: The Hasse diagram for Þ$_{3,4,2}$

The poset $\mathbf{Y}_k^{n-k}$ has two maximal elements. Let m denote the "lower" of these two, by which we main, the maximal element for which $\mathrm{lb}(m)$ is smaller. (If the two maximal elements have the same set of elements below them, they are indistinguishable.) Because $\mathbf{Y}_k^{n-k}$ is $\curlywedge$-free, any poset system of the form $(\mathbf{Y}_k^{n-k}, m, A)$ is reducible. The form of such a poset is relatively straightforward:

Definition 7.3. Define the three-parameter family of posets Þ$_{a,b,c}$ by

$$Þ_{a,b,c} := \mathbf{C}_a + (\mathbf{C}_1 \amalg \mathbf{C}_b) + \mathbf{C}_c.$$

We chose the old-english letter Þ (pronounced "thorn") because the Hasse diagrams of posets Þ$_{a,b,c}$ bear resemblance the character Þ (see Figure 7.5).

There are several simple, yet important observations regarding thorns. Note that when $a = 0$, the poset Þ$_{a,b,c}$ is a $\mathbf{Y}$-poset. When $c = 0$, Þ$_{a,b,c}$ is the dual of a $\mathbf{Y}$-poset. In fact, the family of thorns is self-dual: Þ$^*_{a,b,c}$ = Þ$_{c,b,a}$. When $b = 0$, Þ$_{a,b,c} = \mathbf{C}_{a+c+1}$.

This leads us to the following conjecture:

Conjecture 7.4. *Amongst the posets Þ$_{0,b,k}$, Þ$_{1,b,k-1}$, $\cdots$ Þ$_{k,b,1}$, The function* $\mathbf{k}(-;2)$ *is minimized at Þ$_{0,b,k}$ = Þ$_{k,b,0}$.*

We have verified this conjecture for all Þ$_{a,b,c}$ where $a + b + c \leq 12$. Appendix B contains the values of Þ$_{a,b,c}$ for $a + b + c \leq 9$. If true in general, it would yield a proof of Conjecture 7.1 and in turn provide a satisfactory proof that $\mathfrak{A}_{n+1} \leq \mathbf{k}(U_n(2))$.

Proposition 7.5. *Conjecture 7.4 implies Conjecture 7.1.*

Proof. Label the elements of $\mathbf{Y}_k^{n-k}$ as follows. The lowest k elements are labeled $x_1, \ldots, x_k$ in ascending order. The remaining yet unlabelled elements are of the form $\mathbf{C}_1 \amalg \mathbf{C}_{n-k}$. Label the isolated element m, and label the remaning elements $x_{k+1}, \ldots, x_n$.

As m is maximal, and $\mathbf{Y}_k^{n-k}$ is $\curlywedge$-free, we can apply Theorem 4.11. The anti-chains below m are either $\varnothing$ or a singleton $\{x_i\}$ for $1 \leq i \leq k$. For the poset system $(\mathbf{Y}_k^{n-k}, m, \varnothing)$, Lemma 4.7 asserts that

$$\mathbf{k}(\mathbf{Y}_k^{n-k}, m, \varnothing; 2) = \mathbf{k}(\mathbf{Y}_k^{n-k} - m; 2) = \mathbf{k}(\mathbf{C}_n; 2) = a_{n,n}.$$

For poset systems of the form $S = (\mathbf{Y}_k^{n-k}, m, \{x_i\})$, we may apply Lemma 4.9 to obtain

$$\mathbf{k}(\mathbf{Y}_k^{n-k}, m, \{x_i\}) = \mathbf{k}(\mathcal{D}(S), m, \{x_i\}) = \mathbf{k}(\text{Þ}_{i-1,k-i,n-k}).$$

From the assumed Conjecture 7.4, we have

$$\mathbf{k}(\text{Þ}_{i-1,k-i,n-k}; 2) \geq \mathbf{k}(\text{Þ}_{n-k+i-1,k-i,0}; 2) = \mathbf{k}(\mathbf{Y}_{n-k+i-1}^{k-i}; 2) = a_{n-1,n-k+i}.$$

Thus, using Theorem 4.11 to reduce $\mathbf{k}(\mathbf{Y}_k^{n-k})$ as a sum over the number of classes in these pattern groups, we have

$$a_{n,k} = \mathbf{k}(\mathbf{Y}_{k-1}^{n-k+1}; 2) = \mathbf{k}(\mathbf{C}_n; 2) + \sum_{i=1}^{k} \mathbf{k}(\text{Þ}_{i-1,k-i,n-k}; 2) \leq a_{n,0} + \sum_{i=1}^{k} a_{n-1,n-k+i}.$$

□

Part II

Asymptotic behavior of $\mathbf{k}(U_n(q))$

CHAPTER 8

Asymptotic behavior of the group $U_n(q)$

8.1 Overview

In 1960, Higman proved the following [H1]:

Theorem 8.1 (Higman). *For every prime power q,*

$$q^{\frac{n^2}{12}(1+o(1))} \leq \mathbf{k}(U_n(q)) \leq q^{\frac{n^2}{4}(1+o(1))},$$

where $o(1)$ means a function of n, independent of q, which tends to zero as n tends to infinity.

Higman's original interest was in enumerating finite p-groups of a given order. He obtained an upper bound for the number of groups of order p^n in terms of $\mathbf{k}(U_n(p))$. While the asymptotics of the number of p-groups has since been resolved via different methods (see [Si, BNV]), the gap between the lower and upper bounds for $\mathbf{k}(U_n(q))$ has not been closed.

In their 1992 paper, Arregi and Vera-López used their technique of canonical matrices to improve Higman's upper bound [VA1, Theorem 5.4]. They show

$$\mathbf{k}(U_n(q)) \leq (n-1)!2^{n-1}q^{\frac{n^2+n}{6}}.$$

Note that $(n-1)!2^{n-1} = q^{O(n \log n)}$, and so these terms do not contribute significantly to the asymptotics of $\mathbf{k}(U_n(q))$. Using the theory of supercharacters (see [DI, Mar2]), Marberg obtained an upper bound with the same asymptotics [Mar1, Theorem 5.1].

We improve on these asymptotics, with the following result:

Theorem 1.7. *For every positive integer n and every prime power q, we have*

$$\mathbf{k}(U_n(q)) \leq p(n)^2 n!\ q^{\alpha n^2 + \frac{n}{2}},$$

where $p(n)$ denotes the number of integer partitions of n, and where

$$\alpha = \frac{40\sqrt{2} - 41}{98} \approx 0.15886.$$

Our approach is to estimate the number of pairs (A, B) of commuting matrices in $U_n(q)$ by conjugating A into Jordan canonical form, and determining the possibilities for the image of B under this conjugation. There are many choices for matrices which conjugate A into Jordan form, and the image of B depends on this choice. Section 8.3 defines our canonical choice X_A for conjugation.

For each upper-triangular matrix A we conjugate its centralizer $C_{\mathcal{U}}(A)$ via our canonical choice X_A defined in Section 8.3. The resulting space $X_A C_{\mathcal{U}}(A) X_A^{-1}$ can often be described by a combinatorial object which we call a *gap array*. Section 8.4 introduces gap arrays and proves several structural lemmas about them. While not every space $X_A C_{\mathcal{U}}(A) X_A^{-1}$ can be encoded by a gap array, every such space is a subspace of one encoded by a gap array. Determining the sizes of these subspaces via the combinatorics of gap arrays, we obtain the same upper bound as Marberg and Arregi and Vera-López (see Corollary 8.24). However, the technique of gap arrays is amenable to further optimization. These optimizations are the content of the proof of Theorem 1.7 (see Section 8.5).

8.2 Preliminaries

Upon the first reading of this chapter, we encourage the reader to skip this section. The definitions and results presented here are important for the chapter but are presented without motivation. Moreover, they showcase none of the key ideas from the chapter. They are placed here only as convenient reference. We suggest reading the results as they are referenced in later sections.

Lemma 8.2. *Let $\lambda \vdash a$, $\mu \vdash b$, and let $T_{\lambda,\mu} : \mathcal{M}_{a\times b} \to \mathcal{M}_{a\times b}$ be defined by $T_{\lambda,\mu}(X) = J_\lambda X - X J_\mu$. Then*

$$\dim\ker T_{\lambda,\mu} = \sum_{i,j} \min\{\lambda_i, \mu_j\} = \langle \lambda', \mu' \rangle.$$

In particular, $\dim C_{\mathcal{M}}(J_\lambda) = \|\lambda'\|^2$.

Proof. To compute the rank, we first do so for two Jordan blocks $J_{(a)}$ and $J_{(b)}$. Note that $J_{(a)}X$ is the matrix obtained by removing the top row of X, shifting all other rows upwards by one, and adding a row of zeros at the bottom. Similarly, $XJ_{(b)}$ is the matrix obtained by removing the rightmost column of X, shifting all other columns right by one, and adding a new column of zeros on the left. If $J_{(a)}X = XJ_{(b)}$, then these conditions guarantee that

1. all diagonals (top-left to bottom-right) are constant, and
2. if a diagonal does not touch both the topmost row and the rightmost column, then it must be zero.

Moreover, these conditions exactly describe the solutions to $J_{(a)}X = XJ_{(b)}$. The dimension of this space is given by the quantity $\min\{a, b\}$, the number of diagonals which are not forced to be zero.

In general, if either $\lambda \vdash a$ or $\mu \vdash b$ have more than one part, then we decompose $\mathcal{M}_{a\times b}$ into smaller subspaces by $\mathcal{M}_{a\times b} \cong \bigoplus_{i,j} \mathcal{M}_{\lambda_i\times\mu_j}$, so that the action on $\mathcal{M}_{\lambda_i\times\mu_j}$ becomes $T_{(\lambda_i),(\mu_j)}$. The dimension of $\ker T_{\lambda,\mu}$ is the sum of the dimensions of the kernels of these subspaces, so

$$\dim\ker T_{\lambda,\mu} = \sum_{i,j} \min\{\lambda_i, \mu_j\}.$$

To see that $\sum_{i,j} \min\{\lambda_i, \mu_j\} = \langle\lambda', \mu'\rangle$, we will show that both sides count the number of pairs of cells in the diagrams of λ and μ (one from each diagram) which lie in the same column. On the one hand, we may specify the column from which we chose the cells first. Then any pair of elements from these columns will suffice. Therefore, we obtain $\sum_i \lambda'_i\mu'_i = \langle\lambda', \mu'\rangle$.

On the other hand, we may first chose the row to which each of the cells belong. If the cell in the diagram for λ lies in row i, and the cell in the diagram for μ lies in row j, then there are $\min\{\lambda_i, \mu_j\}$ choices which place these cells in the same column. As these two expressions count the same quantity, we obtain

$$\sum_{i,j} \min\{\lambda_i, \mu_j\} = \langle\lambda', \mu'\rangle.$$

Lastly, taking $\lambda = \mu$, we see that $\ker T_{\lambda,\lambda}$ is the centralizer of J_λ, and therefore

$$\dim C_{\mathcal{M}}(J_\lambda) = \dim \ker T_{\lambda,\lambda} = \|\lambda'\|^2,$$

which completes the proof. □

Remark 8.3. Matrices which satisfy the first condition (constant on top-left to bottom-right diagonals) are called *Toeplitz* matrices. The matrices in $\ker T_{\lambda,\mu}$ are not only Toeplitz on each block, but are also upper-triangular. Specifically, any diagonal which does not touch the rightmost column and topmost row of a block is necessarily zero.

8.3 Jordan forms and conjugation

Recall from Burside's lemma that

$$\mathbf{k}(U_n) = q^{-\binom{n}{2}} \sum_{A \in \mathcal{U}_n} |C_{\mathcal{U}}(A)| \,. \tag{8.1}$$

Thus, one may estimate $\mathbf{k}(U_n)$ by estimating the size of centralizers of each $A \in \mathcal{U}_n$. Our approach to understanding the size of its centralizer $C_{\mathcal{U}}(A)$ will be to conjugate A into its Jordan form J_λ by some X_A. We then analyze $X_A C_{\mathcal{U}}(A) X_A^{-1}$, because it has the same size as $C_{\mathcal{U}} A$. Moreover, $X_A C_{\mathcal{U}}(A) X_A^{-1}$ is a subspace of the well-understood $C_{\mathcal{M}}(J_\lambda)$.

Given $A \in \mathcal{U}_n$, there is more than one choice for a matrix X_A satisfying $X_A A X_A^{-1} = J_\lambda$. Different choices of X_A may yield different subspaces of $C_{\mathcal{M}}(J_\lambda)$. We must therefore specify X_A carefully. We do this inductively, by first assuming that $A|_{n-1} = J_\mu$ for some partition $\mu \vdash (n-1)$.

We begin by giving an overview of the process by which we put A into Jordan form. This overview, along with the example computation below, should provide enough detail for the reader to understand the conjugation process which defines X_A. For completeness, we also provide explicit definitions of the matrices used in the conjugation procedure.

Conjugation procedure

1. Use the non-zero entries from $A|_{n-1} = J_\mu$ to set as many entries as possible in column n to zero. This can be achieved with a product of upper-triangular transvections. The resulting matrix will only have non-zero entries in column n which are at the bottom of a block. That is, in cells of the form $(n, \mu_1 + \cdots + \mu_k)$. If the entire column n is zero in the result, skip ahead to step 5. In this case, the matrices by which we conjugate in the intermediate steps will all be defined as the identity.

2. We may now assume that column n is non-zero. Conjugate by a diagonal matrix which scales the last column and last row in such a way as to set the first non-zero entry in column n to be 1.

3. Use the first non-zero entry in column n (which now contains the value 1) to set every other value in that column to zero. This is achieved via a product of lower-triangular matrices. Each such lower triangular matrix will fix the J_μ in the top-left corner, and set a single cell in column n to be zero.

4. Apply a permutation matrix to move column n so that its non-zero value aligns with a Jordan block, effectively increasing the size of this block by 1.

5. At this point the matrix is in Jordan form, modulo rearranging the blocks to be in descending order. At most one block must be moved to guarantee this ordering. The block that must be moved is the one whose size was increased (if we increased the size of a block at all). Call this block the *current block*. If we did not increase the size of a block, then we created a new block of size 1. In this case, the newly created block will be called the current block. Apply a permutation matrix which moves the current block as far to the top-left as possible, while still guaranteeing that the blocks are in descending order.

Example 8.4. The following is a worked example of the conjugation procedure for a matrix $A \in \mathcal{U}_9(\mathbb{Q})$. We use $\mathbb{Q}$ as the field for simplicity, though using a finite field does not present any extra difficulty. We write $A^{[i]}$ for the matrix obtained after step i.

$$A = \left(\begin{array}{ccc|cc|cc|c||c} & 1 & & & & & & & 3 \\ & & 1 & & & & & & 1 \\ & & & & & & & & 0 \\ \hline & & & & 1 & & & & 2 \\ & & & & & & & & 2 \\ \hline & & & & & & 1 & & 4 \\ & & & & & & & & 0 \\ \hline & & & & & & & & 1 \\ \hline\hline 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{array}\right) \quad A^{[1]} = \left(\begin{array}{ccc|cc|cc|c||c} & 1 & & & & & & & 0 \\ & & 1 & & & & & & 0 \\ & & & & & & & & 0 \\ \hline & & & & 1 & & & & 0 \\ & & & & & & & & 2 \\ \hline & & & & & & 1 & & 0 \\ & & & & & & & & 0 \\ \hline & & & & & & & & 1 \\ \hline\hline 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{array}\right) \quad A^{[2]} = \left(\begin{array}{ccc|cc|cc|c||c} & 1 & & & & & & & 0 \\ & & 1 & & & & & & 0 \\ & & & & & & & & 0 \\ \hline & & & & 1 & & & & 0 \\ & & & & & & & & 1 \\ \hline & & & & & & 1 & & 0 \\ & & & & & & & & 0 \\ \hline & & & & & & & & \frac{1}{2} \\ \hline\hline 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{array}\right)$$

$$A^{[3]} = \left(\begin{array}{ccc|cc|cc|c||c} & 1 & & & & & & & 0 \\ & & 1 & & & & & & 0 \\ & & & & & & & & 0 \\ \hline & & & & 1 & & & & 0 \\ & & & & & & & & 1 \\ \hline & & & & & & 1 & & 0 \\ & & & & & & & & 0 \\ \hline & & & & & & & & 0 \\ \hline\hline 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{array}\right) \quad A^{[4]} = \left(\begin{array}{ccc|ccc|cc|c} & 1 & & & & 0 & & & \\ & & 1 & & & 0 & & & \\ & & & & & 0 & & & \\ \hline & & & & 1 & 0 & & & \\ & & & & & 1 & & & \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ \hline & & & & & 0 & & 1 & \\ & & & & & 0 & & & \\ \hline & & & & & 0 & & & \end{array}\right) \quad A^{[5]} = \left(\begin{array}{ccc|ccc|cc|c} & 1 & 0 & & & & & & \\ & & 1 & & & & & & \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ \hline & & 0 & & 1 & & & & \\ & & 0 & & & 1 & & & \\ & & 0 & & & & & & \\ \hline & & 0 & & & & & 1 & \\ & & 0 & & & & & & \\ \hline & & 0 & & & & & & \end{array}\right)$$

Note that $A^{[4]} = A^{[5]}$, but the explicitly shown zeros are in different locations. This is to emphasize that the last step in our conjugation procedure may fix the Jordan form (as it does in this case). However, when we apply the same procedure to $C_{\mathcal{U}}(A)$, this last action will often have a non-trivial effect.

Remark 8.5. It is tempting to assume that, if A is already in Jordan form, then the conjugating matrix X_A will be the identity. This is not necessarily so. While many of the steps in the conjugation procedure are trivial, the permutation matrices in steps 4 and 5 need not be the identity. For example, if $A = \mathrm{id}_n$, then X_A is the permutation matrix defined by the permutation $w : k \mapsto n+1-k$. It is true however that if A is already in Jordan form, then X_A will be a permutation matrix.

