Phase transitions in the evolution of partially ordered sets

DISSERTATION

zur Erlangung des akademischen Grades
DOCTOR RERUM NATURALIUM

im Fach Informatik

eingereicht an der
Mathematisch–Naturwissenschaftlichen Fakultät II
der Humboldt–Universität zu Berlin

von
Dipl.–Math.
Anuschirawan Ralf Taraz
geboren am 7. März 1971 in Köln

Präsident der Humboldt–Universität zu Berlin
Prof. Dr. Dr. h.c. Hans Meyer

Dekan der Mathematisch–Naturwissenschaftlichen Fakultät II
Prof. Dr. Bodo Krause

Gutachter:
1. Prof. Dr. Hans Jürgen Prömel
2. Prof. Dr. Angelika Steger
3. Prof. Dr. Vojtěch Rödl

Tag der mündlichen Prüfung: 6. Januar 1999

Contents

Preface **v**

Notation **ix**

1 Introduction **1**
1.1 Counting graphs 1
1.2 Counting posets 8
1.3 Counting posets with a given number of relations 12

2 Partitionable posets **21**
2.1 Configurations 23
2.2 Linear configurations 29

3 Layer posets **39**
3.1 Determining candidates 43
3.2 Dismissing candidates 48
3.3 Linearity and monotonicity 56
3.4 Technical computations 60

Bibliography **77**

Preface

Counting is one of the most elementary things in the world. It appears in almost all kinds of human activity, ranging from rhythm in music and poetry, proportion and dimension in arts, statistics in social sciences and experiments in natural sciences.

At the same time, counting, or equivalently, the idea of numbers, is also one of the first levels of abstraction that one seeks – two objects (they could be groups of people, time intervals, or simply sets) are *similar*, if there exists a one–to–one relation between them, which is usually taken as the definition of 'they have the same number' (see Russell's "Principles of Mathematics"). This reduction principle, i.e., giving up the interest in individual identity for the sake of a common property or structure, is one of the underlying mechanisms of science and will be quite visible throughout this thesis.

We will be concerned with *asymptotic enumeration*, by which we mean obtaining good approximations for the numbers of objects in a given class. In general, counting is easy in situations where one can generate all objects by a sequence of *independent* decisions such that two different sequences never produce the same object – for then the number of members is precisely the product of the number of possibilities one had at each step. For example, the number of subsets in an n-element set is 2^n, because for each element we can decide independently whether it should belong to the subset or not. However, the required independence is often not given, on the contrary one usually finds that depending on how earlier decisions are taken later decisions may be forced. Since we are only interested in the approximate number of objects, one way to circumvent this problem is the following: within the class in question, find a subclass which can be enumerated by independent decisions, and show that this subclass contains *almost all* of the objects of the larger class.

This strategy is very attractive since it not only estimates the cardinality of the larger class, but at the same time proves a result about the global structure of the larger class: draw an object uniformly at random from the class, then with high probability it satisfies the conditions that define the subclass. This observation leads us to two important points.

First, it shows that the design of algorithms can profit from good estimates for cardinalities and knowledge of the global structure of the space of inputs. This applies especially when we consider the *average case* analysis of an algorithm: even though it might perform very badly in the worst case, it may well be that it is extremely successful for the

overwhelming majority of input instances, *because* they belong to a large subclass which has a nice structure.

Second, it brings us back to the reduction principle of sacrificing individual identity in the course of reunification under the name of a global structure. This might seem disturbing because it is actually not the common structure of all members, but only that of the majority! Take two results, which we will consider in more detail in this thesis: 'almost all triangle-free graphs are bipartite', and 'almost all partially ordered sets have height 3'. These statements are likely to surprise someone who works with odd cycles or linear orders. In other words, the degree of acceptance or understanding that a result of the above type receives will depend entirely on the subjective view of the audience, namely on whether the audience really regards all instances of the given universe as equally likely or interesting. However, there is a way to accommodate different points of view, and this brings us to the following.

The evolution process of a random structure denotes the following 'experiment': choose a parameter of the objects in the class under consideration, for example the number of edges (in a triangle–free graph) or the number of comparable pairs (in a poset). Now consider only one 'slice' of the universe, namely only those objects where the parameter is equal to a fixed value x and pose the following questions: what does a typical object from this slice look like? How many objects are there in this slice? And how do the answers to the first two questions change when x changes?

We keep to our examples, since this will lead us straight to the main result of this thesis. In the case of triangle–free graphs, Prömel and Steger showed in 1996 that as the number of edges increases, a random triangle–free graph exhibits two *phase transitions*: first it is almost surely bipartite, then it is not, and then it is once again. In the case of posets denote by $\mathcal{P}_{n,d}$ the class of all n-point posets with dn^2 relations. In 1978, Dhar determined $|\mathcal{P}_{n,d}|$ in the range $\frac{1}{8} \leq d \leq \frac{3}{16}$ and showed that here a typical poset consists of three 'layers' (see below for a definition). In 1979, Kleitman and Rothschild determined $|\mathcal{P}_{n,d}|$ in the range $0 < d \leq \frac{1}{8}$ and showed that here a typical poset consists of two layers, i.e. it is bipartite. Their paper concludes with the following remark:

> Above $\frac{3}{16}n^2$ one might guess that the dominant states have three [layers] with equal top and bottom population at least until $d = \frac{2}{9}n^2$ at which point all three [layers] would be equal in size. Eventually four and higher [layer posets] would dominate. Just where and what happens there is not obvious. Presumably phase transitions occur at each such point. One not only lacks a method of proof, but even sensible conjectures in this higher range.

The main result of this thesis is to complete the picture by describing the whole evolution process of $\mathcal{P}_{n,d}$ in the range $0 < d < \frac{1}{2}$. We determine $|\mathcal{P}_{n,d}|$ for any d and show that there exist an infinite number of phase transitions. Finally we describe how the structure of a typical partially ordered set changes during these phases.

This thesis is organized as follows. Chapter 1 contains a detailed introduction to the area of asymptotic enumeration, random structures and evolution processes. The chapter is subdivided into Section 1.1 that deals with graphs, Section 1.2 that deals with partially ordered sets, and finally Section 1.3, where we first describe the previous work by Dhar and Kleitman and Rothschild before we state and explain our main result.

Chapter 2 contains the first part of the proof of the main result. Here we first prove in Section 2.1 that every poset is 'close' to one with a 'partitionable' structure and then show in Section 2.2 that it suffices to consider the case where the partition classes are arranged in a 'linear' way.

Chapter 3 determines the number and size of the partition classes that such a linear partitionable 'configuration' must have in order to describe the typical structure of a poset in $\mathcal{P}_{n,d}$. Here we first develop a number of necessary conditions in Section 3.1 and then concentrate on those configurations satisfying them. In Section 3.2 we show that for each d there can be only one such configuration. Section 3.3 closes with a proof of the main result. Section 3.4 contains the proofs of a number of technical statements needed in Chapter 3.

Both in Chapters 2 and 3 an outline of the proofs can be found towards the beginning of the chapters.