To be specific about this procedure, we now write down explicitly the matrices used in the conjugation procedure. We conjugate A by a product of five matrices, one for each step in the conjugation procedure. As was shown in Example 8.4, for $i = 1, \ldots, 5$ we will use $A^{[i]}$ to denote the matrix obtained after step i.

First, conjugate by $E_A := 1 + \sum_{i=1}^{n-1} A_{i,n} e_{i+1,n}$, where $e_{i,j}$ is the matrix containing a 1 in entry (i, j) and zeros everywhere else. Conjugating A by E_A gives the matrix $A^{[1]} = E_A A E_A^{-1}$ satisfying

1. $A^{[1]}|_{n-1} = J_\mu$,

2. each row of $A^{[1]}$ has at most one non-zero entry.

Thus, the only non-zero entries in the last column of $A^{[1]}$ are in line with the bottom of a Jordan block of J_μ.

Second, let x denote the first non-zero entry in column n of $A^{[1]}$, if it exists (and $x = 1$ otherwise). Define

$$\Delta_A := \operatorname{diag}(\underbrace{1, \ldots, 1}_{n-1}, x).$$

Conjugating $A^{[1]}$ by Δ_A yields the matrix $A^{[2]} = \Delta_A A^{[1]} \Delta_A^{-1}$ which has at most one non-zero entry in each row, and which has a 1 as its first non-zero entry in column n (if column n is has any non-zero entries).

Third, define a lower-triangular matrix L_A which uses the first non-zero entry in the nth column to set all other entries in that column to zero. If column n is already zero, then simply set L_A to be the identity. Otherwise, let $\widetilde{\mu}_s := \sum_{i=1}^{s} \mu_i$. These numbers are the indices of rows which are at the bottom of Jordan blocks. Every non-zero entry in column n of $A^{[2]}$ appears in such a row. Define r to be the integer such that $\widetilde{\mu}_r$ is the index of the row containing the first non-zero entry in the nth column of $A^{[2]}$, and for $\alpha \in \mathbb{F}_q$ define

$$F_{j,r}(\alpha) := 1 + \alpha \sum_{k=1}^{\mu_j} e_{\widetilde{\mu}_{j-1}+k,\, \widetilde{\mu}_r - \mu_j + k}, \tag{8.2}$$

where $e_{i,j}$ denotes the matrix with a 1 in position (i, j) and zeros everywhere else. Left-multiplication by $F_{j,r}(\alpha)$ takes the last μ_j rows in the rth block (rows $\widetilde{\mu}_r - \mu_j + 1$ through $\widetilde{\mu}_r$), and adds them to the rows in the jth block (first scaling them by α). Right-multiplication by $F_{j,r}(\alpha)$ takes the μ_j columns in the jth block, and subtracts them from the last μ_j columns in the rth block (first scaling them by α). Conjugating $A^{[2]}$ by $F_{j,r}(-A^{[2]}_{n,\widetilde{\mu}_j})$ will leave $A^{[2]}$ unchanged in every entry except in the entry indexed by $(n, \widetilde{\mu}_j)$, which will be set to zero. We can therefore define

$$L_A := \prod_{j>r} F_{j,r}(-A^{[2]}_{n,\widetilde{\mu}_j}),$$

so that $A^{[3]} = L_A A^{[2]} L_A^{-1}$ has one non-zero entry in column n, and that value is 1.

Fourth, conjugate by a permutation matrix to make column n align with the correct block. We apply the permutation $\sigma_A := (\widetilde{\mu}_r + 1, \widetilde{\mu}_r + 2, \ldots, n)$. If column n is zero in $A^{[3]}$ then no such permutation is necessary, and we set σ_A to be the identity. Now $A^{[4]} = \sigma_A A^{[3]} \sigma_A^{-1}$ is the direct sum of Jordan blocks, though not necessarily in descending order.

Lastly, we apply a permutation matrix τ_A which moves the block as close to the top-left as possible to put the blocks into descending order, in such a way as to preserve the relative order of all other Jordan blocks.

We are now in a position to define our choice of conjugating matrix X_A, so that $X_A A X_A^{-1}$ is in Jordan form. Define X_A recursively. For the unique $A \in \mathcal{U}_1$, we take $X_A = (1)$ to be the identity matrix. For $n > 1$, let $B = A|_{n-1}$, and define $A' = X_B A X_B^{-1}$, so that $A'|_{n-1}$ is in Jordan form. Then define

$$Y_{A'} := \tau_{A'} \sigma_{A'} L_{A'} \Delta_{A'} E_{A'}, \text{ and} \tag{8.3}$$

$$X_A := Y_{A'} X_B. \tag{8.4}$$

Here we have implicitly identified $X_B \in \mathcal{M}_{(n-1)\times(n-1)}$ with $X_B \oplus 1 \in \mathcal{M}_{n\times n}$. By construction, we have $X_A A X_A^{-1} = J_\lambda$.

8.4 Gap arrays

In this section, we introduce a new combinatorial object called a *gap array*. Gap arrays are used to encode certain subspaces of $C_{\mathcal{M}}(J_\lambda)$. Many algebraic operations we can apply to these subspaces are encoded succinctly by combinatorial operations we can apply to gap arrays. This will be our primary tool for computing upper bounds on $\mathbf{k}(U_n)$.

8.4.1 Definitions

Definition 8.6. Let $\lambda \vdash n$ be a partition of length $\ell := \ell(\lambda)$. A *gap array* of type λ is an $(\ell \times \ell)$-matrix $\mathbf{G} = (\mathbf{G}_{i,j})$ of integers satisfying

$$\max\{0, \lambda_i - \lambda_j\} \leq \mathbf{G}_{i,j} \leq \lambda_i. \tag{8.5}$$

We use gap arrays to define subspaces of $C_{\mathcal{M}}(J_\lambda)$. Recall that $C_{\mathcal{M}}(J_\lambda)$ has a natural block decomposition into blocks of $(\lambda_i \times \lambda_j)$-submatrices. The block with rows corresponding to the ith part of λ and columns corresponding to the jth part of λ is called the (i, j)-*block*. Specifically, it consists of all cells (x, y) in the matrix such that

$$\lambda_1 + \cdots + \lambda_{i-1} < x \leq \lambda_1 + \cdots + \lambda_i, \text{ and}$$
$$\lambda_1 + \cdots + \lambda_{j-1} < y \leq \lambda_1 + \cdots + \lambda_j.$$

Definition 8.7. For a gap array $\mathbf{G}$ of type λ, we let $C(\mathbf{G})$ denote the subspace of $C_{\mathcal{M}}(J_\lambda)$ satisfying the condition that on the (i, j)-block, the lowest $\mathbf{G}_{i,j}$ diagonals touching the right boundary are all zero.

Example 8.8. Let $\lambda = (6, 2, 1, 1) \vdash 10$, and let $\mathbf{G}$ denote the gap array

$$\mathbf{G} = \begin{pmatrix} 2 & 4 & 6 & 5 \\ 1 & 0 & 1 & 2 \\ 0 & 1 & 1 & 0 \\ 1 & 0 & 0 & 0 \end{pmatrix}.$$

Then the corresponding subspace $C(\mathbf{G})$ of $C_{\mathcal{M}}(J_\lambda)$ consists of all 10×10 matrices of the form shown in Figure 8.1. In this example, $C(\mathbf{G})$ is a 16-dimensional subspace of $C_{\mathcal{M}}(J_\lambda)$.

Remark 8.9. The name "gap array" is chosen, because the recorded numbers measure the gap between the bottom of a block and the triangle of potentially non-zero entries it contains. Recall from the proof of Lemma 8.2 that a $C_{\mathcal{M}}(J_\lambda)$ is characterized by, being Toeplitz on each block, and having any diagonal which does not touch both the topmost row and rightmost column of a block set to zero. Thus, $C_{\mathcal{M}}(J_\lambda) = C(\mathbf{G})$, where $\mathbf{G}_{i,j} = \max\{0, \lambda_i - \lambda_j\}$. At the

$$
\left(\begin{array}{cccccc|cc|c|c}
 & & a_3 & a_4 & a_5 & a_6 & b_5 & b_6 & & c_6 \\
 & & & a_3 & a_4 & a_5 & & b_5 & & \\
 & & & & a_3 & a_4 & & & & \\
 & & & & & a_3 & & & & \\
 & & & & & & & & & \\
 & & & & & & & & & \\
\hline
 & & & & & d_2 & e_1 & e_2 & f_2 & \\
 & & & & & & & e_1 & & \\
\hline
 & & & & & g_1 & & & & h_1 \\
\hline
 & & & & & & & i_1 & j_1 & k_1
\end{array}\right)
$$

Figure 8.1: The form of all matrices in the subspace $C(\mathbf{G})$ of $C_{\mathcal{M}}(J_\lambda)$

other extreme, the zero subspace of $\mathcal{M}_{n\times n}$ is $C(\mathbf{G})$, where $\mathbf{G}_{i,j} = \lambda_i$. The inequalities in (8.5), which defines gap arrays, encode these observations.

For gap arrays $\mathbf{G} = (\mathbf{G}_{i,j})$ and $\mathbf{H} = (\mathbf{H}_{i,j})$ of the same type, we say that $\mathbf{G} \leq \mathbf{H}$ if $\mathbf{G}_{i,j} \leq \mathbf{H}_{i,j}$ for all i and j. It follows that $\mathbf{G} \leq \mathbf{H}$ if and only if $C(\mathbf{H}) \subseteq C(\mathbf{G})$. We also write $|\mathbf{G}| = \sum_{i,j} \mathbf{G}_{i,j}$.

Proposition 8.10. *Let* $\mathbf{G}$ *be a gap array of type* $\lambda \vdash n$*, and let* $\ell = \ell(\lambda)$*. Then*

$$\dim C(\mathbf{G}) = n\ell - |\mathbf{G}|\,.$$

Proof. Observe that $\mathbf{G}_{i,j}$ counts the number of diagonals touching the rightmost column of the (i,j)-block which are zero in every element of $C(\mathbf{G})$. Thus, $|\mathbf{G}|$ counts the number of diagonals in all blocks which are necessarily zero. The quantity $\dim C(\mathbf{G})$ counts the number of diagonals in all blocks which may be non-zero. Together, these constitute all diagonals on all blocks, so $|\mathbf{G}| + \dim C(\mathbf{G})$ is a constant depending only on the type λ of $\mathbf{G}$. Taking $\dim C(\mathbf{G}) = 0$, by setting $\mathbf{G}_{i,j} = \lambda_i$, we see that $|\mathbf{G}| = \sum_{i,j} \lambda_i = n\ell$. Hence, for every gap array $\mathbf{G}$ of type λ, we have $\dim C(\mathbf{G}) = n\ell - |\mathbf{G}|$. □

8.4.2 Combinatorics of gap arrays

Recall from Section 8.3 that we can put a matrix A into Jordan form by iteratively conjugating larger and larger submatrices into Jordan form. As we do so, we conjugate $C_{\mathcal{U}}(A)$ by the same process. The purpose of this section is to translate what happens to $C_{\mathcal{U}}(A)$ in this iterative procedure into the combinatorics of gap arrays.

For a partition λ, and for $r \in \{1, \ldots, \ell(\lambda)\}$, let $\phi_r(\lambda)$ denote the partiton obtained by adding a block to the diagram for λ into row r. If $\lambda_r = \lambda_{r-1}$, then the result will be a composition and no longer a partition, so we reorder the rows to obtain a partition. For example, if $\lambda = (3^1 2^2 1^4)$, then $\phi_6(\lambda) = (3^1 2^3 1^3)$. If $r = \ell(\lambda) + 1$, then we define $\phi_r(\lambda)$ to be the partition obtained by adding a part of size 1.

Definition 8.11. For a gap array $\mathbf{G}$ of type $\lambda = (\lambda_1, \ldots, \lambda_\ell)$, we define a new gap array $\psi_r(\mathbf{G})$ of type $\phi_r(\lambda)$. Define $\psi_r(\mathbf{G})$ as follows:

1. If $r = \ell + 1$, add a new row of all zeros at the bottom of $\mathbf{G}$, and a new column at the right of $\mathbf{G}$ with the values $(\lambda_1, \ldots, \lambda_\ell, 0)$.
2. Subtract 1 from each non-zero entry in column r of $\mathbf{G}$.
3. Add 1 to each entry in row r.
4. Permute the rows and columns so that corresponding block sizes are in decreasing order. Specifically, move row and column r to be the very first row and column of their particular block size. Keep all other rows/columns in the same relative order.

Proposition 8.12. *Let $\mathbf{G}$ be a gap array of type λ, and let $r \in \{1, \ldots, \ell(\lambda) + 1\}$. Then $\psi_r(\mathbf{G})$ is a gap array of type $\phi_r(\lambda)$.*

Proof. First note that to obtain $\psi_r(\mathbf{G})$, the rows and columns of $\mathbf{G}$ are rearranged according to the same permutation as the entries in λ when computing $\phi_r(\lambda)$. We can therefore ignore the permutations and check that that the entries in $\psi_r(\mathbf{G})$ satisfy the inequalities given in (8.5).

Note that $\mathbf{G}_{i,j}$ will be increased (by one) if and only if $i = r$. Because λ_r is also increased by one, the upper bound given in (8.5) will still be satisfied. From the definition of ψ_r, we see that no entry in $\psi_r(\mathbf{G})$ can be negative. Hence, the only way an inequality from (8.5) can fail to be satisfied is if $\mathbf{G}_{i,j}$ is decreased (by one). However, in this situation, it must be that $j = r$, and so the lower bound of $\lambda_i - \lambda_j$ is also decreased by one. Thus, the inequalities from (8.5) are always preserved, and so $\psi_r(\mathbf{G})$ is a gap array of type $\phi_r(\mathbf{G})$. □

Example 8.13. Starting with $\lambda = (3, 2, 1) \vdash 6$, we apply ϕ_2, ϕ_4, and ϕ_4 in that order. We obtain

$$\phi_2(\lambda) = (3, 3, 1)$$
$$\phi_4(\phi_2(\lambda)) = (3, 3, 1, 1)$$
$$\phi_4(\phi_4(\phi_2(\lambda))) = (3, 3, 2, 1).$$

Correspondingly, for a gap array, let

$$\mathbf{G} = \begin{pmatrix} 1 & 1 & 2 \\ 1 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix}.$$

Applying ψ_2, ψ_4, and ψ_4 in that order to $\mathbf{G}$, we obtain

$$\psi_2(\mathbf{G}) = \begin{pmatrix} 1 & 2 & 2 \\ 0 & 1 & 2 \\ 0 & 0 & 1 \end{pmatrix}, \quad \psi_4(\psi_2(\mathbf{G})) = \begin{pmatrix} 1 & 2 & 2 & 2 \\ 0 & 1 & 2 & 2 \\ 1 & 1 & 1 & 1 \\ 0 & 0 & 0 & 1 \end{pmatrix}, \text{ and}$$

$$\psi_4(\psi_4(\psi_2(\mathbf{G}))) = \begin{pmatrix} 1 & 2 & 1 & 2 \\ 0 & 1 & 1 & 2 \\ 1 & 1 & 1 & 1 \\ 1 & 1 & 0 & 1 \end{pmatrix}.$$

Definition 8.14. Let $\lambda \vdash n$, and let $r \in \{1, \dots, \ell(\lambda)\}$. Then a gap array $\mathbf{G}$ of type λ is called *r-valid* if the following two conditions hold

1. Row r is element-by-element weakly larger than every row below it. That is, for every $j > r$, and every k,

$$\mathbf{G}_{j,k} \leq \mathbf{G}_{r,k}.$$

2. Column r is element-by-element weakly smaller than every column to the right. That is, for every $j > r$ and every k,

$$\mathbf{G}_{k,r} \leq \mathbf{G}_{k,j}.$$

Validity is a technical condition which will be needed in Lemma 8.15. The most important property to recognize is that a gap array is r-valid for every r if and only if its entries are increasing from left to right and from bottom to top. This is a property that we will use in the proof of Proposition 8.19

We need one more definition to state Lemma 8.15. For a subspace $V \subseteq \mathcal{M}_{(n-1)\times(n-1)}$, define

$$\overline{V} := \{X \in \mathcal{M}_{n\times n} \mid X|_{n-1} \in V, X_{n,i} = 0 \text{ for all } i \leq n-1\}.$$

Graphically, we are considering the subspace shown in Figure 8.2. We are now ready to state and prove our main lemma regarding gap arrays. This subspace construction is particularly relevant for us, because $\mathcal{U}_n = \overline{\mathcal{U}_{n-1}}$.

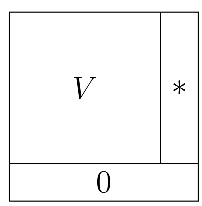


Figure 8.2: A graphical representation of $\overline{V}$, where $V \subseteq \mathcal{M}_{(n-1)\times(n-1)}$.