I am very grateful to Prof. Dr. Hans Jürgen Prömel, who introduced me to the area of asymptotic enumeration in general and the evolution of partially ordered sets in particular. This thesis was written under his supervision and is based on joint work with him and Prof. Dr. Angelika Steger. I would like to thank both of them for their encouraging collaboration. Moreover, I owe many thanks to my friends and colleagues for very helpful discussions and suggestions. Finally, I express my deep appreciation to my family for keeping me in touch with reality.

Notation

In this section we introduce the basic notation used in this thesis. If n is a positive integer, then we let $[n] := \{1, \ldots, n\}$. For a real number x we denote by $\lfloor x \rfloor$ the largest integer that is smaller or equal to x, and by $\lceil x \rceil$ the smallest integer that is larger or equal to x. However we will often be somewhat sloppy and disregard rounding in which case the brackets will be omitted but tacitly assumed. Furthermore, the logarithm to base 2 will be denoted by log.

We will sometimes use the Landau symbols $O, o, \Omega, \Theta, \omega$. Notice that in the case of o we will also use them for functions with a negative range.

The *entropy function* $H : (0, 1) \to (0, 1]$ is defined by

$$H(x) := -x \log x - (1 - x) \log(1 - x).$$

It has the following useful property for any fixed $0 < x < 1$:

$$\binom{n}{x \cdot n} = 2^{H(x) \cdot n + o(n)}.$$

A *graph* is an ordered pair $G = (V(G), E(G))$ where $V = V(G)$ is a set, and $E = E(G)$ is a subset of the set of all unordered pairs of V, i.e. $E \subseteq \left[\begin{smallmatrix} V \\ 2 \end{smallmatrix}\right]$. Here V is called the set of *vertices* and E the set of *edges*. In a *directed graph*, sometimes abbreviated as *digraph*, the edges are directed and hence $E \subseteq V \times V$. A graph H is a *(weak) subgraph* of G if $V(H) \subseteq V(G)$ and $E(H) \subseteq E(G) \cap \left[\begin{smallmatrix} V(H) \\ 2 \end{smallmatrix}\right]$. If equality holds in the latter condition, then H is an *induced subgraph* of G. If H is not a subgraph of G, then G is called *H-free*.

Two vertices v, w of a graph G are called *adjacent* if $\{v, w\} \in E(G)$. A vertex w is a *neighbour* of a vertex v if v and w are adjacent. The *neighbourhood* $\Gamma(v)$ is the set of all neighbours of v, and the *degree* $deg(v)$ of v is the cardinality of its neighbourhood. A *k-colouring* of a graph G is a function $f : V(G) \to [k]$ such that for every edge $\{v, w\} \in E(G)$ we have $f(v) \neq f(w)$. The chromatic number $\chi(G)$ is the least integer k for which there exists a k-colouring of G. If $\chi(G) = 2$, then G is called *bipartite*.

A *partially ordered set*, often abbreviated as *poset*, is an ordered pair $P = (X, P)$ where X is a set (usually called the *point set* or the *ground set*) and P is a reflexive,

antisymmetric, and transitive binary relation on X. For two points $x, y \in X$ we write $x \leq y$ if $(x, y) \in P$, and $x < y$ if $x \leq y$ and $x \neq y$.

If $x < y$ then we say that x, y form a *comparable pair* or, in abuse of notation, a *relation*. If neither $x \leq y$ nor $y \leq x$ then we say that x and y are *incomparable* and write $x \| y$. We denote by $inc(x)$ the set of all points that are incomparable to x. Moreover we say that x is *covered* by y (also y *covers* x, or (x, y) is a *cover relation*) if $x < y$ and there is no point z for which $x < z$ and $z < y$ holds. In this case we write $x <: y$. On the other hand, if $x < y$ but (x, y) is *not* a cover relation, we write $x \ll y$ and call it a *forced relation*.

A subset $\{x_1, \ldots, x_k\} \subseteq X$ is called a *chain* if all pairs x_i, x_j are comparable. It is called an *antichain* if all pairs are incomparable. In the case of a chain we write $[x_1, \ldots, x_k]$ if $x_1 > \cdots > x_k$. If the complete point set X is a chain, then P is called a *linear* (or *total*) order.

A point x is called *maximal* (respectively, *minimal*) if there is no point y with $x < y$ (respectively, $y < x$). A chain is called *maximal* if it cannot be extended to a larger chain. It is called *maximum* if no other chain contains more points. The same terminology applies for antichains. The *height* of a poset is the number of points in a maximum chain, the *width* is the number of points in a maximum antichain.

A poset is called *ranked* if there exists a *rank function* r that assigns non–negative integers to the points in such a way that for some minimal point x we have $r(x) = 0$, and whenever $x <: y$ we have $r(x) + 1 = r(y)$. If furthermore all minimal points and all maximal points have the same rank respectively (which implies that every maximal chain is also maximum) then we say that the poset is *graded*.

With a poset $P = (X, P)$ we associate the *comparability graph* G and the *cover graph* G'. The vertex sets of both graphs are given by X, the edges in G are formed by the comparable pairs in P, while the edges in G' are formed by the cover pairs in P. The degree of a point is defined as the degree of the corresponding vertex in the comparability graph.

The *dimension* of a poset P, as defined by Dushnik and Miller [17], is the least integer k such that the relation on P is equal to the intersection of k linear orders on the same point set.

Unless stated otherwise, we will be dealing with *labeled* graphs and posets. Without loss of generality, every graph on n vertices will have vertex set $[n]$ and every poset on n points will have point set $[n]$.

Chapter 1

Introduction

In this chapter we give a detailed introduction to the various themes that we merely touched upon in the preface. In the first section we focus on graphs. Here we will consider the enumeration of graph properties as well as results about the structure of random graphs. We shall explain how they can be used in the design and analysis of algorithms. Finally we will deal with evolution processes. In the second section we focus on partially ordered sets and their enumeration, as well as several models of random posets. Let us mention at this point that although we give far more details here than in the preface, there is still not enough space for a truly adequate treatment, so we refer the interested reader to the works by Steger [59] and Brightwell [11].

The reader who is familiar with the above material should skip the first two sections and start with the third section. It begins with an account of the results of Dhar and Kleitman and Rothschild that represent the immediate surroundings of our results. The latter are then stated, explained and illustrated.

1.1 Counting graphs

Hereditary properties of graphs. Many interesting families of graphs are defined by requiring that every member of the family must (or, to the contrary, must not) have one or several particular substructures. As an illustration, think of graphs possessing a path which runs through all the vertices, graphs which contain no odd cycles as subgraphs or graphs that contain no K_5 and $K_{3,3}$ as minors. If one now tries to enumerate such a family of graphs, then one realizes that the decision whether a particular edge belongs to the graph or not is in fact highly dependent on previous decisions, because the structural properties defining the family have to be satisfied.

Let us consider such families of graphs in more detail. A *graph property* $\mathcal{P}$ is a union of isomorphism classes of finite graphs, and in order to avoid trivialities, we will always

assume that the properties we deal with contain infinitely many non-isomorphic graphs, but, at least for all large enough n, not all graphs on n vertices. A property is called *hereditary* if it is closed under taking induced subgraphs.

When considering a graph property, two natural questions arise, the first of which is: *how many graphs 'belong' to the property?* Again, for more precision, let $\mathcal{P}_n$ be the set of those graphs in $\mathcal{P}$ with vertex set $[n]$, then this first question asks to determine the sequence

$$(|\mathcal{P}_n|)_{n=1}^{\infty}.$$

Scheinerman and Zito [57] were able to show that for every hereditary property the rate of growth (i.e. the above sequence) belongs to one of the following categories: constant, polynomially, exponentially, factorially, or superfactorially.

We will be interested in properties that are very rich, by which we mean that their rate of growth is close to maximal. Following the terminology used in [6], we define the *logarithmic density* of $\mathcal{P}$ to be

$$c_n(\mathcal{P}) := \frac{\log|\mathcal{P}_n|}{\binom{n}{2}}, \quad \text{in other words,} \quad 2^{c_n(\mathcal{P})\cdot\binom{n}{2}} = |\mathcal{P}_n|, \tag{1.1}$$

and denote the *asymptotic logarithmic density* of $\mathcal{P}$ by

$$c(\mathcal{P}) := \lim_{n\to\infty} c_n(\mathcal{P}), \tag{1.2}$$

provided that the limit exists.

If $\mathcal{P}$ is also closed under taking (weak) subgraphs, then $\mathcal{P}$ is called *monotone* and here it makes sense to consider the following, second question: *how many edges can a graph in $\mathcal{P}$ possibly have?* For a hereditary property it might very well be, for example, that $K_n \in \mathcal{P}_n$, so in order to pose a generally sensible question we need to rephrase it a bit: try to maximize the number of edges a graph $G = ([n], E)$ can have such that there exists a graph $\tilde{G} = ([n], \tilde{E})$ with $E \cap \tilde{E} = \emptyset$ so that for all $E' \subseteq E$ the graph $G' = ([n], \tilde{E} \cup E')$ always lies in $\mathcal{P}$. Then $e_n(\mathcal{P}) := |E|/\binom{n}{2}$ is called the *(normalized) size* of $\mathcal{P}$ while

$$e(\mathcal{P}) := \lim_{n\to\infty} e_n(\mathcal{P}), \tag{1.3}$$

denotes the *asymptotic size* of $\mathcal{P}$, provided that the limit exists. $\tilde{G}$ is called the *supplemental graph* to the *extremal graph* G. Since there are $e_n(\mathcal{P})\binom{n}{2}$ different edges that can be added to $\tilde{G}$ while still remaining in $\mathcal{P}$, we have

$$|\mathcal{P}_n| \geq 2^{e_n(\mathcal{P})\binom{n}{2}}, \quad \text{hence} \quad c(\mathcal{P}) \geq e(\mathcal{P}). \tag{1.4}$$

Trivially, every hereditary property $\mathcal{P}$ can be described in terms of forbidden induced subgraphs: from the set of all graphs not contained in $\mathcal{P}$ it suffices to forbid exactly those

that are minimal (w.r.t. taking induced subgraphs). In complete analogy, a monotone property can be defined by forbidden weak subgraphs. We will first concentrate on the case where only one subgraph is forbidden. Denote by $\mathcal{F}orb_n(H)$ the family of all graphs with vertex set $[n]$ that do not contain H as a (weak) subgraph, and by $\mathcal{F}orb_n^*(H)$ the family of all graphs with vertex set $[n]$ that do not contain H as an induced subgraph. Also define $ex_n(H)$ and $ex_n^*(H)$ as follows:

$$ex_n(H) := e_n(\mathcal{F}orb_n(H)) \cdot \binom{n}{2}, \qquad ex_n^*(H) := e_n(\mathcal{F}orb_n^*(H)) \cdot \binom{n}{2}.$$

Turán's theorem [63, 64], which is widely considered to be the starting point of extremal graph theory, states that

$$ex_n(K_\ell) = \left(1 - \frac{1}{\ell - 1}\right) \cdot \frac{n^2}{2} + O(n).$$

Generalizing this, Erdős, Stone and Simonovits [31, 30] discovered that it is actually the chromatic number of H which determines the magnitude of $ex_n(H)$:

$$ex_n(H) = \left(1 - \frac{1}{\chi(H) - 1}\right) \cdot \frac{n^2}{2} + o(n^2).$$

The following result brings us back to the first question of determining $c(\mathcal{F}orb_n(H))$ and shows how closely extremal graph theory is connected with enumeration. In 1986, Erdős, Frankl and Rödl [21] proved that

$$|\mathcal{F}orb_n(H)| = 2^{(1-\frac{1}{\chi(H)-1})\cdot\frac{n^2}{2}+o(n^2)},$$

hence for any graph H with $\chi(H) \geq 3$

$$|\mathcal{F}orb_n(H)| = 2^{ex_n(H)\cdot(1+o(1))}, \tag{1.5}$$

or, equivalently,

$$c(\mathcal{F}orb_n(H)) = e(\mathcal{F}orb_n(H)) = 1 - \frac{1}{\chi(H) - 1}.$$

Analogous results for the case when H is bipartite are not known. Partial results for the case of $H = C_k$ were given by Kleitman and Wilson [41] and Kreuter [45].

In the case of forbidden *induced* subgraphs Prömel and Steger discovered similar results. They first defined the graph parameter $\tau(H)$ which is a generalization of the chromatic number $\chi(H)$ and the clique covering number $\sigma(H)$. Namely, denote by $\tau(H)$ the greatest positive integer k for which there exists an integer $k' \in \{0, \ldots, k-1\}$ such that for every $(k-1)$–colourable graph H' in which k' colour classes are replaced by cliques we have

$H' \in \mathcal{F}orb_n^*(H)$. Now we can state the results by Prömel and Steger [52, 53] where τ plays the role of χ:

$$ex_n^*(H) = \left(1 - \frac{1}{\tau(H)-1}\right) \cdot \frac{n^2}{2} + o(n^2),$$

$$|\mathcal{F}orb_n^*(H)| = 2^{(1-\frac{1}{\tau(H)-1})\cdot\frac{n^2}{2}+o(n^2)},$$

thus for any graph H with $\tau(H) \geq 3$

$$|\mathcal{F}orb_n^*(H)| = 2^{ex_n^*(H)\cdot(1+o(1))}, \tag{1.6}$$

and again, this is equivalent to

$$c(\mathcal{F}orb_n^*(H)) = e(\mathcal{F}orb_n^*(H)) = 1 - \frac{1}{\tau(H)-1}.$$

Later, Alekseev [3] and Bollobás and Thomason [8] independently proved that this also holds when forbidding whole families of induced subgraphs. Thus for every hereditary property $\mathcal{P}$ the asymptotic logarithmic density and the asymptotic size exist and are in fact equal.

Counting and typical structure. One of the most interesting aspects of the above results as expressed in (1.5) or (1.6) is that not only do they compute the asymptotic number of elements but they also show that a substantial proportion of the class $\mathcal{P}$ in question are from a well–structured subclass $\mathcal{P}'$, namely the class of those graphs that are generated by combining a subgraph of the extremal graph with the supplemental graph. These graphs will obviously be very similar in structure, and hence we are tempted to believe that this structure is not only representative for $\mathcal{P}'$ but also for $\mathcal{P}$.