Lemma 8.15. *Let $A \in \mathcal{U}_n$ and let $\mu \vdash (n-1)$ such that $A|_{n-1} = J_\mu$. Moreover, let $\mathbf{G}$ be an r-valid gap array of type μ. Then*

$$Y_A \left(\overline{C(\mathbf{G})} \cap C_{\mathcal{M}}(A)\right) Y_A^{-1} = C(\psi_r(\mathbf{G})).$$

Proof. Recall the definition $Y_A := \tau_A \sigma_A L_A \Delta_A E_A$ given in (8.3). We first consider $Y_A\overline{C(\mathbf{G})}Y_A^{-1}$. Note that for any subspace $V \subseteq \mathcal{M}_{(n-1)\times(n-1)}$, the matrix E_A normalizes $\overline{V}$, as its action only changes the last column, and fixes the bottom-right entry. Similarly, Δ_A normalizes $\overline{V}$. Hence,

$$Y_A\overline{C(\mathbf{G})}Y_A^{-1} = (\tau_A\sigma_A L_A)\overline{C(\mathbf{G})}(\tau_A\sigma_A L_A)^{-1}.$$

Next, we claim that L_A normalizes $\overline{C(\mathbf{G})}$. Let $A_{[i,j]}$ denote the $(\mu_i \times \mu_j)$-matrix obtained by restricting A to the (i,j)-block. Recall that L_A is defined in (8.2) to be the product of the lower-triangular matrices of the form $F_{j,r}(\alpha)$. Let $B \in \overline{C(\mathbf{G})}$. Then by construction, $F_{j,r}(\alpha) \cdot B$ agrees with B on all blocks except for those of the form $B_{[j,k]}$ for some k. On such blocks, the action of left-multiplication by $F_{j,r}(\alpha)$ is to add the bottom λ_j rows of $\alpha \cdot B_{[r,k]}$ to $B_{[j,k]}$. Because $\mathbf{G}$ is r-valid, and $j > r$, we know that $\mathbf{G}_{j,k} \leq \mathbf{G}_{r,k}$ for all k. Thus, every diagonal of non-zero entries in $\alpha \cdot B_{[r,k]}$ gets added to a diagonal in $B_{[j,k]}$ which is allowed to be non-zero. As a result, left-multiplication by $F_{j,r}(\alpha)$ stabilizes $\overline{C(\mathbf{G})}$.

The fact that left-multiplication by $F_{j,r}(\alpha)$ stabilizes $\overline{C(\mathbf{G})}$ is shown graphically by the left diagram in Figure 8.3. The light-gray strips represent the rows and columns of the relevant blocks. The vertical strip represents the $(*,k)$-blocks, and the horizontal strips represent the $(r,*)$- and $(j,*)$-blocks. The dark-gray triangles represent those diagonals in each block which may be non-zero. The black triangle shows the cells in the (r,k)-block which get carried to the (j,k)-block via the left-multiplication by $F_{j,r}(\alpha)$.

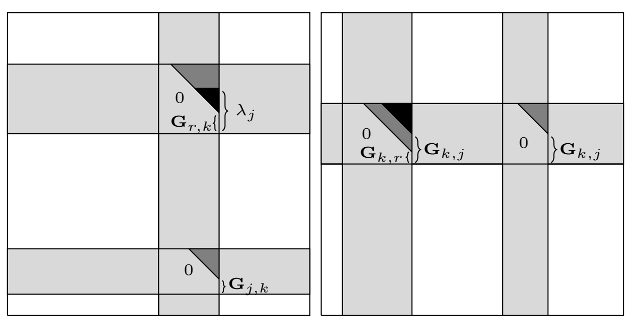


Figure 8.3: A graphical representation of how r-validity implies that conjugation by $F_{j,r}(\alpha)$ stabilizes $\overline{C(\mathbf{G})}$.

Similarly, right-multiplication by $F_{j,r}(\alpha)^{-1}$ maps $\overline{C(\mathbf{G})}$ to itself as a consequence of the column-condition of r-validity. For each k, the action of right-multiplication by $F_{j,r}(\alpha)^{-1}$ subtracts $\alpha \cdot B_{[k,j]}$ from $B_{[k,r]}$, where the right edges of each block are aligned, and $\alpha \cdot B_{[k,j]}$ is extended to the left with zeros if necessary. Because $\mathbf{G}$ is r-valid, and $j > r$, we know that $\mathbf{G}_{k,r} \leq \mathbf{G}_{k,j}$ for all k. Thus, every diagonal of non-zero entries in $B_{[k,j]}$ gets subtracted from a diagonal in $B_{[k,r]}$ which is allowed to be non-zero.

The fact that right-multiplication by $F_{j,r}(\alpha)^{-1}$ stabilizes $\overline{C(\mathbf{G})}$ This is shown graphically by the diagram on the right in Figure 8.3. The dark-gray triangles represent those diagonals in each block which may be non-zero. The black triangle shows where the right-multiplication by $F_{j,r}(\alpha)^{-1}$ carries the triangle from the (k,j)-block. Because $\mathbf{G}_{k,j} \geq \mathbf{G}_{k,r}$, the black triangle lies inside the dark-gray triangle in the (k,r)-block. Hence, right-multiplication by $F_{j,r}(\alpha)^{-1}$ stabilizes $\overline{C(\mathbf{G})}$. In this way, we see that r-validity is a combinatorial description of the fact that $F_{j,r}(\alpha)$ normalizes $\overline{C(\mathbf{G})}$. It follows that $Y_A\overline{C(\mathbf{G})}Y_A^{-1} = (\tau_A\sigma_A)\overline{C(\mathbf{G})}(\tau_A\sigma_A)^{-1}$.

Next, recognize that σ_A moves the last row and column into the rth block, effectively increasing the size of the rth block by 1. Then τ_A acts by rearranging these blocks, guaranteeing that their sizes are in descending order. These permutations have the same action on the block sizes as the function ϕ_r does to the parts of μ. Thus, the block sizes are now described by the partition $\phi_r(\mu)$. Hence, because $Y_A C_{\mathcal{M}}(A) Y_A^{-1} = C_{\mathcal{M}}(Y_A A Y_A^{-1}) = C_{\mathcal{M}}(J_{\phi_r(\mu)})$, we have

$$Y_A\left(\overline{C(\mathbf{G})} \cap C_{\mathcal{M}}(A)\right)Y_A^{-1} = (\tau_A\sigma_A)\overline{C(\mathbf{G})}(\tau_A\sigma_A)^{-1} \cap C_{\mathcal{M}}(J_{\phi_r(\mu)}). \tag{8.6}$$

The left-hand side of (8.6) is a subspace of $C_{\mathcal{M}}(J_{\phi_r(\mu)})$. Therefore every block must be Toeplitz (constant on diagonals), and any diagonal not touching both the topmost row and rightmost column of a block must be zero. We now consider how the action of $\tau_A\sigma_A$ affects each block. Because τ_A only permutes the blocks, we focus our attention on the action of σ_A. If $i \neq r$ and $j \neq r$, then σ_A has no affect on an (i,j)-block. This coincides with the fact that ψ_r does not affect cells (i,j) in a gap array when $i \neq r$ and $j \neq r$.

Considering an (r,j)-block for $j \neq r$, we see that σ_A moves the last row to the bottom of such a block. As this row is necessarily zero, it increases the gap on an (r,j)-block by 1. This is precisely what the map ψ_r encodes by adding 1 to each entry in row r of the gap array.

Considering an (i,r)-block, for $i \neq r$, we see that σ_A moves the last column to the right edge of this block. Because the result is guaranteed to be Toeplitz, the size of the gap can only decrease by 1, as shown in Figure 8.4. If the gap size on an (i,r)-block was already zero in $C(\mathbf{G})$, then such matrices cannot possibly lie in $Y_A\left(\overline{C(\mathbf{G})} \cap C_{\mathcal{M}}(A)\right)Y_A^{-1}$, because the action of σ_A would force the (i,r)-block to no longer satisfy the upper-triangular conditions necessary to lie in $C_{\mathcal{M}}(J_{\phi_r(\mu)})$. Thus, the resulting block will have its gap size decreased by 1 unless it is already zero, in which case, the gap size remains zero. This is encoded by the action of ψ_r on the rth column of $\mathbf{G}$.

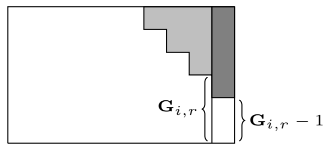


Figure 8.4: A graphical representation of how an (i,r)-block is affected by the action of σ_A for $i \neq r$.

Regarding the (r,r)-block, both its width and height are increased by 1. As the resulting space is guaranteed to be Toeplitz, the gap size does not change. This is consistent with $\psi_r(\mathbf{G})$. Thus, it follows that $Y_A\left(\overline{C(\mathbf{G})} \cap C_{\mathcal{M}}(A)\right) = C(\psi_r(\mathbf{G}))$ as desired. $\square$

8.5 Proof of upper bound on $\mathbf{k}(U_n)$

In this section, we use the techniques of gap arrays developed earlier in the chapter to prove the following upper bound on $\mathbf{k}(U_n)$.

Theorem 1.7. *For every positive integer n and every prime power q, we have*

$$\mathbf{k}(U_n(q)) \leq p(n)^2 n!\; q^{\alpha n^2 + \frac{n}{2}},$$

where $p(n)$ denotes the number of integer partitions of n, and where

$$\alpha = \frac{40\sqrt{2} - 41}{98} \approx 0.15886.$$

Remark 8.16. Notice that even when q is as small as 2, the term $p(n)^2 n!$ is $O(q^{n \log n})$. The dominant term in the preceding theorem is $q^{\alpha n^2}$.

We begin with an overview of our proof methods. We first construct, for each $\lambda \vdash n$, a gap array $\mathbf{G}^\lambda$ that encodes a particular subspace of $C_{\mathcal{M}}(J_\lambda)$. The space $C(\mathbf{G}^\lambda)$ contains every space of the form $X_A C_{\mathcal{U}}(A) X_A^{-1}$, where A is conjugate to J_λ. In this sense, G^λ is the worst-case gap matrix corresponding to an upper-triangular matrix of Jordan type λ. This is statement is made precise in Theorem 8.21.

The gap array $\mathbf{G}^\lambda$ is particularly easy to analyze combinatorially. We compute $\left|\mathbf{G}^\lambda\right|$ in Proposition 8.20. We then show that the worst possible choice λ gives an exponent of $\frac{1}{6}n^2$. By itself, this is already asymptotically equivalent to the bounds given by Marberg in [Mar1] and by Vera-López and Arregi [VA1]. The content of the proof of Theorem 1.7 is to combine the technique of gap arrays with a simpler bound in order to improve this exponent.

Recall the notation $\mathrm{Comm}(\mathcal{U}_n) = \{(A, B) \in \mathcal{U}_n \times \mathcal{U}_n \mid AB = BA\}$. We introduce two families of subspaces of $\mathrm{Comm}(\mathcal{U}_n)$ which we parameterize by partitions. For partitions $\lambda, \mu \vdash n$, define

$$\mathrm{Comm}(\lambda) := \{(A, B) \in \mathcal{U}_n \times \mathcal{U}_n \mid AB = BA,\ \mathrm{sh}(A) = \lambda\},\ \text{and}$$
$$\mathrm{Comm}(\lambda, \mu) := \{(A, B) \in \mathcal{U}_n \times \mathcal{U}_n \mid AB = BA,\ \mathrm{sh}(A) = \lambda,\ \mathrm{sh}(B) = \mu\}.$$

We improve on the $\frac{1}{6}n^2$ exponent by first proving a technical lemma (Lemma 8.23) regarding the bounds on each $\mathrm{Comm}(\lambda)$. This lemma allows us to determine when gap arrays will be useful, and when we should bound $\mathbf{k}(U_n)$ via different (but simpler) techniques. The proof of Theorem 1.7 combines these techniques.

8.5.1 Bounds on gap arrays

For any partition $\lambda \vdash n$, there is a specific gap array whose corresponding subspace of $C_{\mathcal{M}}(J_\lambda)$ contains all possible subspaces coming from upper-triangular matrices. Moreover, we can construct this gap array explicitly

Definition 8.17. For $\lambda \vdash n$, we define the gap array $\mathbf{G}^\lambda$ by

$$\mathbf{G}^\lambda_{i,j} = \begin{cases} \lambda_i - \lambda_j & \text{if } \lambda_i > \lambda_j \\ 1 & \text{if } \lambda_i = \lambda_j \text{ and } i \leq j \\ 0 & \text{otherwise.} \end{cases}$$

Example 8.18. For the partition $\lambda = (6, 3, 1, 1, 1)$, we have

$$\mathbf{G}^\lambda = \begin{pmatrix} 1 & 3 & 5 & 5 & 5 \\ 0 & 1 & 2 & 2 & 2 \\ 0 & 0 & 1 & 1 & 1 \\ 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix}.$$

Proposition 8.19. *For every $\lambda \vdash n$, the gap array $\mathbf{G}^\lambda$ is r-valid for every r, and*

$$\mathbf{G}^{\phi_r(\lambda)} \leq \psi_r(\mathbf{G}^\lambda).$$

Proof. A gap array is r-valid for every r if and only if the entries are weakly increasing from left to right and from bottom to top. From the definition of $\mathbf{G}^\lambda$, it is clear that $\mathbf{G}^\lambda$ satisfies this property. It therefore suffices to prove that $\mathbf{G}^{\phi_r(\lambda)} \leq \psi_r(\mathbf{G}^\lambda)$.

First suppose that $r = \ell + 1$, where $\ell = \ell(\lambda)$. Recalling the definition of ψ_r from Subsection 8.4.2, we see that that $\psi_r(\mathbf{G}^\lambda)$ adds a new column containing the value $\lambda_i - 1$ in entry $(i, \ell + 1)$, and a new row of all 1s in row $\ell + 1$. This row and column are then permuted to be the first row and column corresponding to a part of $\phi_r(\lambda)$ of size 1. We may obtain $\mathbf{G}^{\phi_r(\lambda)}$ can be obtained from $\mathbf{G}^\lambda$ by adding a new column containing the number $\lambda_i - 1$ in the cell associated to the ith part of λ, and a new row containing 1s and 0s. The new columns and rows in these two gap arrays align by construction. If the new values in these rows and columns do not agree, it is because a cell has the value 0 in $\mathbf{G}^{\phi_r(\lambda)}$, and the value 1 in $\psi_r(\mathbf{G}^\lambda)$. Thus, we obtain $\mathbf{G}^{\phi_r(\lambda)} \leq \psi_r(\mathbf{G}^\lambda)$.

Now suppose $1 \leq r \leq \ell(\lambda)$. The map ψ_r acts by subtracting 1 from each non-zero entry in column r, adding 1 to each entry in row r, and then permuting the rows and columns so that this row and column are the first corresponding to a part of size $\lambda_r + 1$. Let w denote this permutation. Let s denote the new index of row/column r after this permutation, and let $\mathbf{H} = \psi_r(\mathbf{G}^\lambda)$. Then from the definition of the map ψ_r in Subsection 8.4.2, we see that

1. $\mathbf{H}_{w(r),w(r)} = 1$,

2. $\mathbf{H}_{w(r),w(j)} = \mathbf{G}^\lambda_{r,j} + 1$ for $j \neq r$,

3. $\mathbf{H}_{w(i),w(r)} = \max\{0, \mathbf{G}^{\lambda}_{i,r} - 1\}$ for $i \neq r$,

4. $\mathbf{H}_{w(i),w(j)} = \mathbf{G}^{\lambda}_{i,j}$ for all $i, j \neq r$.

Now consider $\mathbf{G}^{\phi_r(\lambda)}$. Recall that ϕ_r acts on λ by increasing the rth part, and then permuting the parts so that the resulting composition becomes a partition again. Specifically, this can be done by applying the same permutation w as above. In other words, let $\mu := \phi_r(\lambda)$. Then

$$\mu_{w(i)} = \begin{cases} \lambda_r + 1 & \text{if } i = r \\ \lambda_i & \text{if } i \neq r. \end{cases}$$

We now show that for all i and j, we have $\mathbf{G}^{\mu}_{w(i),w(j)} \leq \mathbf{H}_{w(i),w(j)}$. We break this into four cases to coincide with the four items defining $\mathbf{H}_{w(i),w(j)}$ above.

Case 1: $i = j = r$

It is immediate from the definition of $\mathbf{G}^{\mu}$ that every entry on the main diagonal is equal to 1. Thus,

$$\mathbf{G}^{\mu}_{w(r),w(r)} = 1 = \mathbf{H}_{w(r),w(r)}.$$

Case 2: $i = r$**,** $j \neq r$

If $\mu_{w(r)} > \mu_{w(j)}$, then $\mathbf{G}^{\mu}_{w(r),w(j)} = \mu_{w(r)} - \mu_{w(j)}$. We therefore obtain

$$\begin{aligned} \mathbf{G}^{\mu}_{w(r),w(j)} &= \mu_{w(r)} - \mu_{w(j)} \\ &= \lambda_r + 1 - \lambda_j = \mathbf{G}^{\lambda}_{r,j} + 1 = \mathbf{H}^{\lambda}_{w(r),w(j)}. \end{aligned}$$

Otherwise, we see that $\mathbf{G}^{\mu}_{w(r),w(j)} \leq 1$, and $\mathbf{H}_{w(r),w(j)} = \mathbf{G}^{\lambda}_{r,j} + 1 \geq 1$. Thus, we have $\mathbf{G}^{\mu}_{w(r),w(j)} \leq \mathbf{H}_{w(r),w(j)}$ as desired.

Case 3: $i \neq r$, $j = r$

First suppose $\mu_{w(i)} > \mu_{w(r)}$, so that $\mathbf{G}^{\mu}_{w(i),w(r)} = \mu_{w(i)} - \mu_{w(r)}$. Thus, we see that $\lambda_i - (\lambda_r + 1) > 0$, so $\mathbf{H}_{w(i),w(r)} = \mathbf{G}^{\lambda}_{i,r} - 1$. We therefore obtain

$$\mathbf{G}^{\mu}_{w(i),w(r)} = \mu_{w(i)} - \mu_{w(r)} = \lambda_i - (\lambda_r + 1) = \mathbf{H}_{w(i),w(r)}.$$

Next, suppose that $\mu_{w(i)} = \mu_{w(r)}$, and $w(i) \leq w(r)$. Recall that w is defined to have the property that $w(r)$ is the first part in μ of size $\mu_{w(r)}$. It follows that if $\mu_{w(i)} = \mu_{w(r)}$ and $w(i) \leq w(r)$, then in fact $w(i) = w(r)$. This is not possible, as $i \neq r$ by hypothesis.