Naturally the subject of counting elements in a set and determining the structure of a 'typical' element are closely related: if we know that a typical element possesses a certain structure, then this must mean that most elements in $\mathcal{P}$ have this structure, hence it might suffice to enumerate the subclass of elements with exactly this structure – and enumerating them might actually be a lot easier because of their structure. Or, vice versa, if we find a 'big' subclass $\mathcal{P}'$ within the class $\mathcal{P}$, then the structure defining $\mathcal{P}'$ might well be that of a typical element. Obviously here the terms *typical* and *big* are still far too vague and we shall make them precise as soon as possible.

One way to prove that an element typically possesses a certain structure is to choose an element at random from the set in question and to show that with very high probability it has the property – this is the underlying theme of the theory of random graphs. It first appeared as a tool for a probabilistic construction, where Erdős [20] proved that for every

n there exists a graph on n vertices which has neither a clique nor an independent set on $2\log n$ vertices, and in fact the proof shows that *almost all* graphs on n vertices satisfy this condition. Since this is the first result of this type that we are mentioning, let us state it again with more precision: denote by $\mathcal{G}_n$ the set of all labeled graphs with vertex set $[n]$ and let $\mathcal{G}'_n$ be the set of those graphs in $\mathcal{G}_n$ that have neither a clique nor an independent set of order $2\log n$. Then

$$\frac{|\mathcal{G}'_n|}{|\mathcal{G}_n|} \longrightarrow 1 \qquad \text{as } n \to \infty.$$

Equivalently, introducing the language of classical random graph theory, let $G_{n,p}$ be the random graph obtained by starting with the empty graph on the vertex set $[n]$ and then inserting each edge *independently* with probability p. Obviously for a fixed graph G the chance that $G_{n,\frac{1}{2}} = G$ is the same for all graphs G, namely $2^{-\binom{n}{2}}$, which is the reason why $G_{n,\frac{1}{2}}$ is referred to as the *uniform model*. So Erdős' result stands as

$$\text{Prob}\Big[\ \omega(G_{n,\frac{1}{2}}) < 2\log n \text{ and } \alpha(G_{n,\frac{1}{2}}) < 2\log n\Big] \longrightarrow 1 \qquad \text{as } n \to \infty.$$

We will later come back to the random graph $G_{n,p}$ when we deal with the evolution process of $G_{n,p}$, but let us already at this point mention that the theory of random graphs and in particular the $G_{n,p}$ model has proved to be a great success (for an overview we refer the reader to the monograph by Bollobás [5] or the survey article by Karoński [34]), provoking the development of new techniques on the boundary between combinatorics and probability theory such as large deviation inequalities in general and martingale inequalities in particular. However, most of these methods rely heavily on the fact that the random structures are generated by a series of independent events. As we saw already at the beginning of this chapter, unfortunately this is no longer true when seeking random elements from classes such as $\mathcal{F}orb_n(H)$ or $\mathcal{F}orb^*_n(H)$, hence most of the above mentioned methods remain ineffective here.

Nevertheless there are results about the structure of almost all graphs in these classes. The starting point here is the class $\mathcal{F}orb_n(K_3)$. In 1976, Erdős, Kleitman and Rothschild [23] proved that

$$\textit{almost all triangle–free graphs are 2-colourable.} \tag{1.7}$$

This result was generalized in two ways: for any constant ℓ, Lamken and Rothschild [47] showed that almost all graphs in $\mathcal{F}orb_n(C_{2\ell+1})$ are 2-colourable, and Kolaitis, Prömel and Rothschild [44] proved that almost all graphs in $\mathcal{F}orb_n(K_{\ell+1})$ are ℓ-colourable. A further extension, re–uniting the latter two results, was obtained by Prömel and Steger [54] who characterized the family of forbidden graphs H for which such a statement holds. Denote

by $\mathcal{C}ol_n(\ell)$ the family of all labeled ℓ-colourable graphs with vertex set $[n]$. Then for a graph H with $\chi(H) = \ell + 1 \geq 3$

$$|\mathcal{F}orb_n(H)| = (1 + o(1)) \cdot |\mathcal{C}ol_n(\ell)|$$

if and only if H contains an edge whose deletion from H causes $\chi(H)$ to decrease.

Using typical structure. It is clear that information about the space of input instances can be an advantage when designing algorithms that take inputs from that space. On the one hand, the *worst–case analysis* of an algorithm will always concentrate on the worst–possible problem instance when computing the running time, and we know from the theory of NP–completeness that (under the assumption P≠NP) with this measure of performance most of the optimization problems cannot be solved in polynomial time. On the other hand, such worst–case pessimism might be misleading and it might in fact be much more realistic to attempt a *probabilistic analysis* of an algorithm. This analysis assigns some weight to every instance, turning our input space into a probability space and the running time into an *expected running time*. For practical applications, the assigned weights would be chosen to correspond to the frequency of occurrence of the problem instance, and the success depends on how well the probability space can model the space of real problem instances.

One of the first examples of this approach is the simplex algorithm. The well-known article by Klee and Minty [37] shows that its worst–case running time is exponential, but using a certain pivot-rule and considering a natural probability distribution it can be shown to have a polynomial expected running time [10]. Another example is the graph colouring problem. In 1984 Karp [35] asked for an algorithm that optimally colours $G_{n,p}$ in polynomial expected time. This central question is still open, but partial results by Kučera [46], Dyer and Frieze [18], and Prömel and Steger [55] are known. For example, the aforementioned paper presents an algorithm that, for some constant ℓ, colours any graph in $\mathcal{F}orb_n(K_\ell)$ with the optimal number of colours in expected linear time.

The point we want to make in this section, namely how knowledge about the structure of typical inputs helps to design algorithms, is the basis of all the algorithms mentioned above. It is probably best illustrated by yet another example: an algorithm by Wilf [66] which, for some constant ℓ, takes as input a graph G drawn uniformly at random from $\mathcal{G}_n$ and either produces a correct ℓ-colouring of G or correctly decides that G is not ℓ-colourable – and does so in constant(!) expected time. The algorithm is extremely simple: while reading the input vertex by vertex, it keeps looking for cliques on $\ell + 1$ vertices. As soon as it finds one, it stops and returns the answer "no". If, even after the whole input has been read, no such clique has been found, it uses an arbitrary brute–force method to come to a conclusion. The reason why the algorithm can be expected to be so fast, is also simple: we know that the number of $K_{\ell+1}$-free graphs is small when compared to the number of all graphs, hence with a high chance the input graph has a clique of size