In the remaining cases, $\mathbf{G}^{\mu}_{w(i),w(r)} = 0$. We know from the definition of gap arrays that $\mathbf{H}_{w(i),w(r)} \geq 0$ as desired.

Case 4: $i \neq r$, $j \neq r$

Recalling that for $i \neq r$, we have $\mu_{w(i)} = \lambda_i$, we see that

$$\mathbf{G}^{\mu}_{w(i),w(j)} = \mathbf{G}^{\lambda}_{i,j} = \mathbf{H}_{w(i),w(j)}.$$

Thus, for every i and j, we have $\mathbf{G}^{\mu}_{w(i),w(j)} \leq \mathbf{H}_{w(i),w(j)}$, so we obtain the relationship $\mathbf{G}^{\phi_r(\lambda)} \leq \psi_r(\mathbf{G}^{\lambda})$ as desired. □

Proposition 8.20. *For any $\lambda \vdash n$ with $\ell = \ell(\lambda)$, we have*

$$\left|\mathbf{G}^{\lambda}\right| = n\ell - n - 2\,\mathbf{n}(\lambda) + \sum_i \frac{m_i^2}{2} + \frac{\ell}{2}.$$

Proof. Recall that $\mathbf{G}^{\lambda}$ is an upper-triangular gap array, and that for $i \leq j$ and $\lambda_i = \lambda_j$, we have $\mathbf{G}^{\lambda}_{i,j} = 1$. Such cells in $\mathbf{G}^{\lambda}$ contribute $\sum_i \binom{m_i+1}{2}$. For $\lambda_i > \lambda_j$, we have $\mathbf{G}^{\lambda}_{i,j} = \lambda_i - \lambda_j$. Altogether, we have

$$\left|\mathbf{G}^{\lambda}\right| = \sum_{i>j} \lambda_i - \sum_{i>j} \lambda_j + \sum_i \binom{m_i + 1}{2}.$$

To compute $\sum_{i>j} \lambda_i$, we instead sum over all possible pairs (i, j), and subtract those for which $i \leq j$. Summing the term λ_i over all possible pairs (i, j) we obtain $n\ell$. Those terms where

$i = j$ contribute $\sum_i \lambda_i = n$. Thus, we have

$$\left|\mathbf{G}^{\lambda}\right| = n\ell - n - 2\sum_{i>j} \lambda_j + \sum_i \frac{m_i^2}{2} + \sum_i \frac{m_i}{2}.$$

From (1.1), we see that $\sum_{i>j} \lambda_j = \sum_j (j-1)\lambda_j = \mathbf{n}(\lambda)$. Furthermore, $\sum_i m_i = \ell$, hence

$$\left|\mathbf{G}^{\lambda}\right| = n\ell - n - 2\,\mathbf{n}(\lambda) + \sum_i \frac{m_i^2}{2} + \frac{\ell}{2}.$$

□

Theorem 8.21. *For any $A \in \mathcal{U}_n$ with Jordan form J_λ, we have*

$$X_A C_{\mathcal{U}}(A) X_A^{-1} \subseteq C(\mathbf{G}^{\lambda}).$$

Proof. We proceed by induction on n. In the base case, when $n = 1$, we see that A must be the 1×1 zero-matrix, and X_A is the 1×1 identity matrix. Thus, $X_A C_{\mathcal{U}}(A) X_A^{-1}$ consists of only the zero matrix, which is equal to $C(\mathbf{G}^{(1)})$.

For $n > 1$, let $A \in \mathcal{U}_n$ be GL-conjugate to J_λ. Let $B := A|_{n-1}$ be GL-conjugate to J_μ. The key to our inductive step is the equality

$$C_{\mathcal{U}}(A) = \overline{C_{\mathcal{U}}(B)} \cap C_{\mathcal{M}}(A). \tag{8.7}$$

To see this equality, first note that for two upper-triangular matrices X and Y to commute, it must also be that $X|_k$ and $Y|_k$ commute for every k. Hence, we have $C_{\mathcal{U}}(A) \subseteq \overline{C_{\mathcal{U}}(B)}$. It is immediate that $C_{\mathcal{U}}(A) \subseteq C_{\mathcal{M}}(A)$, proving that the left-hand side is contained in the right-hand side. In the other direction, observe that $\overline{C_{\mathcal{U}}(B)}$ consists only of upper-triangular matrices, and that $C_{\mathcal{M}}(A)$ consists only of matrices that commute with A.

We implicitly embed X_B into $\mathcal{M}_{n\times n}$ by identifying X_B with $X_B \oplus 1$. With this identification, we have

$$X_B \overline{V} X_B^{-1} \subseteq \overline{X_B V X_B^{-1}}, \text{ for any subspace } V \subseteq \mathcal{M}_{(n-1)\times(n-1)}.$$

Now consider conjugating both sides of (8.7) by X_B. We obtain

$$X_B C_{\mathcal{U}}(A) X_B^{-1} \;\subseteq\; \overline{X_B C_{\mathcal{U}}(B) X_B^{-1}} \;\cap\; (X_B C_{\mathcal{M}}(A) X_B^{-1}).$$

By inductive hypothesis, we may assume that $X_B C_{\mathcal{U}}(B) X_B^{-1} \subseteq C(\mathbf{G}^\mu)$. Thus,

$$X_B C_{\mathcal{U}}(A) X_B^{-1} \subseteq \overline{C(\mathbf{G}^\mu)} \cap C_{\mathcal{M}}(A'),$$

where $A' = X_B A X_B^{-1}$. Now A' satisfies the hypotheses of Lemma 8.15, so we may further conjugate by $Y_{A'}$ to obtain

$$Y_{A'} X_B C_{\mathcal{U}}(A) X_B^{-1} Y_{A'}^{-1} \subseteq Y_{A'} \left(\overline{C(\mathbf{G}^\mu)} \cap C_{\mathcal{M}}(A')\right) Y_{A'}^{-1} = C(\psi_r(\mathbf{G}^\mu)).$$

From Proposition 8.19, we know that $\mathbf{G}^\lambda \leq \psi_r(\mathbf{G}^\mu)$, from which it follows immediately that $C(\psi_r(\mathbf{G}^\mu)) \subseteq C(\mathbf{G}^\lambda)$ as desired. □

8.5.2 Bounds on $\mathbf{k}(U_n)$

We will bound $\mathbf{k}(U_n)$ by using a map $h : \ell^2(\mathbb{Z}^+) \to \mathbb{R}$. Specifically, define

$$h(v) := \|v\|^2 - \|v - Lv\|^2.$$

The following theorem highlights how we use this function.

Theorem 8.22. *For every $\lambda \vdash n$, we have*

$$|\mathrm{Comm}(\lambda)| \leq \sqrt{n!}\, q^{\frac{1}{2}(n^2 + h(\lambda'))}.$$

Proof. Observe that

$$|\mathrm{Comm}(\lambda)| = \sum_{A : \mathrm{sh}(A) = \lambda} |C_{\mathcal{U}}(A)|.$$

From Theorem 8.21 and Proposition 8.10, we know that for every A,

$$|C_{\mathcal{U}}(A)| \leq \left|C(\mathbf{G}^\lambda)\right| = q^{n\ell - \left|\mathbf{G}^\lambda\right|}. \tag{8.8}$$

Combining (8.8) with (2.1), we see that

$$\begin{aligned} |\mathrm{Comm}(\lambda)| &= \sum_{A : \mathrm{sh}(A) = \lambda} |C_{\mathcal{U}}(A)| \leq \sum_{A : \mathrm{sh}(A) = \lambda} q^{n\ell - \left|\mathbf{G}^\lambda\right|} \\ &= F^\lambda(q)\, q^{n\ell - \left|\mathbf{G}^\lambda\right|} \leq f^\lambda q^{\binom{n}{2} - \mathbf{n}(\lambda) + n\ell - \left|\mathbf{G}^\lambda\right|} \end{aligned}$$

It is a basic result from the representation theory of the symmetric group that $f^\lambda \le \sqrt{n!}$. It therefore suffices to show that $\binom{n}{2} - \mathbf{n}(\lambda) + n\ell - \left|\mathbf{G}^\lambda\right| \le \frac{1}{2}(n^2 + h(\lambda'))$, or equivalently, that

$$n\ell - \tfrac{n}{2} - \mathbf{n}(\lambda) - \left|\mathbf{G}^\lambda\right| \le \tfrac{1}{2}h(\lambda'). \tag{8.9}$$

Expanding $\left|G^\lambda\right|$ according to Proposition 8.20, and rewriting $\mathbf{n}(\lambda)$ and $m_i(\lambda)$ according to (1.2) and (1.3), we obtain

$$\begin{aligned} n\ell - \tfrac{n}{2} - \mathbf{n}(\lambda) - \left|\mathbf{G}^\lambda\right| &= \tfrac{n}{2} - \tfrac{\ell}{2} + \mathbf{n}(\lambda) - \sum_i \tfrac{m_i^2}{2} \\ &= -\tfrac{\ell}{2} + \tfrac{1}{2}\left(\|\lambda'\|^2 - \|\lambda' - L\lambda'\|^2\right) \\ &\le \tfrac{1}{2}h(\lambda'), \end{aligned}$$

which completes the proof. □

Lemma 8.23. *Let $D \subseteq \ell^1(\mathbb{Z}^+)$ denote the vector space of all $v \in \ell^1(\mathbb{Z}^+)$ satisfying the condition[8] that $v_1 \ge v_2 \ge \cdots \ge 0$. Then the following hold:*

1. *For every $c \in \mathbb{R}$, $h(cv) = c^2 h(v)$.*
2. *$\frac{dh}{dv_1} \ge \frac{dh}{dv_k}$ for all positive integers $k \ge 3$.*
3. *For all $v \in D$, we have $h(v) \le 2\|v\|_1 v_2 - 3v_2^2$.*
4. *For all $v \in D$ satisfying $v_1 \ge \frac{1}{2}\|v\|_1$, and for all $k \ge 4$, we have $\frac{dh}{dv_2} \ge \frac{dh}{dv_k}$.*
5. *For $v \in D$ with $v_1 \ge \frac{1}{2}\|v\|_1$, we have $h(v) \le \|v\|_1 v_1 - \frac{3}{4}v_1^2$.*

Proof. Part 1 follows from the fact that h is a homogenous of degree two on $\ell^2(\mathbb{Z}^+)$. For part 2, we compute $\frac{dh}{dv_i}$. For $i = 1$, we see that $\frac{dh}{dv_1} = 2v_2$. For $i > 1$,

$$\frac{dh}{dv_i} = 2v_{i-1} - 2v_i + 2v_{i+1}.$$

[8] *The condition that all v_i be non-negative is unnecessary. Because the terms are decreasing, if any term is strictly less than zero, the entire sequence would not be in $\ell^1(\mathbb{Z}^+)$. We include it as part of the condition for the sake of clarity.*

Thus, for $k \geq 3$, we have

$$\frac{dh}{dv_1} - \frac{dh}{dv_k} = (2v_2 - 2v_{k-1}) + (2v_k - 2v_{k+1}).$$

This is positive, because v is a weakly decreasing sequence.

For part 3, define $w = w(v) = (\|v\|_1 - v_2, v_2, 0, \dots) \in D$. Let

$$D_{\leq k} = \{v \in D \mid v_{k+1} = 0\}.$$

Furthermore, define $e_i \in \ell^1(\mathbb{Z}^+)$ to be 1 in the ith position, and zero elsewhere. First suppose that $v \in D_{\leq k}$ for some $k \geq 3$. From part 2, we know that

$$h(v + v_k(e_1 - e_k)) \geq h(v).$$

Moreover, $v + v_k(e_1 - e_k) \in D_{\leq k-1}$. Iterating this process, we see that if $v \in D_{\leq k}$ for some $k \geq 3$, then we have

$$h(v) \leq h(\|v\|_1 - v_2, v_2, 0, \dots) = 2\|v\|_1 v_2 - 3v_2^2.$$

Because $\bigcup_{k=2}^{\infty} D_{\leq k}$ is dense in D, and both the left- and right-hand sides are continuous, we see that $v \leq 2\|v\|_1 v_2 - 3v_2^2$ for every $v \in D$.

For part 4, note that

$$\frac{dh}{dv_2} - \frac{dh}{dv_k} = 2[v_1 - (v_2 - v_3) - (v_{k-1} - v_k) - v_{k+1}].$$

If $k \geq 4$, the quantity subtracted from v_1 is less than $\|Lv\|_1 = \|v\|_1 - v_1$. Because $v_1 > \frac{1}{2}\|v\|_1$, the quantity $\frac{dh}{dv_2} - \frac{dh}{dv_k}$ must be non-negative.

Lastly, for part 5, we again consider $D_{\leq k} = \{v \in D \mid v_{k+1} = 0\}$. For $v \in D_{\leq k}$, note that part 4 implies $h(v + v_k(e_2 - e_k)) \geq h(v)$. Because $v_1 \geq \frac{1}{2}\|v\|_1$, we see $v + v_k(e_2 - e_k) \in D_{\leq k-1}$. Iterating this process, we eventually reach

$$w = w(v) := (v_1, \|v\|_1 - v_1 - v_3, v_3, 0, \dots).$$

Thus, if $v \in D_{\leq k}$ for some $k \geq 4$, we have

$$h(v) \leq h(w) = -\|v\|_1^2 + 4\|v\|_1 v_1 - 3v_1^2 + 4\|v\|_1 v_3 - 6v_1 v_3 - 4v_3^2.$$

As both the left- and right-hand sides are continuous, and $\bigcup_{k=3}^{\infty} D_{\leq k}$ is dense in D, this bound holds for all $v \in D$. From here it suffices to show that $\|v\|_1 v_1 - \frac{3}{4}v_1^2 - h(w) \geq 0$ for all v. To this end, note that

$$\|v\|_1 v_1 - \tfrac{3}{4}v_1^2 - h(w) = \tfrac{1}{4}\left(2\|v\|_1 - 3v_1 - 4v_3\right)^2 \geq 0.$$

Hence, $h(v) \leq h(w) \leq \|v\|_1 v_1 - \frac{3}{4}v_1^2$ as desired. □

Corollary 8.24. *The number of conjugacy classes in the group of upper-triangular matrices is bounded by*

$$\mathbf{k}(U_n) \leq p(n)\sqrt{n!}\; q^{\frac{n^2}{6}+\frac{n}{2}},$$

where $p(n)$ is the number of integer partitions of n.

Proof. Because $|\mathrm{Comm}(\mathcal{U}_n)| = q^{\binom{n}{2}}\mathbf{k}(U_n)$, it suffices to show that

$$|\mathrm{Comm}(\mathcal{U}_n)| \leq p(n)\sqrt{n!}\; q^{\frac{2n^2}{3}}.$$

To this end, we stratify $\mathrm{Comm}(\mathcal{U}_n)$ by the Jordan type of the first matrix in each pair $(A, B) \in \mathrm{Comm}(\mathcal{U}_n)$. From Theorem 8.22, we have

$$|\mathrm{Comm}(\mathcal{U}_n)| = \sum_{\lambda \vdash n} |\mathrm{Comm}(\lambda)| \leq \sum_{\lambda \vdash n} \sqrt{n!} q^{\frac{1}{2}(n^2 + h(\lambda'))}.$$

The sum is over $p(n)$ terms, so it suffices to show that each term is bounded $\sqrt{n!}q^{\frac{2}{3}n^2}$. In other words, it suffices to show that for every $\lambda \vdash n$, we have

$$h(\lambda') \leq \frac{n^2}{3}.$$

Let $D = \{v \in \ell^1(\mathbb{Z}^+) \mid v_1 \geq v_2 \geq \cdots \geq 0\}$. Let $c \in \mathbb{R}$ such that $v_2 = c\|v\|_1$. Then from part 3 of Lemma 8.23, we have $h(v) \leq (2c - 3c^2)\|v\|_1^2$. This quantity is maximized at a value of $\frac{1}{3}\|v\|_1^2$, by taking $c = \frac{1}{3}$. For any $\lambda \vdash n$, we therefore have $h(\lambda') \leq \frac{n^2}{3}$, as desired. Thus, our upper bound for the number of pairs of commuting upper-triangular matrices is

$$|\mathrm{Comm}(\mathcal{U}_n)| \leq p(n)\sqrt{n!}\; q^{\frac{2n^2}{3}},$$

which implies the desired result. □

We are now ready to prove Theorem 1.7, which we restate here for the reader's convenience.

Theorem 1.7. *For every positive integer n and every prime power q, we have*

$$\mathbf{k}(U_n(q)) \leq p(n)^2 n!\ q^{\alpha n^2 + \frac{n}{2}},$$

where $p(n)$ denotes the number of integer partitions of n, and where

$$\alpha = \frac{40\sqrt{2} - 41}{98} \approx 0.15886.$$

Proof. Because $|\mathrm{Comm}(\mathcal{U}_n)| = \sum_{\lambda,\mu \vdash n} |\mathrm{Comm}(\lambda, \mu)|$, it suffices to show that

$$|\mathrm{Comm}(\lambda, \mu)| \leq \sqrt{n!} q^{cn^2 + \frac{n}{2}}.$$

Let δ and ε denote small positive quantities, each less than $\frac{1}{6}$, to be optimized later. We bound $\mathrm{Comm}(\lambda, \mu)$ in one of three ways, depending on the shapes of λ and μ. Let $v \in \ell^1(\mathbb{Z}^+)$ be defined by $v := \lambda'/n$, so that $\|v\|_1 = 1$.