$\ell+1$ already at the beginning of the input string. To be a bit more precise, calculate the probability that the above algorithm has *not* found the required clique after reading the first i vertices and the edges between them. This can only happen if the subgraph induced by these first i vertices is in $\mathcal{F}orb_i(K_{\ell+1})$, and there are $2^{\binom{n}{2}-\binom{i}{2}}$ different ways in which such an initial subgraph can be extended to a graph on n vertices. Using the bound given in (1.5), the probability is given by

$$2^{-\binom{n}{2}} \cdot |\mathcal{F}orb_i(K_{\ell+1})| \cdot 2^{\binom{n}{2}-\binom{i}{2}} \leq 2^{(1-\frac{1}{\ell})\cdot\binom{i}{2}+o(i^2)-\binom{i}{2}}$$

Hence if we allow $O(i^{\ell+1})$ time to look for the clique on the first i vertices and $O(\ell^n) = O(2^{n\log n})$ time for the brute–force procedure at the end, the expected running time is bounded from above by

$$\sum_{i=0}^{n-1} O((i+1)^{\ell+1}) \cdot 2^{-\frac{1}{\ell}\binom{i}{2}+o(i^2)} + O(2^{n\log n}) \cdot 2^{-\frac{1}{\ell}\binom{n}{2}+o(n^2)} = O(1).$$

Evolving typical structure. In the last paragraph we saw how information about the structure of typical elements from a given probability space can be used when designing algorithms that take inputs from that space. This brings us to the following question: *how does the structure of a typical graph gradually change as the probability distribution changes?* This is the key question when investigating the evolution process of the random graph $G_{n,p}$.

In a series of seminal papers Erdős und Rényi [24, 25, 26, 27, 28, 29] investigated the evolution of random graphs, thereby marking the starting point of the theory of random graphs. For many graph properties Erdős und Rényi noticed so-called phase transition phenomena: *until p reaches a certain threshold, $G_{n,p}$ will almost surely not have the property in question, whereas beyond this threshold suddenly it will almost surely possess it.* Here *almost surely* stands for "with probability tending to 1, as $n \to \infty$". (For brevity we will omit this phrase in the following.) This area has been and still is a flourishing field of research. Here we only briefly outline the main evolutionary phases as described in [25] and again refer the reader for more details and further references to [5, 13, 34].

Phase 1: $p = o(1/n)$. Here $G_{n,p}$ is a forest. Trees on k vertices appear for the first time when $p = \Theta(n^{-k/(k-1)})$.

Phase 2: $p \sim c/n$ where $0 < c < 1$. All connected components of $G_{n,p}$ are either trees or trees with one additional edge. The largest component is a tree of order $\Theta(\log n)$.

Phase 3: $p \sim 1/n$. Here the largest component, now called the *giant component*, suddenly has order $n^{2/3}$. However by 'slowing down' the growth of p one sees that the evolution of $G_{n,p}$ is rapid but continuous.

Phase 4: $p \sim c/n$ where $c > 1$. This phase can roughly be described as the merging of the small components (most of which are trees) into the giant component, which now has $\Theta(n)$ vertices.
Phase 5: $p \sim c \log n/n$ where $c \geq 1$. Here $G_{n,p}$ is about to become connected.
Phase 6: $p \sim \omega(n) \log n/n$ where $\omega(n) \to \infty$. Here $G_{n,p}$ is connected and all the vertices have asymptotically the same degree.

Now we turn our attention to what could be called *constrained* graphical evolution. We have already seen in (1.7) that almost all triangle-free graphs are bipartite. But is this also true if we focus our attention on an early evolutionary stage? Let us choose our random graph uniformly, not from $\mathcal{F}orb_n(K_3)$ but from $\mathcal{F}orb_{n,m}(K_3)$, the set of all triangle–free graphs with vertex set $[n]$ that have m edges. As already pointed out before, due to the lack of independence most of the methods used to obtain results about $G_{n,p}$ are ineffective when one considers random graphs from $\mathcal{F}orb_n(H)$. In spite of this, Prömel and Steger [56] succeeded in showing that there exist constants $c_1, c_2, c_3 > 0$ such that

$$\text{Prob}\Big[\ G \text{ is bipartite} \mid G \in \mathcal{F}orb_{n,m}(K_3)\Big] = 1 - o(1), \qquad \text{for all } m \geq c_1 n^{7/4} \log n,$$

$$\text{Prob}\Big[\ G \text{ is bipartite} \mid G \in \mathcal{F}orb_{n,m}(K_3)\Big] = o(1), \qquad \text{for all } c_2 n \leq m \leq c_3 n^{3/2}. \quad (1.8)$$

In the above discussion of the evolution of $G_{n,p}$ we saw that $p = 1/n$ marks the threshold of any cycle in $G_{n,p}$, from which one deduces that there exists a constant c_4 such that

$$\text{Prob}\Big[\ G \text{ is bipartite} \mid G \in \mathcal{F}orb_{n,m}(K_3)\Big] = 1 - o(1), \qquad \text{for all } m \leq c_4 n.$$

Interestingly, this indicates that the evolution of the random triangle-free graph exhibits *two* phase transitions with respect to being bipartite: first it is almost surely bipartite, then it is not, and then it is once again. The question remains open in the interval $n^{3/2} \ll m \ll n^{7/4} \log n$. Recently Łuczak [49] made some progress by showing that in this range with high probability a random graph is nearly bipartite (i.e. it can be made bipartite by removing very few edges).

1.2 Counting posets

Now we move from counting graphs to counting partially ordered sets. Recall that a graph is nothing but a set with a binary relation. We increase the complexity of the setting ever so slightly by requiring that the relation be reflexive, antisymmetric and transitive. The result is a *partially ordered set* or, as abbreviated, a *poset*. But already this small increase in structure makes counting a much harder job: When trying to count posets in the same

way as we counted graphs, we immediately notice the same obstacle as when counting graphs with forbidden subgraphs, namely the loss of independence: once we have decided that the relations $x < y$ and $y < z$ are to be present, we have no choice but to insert the relation $x < z$ in order to guarantee transitivity.

To begin with, let $\mathcal{P}$ be an arbitrary family of labeled posets, and denote by $\mathcal{P}_n$ the set of those posets in $\mathcal{P}$ with point set $[n]$. Again, just as with hereditary graph properties, we will be interested in determining the structure of a typical element from $\mathcal{P}_n$ as well as finding an estimate for the cardinality[1] of $\mathcal{P}_n$. Again we will content ourselves with determining the logarithm of $|\mathcal{P}_n|$, however this time with a different normalizing factor. Let

$$c_n(\mathcal{P}) := \frac{\log |\mathcal{P}_n|}{n^2}, \quad \text{in other words,} \quad 2^{c_n(\mathcal{P})\cdot n^2} = |\mathcal{P}_n| \tag{1.9}$$

and define

$$c(\mathcal{P}) := \lim_{n\to\infty} c_n(\mathcal{P}). \tag{1.10}$$

Now denote by $\mathcal{P}$ the family of *all* labeled posets. Let us first look for an easy lower bound on $|\mathcal{P}_n|$. It will suffice to consider posets with height 2, so fix two antichains X and Y, each on $\frac{n}{2}$ points. In this case there are no problems with transitivity, so decide *independently* for each pair $(x, y) \in X \times Y$ whether or not $x < y$ should hold. Hence

$$|\mathcal{P}_n| \;\geq\; 2^{\frac{n^2}{4}}.$$

Notice that here we didn't even care for the $\binom{n}{n/2}$ ways in which we could have chosen the points which go to, say, X.