First consider the case where $\left|v_2 - \frac{1}{3}\right| \geq \delta$. From Lemma 8.23, we use the bound $h(v) \leq 2v_2 - 3v_2^2$. As a function of v_2, the right-hand side attains its maximum at $v_2 = \frac{1}{3}$. Thus, for v_2 satisfying $\left|v_2 - \frac{1}{3}\right| \geq \delta$, the right-hand side is maximized at $v_2 = \frac{1}{3} \pm \delta$ with a value of $\frac{1}{3} - 3\delta^2$. Thus,

$$|\mathrm{Comm}(\lambda, \mu)| \leq \sqrt{n!}\ q^{\frac{n^2}{2}(1+h(v))} \leq \sqrt{n!}\ q^{\frac{n^2}{2}(\frac{4}{3} - 3\delta^2)}.$$

Because $|\mathrm{Comm}(\lambda, \mu)| = |\mathrm{Comm}(\mu, \lambda)|$, we also have this bound if $w = \mu'/n$ satisfies $\left|w_2 - \frac{1}{3}\right| \geq \delta$. Henceforth we may assume that both v and w have their second entry in the open interval $(\frac{1}{3} - \delta, \frac{1}{3} + \delta)$.

Second, if $v_1 > \frac{2}{3} - \varepsilon$, then from Lemma 8.23, we use the bound $h(v) \leq v_1 - \frac{3}{4}v_1^2$. Because $\varepsilon \leq \frac{1}{6}$, we know that $v_1 > \frac{1}{2}$, and so v satisfies the hypotheses of part 5 of Lemma 8.23. The polynomial $x - \frac{3}{4}x^2$ is maximized at $\frac{2}{3}$, and so for $v_1 < \frac{2}{3} - \varepsilon$, we see that $h(v) \leq \frac{1}{3} - \frac{3}{4}\varepsilon^2$. Thus,

$$|\mathrm{Comm}(\lambda, \mu)| \leq f^\lambda q^{\frac{n^2}{2}(1+h(v))} \leq \sqrt{n!}\ q^{\frac{n^2}{2}(\frac{4}{3} - \frac{3}{4}\varepsilon^2)}.$$

By symmetry, this bound also applies if $\mu_1' \leq (\frac{2}{3} - \varepsilon)n$, so we may assume that both λ_1'/n and μ_1'/n are at least $\frac{2}{3} - \varepsilon$.

Lastly, consider the case where

$$\frac{\lambda'_1}{n} \geq \frac{2}{3} - \varepsilon, \text{ and } \left|\frac{\lambda'_2}{n} - \frac{1}{3}\right| \leq \delta,$$

and similarly for μ. We obtain an upper bound in the last case by disregarding commutativity, and using the bound

$$|\mathrm{Comm}(\lambda, \mu)| \leq F^\lambda F_\mu = f^\lambda f^\mu q^{n^2 - n - \mathbf{n}(\lambda) - \mathbf{n}(\mu)}.$$

To maximize the exponent, we take $\lambda' = \mu' = \left((\frac{2}{3} - \varepsilon)n, (\frac{1}{3} - \delta)n, 1, 1, \ldots, 1\right)$. The exponent $n^2 - n - \mathbf{n}(\lambda) - \mathbf{n}(\mu)$ is therefore bounded above by

$$n^2 \left[1 - (\tfrac{2}{3} - \varepsilon)^2 - (\tfrac{1}{3} - \delta)^2\right].$$

Thus, our three bounds on the exponent are given by

$$n^2 \left[\tfrac{2}{3} - \tfrac{3}{2}\delta^2\right], \ n^2 \left[\tfrac{2}{3} - \tfrac{3}{8}\varepsilon^2\right] \text{ and, } n^2 \left[1 - (\tfrac{2}{3} - \varepsilon)^2 - (\tfrac{1}{3} - \delta)^2\right].$$

Taking $\varepsilon = 2\delta = \frac{4}{21}(5 - 3\sqrt{2}) \approx 0.1443$, it is easy to verify that all three of these quantities are equal to βn^2, where

$$\beta := \frac{4 + 20\sqrt{2}}{49} \approx 0.65886.$$

Thus, we obtain the bound $\mathrm{Comm}(\mathcal{U}_n) \leq \sum_{\lambda,\mu} n! q^{\beta n^2}$, proving that

$$\mathbf{k}(U_n) \leq p(n)^2 \ n! \ q^{(\alpha - \frac{1}{2})n^2 + \frac{n}{2}} = p(n)^2 \ n! \ q^{\alpha n^2 + \frac{n}{2}}.$$

□

CHAPTER 9

Asymptotic behavior of the lower central series of $U_n(q)$

9.1 Overview

This chapter contains a proof of an upper bound on the number of conjugacy classes in groups in the lower central series for U_n. The results here are somewhat technical, and have no reliance on the combinatorial structures of poset systems or gap arrays defined in previous chapters. We begin with an overview of the tools we use in this chapter.

Recall that for $A \in \mathcal{U}_n$, we define $A|_k \in \mathcal{U}_k$ to be the $k \times k$ submatrix in the top-left corner of A. One key feature of upper-triangular matrices is that in order for $A, B \in \mathcal{U}_n$ to commute, it must be that $A|_k$ and $B|_k$ commute. Moreover, the same is true if you restrict to the submatrices in the bottom-right corners. This provides the possibility for an inductive technique to create bounds on $\mathbf{k}(U_n)$. Specifically, one could dissect each matrix in $\mathcal{U}_n$ into three pieces: Two of these pieces would be smaller upper-triangular, and one would be some rectangular matrix (see Figure 9.1 for an example decomposition).

If one could analyze the conditions which yield commutativity, a bound could be built up inductively based on the conditions on the rectangular piece as well as the two smaller instances of upper triangular matrices. In this chapter, we develop such an inductive tool, not only for U_n, but also for all subgroups in the lower-central series. Concretely, the kth term in the lower central series for U_n is denoted $U_{n,k}$ and is the subgroup consisting of matrices for which the first k diagonals above the main diagonal are be zero. As always, we write the corresponding nilpotent algebra as $\mathcal{U}_{n,k} = \{X \in \mathcal{U}_n : 1 + X \in U_{n,k}\}$.

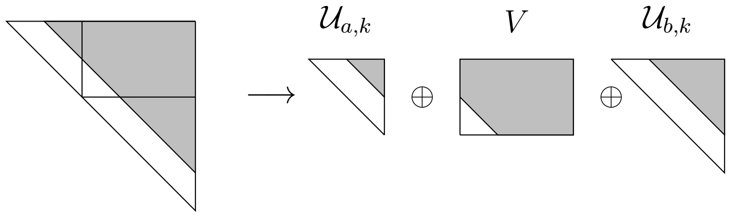


Figure 9.1: A graphical representation of the decomposition of $\mathcal{U}_{a+b,k}$ into $\mathcal{U}_{a,k} \oplus V \oplus \mathcal{U}_{b,k}$. The gray regions represent the cells which can be non-zero. The white regions represent the cells which must contain zeros.

Theorem 1.8. *For each $m \in \mathbb{N}$, define*

$$\gamma_m := \frac{1}{6} - \frac{13}{24} \cdot 4^{-m} + 2^{-(m+1)} - 4^{-(m+1)}m.$$

Then for every q,

$$\mathbf{k}(U_{n,k}(q)) \leq q^{\gamma_m n^2(1+o_m(1))},$$

where $m = \left\lfloor \log_2 \left(\frac{n}{k}\right) \right\rfloor$, and $o_m(1)$ denotes a function which, for each fixed m, tends to zero as n tends to infinity.

Note that as $m \to \infty$, The constant γ_m tends to $\frac{1}{6}$. Large values of m correspond to fractions k/n which are to 1, recovering the fact that $\mathbf{k}(U_n) \leq q^{\frac{n^2}{6}(1+o(1))}$. Of course, one must be careful about the details of the limit, but making this precise is straightforward. This yields a fourth proof of the one-sixth upper bound, the first three being given in [VA1], [Mar1], and Corollary 8.24.

9.2 Notation

For a nonnegative integer $k \leq n-1$, let $U_{n,k}(q)$ denote the kth term in the lower central series of $U_n(q)$. Explicitly,

$$U_{n,k}(q) = \{A \in U_n(q) : A_{i,j} = 0 \text{ whenever } 0 < i - j \leq k\}.$$

By $\mathcal{U}_{n,k}(q)$, we mean the $\mathbb{F}_q$-subalgebra of $\mathcal{U}_n(q)$ defined by

$$\mathcal{U}_{n,k}(q) := \{X : 1 + X \in U_{n,k}(q)\}.$$

Note that $\mathcal{U}_n(q) = \mathcal{U}_{n,0}(q)$. As always, we omit the parameter q when the field is clear from context.

For a finite group G, we define the *commuting probability* $\text{cp}(G)$ to be the probability that two elements chosen uniformly at random from G commute. Clearly

$$\text{cp}(G) = \frac{|\text{Comm}(G)|}{|G|^2} = \frac{\mathbf{k}(G)}{|G|}.$$

We will also use the shorthand $\text{cp}(n,k) = \text{cp}(U_{n,k}(q))$.

Lastly, for $A \in \mathcal{U}_a$ and $B \in \mathcal{U}_b$, let $T_{A,B} : \mathcal{M}_{a\times b} \to \mathcal{M}_{a\times b}$ by $T_{A,B}(X) := AX - XB$. Recall the map $T_{\lambda,\mu}$ defined in Lemma 8.2, as $T_{\lambda,\mu}(X) = J_\lambda X - Y J_\mu$. In this way, $T_{\lambda,\mu}$ was shorthand for T_{J_λ,J_μ}. We will use both of these notations in this Chapter.

9.3 Key lemmas

The following lemma is the key tool used in the proof of Theorem 1.8.

Lemma 9.1. *Let a and b be positive integers, and let k be an integer such that $0 \le k < a+b$. Then*

$$\text{cp}(a+b,k) = \frac{1}{|\mathcal{U}_{a,k}|^2 \cdot |\mathcal{U}_{b,k}|^2} \sum q^{-\operatorname{rank}(T_{A_1,B_1},T_{A_2,B_2})}$$

where $\operatorname{rank}(X,Y) := \dim(\operatorname{Im} X + \operatorname{Im} Y)$, *and where the sum is over all* $(A_1,A_2) \in \text{Comm}(\mathcal{U}_{a,k})$ *and all* $(B_1,B_2) \in \text{Comm}(\mathcal{U}_{b,k})$.

Proof. Let $V = \{X \in \mathcal{M}_{a\times b} : X_{i,j} = 0 \text{ whenever } j - i \ge a-k\}$. We begin with the vector space isomorphism $\mathcal{U}_{a+b,k} \to \mathcal{U}_{a,k} \oplus V \oplus \mathcal{U}_{b,k}$ given by

$$e_{i,j} \mapsto \begin{cases} (e_{i,j},0,0) & \text{if } i,j \le a \\ (0,e_{i,j-a},0) & \text{if } i \le a, j > a \\ (0,0,e_{i-a,j-a}) & \text{if } i,j > a. \end{cases}$$

Graphically, the isomorphism is shown in Figure 9.1.

The multiplicative structure on $\mathcal{U}_{a+b,k}$ can be pushed through this isomorphism to make $\mathcal{U}_{a,k} \oplus V \oplus \mathcal{U}_{b,k}$ an $\mathbb{F}_q$-algebra, and is given by

$$(A_1, X_1, B_1) \cdot (A_2, X_2, B_2) = (A_1 B_1,\ A_1 X_2 + X_1 B_2,\ A_2 B_2).$$

We count pairs of commuting elements in $\mathcal{U}_{a,k} \oplus V \oplus \mathcal{U}_{b,k}$. Note that (A_1, X_1, B_1) and (A_2, X_2, B_2) commute if and only all three of the following conditions hold:

1. A_1 and A_2 commute in $\mathcal{U}_{a,k}$,
2. B_1 and B_2 commute in $\mathcal{U}_{b,k}$,
3. $T_{A_1,B_1}(X_2) = T_{A_2,B_2}(X_1)$.

For $i = 1, 2$, let $T_i = T_{A_i,B_i}$. We may count pairs of commuting elements in $\mathcal{U}_{a+b,k}$ by counting how many $(X_1, X_2) \in V \oplus V$ satisfy $T_1(X_2) = T_2(X_1)$, and summing over all $(A_1, A_2) \in \mathrm{Comm}(\mathcal{U}_{a,k})$ and all $(B_1, B_2) \in \mathrm{Comm}(\mathcal{U}_{b,k})$. Therefore,

$$\mathrm{cp}(a + b, k) = \frac{1}{|\mathcal{U}_{a+b,k}|^2} \sum \#\{(X_1, X_2) \in V \times V : T_1(X_2) = T_2(X_1)\},$$

where the sum is over

$$S := \{(A_1, A_2, B_1, B_2) : (A_1, A_2) \in \mathrm{Comm}(\mathcal{U}_{a,k}), (B_1, B_2) \in \mathrm{Comm}(\mathcal{U}_{b,k})\}.$$

It is important to remember that T_i is shorthand for T_{A_i,B_i}, so T_i has an implicit dependence on the summand. Note that $\{(X_1, X_2) : T_1(X_2) = T_2(X_1)\}$ is the kernel of the map $\Phi : V \oplus V \to V$ defined by

$$\Phi(X, Y) := T_1(X) - T_2(Y).$$

Obviously, $\operatorname{Im}\Phi = \operatorname{Im} T_1 + \operatorname{Im} T_2$. By the rank-nullity theorem, we have

$$\mathrm{cp}(a + b, k) = \left(\frac{q^{\dim V}}{|\mathcal{U}_{a+b,k}|}\right)^2 \sum_S q^{-\dim(\operatorname{Im} T_1 + \operatorname{Im} T_2)}. \tag{9.1}$$

Recalling that $\dim \mathcal{U}_{a+b,k} = \dim \mathcal{U}_{a,k} + \dim V + \dim \mathcal{U}_{b,k}$, we see that

$$\frac{q^{\dim V}}{|\mathcal{U}_{a+b,k}|} = \frac{1}{|\mathcal{U}_{a,k}|\, |\mathcal{U}_{b,k}|}. \tag{9.2}$$

Putting together equations (9.1) and (9.2), we obtain the desired result. □

Lemma 9.2. *Let* $N_r(a,b) := \{(A,B) \in \mathcal{U}_a \times \mathcal{U}_b : \operatorname{rank} T_{A,B} \le r\}$. *Then*

$$|N_{a,b}(r)| \le p(a)p(b)\sqrt{a!b!}\; q^{\frac{(a-b)^2}{2}+r}.$$

Proof. Because $\dim\ker T_{A,B} = \dim\ker T_{C,D}$ whenever A is GL_a-conjugate to C and B is GL_b-conjugate to D, we may safely assume that A and B are in Jordan canonical form. From Lemma 8.2, we know $\dim\ker T_{\lambda,\mu} = \langle \lambda', \mu' \rangle$. Applying the Cauchy-Schwarz inequality followed by the classical arithmetic-geometric mean inequality, we obtain

$$\dim\ker T_{\lambda,\mu} = \langle \lambda', \mu' \rangle \le \|\lambda'\| \cdot \|\mu'\| \le \frac{\|\lambda'\|^2 + \|\mu'\|^2}{2} = \frac{a+b}{2} + \mathbf{n}(\lambda) + \mathbf{n}(\mu). \tag{9.3}$$

We stratify $N_{a,b}(r)$ by the Jordan forms of the pairs (A,B) which appear in $N_{a,b}(r)$. Let $S := \{(\lambda,\mu) : \langle \lambda', \mu' \rangle \ge ab - r\}$. The pairs of partitions in S are precisely the partitions indexing Jordan forms of pairs of matrices in $N_{a,b}(r)$. Thus,

$$|N_{a,b}(r)| = \sum_{(\lambda,\mu)\in S} F^\lambda(q) F^\mu(q).$$

Recall from (2.1) that $F^\lambda(q) \le f^\lambda q^{\binom{a}{2} - \mathbf{n}(\lambda)}$, so

$$|N_{a,b}(r)| \le \sum_{(\lambda,\mu)\in S} f^\lambda f^\mu q^{\frac{a^2+b^2}{2} - \left(\frac{a+b}{2} + \mathbf{n}(\lambda) + \mathbf{n}(\mu)\right)},$$

where $F^\lambda(q)$ counts the number of matrices in $\mathcal{U}_n(q)$ which are conjugate to J_λ (defined in Section 2.3). Now applying (9.3) we get

$$|N_{a,b}(r)| \le \sum_{(\lambda,\mu)\in S} f^\lambda f^\mu q^{\frac{a^2+b^2}{2} - (ab-r)} = q^{(a-b)^2/2 + r} \sum_{(\lambda,\mu)\in S} f^\lambda f^\mu.$$

Noting from the representation theory of $\mathfrak{S}_n$ (see, e.g. [Sa]) that $f^\lambda \le \sqrt{a!}$ and that $|S| \le p(a)p(b)$, the result follows. $\square$

We now present the proof of Theorem 1.8. Define $\beta_0 := 0$, and for each positive integer m, define β_m inductively by

$$\beta_m := \tfrac{1}{4}\left(\beta_{m-1} - (1 - 2^{-m})^2\right).$$

A routine inductive argument shows that for all m, we have

$$\beta_m = -\frac{1}{3} - \frac{2}{3} \cdot 4^{-m} + 2^{-m} - 4^{-(m+1)}m.$$

Theorem 9.3. *Let m be a non-negative integer such that $2^{-(1+m)} \le \frac{k}{n} \le 2^{-m}$. Then*

$$\mathrm{cp}(n,k) \le q^{\beta_m n^2(1+o_m(1))},$$

where $o_m(1)$ is a function of m and n which, for each fixed m, tends to zero as n tends to infinity.

Proof. We proceed with the proof that $\mathrm{cp}(n,k) \le q^{\beta_m n^2(1+o_m(1))}$ by induction on m. When $m = 0$, we have $\frac{1}{2} \le \frac{k}{n} \le 1$, and so $U_{n,k}$ is abelian, proving that $\mathrm{cp}(n,k) = 1 = q^0$ as desired.

Now let $V = \{X \in \mathcal{M}_{a\times b} : X_{i,j} = 0$ whenever $j - i \ge a - k\}$, and stratify $\mathrm{Comm}(\mathcal{U}_{a,k}) \times \mathrm{Comm}(\mathcal{U}_{b,k})$ by the the value of $\mathrm{rank}(T_{A_1,B_1}, T_{A_2,B_2})$. Specifically

$$C^k_{a,b}(r) := \{(A_1, A_2, B_1, B_2) \in \mathrm{Comm}(\mathcal{U}_{a,k}) \times \mathrm{Comm}(\mathcal{U}_{b,k}) : \\ \mathrm{rank}_V(T_{A_1,B_1}, T_{A_2,B_2}) = r\}.$$

From Lemma 9.1, we have

$$\mathrm{cp}(n,k) = \frac{1}{|\mathcal{U}_{a,k}|^2\,|\mathcal{U}_{b,k}|^2} \sum_{j=0}^{ab-\binom{k+1}{2}} q^{-j} \left|C^k_{a,b}(j)\right|.$$

Pick some number r to be optimized later, and split the sum based on the relationship between r and j. First considering the case when $j \ge r$ (and thus $q^{-j} \le q^{-r}$), we have

$$\sum_{j\ge r} q^{-j}\left|C^k_{a,b}(j)\right| \le q^{-r}\sum_{j\ge r}\left|C^k_{a,b}(j)\right| \le q^{-r}\left|\mathrm{Comm}(\mathcal{U}_{a,k})\right|\left|\mathrm{Comm}(\mathcal{U}_{b,k})\right|. \tag{9.4}$$

On the other hand, if $j < r$, we forget about the commutativity relation, and remember only that the rank of each map T_{A_i,B_i} must be bounded by r. Thus, for $j < r$ we have

$C^j_{a,b}(k) \le |N_{a,b}(j)|^2$. Combining this with (9.4) and Lemma 9.2, we obtain

$$\begin{aligned}
\mathrm{cp}(n,k) &\le q^{-r}\,\mathrm{cp}(a,k)\,\mathrm{cp}(b,k) + \frac{1}{|\mathcal{U}_{a,k}|^2\,|\mathcal{U}_{b,k}|^2}\sum_{j=0}^{r-1} q^{-r}\,|N_{a,b}(j)|^2 \\
&\le q^{-r}\,\mathrm{cp}(a,k)\,\mathrm{cp}(b,k) + \frac{p(a)^2p(b)^2a!b!}{|\mathcal{U}_{a,k}|^2\,|\mathcal{U}_{b,k}|^2}\sum_{j=0}^{r-1} q^{(a-b)^2+2j-r} \\
&\le q^{-r}\,\mathrm{cp}(a,k)\,\mathrm{cp}(b,k) + \frac{p(a)^2p(b)^2a!b!}{|\mathcal{U}_{a,k}|^2\,|\mathcal{U}_{b,k}|^2} q^{(a-b)^2+r-1}.
\end{aligned}$$

To optimize these parameters, we take $a = \lfloor n/2 \rfloor$ and $b = \lceil n/2 \rceil$, and

$$r = \left[\frac{1}{4}\left(\beta_{m-1} + \left(1-2^{-m}\right)^2\right)n^2\right], \tag{9.5}$$

where the square brackets denote the nearest integer function. In the calculations to follow, we omit the $[\cdot]$, as it complicates the computation, and does not contribute to the leading term in the exponent. The difference is absorbed into the $o_m(1)$ term.

Because $2^{-m} \le \frac{2k}{n} \le 2^{1-m}$, by inductive hypothesis, we have

$$\mathrm{cp}(a,k) \le q^{\beta_{m-1}a^2(1+o_m(1))} = q^{\frac{\beta_{m-1}n^2}{4}(1+o_m(1))}.$$

Note that $p(a)^2p(b)^2a!b! \le q^{o(n^2)}$, and that

$$|\mathcal{U}_{a,k}| = q^{\frac{(a-k)^2}{2}(1+o_m(1))} = q^{\frac{1}{8}\left(1-\frac{2k}{n}\right)^2 n^2(1+o_m(1))}.$$

Combining these facts, we obtain

$$\mathrm{cp}(n,k) \le q^{-r+\frac{\beta_{m-1}n^2}{2}(1+o_m(1))} + q^{r-\frac{1}{2}(1-2^{-m})^2n^2(1+o_m(1))}.$$

Substitiuting in (9.5), both exponents on the right-hand side become

$$\tfrac{1}{4}\left(\beta_{m-1} - (1-2^{-m})^2\right)n^2(1+o_m(1)),$$

which is equal to $\beta_m n^2(1+o_m(1))$ by the recursive definition of β_m. Thus,

$$\mathrm{cp}(n,k) \le 2q^{\beta_m n^2(1+o_m(1))} = q^{\beta_m n^2(1+o_m(1))},$$

which completes the proof. $\square$

Finally, Theorem 1.8 follows quickly as a corollary to Theorem 9.3.

Theorem 1.8. *For each* $m \in \mathbb{N}$*, define*

$$\gamma_m := \frac{1}{6} - \frac{13}{24} \cdot 4^{-m} + 2^{-(m+1)} - 4^{-(m+1)}m.$$

Then for every q*,*

$$\mathbf{k}(U_{n,k}(q)) \leq q^{\gamma_m n^2(1+o_m(1))},$$

where $m = \left\lfloor \log_2 \left(\frac{n}{k}\right) \right\rfloor$*, and* $o_m(1)$ *denotes a function which, for each fixed* m*, tends to zero as* n *tends to infinity.*

Proof. Recalling that $\mathbf{k}(G) = \mathrm{cp}(G)\,|G|$, we see that Theroem 9.3 implies that

$$\mathbf{k}(U_{n,k}) \leq q^{\binom{n-k}{2}+\beta_m n^2(1+o_m(1))}.$$

Combining this with the fact that $\binom{n-k}{2} = \frac{1}{2}\left(1 - \frac{k}{n}\right)^2 n^2(1 + o_m(1))$, and also the fact that $\frac{k}{n} \geq 2^{-(1+m)}$, we obtain

$$\mathbf{k}(G) \leq q^{\left(\frac{1}{2}\left(1-2^{-(1+m)}\right)^2+\beta_m\right)n^2(1+o_m(1))} \leq q^{\gamma_m n^2(1+o_m(1))},$$

where $\gamma_m = \frac{1}{6} - \frac{13}{24} \cdot 4^{-m} + 2^{-(m+1)} - 4^{-(m+1)}m.$ ∎

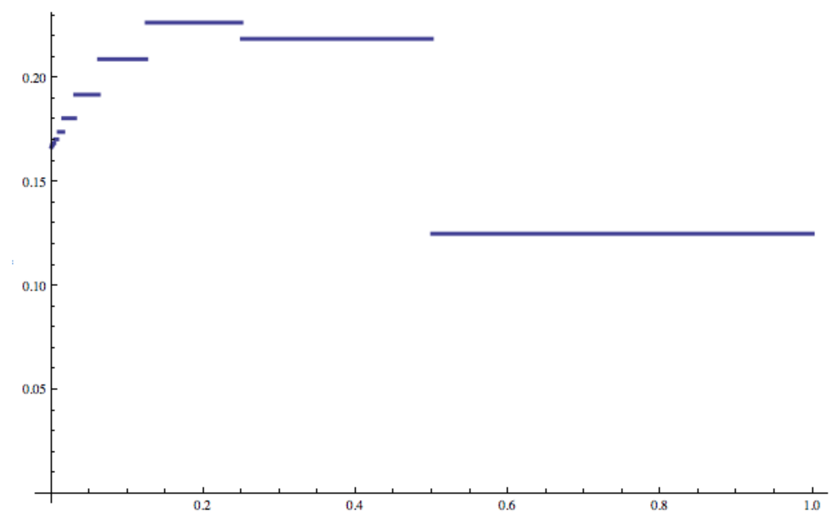


Figure 9.2: A graph of the value in the exponent of the upper-bound given in Theorem 1.8. The x-axis is the value $x = k/n$.

CHAPTER 10

Concluding remarks

10.1 Connection with the orbit method

Our approach is motivated by the philosophy of Kirillov's orbit method (see [K1]). In the case of $U_n(\mathbb{R})$, the orbit method provides a correspondence between the irreducible unitary representations of $U_n(\mathbb{R})$ and the co-adjoint orbits. Moreover, the co-adjoint orbits enjoy the structure of a symplectic manifold. The unitary characters can actually be recovered by integrating a particular form against the corresponding orbit.

Over finite fields, a manifold structure is not possible, but some of the philosophy of the orbit method seems to still be relevant and some formulas translate without difficulty. For example, the number of conjugacy classes (and therefore irreducible representations), is equal to the number of co-adjoint orbits (Lemma 4.1). However, the naturally analogous character formula does not hold [IK].

10.2 Isaacs work with characters

In [Is], Isaacs introduced pattern groups and explained that one can count characters in $U_n(q)$ by counting characters in stabilizers of a certain group action (see also [DT]). These stabilizers are themselves pattern groups, and lend themselves to a similar recursion, but over characters, rather than co-adjoint orbits. There is more than superficial difference between the recursion in [Is] and in this paper. In fact, it follows from [IK], that the characters cannot correspond to co-adjoint orbits via the natural analogue of Kirillov's orbit method.

10.3 Computational history of Higman's conjecture

Higman originally stated Conjecture 1.3 in the form of an open problem [H1]; it received the name "Higman's Conjecture" more recently. Higman originally checked that the conjecture holds for $n \leq 5$. The calculation of the number of conjugacy classes was later extended to $n \leq 8$ by Gudivok et al. in [G$^+$]. [9] The authors use a variation on the brute force algorithm.

Later, Arregi and Vera-López verified Higman's conjecture for $n \leq 13$ in [VA4] by a clever application of a brute force algorithm for counting adjoint orbits. They also proved that the number of conjugacy classes of cardinality q^s is polynomial for $s \leq n-3$ [VA3]. Moreover, they verified that, as a polynomial in $(q-1)$, the number of conjugacy classes of cardinality q^s has non-negative integral coefficients (for $s \leq n-3$). For other partial results Higman's conjecture see also [ABT, Is, Mar1].

10.4 Asymptotics of $\mathbf{k}(U_n(q))$

In recent years, much effort has been made to improve Higman's upper bounds for the asymptotics of $\mathbf{k}(U_n(q))$, as $n \to \infty$, see [Mar1, VA1]. For a fixed q, it is conjectured that

$$\mathbf{k}(U_n(q)) = q^{\frac{n^2}{12}(1+o(1))} \text{ as } n \to \infty. \tag{10.1}$$

The lower bound is known and due to Higman in the original paper [H1], while the best upper bound to date is presented in Chapter 8 of this dissertation.

The above asymptotics have curious connection to the enumerative and computational work in Chapter 6 Arregi and Vera-López conjectured in [VA4] a refinement of Higman's Conjecture 1.3 stating that not only is $\mathbf{k}(U_n(q))$ a polynomial, but that the degree of the polynomials $\mathbf{k}(U_n)$ are equal to $\lfloor n(n+6)/12 \rfloor$. If true, this would confirm the asymptotics (10.1) as well. While we do not believe Higman's conjecture, the degree formula continues to hold for new values, so it is now known for all $n \leq 16$.

[9] Their paper claims the computation for $n \leq 9$, however they madea minor mistake in the computation of $U_9(q)$. The polynomial they computed was however still correct when evaluated at 2.

10.5 Posets with non-polynomial behavior

In [HP], Halasi and Pálfy exhibit a pattern group for which the number of conjugacy classes is not a polynomial in the size of the field. Though they do not provide explicit bounds, their construction yields a 5,592,412-element poset. We obtained the 13-element poset $\mathfrak{P}$ shown in Figure 5.1 by modifying their construction.

It would be interesting to see if the poset $\mathfrak{P}$ is in fact the smallest poset for which $\mathbf{k}(\mathfrak{P}) \notin \mathbb{Z}[q]$. By Theorem 3.1, such posets must have at least 10 elements. Unfortunately, even this computation might be difficult since the total number of connected posets is rather large. For example, there are roughly $1.06 \cdot 10^9$ connected posets on 12 elements, see e.g. [BM] and [OEIS, A000608].

10.6 Computation time

When our algorithm falls back on the VA-algorithm, the poset systems it must compute have minimal shared computational resources. For this reason, our technique lends itself well towards parallelization. This, along with several optimization techniques we believe could be used to compute $\mathbf{k}(U_{17}(q))$ and $\mathbf{k}(U_{18}(q))$. However, due to the super-exponential growth rate of $\mathbf{k}(U_n(q))$, pushing the computation significantly further will likely require different techniques.

Based on our computations, one can try to give a conservative lower bound to the cost of computing $\mathbf{k}(U_{59}(q))$. Assuming the current rate of increase in timing, we estimate our algorithm to need about 10^{66} years of CPU time. Alternatively, if we assume Moore's law[10] will continue to hold indefinitely, this computation will not become feasible until the year 2343.

[10] Moore's law is the observation that the number of transistors per square inch on an integrated circuit has been doubling roughly every 18 months. Quite roughly, this can be interpreted as computer performance increase.

10.7 Embedding non-polynomial posets into chains

There are two directions in which the bound $n \geq 59$ in Conjecture 3.3 can be decreased. First, it is perhaps possible that $\mathfrak{P}$ embeds into a smaller chain. This is a purely combinatorial problem which perhaps can be attacked computationally. We would be interested to see if such improvement is possible.

Second, it is conceivable and perhaps likely that there are posets P with more than 13 elements which embed into $\mathbf{C}_n$ with $n < 59$, and have non-polynomial $\mathbf{k}(U_P)$. Since $\mathfrak{P}$ really encodes the variety $x^2 = 1$, it would be natural to consider other algebraic varieties which have different point counts depending on the characteristic. This is a large project which goes beyond the scope of this work.

10.8 Murphy's law and universality

Recall that by the Halasi–Pálfy theorem, the functions $\mathbf{k}(P)$ be as bad as any algebraic variety [HP]. Theorem 5.5 suggests that $\mathbf{k}(U_n(q))$ is also this bad. This would be in line with other universality results in algebra and geometry, see e.g. [**?**, Mn, Vak].

In a different direction, Alperin showed that the action of U_n by conjugation on GL_n does have polynomial behavior [Al]. Specifically, he showed

$$|\mathrm{GL}_n / U_n| \in \mathbb{Z}[q]$$

for all $n > 0$. Moreover, because U_n acts by conjugation on each cell of the Bruhat decomposition of GL_n, we have

$$|\mathrm{GL}_n / U_n| = \sum_{w \in \mathfrak{S}_n} |B_n w B_n / U_n| \, .$$

The term in the summation corresponding to the identity element of $\mathfrak{S}_n$ is $|B_n/U_n|$, which bears resemblance to $\mathbf{k}(U_n)$. Complementary to our heuristic in Remark 5.10, Alperin noted that it seems unlikely that the summation on the right-hand side has even one non-polynomial term, given that the left-hand side is a polynomial.

For another similar phenomenon, let us mention that there are many moduli spaces which satisfy *Murphy's law*, a version of Mnëv's Universality Theorem [Vak]. Over $\mathbb{F}_q$, these moduli spaces have a non-polynomial number of points. But of course, when summed over all possible configurations these functions of q add up to a polynomial, the size of the Grassmannian or other flag varieties.

To reconcile these examples with our main approach in Chapter 4, think of them as different examples of counting points on orbifolds. Apparently, both the Grassmannian and Alperin's actions are *nice*, while conjugation on $U_n(q)$ is not. This is not very surprising. For example, both binomial coefficients $\binom{n}{k}$ and the number of integer partitions $p(n)$ count the orbits of certain combinatorial actions (see the *twelvefold way* [St]). However, while the former are "nice" indeed, the latter are notoriously complicated. Despite a large body of work on partitions, from Euler to modern times, little is known about divisibility of $p(n)$; for example, the *Erdős conjecture* that every prime s is a divisor of some $p(n)$ remains wide open (see e.g. [AO]).[11] This suggests that certain numbers of orbits are so wild, that even proving that they are wild is a great challenge.

10.9 Families of posets

There is little hope of finding interesting classes of posets for which $\mathbf{k}(U_P)$ is always a polynomial. If such a class $\mathcal{P}$ contains posets of arbitrary height, then one can show that all chains embed into some member of $\mathcal{P}$. As embedding is a transitive property, the family $\mathcal{P}$ shares the same universality properties that $\{\mathbf{C}_n\}$ has. Even the posets of height no more than three can be as bad as arbitrary algebraic varieties [HP]. Of course, if P is a poset of height two, then U_P is abelian and therefore $\mathbf{k}(U_P)$ is a polynomial in q, but such a family could hardly be considered interesting.

[11] Naturally, one would assume that asymptotically, we have $s|p(n)$ or a positive fraction of n. This is known for some primes s, such as $5, 7$ and 11 due to *Ramanujan's congruences*, but is open for 2 and 3, see e.g. [AO].

References

[ABT] M. Aguiar, N. Bergeron and N. Thiem, Hopf monoids from class functions on unitriangular matrices, *Algebra Number Theory* **7** (2013), 1743–1779.

[AO] S. Ahlgren and K. Ono, Addition and counting: the arithmetic of partitions, *Notices AMS* **48** (2001), 978–984.

[Al] J. Alperin, Unipotent conjugacy in general linear groups, *Comm. Algebra* **34** (2006), 889–891.

[BB] P. Belkale and P. Brosnan, Matroids, motives, and a conjecture of Kontsevich, *Duke Math. J.* **116** (2003), 147–188.

[BM] G. Brinkmann and B. D. McKay, Posets on up to 16 Points, *Order* **19** (2002), 147–179.

[BNV] S. R. Blackburn, Peter M. Neumann, and G. Venkataraman, *Enumeration of finite groups*, Cambridge Univ. Press, Cambridge, UK, 2007.