Upper bounds are much harder to obtain. In 1970, Kleitman and Rothschild [38] first gave the following bound:

$$|\mathcal{P}_n| \;\leq\; 2^{n^2/4+O(n^{3/2}\log n)}, \tag{1.11}$$

hence

$$c(\mathcal{P}) = \frac{1}{4}.$$

A few years later [39], they were able to compute the exponent much more precisely:

$$\log |\mathcal{P}_n| \;=\; \frac{n^2}{4} + \frac{3n}{2} + O(\log n). \tag{1.12}$$

[1] It is worth mentioning that $|\mathcal{P}_n|$ is equal to the number of labeled T_0-topologies on n elements – see e.g. Birkhoff [4].

The technique used by Kleitman and Rothschild in the proofs of these two papers has become known as *the Kleitman-Rothschild method*. Again, for some class $\mathcal{C}$, it emphasizes the close connection between enumerating $\mathcal{C}$ and determining the structure of a typical element in $\mathcal{C}$. In rough terms, it can be described as follows: partition $\mathcal{C}$ into a number of subclasses such that (i) all but one class (which we will denote by $\mathcal{C}_0$) are negligible in size and (ii) $\mathcal{C}_0$ has a nice structure which easily allows enumeration. This method was in fact used in the proofs of most of the results that we mentioned in Section 1.1.

Let us now turn our attention to the particular class $\mathcal{C}_0$ in the paper mentioned last which established the bound in (1.12). The posets in $\mathcal{C}_0$ all consist of three disjoint antichains X_1, X_2 and X_3, where both X_1 and X_3 have roughly $n/4$ points, all points in X_3 are above all points in X_1, and whenever $x < y$ for two points $x \in X_i$ and $y \in X_j$ we must have $i < j$. Following the terminology used in [12], such a poset is called a 3-*layer poset*. So Kleitman and Rothschild [39] not only proved (1.12) but also the following theorem.

Theorem 1.2.1 *Almost all posets are* 3*-layer posets.* 2

Admittedly the 3-layer structure seems at first glance surprisingly 'narrow minded' – we shall demonstrate later (see 1.14) that there really do exist very many such posets. The above result is nevertheless a bit of a disappointment, since it somewhat questions the validity of the uniform model for random posets as an interesting area of investigation. As Brightwell [11] puts it:

> [Theorem 1.2.1] has been widely viewed as a negative result. It says that almost every partial order has height three, and even more, is *ranked*, that is, all the maximal chains have the same length (namely, three). Somehow this does not conform to the practising mathematician's view of what a "typical" partial order ought to look like.

This explains, at least in parts, the development of other models of random partially ordered sets, which we briefly introduce in the following. But before we get there let us close this paragraph with a few tangential remarks. Prömel [51] was able to prove that almost all posets have no nontrivial automorphisms, and that hence dividing the number of labeled posets by $n!$ gives the approximate number of *unlabeled* posets. Möhring [50] showed that almost all comparability graphs are uniquely partially orderable, from which he deduced that the number of labeled comparability graphs is approximately equal to half of the number of labeled posets (because for every poset its dual gives rise to the same comparability graph). Moreover, it is worth mentioning that if one considers special classes of posets, then it is often possible to enumerate them more precisely; especially in those cases where one has a sufficiently strong structural theorem. We mention a few

examples: Klotz and Lucht [43] and Kleitman and Winston [42] give estimates for the number of lattices. Klarner [36] uses the generating function technique to enumerate graded posets. Interval orders have been counted by Hanlon [33], while Stanley [58] has enumerated series–parallel posets. An overview can be found in the survey article by El-Zahar [19].

Random partially ordered sets. Here we merely sketch an introduction to the three most prominent models of random partially ordered sets other than the uniform model: random bipartite orders, random graph orders and random k-dimensional orders. More details and many more references can be found in the survey articles by Brightwell [11], Trotter [62], and Winkler [67].

We start by pointing out that there is at least one advantage in the uniform model: we can now obtain results about the uniform model by working in a different model of a random poset where relations are mutually independent. Namely, construct a random poset by taking a 3-layer poset and letting each relation $x_1 < x_2$ or $x_2 < x_3$ (where $x_i \in X_i$) exist with probability $\frac{1}{2}$. Then Theorem 1.2.1 states that if a property is possessed almost surely by this random poset, then the same will hold for a random poset from the uniform model. Having recovered independence, the model of a random layer poset can of course be considered more generally with each relation holding with probability $p = p(n)$.

Intensive research has been conducted on the case of random 2-layer posets, which are also called *random bipartite orders*. Here the two antichains X and Y have the same number of elements, say n, and each relation $x < y$ holds with probability p. We restrict ourselves to mentioning just one aspect in this area. Determining the height of a random bipartite order is of course trivial, and the same applies for the width if one focuses attention on the range of p where $n \cdot p - \log n \to \infty$ since here the random bipartite graph almost surely has a complete matching. Far more interesting is the question about the dimension of a random bipartite order. It is known [16] that for every poset P one has $\dim(P) \leq \text{width}(P)$. Furthermore a result by Füredi and Kahn [32] states that $\dim(P) \leq 50\Delta(\log \Delta)^2$ where Δ denotes the maximum degree of P. However, for quite some time it was not known whether the dimension of a poset could exceed $\Delta + 1$, but then Erdős, Kierstead and Trotter [22] proved that for almost all random bipartite orders this is indeed the case, provided p lies in the right range.

A random bipartite order can of course be considered as induced by a random bipartite graph. This approach, i.e. the connection between random graphs and random posets, can be extended to general graphs in the following way. Suppose we were given a graph G together with a linear order $\prec$ on the vertex set of G. Consider first only those relations $i \prec_G j$ where i and j are vertices, $i \prec j$ and $\{i, j\}$ is an edge in G. The associated poset $<_G$ on the vertex set is then defined as the transitive closure of $\prec_G$. Then a *random graph order*, usually denoted by $P_{n,p}$, is the poset associated with $G_{n,p}$ and an arbitrary linear order in the above way. From the earlier discussion of $G_{n,p}$, it is intuitively obvious, at

least if p is a constant, that there is a big difference between a typical random poset $P_{n,p}$ and a typical random poset from the uniform model: the height of $P_{n,p}$ is much larger and its width is much smaller. Random graph orders were introduced as such by Albert and Frieze [2] even though they had implicitly appeared earlier elsewhere. As an example of an interesting problem on random graph orders, we briefly address what seems to be an elementary question: for a given p, *how many relations does $P_{n,p}$ have?* We use this question to illustrate the already mentioned fact that knowledge of the structure of random elements helps with the running time analysis of algorithms. Suppose one wanted to push the idea of "sorting in parallel" even further than the well–known AKS sorting–networks [1] which take $O(\log n)$ rounds, each with $O(n)$ comparisons. Namely, one attempts to make m comparisons in each round for some number $n \ll m \ll \binom{n}{2}$. Suppose that the keys to be sorted are given by $x_1 < \cdots < x_n$. Choose p such that $p \cdot \binom{n}{2} = m$, pick a random graph $G_{n,p}$ on the keys and compare adjacent vertices. Now it is the underlying linear order that we want to recover from the information obtained via (i) the comparisons *and* (ii) the transitive information that can be deduced from them. Hence the crucial point that determines the expected running time of this approach is: *the more relations we have in* $<_G$ the faster the algorithm will be. In this context, Bollobás and Brightwell [7] could show that for some range of p the number of relations in $P_{n,p}$ is $\binom{n}{2} - o(n^2)$.