[DI] P. Diaconis and I. M. Isaacs, Supercharacters and Superclasses for Algebra Groups, *Trans. AMS* **360** (2008), 2359–2392.

[DT] P. Diaconis and N. Thiem, Supercharacter formulas for pattern groups, *Trans. AMS* **361** (2009), 3501–3533.

[DV] M. du Sautoy and M. Vaughan-Lee, Non-PORC behaviour of a class of descendant p-groups, *J. Algebra* **361** (2012), 287–312.

[Ga] V. Gasharov, Incomparability graphs of $(3+1)$-free posets are s-positive *Discrete Math.* **157** (1996), 193–197.

[GP] C. Gann and R. Proctor, *Chapel Hill Poset Atlas*, `http://www.unc.edu/~rap/Posets/`.

[G$^+$] P. Gudivok, Yu. Kapitonova, S. Polyak, V. Rud′ko, and A. Tsitkin, Classes of conjugate elements of a unitriangular group, *Kibernetika* **1** (1990), 40–48.

[HP] Z. Halasi and P. Pálfy, The number of conjugacy classes in pattern groups is not a polynomial function, *J. Group Theory* **14** (2011), 841–854.

[H1] G. Higman, Enumerating p-groups. I. Inequalities, *Proc. LMS* **3** (1960), 24–30.

[H2] G. Higman, Enumerating p-groups. II. Problems whose solution is PORC, *Proc. LMS* **3** (1960), 566–582.

[Is] I. M. Isaacs, Counting characters of upper triangular groups, *J. Algebra* **315** (2007), 698–719.

[IK] I. Isaacs and D. Karagueuzian, Conjugacy in groups of upper triangular matrices, *J. Algebra* **202** (1998), 704–711.

[K1] A. Kirillov, *Lectures on the orbit method*, AMS, Providence, RI, 2004.

[K2] A. Kirillov, On the combinatorics of coadjoint orbits, *Funct. Anal. Appl.* **27** (1993), 62–64.

[K3] A. Kirillov, Variations on the triangular theme, in *Lie groups and Lie algebras*, AMS, Prividence, RI, 1995, 43–73.

[K4] A. Kirillov, Two more variations on the triangular theme, in *The Orbit Method in Geometry and Physics*, Birkhäuser, Boston, MA, 2003, 243–258.

[LN] C. R. Leedham-Green and M. F. Newman, Space groups and groups of prime power order I, *Arch. Math. (Basel)* **35** (1980), 293–302.

[Mac] I. G. Macdonald, *Symmetric Functions and Hall Polynomials*, (Second Ed.), Oxford Univ. Press, New York, USA, 1995.

[Mar1] E. Marberg, Constructing Modules of Algebra Groups, Undergraduate thesis, Stanford Univ. (2008).

[Mar2] E. Marberg, Combinatorial methods of character enumeration for the unitriangular group, *J. Algebra* **345** (2011), 295–323.

[Mn] N. E. Mnëv, The universality theorems on the classification problem of configuration varieties and convex polytopes varieties, in *Lecture Notes in Math.* **1346**, Springer, Berlin, 1988, 527–543.

[R] G.-C. Rota, On the foundations of combinatorial theory. I. Theory of Möbius functions, *Z. Wahrsch.* **2** (1964), 340–368.

[OEIS] N. J. A. Sloane, *The On-Line Encyclopedia of Integer Sequences*, `http://oeis.org`.

[Sa] B. E. Sagan, *The Symmetric Group: Representations, Combinatorial Algorithms, and Symmetric Functions*, Springer, 2001.

[Si] C. C. Sims, Enumerating p-groups, *Proc. LMS* **3** (1965), 151–166.

[Sp] E. Spiegel and C. O'Donnell, *Incidence Algebras* CRC Press, 1997.

[St] R. P. Stanley, *Enumerative Combinatorics*, Vol. 1 (Second Ed.), Cambridge Univ. Press, Cambridge, UK, 2011.

[Th] J. G. Thompson, $k(U_n(\mathbb{F}_q))$, preprint (2004), 84 pp.; available at `http://tinyurl.com/m2h24nm`.

[Tr] W. Trotter, New perspectives on interval orders and interval graphs, in *Surveys in Combinatorics*, Cambridge Univ. Press, Cambridge, UK, 1997, 237–286.

[Vak] R. Vakil, Murphy's law in algebraic geometry: badly-behaved deformation spaces, *Invent. Math.* **164** (2006), 569–590.

[VA1] A. Vera-López and J. M. Arregi, Conjugacy classes in unitriangular matrices, *J. Algebra* **152** (1992), 1–19.

[VA2] A. Vera-López and J. M. Arregi, Some algorithms for the calculation of conjugacy classes in the Sylow p-subgroups of $\mathrm{GL}(n, q)$, *J. Algebra* **177** (1995), 899–925.

[VA3] A. Vera-López and J. M. Arregi, Polynomial properties in unitriangular matrices, *J. Algebra* **244** (2001), 343–351.

[VA4] A. Vera-López and J. M. Arregi, Conjugacy classes in unitriangular matrices, *Linear Algebra Appl.* **370** (2003), 85–124.

[VL] M. Vaughan-Lee, Graham Higman's PORC conjecture, *Jahresber. Dtsch. Math.-Ver.* **114** (2012), 89–106.

[W] A. J. Weir, Sylow p-subgroups of the classical groups over finite fields with characteristic prime to p, *Proc. AMS* **6** (1955), 529–533.

[Y] M. Yip, q-Rook placements and Jordan forms of upper-triangular nilpotent matrices, *DMTCS Proc.* (2013), 1017–1028.

APPENDIX A

The polynomials $\mathbf{k}(U_n(q))$

In this appendix, we list the values of $\mathbf{k}(U_n)$ for $n = 1, \ldots, 16$. These values were computed with the methods described in Chapter 6.

$$
\begin{aligned}
\mathbf{k}(U_1) &= 1 \\
\mathbf{k}(U_2) &= 1 + t \\
\mathbf{k}(U_3) &= 1 + 3t + t^2 \\
\mathbf{k}(U_4) &= 1 + 6t + 7t^2 + 2t^3 \\
\mathbf{k}(U_5) &= 1 + 10t + 25t^2 + 20t^3 + 5t^4 \\
\mathbf{k}(U_6) &= 1 + 15t + 65t^2 + 105t^3 + 70t^4 + 18t^5 + t^6 \\
\mathbf{k}(U_7) &= 1 + 21t + 140t^2 + 385t^3 + 490t^4 + 301t^5 + 84t^6 + 8t^7 \\
\mathbf{k}(U_8) &= 1 + 28t + 266t^2 + 1120t^3 + 2345t^4 + 2604t^5 + 1568t^6 + 496t^7 + 74t^8 + 4t^9 \\
\mathbf{k}(U_9) &= 1 + 36t + 462t^2 + 2772t^3 + 8715t^4 + 15372t^5 + 15862t^6 + 9720t^7 + 3489t^8 \\
&\quad + 701t^9 + 72t^{10} + 3t^{11} \\
\mathbf{k}(U_{10}) &= 1 + 45t + 750t^2 + 6090t^3 + 26985t^4 + 69825t^5 + 110530t^6 + 110280t^7 \\
&\quad + 70320t^8 + 28640t^9 + 7362t^{10} + 1170t^{11} + 110t^{12} + 5t^{13} \\
\mathbf{k}(U_{11}) &= 1 + 55t + 1155t^2 + 12210t^3 + 72765t^4 + 261261t^5 + 592207t^6 + 877030t^7 \\
&\quad + 868725t^8 + 583550t^9 + 267542t^{10} + 83909t^{11} + 18007t^{12} + 2618t^{13} \\
&\quad + 242t^{14} + 11t^{15}
\end{aligned}
$$

$$\begin{aligned}\mathbf{k}(U_{12}) = 1 &+ 66t + 1705t^2 + 22770t^3 + 176055t^4 + 841302t^5 + 2600983t^6 + 5387646t^7 \\ &+ 7680310t^8 + 7684820t^9 + 5473050t^{10} + 2803182t^{11} + 1042181t^{12} + 284109t^{13} \\ &+ 57256t^{14} + 8484t^{15} + 890t^{16} + 60t^{17} + 2t^{18}\end{aligned}$$

$$\begin{aligned}\mathbf{k}(U_{13}) = 1 &+ 78t + 2431t^2 + 40040t^3 + 390390t^4 + 2403258t^5 + 9766471t^6 + 27116232t^7 \\ &+ 52873678t^8 + 74012653t^9 + 75670881t^{10} + 57294120t^{11} + 32515314t^{12} \\ &+ 14000495t^{13} + 4635125t^{14} + 1195116t^{15} + 241436t^{16} + 37778t^{17} + 4381t^{18} \\ &+ 338t^{19} + 13t^{20}\end{aligned}$$

$$\begin{aligned}\mathbf{k}(U_{14}) = 1 &+ 91t + 3367t^2 + 67067t^3 + 805805t^4 + 6225219t^5 + 32296264t^6 + 116332645t^7 \\ &+ 298956658t^8 + 560602042t^9 + 781499719t^{10} + 822549728t^{11} + 662497381t^{12} \\ &+ 413509705t^{13} + 202666910t^{14} + 79124292t^{15} + 24968979t^{16} + 6441876t^{17} \\ &+ 1362732t^{18} + 233758t^{19} + 31542t^{20} + 3159t^{21} + 210t^{22} + 7t^{23}\end{aligned}$$

$$\begin{aligned}\mathbf{k}(U_{15}) = 1 &+ 105t + 4550t^2 + 107835t^3 + 1566565t^4 + 14864850t^5 + 96136040t^6 + 437680815t^7 \\ &+ 1440259535t^8 + 3502779995t^9 + 6416611201t^{10} + 8998108665t^{11} + 9796436195t^{12} \\ &+ 8387410675t^{13} + 5718426690t^{14} + 3145744973t^{15} + 1416179446t^{16} + 529371274t^{17} \\ &+ 166405370t^{18} + 44325415t^{19} + 9997955t^{20} + 1887955t^{21} + 291345t^{22} + 35270t^{23} \\ &+ 3130t^{24} + 180t^{25} + 5t^{26}\end{aligned}$$

$$\begin{aligned}\mathbf{k}(U_{16}) = 1 &+ 120t + 6020t^2 + 167440t^3 + 2894710t^4 + 33137104t^5 + 261929668t^6 \\ &+ 1475199440t^7 + 6072906125t^8 + 18674026800t^9 + 43703418616t^{10} \\ &+ 79124540872t^{11} + 112420822696t^{12} + 126975887444t^{13} + 115398765556t^{14} \\ &+ 85415064915t^{15} + 52146190588t^{16} + 26615252562t^{17} + 11515549082t^{18} \\ &+ 4278222573t^{19} + 1378103758t^{20} + 386616800t^{21} + 94259304t^{22} + 19784488t^{23} \\ &+ 3513854t^{24} + 514128t^{25} + 59504t^{26} + 5104t^{27} + 288t^{28} + 8t^{29}\end{aligned}$$

APPENDIX B

The polynomials $\mathbf{k}(\text{Þ}_{a,b,c})$

In this appendix, we list the values of $\mathbf{k}(\text{Þ}_{a,b,c})$ for $a+b+c \leq 9$. These values were computed with the methods described in Chapter 6. Recall that $\mathbf{k}(\text{Þ}_{a,b,c}) = \mathbf{k}(\text{Þ}_{c,b,a})$, so we only provide the values $\mathbf{k}(\text{Þ}_{a,b,c})$ where $a \geq c$. Moreover, when either $b = 0$ or we have $\text{Þ}_{a,b,c} = \mathbf{C}_{a+b+c+1}$, and when $a = c = 0$, we have $\text{Þ}_{a,b,c} = \mathbf{C}_b \amalg \mathbf{C}_1$. The number of classes in their associated pattern groups is already provided in Appendix A, so we skip these as well.

Values of $\mathbf{k}(\text{Þ}_{a,b,c})$ where $\mathbf{a}+\mathbf{b}+\mathbf{c} = \mathbf{3}$:

$$\mathbf{k}(\text{Þ}_{2,1,0}) = 1 + 5t + 6t^2 + 2t^3$$

$$\mathbf{k}(\text{Þ}_{1,1,1}) = 1 + 5t + 6t^2 + 4t^3 + t^4$$

$$\mathbf{k}(\text{Þ}_{1,2,0}) = 1 + 4t + 4t^2 + t^3$$

Values of $\mathbf{k}(\text{Þ}_{a,b,c})$ where $\mathbf{a}+\mathbf{b}+\mathbf{c} = \mathbf{4}$:

$$\mathbf{k}(\text{Þ}_{3,1,0}) = 1 + 9t + 22t^2 + 19t^3 + 7t^4 + t^5$$

$$\mathbf{k}(\text{Þ}_{2,1,1}) = 1 + 9t + 22t^2 + 23t^3 + 11t^4 + 2t^5$$

$$\mathbf{k}(\text{Þ}_{2,2,0}) = 1 + 8t + 18t^2 + 15t^3 + 4t^4$$

$$\mathbf{k}(\text{Þ}_{1,2,1}) = 1 + 8t + 18t^2 + 19t^3 + 10t^4 + 2t^5$$

$$\mathbf{k}(\text{Þ}_{1,3,0}) = 1 + 7t + 13t^2 + 9t^3 + 2t^4$$

Values of $\mathbf{k}(\text{Þ}_{a,b,c})$ where $\mathbf{a+b+c=5}$:

$$\mathbf{k}(\text{Þ}_{4,1,0}) = 1 + 14t + 59t^2 + 98t^3 + 76t^4 + 28t^5 + 4t^6$$

$$\mathbf{k}(\text{Þ}_{3,1,1}) = 1 + 14t + 59t^2 + 104t^3 + 91t^4 + 42t^5 + 10t^6 + t^7$$

$$\mathbf{k}(\text{Þ}_{2,1,2}) = 1 + 14t + 59t^2 + 106t^3 + 96t^4 + 42t^5 + 7t^6$$

$$\mathbf{k}(\text{Þ}_{3,2,0}) = 1 + 13t + 52t^2 + 85t^3 + 63t^4 + 20t^5 + 2t^6$$

$$\mathbf{k}(\text{Þ}_{2,2,1}) = 1 + 13t + 52t^2 + 93t^3 + 87t^4 + 44t^5 + 11t^6 + t^7$$

$$\mathbf{k}(\text{Þ}_{2,3,0}) = 1 + 12t + 44t^2 + 67t^3 + 44t^4 + 12t^5 + t^6$$

$$\mathbf{k}(\text{Þ}_{1,3,1}) = 1 + 12t + 44t^2 + 73t^3 + 65t^4 + 29t^5 + 5t^6$$

$$\mathbf{k}(\text{Þ}_{1,4,0}) = 1 + 11t + 35t^2 + 45t^3 + 25t^4 + 5t^5$$

Values of $\mathbf{k}(\text{Þ}_{a,b,c})$ where $\mathbf{a+b+c=6}$:

$$\mathbf{k}(\text{Þ}_{5,1,0}) = 1 + 20t + 130t^2 + 360t^3 + 490t^4 + 356t^5 + 139t^6 + 27t^7 + 2t^8$$

$$\mathbf{k}(\text{Þ}_{4,1,1}) = 1 + 20t + 130t^2 + 368t^3 + 530t^4 + 426t^5 + 195t^6 + 48t^7 + 5t^8$$

$$\mathbf{k}(\text{Þ}_{3,1,2}) = 1 + 20t + 130t^2 + 372t^3 + 550t^4 + 447t^5 + 199t^6 + 45t^7 + 4t^8$$

$$\mathbf{k}(\text{Þ}_{4,2,0}) = 1 + 19t + 119t^2 + 325t^3 + 441t^4 + 315t^5 + 117t^6 + 20t^7 + t^8$$

$$\mathbf{k}(\text{Þ}_{3,2,1}) = 1 + 19t + 119t^2 + 337t^3 + 507t^4 + 439t^5 + 226t^6 + 69t^7 + 12t^8 + t^9$$

$$\mathbf{k}(\text{Þ}_{2,2,2}) = 1 + 19t + 119t^2 + 341t^3 + 529t^4 + 471t^5 + 241t^6 + 69t^7 + 11t^8 + t^9$$

$$\mathbf{k}(\text{Þ}_{3,3,0}) = 1 + 18t + 107t^2 + 281t^3 + 364t^4 + 242t^5 + 81t^6 + 13t^7 + t^8$$

$$\mathbf{k}(\text{Þ}_{2,3,1}) = 1 + 18t + 107t^2 + 293t^3 + 436t^4 + 375t^5 + 185t^6 + 48t^7 + 5t^8$$

$$\mathbf{k}(\text{Þ}_{2,4,0}) = 1 + 17t + 94t^2 + 229t^3 + 273t^4 + 164t^5 + 47t^6 + 5t^7$$

$$\mathbf{k}(\text{Þ}_{1,4,1}) = 1 + 17t + 94t^2 + 237t^3 + 325t^4 + 252t^5 + 105t^6 + 20t^7 + t^8$$

$$\mathbf{k}(\text{Þ}_{1,5,0}) = 1 + 16t + 80t^2 + 170t^3 + 175t^4 + 88t^5 + 19t^6 + t^7$$