The last model of random posets that we are going to mention here is that of *random k-dimensional orders*. The (original) definition of dimension lends itself rather nicely to the construction of a probabilistic model: choose k linear orders (on the set $[n]$) uniformly and independently at random from the set of all linear orders and let $P_k(n)$ be the intersection of these linear orders. Using equivalent definition of dimension one obtains equivalent models of k-dimensional orders, some of which are very attractive because they exhibit some kind of independence. Random k-dimensional orders were introduced by Winkler, but again, they had appeared earlier in disguise. We only mention one example: determining the expected height (or width) of $P_2(n)$ is known as Ulam's problem; it is easily seen to be equivalent to asking for the expected length of the longest increasing subsequence in a random permutation. This problem turned out to be surprisingly difficult, but in 1988 Bollobás and Winkler [9] proved that for every k there exists a constant c_k such that the expected height of $P_k(n)$ is asymptotic to $c_k n^{1/k}$. It is known [48, 65] that $c_2 = 2$ and that $\lim_{k\to\infty} c_k = e$, but for $k > 2$ the value of c_k is not known.

1.3 Counting posets with a given number of relations

We are now in the position where we take the last turn and move to the main theme of this thesis. Recall Theorem 1.2.1 which states that almost all posets are 3-layer posets, and compare this with the situation we encountered when confronted with results on the typical structure of triangle-free graphs for the first time in this chapter. Here we first

saw in (1.7) that almost every triangle-free graph is bipartite. But then (1.8) provided further insight by showing that this is "only" true because it is true for dense graphs and because almost every triangle-free graph is dense. And, most importantly, it showed that the situation does indeed look very different if one looks at earlier evolutionary stages of the class of triangle-free graphs, i.e. at typical graphs in $\mathcal{F}orb_{n,m}(K_3)$ with $m \ll n^{3/2}$.

Now view the situation of partially ordered sets from the same angle: almost all n-point posets are 3-layer posets. But what does a typical poset on n points *with a given number of relations* look like? And: *how many such posets are there?*

Definitions. Let us introduce some notation. Denote by $\mathcal{P}_{n,d}$ the family of all labeled posets with point set $[n]$ and $d \cdot n^2$ relations, and let

$$\mathrm{c}(d) := \lim_{n\to\infty} \frac{\log |\mathcal{P}_{n,d}|}{n^2},$$

provided the limit exists. In order to compute $\mathrm{c}(d)$ for every $0 < d < \frac{1}{2}$ we will construct a family $\mathcal{P}'$ of posets, such that

$$\mathcal{P}'_n \subseteq \mathcal{P}_{n,d}$$

and

$$\lim_{n\to\infty} \frac{\log |\mathcal{P}'_n|}{n^2} = \lim_{n\to\infty} \frac{\log |\mathcal{P}_{n,d}|}{n^2} = \mathrm{c}(d).$$

A family $\mathcal{P}'$ satisfying the above properties will be called *d-significant* or, if the value of d is either irrelevant or obvious from the context, simply *significant*. It is important to note that the methods used to determine $\mathrm{c}(d)$ don't seem to be strong enough to give *almost–all*–type results for $\mathcal{P}_{n,d}$. We conjecture that in fact almost all posets in $\mathcal{P}_{n,d}$ lie in $\mathcal{P}'$ but will not attempt to prove it.

Strangely enough, the question of determining $\mathrm{c}(d)$ seems to have first arisen in quite a different context: in 1978, Dhar [14] suggested that partial orders can represent the states of a certain model of lattice gas with energy proportional to the number of comparable pairs in the order:

> [$\mathrm{c}(d)$] may be considered to be the entropy function of an interacting "lattice gas" with long–range three–body interaction, in which case, the lattice gas undergoes a first order phase transition as a function of the "chemical activity" of the gas molecules, the value of the chemical activity at the phase transition being 1.

Returning to the mathematics, recall from (1.11) that for any $d \in (0, \frac{1}{2})$,

$$\mathrm{c}(d) \;\le\; \frac{1}{4}. \tag{1.13}$$

Dhar was able to show that the function $\mathrm{c}(d)/d$ is monotone nonincreasing while the function $\mathrm{c}(d)/(1-d)$ is monotone nondecreasing, from which he deduced that $\mathrm{c}(d)$ is continuous. Moreover he proved that

$$\text{for } \frac{1}{8} \le d \le \frac{3}{16} \ : \qquad \mathrm{c}(d) \equiv \frac{1}{4}. \tag{1.14}$$

This is in fact straightforward: in view of the upper bound in (1.13) it suffices to construct $2^{n^2/4}$ posets with $d \cdot n^2$ relations for each d in the given range. This can be done as follows. Let x be the smaller of the two solutions of

$$d = -x^2 + \frac{1}{2} \cdot x + \frac{1}{8},$$

$$\text{i.e.} \quad x = \frac{1}{4} - \sqrt{\frac{3}{16} - d}. \tag{1.15}$$

If $d \in [\frac{1}{8}, \frac{3}{16}]$, then clearly x lies in the range $[0, \frac{1}{4}]$. Now consider the family of 3-layer posets with layers X_1, X_2 and X_3 where

$$|X_1| := \left(\frac{1}{2} - x\right) \cdot n, \quad |X_2| := \frac{1}{2} \cdot n, \quad |X_3| := x \cdot n.$$

As to the relations, we insert $\frac{1}{2}|X_i||X_{i+1}|$ relations between X_i and X_{i+1} for $i = 1, 2$, and (as always in a 3-layer poset) all possible relations between X_1 and X_3. This gives a total of

$$\left[\frac{1}{2} \cdot \left(\frac{1}{2}x + \frac{1}{2}\left(\frac{1}{2} - x\right)\right) + x \cdot \left(\frac{1}{2} - x\right)\right] \cdot n^2 = d \cdot n^2$$

relations, and there are

$$\binom{\frac{1}{2}x \cdot n^2}{\frac{1}{4}x \cdot n^2} \cdot \binom{\frac{1}{2}(\frac{1}{2} - x) \cdot n^2}{\frac{1}{4}(\frac{1}{2} - x) \cdot n^2} = 2^{[\frac{1}{2}x + \frac{1}{2}(\frac{1}{2} - x)]n^2 + o(n^2)} = 2^{n^2/4 + o(n^2)}$$

ways to choose the relations between consecutive layers.