Values of $\mathbf{k}(\text{Þ}_{a,b,c})$ where $\mathbf{a}+\mathbf{b}+\mathbf{c}=\mathbf{7}$:

$$\mathbf{k}(\text{Þ}_{6,1,0}) = 1+27t+251t^2+1055t^3+2280t^4+2764t^5+1965t^6+822t^7+194t^8+23t^9+t^{10}$$

$$\mathbf{k}(\text{Þ}_{5,1,1}) = 1+27t+251t^2+1065t^3+2365t^4+3014t^5+2320t^6+1093t^7+307t^8+47t^9+3t^{10}$$

$$\mathbf{k}(\text{Þ}_{4,1,2}) = 1+27t+251t^2+1071t^3+2416t^4+3136t^5+2439t^6+1140t^7+307t^8+42t^9+2t^{10}$$

$$\mathbf{k}(\text{Þ}_{3,1,3}) = 1+27t+251t^2+1073t^3+2433t^4+3172t^5+2463t^6+1141t^7+307t^8+45t^9+3t^{10}$$

$$\mathbf{k}(\text{Þ}_{5,2,0}) = 1+26t+235t^2+975t^3+2110t^4+2579t^5+1842t^6+764t^7+175t^8+20t^9+t^{10}$$

$$\mathbf{k}(\text{Þ}_{4,2,1}) = 1+26t+235t^2+991t^3+2254t^4+3027t^5+2532t^6+1360t^7+476t^8+108t^9 + 15t^{10}+t^{11}$$

$$\mathbf{k}(\text{Þ}_{3,2,2}) = 1+26t+235t^2+999t^3+2326t^4+3223t^5+2767t^6+1493t^7+508t^8+110t^9 + 15t^{10}+t^{11}$$

$$\mathbf{k}(\text{Þ}_{4,3,0}) = 1+25t+218t^2+881t^3+1865t^4+2212t^5+1513t^6+598t^7+135t^8+17t^9+t^{10}$$

$$\mathbf{k}(\text{Þ}_{3,3,1}) = 1+25t+218t^2+899t^3+2036t^4+2755t^5+2332t^6+1246t^7+408t^8+75t^9+6t^{10}$$

$$\mathbf{k}(\text{Þ}_{2,3,2}) = 1+25t+218t^2+905t^3+2093t^4+2922t^5+2542t^6+1394t^7+485t^8+108t^9 + 15t^{10}+t^{11}$$

$$\mathbf{k}(\text{Þ}_{3,4,0}) = 1+24t+200t^2+774t^3+1568t^4+1762t^5+1126t^6+404t^7+76t^8+6t^9$$

$$\mathbf{k}(\text{Þ}_{2,4,1}) = 1+24t+200t^2+790t^3+1728t^4+2266t^5+1832t^6+902t^7+256t^8+37t^9+2t^{10}$$

$$\mathbf{k}(\text{Þ}_{2,4,0}) = 1+23t+181t^2+655t^3+1235t^4+1281t^5+741t^6+231t^7+35t^8+2t^9$$

$$\mathbf{k}(\text{Þ}_{1,4,1}) = 1+23t+181t^2+665t^3+1340t^4+1601t^5+1146t^6+472t^7+100t^8+8t^9$$

$$\mathbf{k}(\text{Þ}_{1,5,0}) = 1+22t+161t^2+525t^3+875t^4+791t^5+385t^6+92t^7+8t^8$$

Values of $\mathbf{k}(\text{Þ}_{a,b,c})$ where $\mathbf{a+b+c=8}$:

$$\mathbf{k}(\text{Þ}_{7,1,0}) = 1 + 35t + 441t^2 + 2632t^3 + 8400t^4 + 15547t^5 + 17626t^6 + 12611t^7 + 5735t^8 + 1633t^9 + 279t^{10} + 26t^{11} + t^{12}$$

$$\mathbf{k}(\text{Þ}_{6,1,1}) = 1 + 35t + 441t^2 + 2644t^3 + 8556t^4 + 16262t^5 + 19246t^6 + 14657t^7 + 7253t^8 + 2299t^9 + 445t^{10} + 47t^{11} + 2t^{12}$$

$$\mathbf{k}(\text{Þ}_{5,1,2}) = 1 + 35t + 441t^2 + 2652t^3 + 8660t^4 + 16692t^5 + 20046t^6 + 15428t^7 + 7648t^8 + 2396t^9 + 451t^{10} + 46t^{11} + 2t^{12}$$

$$\mathbf{k}(\text{Þ}_{4,1,3}) = 1 + 35t + 441t^2 + 2656t^3 + 8712t^4 + 16893t^5 + 20362t^6 + 15654t^7 + 7718t^8 + 2405t^9 + 454t^{10} + 47t^{11} + 2t^{12}$$

$$\mathbf{k}(\text{Þ}_{6,2,0}) = 1 + 34t + 419t^2 + 2471t^3 + 7885t^4 + 14747t^5 + 17000t^6 + 12385t^7 + 5730t^8 + 1669t^9 + 303t^{10} + 34t^{11} + 2t^{12}$$

$$\mathbf{k}(\text{Þ}_{5,2,1}) = 1 + 34t + 419t^2 + 2491t^3 + 8155t^4 + 16037t^5 + 20100t^6 + 16687t^7 + 9405t^8 + 3654t^9 + 983t^{10} + 179t^{11} + 20t^{12} + t^{13}$$

$$\mathbf{k}(\text{Þ}_{4,2,2}) = 1 + 34t + 419t^2 + 2503t^3 + 8317t^4 + 16755t^5 + 21592t^6 + 18371t^7 + 10504t^8 + 4079t^9 + 1082t^{10} + 193t^{11} + 21t^{12} + t^{13}$$

$$\mathbf{k}(\text{Þ}_{3,2,3}) = 1 + 34t + 419t^2 + 2507t^3 + 8371t^4 + 16985t^5 + 22028t^6 + 18792t^7 + 10721t^8 + 4142t^9 + 1095t^{10} + 195t^{11} + 21t^{12} + t^{13}$$

$$\mathbf{k}(\text{Þ}_{5,3,0}) = 1 + 33t + 396t^2 + 2290t^3 + 7210t^4 + 13312t^5 + 15084t^6 + 10752t^7 + 4854t^8 + 1377t^9 + 240t^{10} + 24t^{11} + t^{12}$$

$$\begin{aligned}\mathbf{k}(\text{Þ}_{4,3,1}) &= 1 + 33t + 396t^2 + 2314t^3 + 7546t^4 + 14958t^5 + 19072t^6 \\ &\quad + 16193t^7 + 9319t^8 + 3649t^9 + 963t^{10} + 167t^{11} + 18t^{12} + t^{13}\end{aligned}$$

$$\begin{aligned}\mathbf{k}(\text{Þ}_{3,3,2}) &= 1 + 33t + 396t^2 + 2326t^3 + 7714t^4 + 15739t^5 + 20766t^6 \\ &\quad + 18219t^7 + 10800t^8 + 4381t^9 + 1234t^{10} + 242t^{11} + 31t^{12} + 2t^{13}\end{aligned}$$

$$\begin{aligned}\mathbf{k}(\text{Þ}_{4,4,0}) &= 1 + 32t + 372t^2 + 2090t^3 + 6408t^4 + 11484t^5 + 12547t^6 \\ &\quad + 8548t^7 + 3668t^8 + 996t^9 + 171t^{10} + 18t^{11} + t^{12}\end{aligned}$$

$$\begin{aligned}\mathbf{k}(\text{Þ}_{3,4,1}) &= 1 + 32t + 372t^2 + 2114t^3 + 6756t^4 + 13208t^5 + 16627t^6 \\ &\quad + 13819t^7 + 7600t^8 + 2711t^9 + 598t^{10} + 74t^{11} + 4t^{12}\end{aligned}$$

$$\begin{aligned}\mathbf{k}(\text{Þ}_{2,4,2}) &= 1 + 32t + 372t^2 + 2122t^3 + 6872t^4 + 13764t^5 + 17843t^6 \\ &\quad + 15310t^7 + 8776t^8 + 3355t^9 + 838t^{10} + 127t^{11} + 9t^{12}\end{aligned}$$

$$\begin{aligned}\mathbf{k}(\text{Þ}_{3,5,0}) &= 1 + 31t + 347t^2 + 1872t^3 + 5505t^4 + 9424t^5 + 9765t^6 \\ &\quad + 6226t^7 + 2427t^8 + 560t^9 + 71t^{10} + 4t^{11}\end{aligned}$$

$$\begin{aligned}\mathbf{k}(\text{Þ}_{2,5,1}) &= 1 + 31t + 347t^2 + 1892t^3 + 5805t^4 + 10909t^5 + 13125t^6 \\ &\quad + 10263t^7 + 5169t^8 + 1622t^9 + 297t^{10} + 28t^{11} + t^{12}\end{aligned}$$

$$\begin{aligned}\mathbf{k}(\text{Þ}_{2,6,0}) &= 1 + 30t + 321t^2 + 1637t^3 + 4520t^4 + 7229t^5 + 6936t^6 \\ &\quad + 4029t^7 + 1392t^8 + 272t^9 + 27t^{10} + t^{11}\end{aligned}$$

$$\begin{aligned}\mathbf{k}(\text{Þ}_{1,6,1}) &= 1 + 30t + 321t^2 + 1649t^3 + 4706t^4 + 8139t^5 + 8876t^6 \\ &\quad + 6130t^7 + 2614t^8 + 647t^9 + 82t^{10} + 4t^{11}\end{aligned}$$

$$\begin{aligned}\mathbf{k}(\text{Þ}_{1,7,0}) &= 1 + 29t + 294t^2 + 1386t^3 + 3465t^4 + 4949t^5 + 4172t^6 \\ &\quad + 2064t^7 + 570t^8 + 78t^9 + 4t^{10}\end{aligned}$$

Values of $\mathbf{k}(\text{Þ}_{a,b,c})$ where $\mathbf{a+b+c=9}$:

$$\mathbf{k}(\text{Þ}_{8,1,0}) = 1 + 44t + 722t^2 + 5824t^3 + 25977t^4 + 69076t^5 + 115696t^6 + 126800t^7 + 93022t^8 + 46120t^9 + 15433t^{10} + 3447t^{11} + 502t^{12} + 45t^{13} + 2t^{14}$$

$$\mathbf{k}(\text{Þ}_{7,1,1}) = 1 + 44t + 722t^2 + 5838t^3 + 26236t^4 + 70812t^5 + 121541t^6 + 138056t^7 + 106308t^8 + 56063t^9 + 20187t^{10} + 4875t^{11} + 759t^{12} + 70t^{13} + 3t^{14}$$

$$\mathbf{k}(\text{Þ}_{6,1,2}) = 1 + 44t + 722t^2 + 5848t^3 + 26421t^4 + 71982t^5 + 125041t^6 + 143812t^7 + 111924t^8 + 59436t^9 + 21464t^{10} + 5199t^{11} + 824t^{12} + 81t^{13} + 4t^{14}$$

$$\mathbf{k}(\text{Þ}_{5,1,3}) = 1 + 44t + 722t^2 + 5854t^3 + 26532t^4 + 72656t^5 + 126871t^6 + 146398t^7 + 113965t^8 + 60318t^9 + 21615t^{10} + 5157t^{11} + 791t^{12} + 72t^{13} + 3t^{14}$$

$$\mathbf{k}(\text{Þ}_{4,1,4}) = 1 + 44t + 722t^2 + 5856t^3 + 26569t^4 + 72876t^5 + 127436t^6 + 147120t^7 + 114446t^8 + 60546t^9 + 21780t^{10} + 5274t^{11} + 838t^{12} + 82t^{13} + 4t^{14}$$

$$\mathbf{k}(\text{Þ}_{7,2,0}) = 1 + 43t + 693t^2 + 5530t^3 + 24619t^4 + 65975t^5 + 112322t^6 + 125890t^7 + 94829t^8 + 48479t^9 + 16878t^{10} + 4004t^{11} + 640t^{12} + 64t^{13} + 3t^{14}$$

$$\mathbf{k}(\text{Þ}_{6,2,1}) = 1 + 43t + 693t^2 + 5554t^3 + 25075t^4 + 69127t^5 + 123397t^6 + 148727t^7 + 124653t^8 + 74225t^9 + 31927t^{10} + 10037t^{11} + 2304t^{12} + 374t^{13} + 39t^{14} + 2t^{15}$$

$$\mathbf{k}(\text{Þ}_{5,2,2}) = 1 + 43t + 693t^2 + 5570t^3 + 25379t^4 + 71135t^5 + 129827t^6 + 160399t^7 + 137692t^8 + 83588t^9 + 36342t^{10} + 11414t^{11} + 2585t^{12} + 409t^{13} + 41t^{14} + 2t^{15}$$

$$\mathbf{k}(\text{Þ}_{4,2,3}) = 1 + 43t + 693t^2 + 5578t^3 + 25531t^4 + 72111t^5 + 132756t^6 + 165206t^7 + 142360t^8 + 86414t^9 + 37488t^{10} + 11763t^{11} + 2668t^{12} + 422t^{13} + 42t^{14} + 2t^{15}$$

$$\mathbf{k}(\text{Þ}_{6,3,0}) = 1 + 42t + 663t^2 + 5209t^3 + 22962t^4 + 61145t^5 + 103483t^6 + 115130t^7 + 85919t^8 + 43456t^9 + 14972t^{10} + 3527t^{11} + 565t^{12} + 58t^{13} + 3t^{14}$$

$$\mathbf{k}(\text{Þ}_{5,3,1}) = 1 + 42t + 663t^2 + 5239t^3 + 23547t^4 + 65280t^5 + 118248t^6 + 145808t^7 + 125938t^8 + 77659t^9 + 34559t^{10} + 11118t^{11} + 2558t^{12} + 407t^{13} + 41t^{14} + 2t^{15}$$

$$\mathbf{k}(\text{Þ}_{4,3,2}) = 1 + 42t + 663t^2 + 5257t^3 + 23898t^4 + 67677t^5 + 126201t^6 + 160796t^7 + 143500t^8 + 91418t^9 + 42430t^{10} + 14768t^{11} + 4001t^{12} + 866t^{13} + 146t^{14} + 17t^{15} + t^{16}$$

$$\begin{aligned}\mathbf{k}(\text{Þ}_{3,3,3}) &= 1 + 42t + 663t^2 + 5263t^3 + 24015t^4 + 68462t^5 + 128701t^6 + 165236t^7 + 148364t^8 \\ &\quad + 94817t^9 + 43765t^{10} + 14757t^{11} + 3636t^{12} + 636t^{13} + 72t^{14} + 4t^{15} \\ \mathbf{k}(\text{Þ}_{5,4,0}) &= 1 + 41t + 632t^2 + 4862t^3 + 21050t^4 + 55070t^5 + 91324t^6 + 99155t^7 + 71975t^8 \\ &\quad + 35399t^9 + 11902t^{10} + 2745t^{11} + 428t^{12} + 42t^{13} + 2t^{14} \\ \mathbf{k}(\text{Þ}_{4,4,1}) &= 1 + 41t + 632t^2 + 4894t^3 + 21690t^4 + 59662t^5 + 107680t^6 + 132419t^7 + 113565t^8 \\ &\quad + 68717t^9 + 29417t^{10} + 8877t^{11} + 1870t^{12} + 268t^{13} + 24t^{14} + t^{15} \\ \mathbf{k}(\text{Þ}_{3,4,2}) &= 1 + 41t + 632t^2 + 4910t^3 + 22010t^4 + 61902t^5 + 115222t^6 + 146781t^7 + 130705t^8 \\ &\quad + 82501t^9 + 37277t^{10} + 12119t^{11} + 2813t^{12} + 448t^{13} + 44t^{14} + 2t^{15} \\ \mathbf{k}(\text{Þ}_{4,5,0}) &= 1 + 40t + 600t^2 + 4490t^3 + 18920t^4 + 48112t^5 + 77275t^6 + 80812t^7 + 56076t^8 \\ &\quad + 26097t^9 + 8200t^{10} + 1749t^{11} + 252t^{12} + 23t^{13} + t^{14} \\ \mathbf{k}(\text{Þ}_{3,5,1}) &= 1 + 40t + 600t^2 + 4520t^3 + 19535t^4 + 52562t^5 + 92870t^6 + 111360t^7 + 92100t^8 \\ &\quad + 52642t^9 + 20570t^{10} + 5364t^{11} + 895t^{12} + 88t^{13} + 4t^{14} \\ \mathbf{k}(\text{Þ}_{2,5,2}) &= 1 + 40t + 600t^2 + 4530t^3 + 19740t^4 + 54022t^5 + 97790t^6 + 120666t^7 + 103213t^8 \\ &\quad + 61674t^9 + 25720t^{10} + 7381t^{11} + 1399t^{12} + 158t^{13} + 8t^{14} \\ \mathbf{k}(\text{Þ}_{3,6,0}) &= 1 + 39t + 567t^2 + 4094t^3 + 16602t^4 + 40528t^5 + 62243t^6 + 61828t^7 + 40251t^8 \\ &\quad + 17173t^9 + 4741t^{10} + 823t^{11} + 84t^{12} + 4t^{13} \\ \mathbf{k}(\text{Þ}_{2,6,1}) &= 1 + 39t + 567t^2 + 4118t^3 + 17106t^4 + 44184t^5 + 74718t^6 + 85075t^7 + 65906t^8 \\ &\quad + 34607t^9 + 12115t^{10} + 2744t^{11} + 383t^{12} + 30t^{13} + t^{14} \\ \mathbf{k}(\text{Þ}_{2,7,0}) &= 1 + 38t + 533t^2 + 3675t^3 + 14119t^4 + 32487t^5 + 46781t^6 + 43208t^7 + 25820t^8 \\ &\quad + 9925t^9 + 2406t^{10} + 354t^{11} + 29t^{12} + t^{13} \\ \mathbf{k}(\text{Þ}_{1,7,1}) &= 1 + 38t + 533t^2 + 3689t^3 + 14420t^4 + 34664t^5 + 53956t^6 + 55724t^7 + 38322t^8 \\ &\quad + 17299t^9 + 4955t^{10} + 848t^{11} + 78t^{12} + 3t^{13} \\ \mathbf{k}(\text{Þ}_{1,8,0}) &= 1 + 37t + 498t^2 + 3234t^3 + 11487t^4 + 24087t^5 + 31234t^6 + 25582t^7 + 13209t^8 \\ &\quad + 4190t^9 + 773t^{10} + 75t^{11} + 3t^{12}\end{aligned}$$