Dhar [14, 15] also obtained several upper bounds for $\mathrm{c}(d)$, from which he concluded on the one hand that

$$\text{whenever either } d < 0.077 \text{ or } d > 0.35 : \quad \mathrm{c}(d) < \frac{1}{4}$$

and on the other hand that $\mathrm{c}(d)$ was nonanalytic. Dhar conjectured that the family of k-layer posets (see below for a formal definition) were rich enough as to dominate the set $\mathcal{P}_{n,d}$, provided that the number of layers (k), the layer sizes ($|X_i|$), and the density of relations between consecutive layers were chosen optimally and depending on d. In this

case, so he predicted, there would be intervals where p as a function of d would remain constant and $\mathrm{c}(d)$ would be a linear function in d. As we shall see, he was quite right.

In 1979, Kleitman and Rothschild [40] completed the description of the evolution process in the range of $0 < d \leq \frac{1}{8}$. They showed that in this range the family of bipartite posets consisting of two disjoint antichains of cardinality $\frac{n}{2}$ and dn^2 relations between the antichains is d-significant and deduced that

$$\text{for } 0 < d \leq \frac{1}{8}: \qquad \mathrm{c}(d) = \frac{1}{4} \cdot H(4 \cdot d), \tag{1.16}$$

where $H(x)$ denotes the entropy function (see page ix for an exact definition). Figure 1.1 illustrates both their result and that by Dhar; here the dashed line represents (1.14).

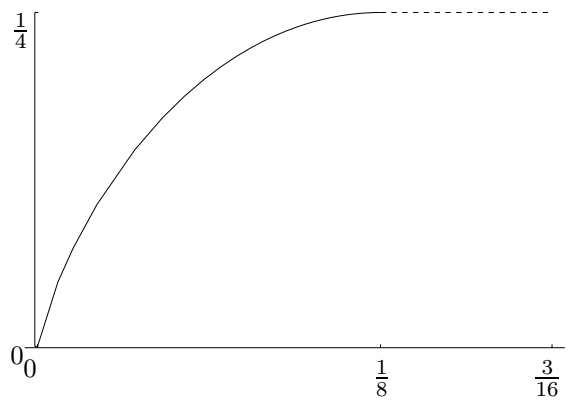

Figure 1.1: $\mathrm{c}(d)$ in the range $(0, \frac{3}{16})$

The concluding remarks in the paper by Kleitman and Rothschild as quoted in the preface now mark our starting point. In the remaining part of this chapter we explain the main results of this thesis, namely the description of the evolution process in terms of significant families and the behaviour of $\mathrm{c}(d)$ in the full range $0 < d < \frac{1}{2}$.

Definitions. Even though it seems clear how to extend the definition of a 3-layer poset to a k-layer poset in a natural way, let us state the formal definition. A poset $P = (X, P)$ is a *k-layer poset*, if there exists a partition of its point set $X = X_1 \cup \cdots \cup X_k$ into k disjoint antichains (*layers*) such that

$$\begin{aligned} x < y \text{ with } x \in X_i \text{ and } y \in X_j \quad &\Longrightarrow \quad i < j \\ \text{for every } i, j \text{ with } j > i+1: \quad x \in X_i,\ y \in X_j \quad &\Longrightarrow \quad x < y. \end{aligned}$$

For some constants $\lambda_1, \ldots, \lambda_k$ with $0 < \lambda_i < 1$ and $\sum_i \lambda_i = 1$ and a constant $0 \leq p \leq 1$, a k-layer poset is said to have *configuration* $(\lambda_1, \ldots, \lambda_k; p)$ if

$$|X_i| = \lambda_i \cdot n \qquad \text{for all } i \in [k]$$

and the number of relations between X_i and X_{i+1} is equal to

$$p \cdot |X_i| \cdot |X_{i+1}|$$

for all $i \in [k-1]$. A *configuration* is called *d-significant* if the family of all posets with this configuration is d-significant. Practising this terminology, we recall the results of Dhar and Kleitman and Rothschild:

$$\left(\frac{1}{2} - x, \frac{1}{2}, x;\ \frac{1}{2}\right) \text{ is } d\text{-significant for } d \in \left[\frac{1}{8}, \frac{3}{16}\right] \text{ where } x \text{ is chosen as in (1.15)},$$
$$\left(\frac{1}{2}, \frac{1}{2};\ 4d\right) \text{ is } d\text{-significant for } d \in \left(0, \frac{1}{8}\right].$$

Before we formally state our result let us explain it in a few words by illustrating what happens after d passes the point $d_3 := \frac{3}{16} = 0.1875$.

At this point our d-significant configuration is given by $(\frac{1}{4}, \frac{1}{2}, \frac{1}{4}; \frac{1}{2})$. As d increases, the density p of relations between consecutive layers increases, the layers X_1 and X_3 grow at the same rate while X_2 shrinks, and the function $\mathrm{c}(d)$ immediately begins to decrease.

When d reaches the point $d_4' := 0.2204..$, the pattern changes dramatically: p, having reached the value 0.5827.., now remains constant. A new fourth layer X_4 begins to grow, X_1 and X_2 shrink while X_3 grows. Perhaps most surprisingly, $\mathrm{c}(d)$ is in this range a *linear* decreasing function. This phase ends at the point $d_4 := 0.2536..$, when $|X_1| = |X_4|$ and $|X_2| = |X_3|$. Interestingly X_4 has now half the size of both X_1 and X_3 at the point d_4'.

We have tried to illustrate this in Figures 1.2, 1.3, and 1.4. After the point d_4' the evolution process repeats the same two–step manner an infinite number of times, as illustrated in Figure 1.5. Our main result can now be stated as follows.

Theorem 1.3.1 *There is an infinite sequence of constants*

$$0 = d_2 < d_3' < d_3 < d_4' < d_4 < \cdots < \frac{1}{2}$$

that describe the evolution process in the following way.

For every $d \in (d_k, d_{k+1}']$ there exists a d-significant configuration $(\lambda_1, \ldots, \lambda_k; p)$. As d increases from d_k to d_{k+1}'
(i) the density p increases,
(ii) the layer sizes λ_i change but remain symmetric: $\lambda_i = \lambda_{k-i+1}$.
(iii) For $k \geq 3$ the function $\mathrm{c}(d)$ is monotone decreasing.

For every $d \in (d_{k+1}', d_{k+1}]$ there exists a d-significant configuration $(\lambda_1, \ldots, \lambda_{k+1}; p)$. As d increases from d_{k+1}' to d_{k+1}
(iv) the density p remains constant,
(v) a new $(k+1)$st layer begins to grow.
(vi) At the end of this second phase the $(k+1)$st layer is half as big as the k-th layer at

the beginning of the phase. In general, the layer sizes at the end are the mean of those at the beginning: $\lambda'_i = (\lambda_{i-1} + \lambda_i)/2$, *denoting the layer sizes at* $d = d'_{k+1}, d_{k+1}$ *by* λ_i, λ'_i *respectively and setting* $\lambda_0 = \lambda_{k+1} = 0$.
(vii) The function $\mathrm{c}(d)$ *is linear in* d *and, for* $k \geq 3$, *monotone decreasing.*

In the statement of the above theorem we have been somewhat vague in that we merely stated that the constants d'_k and d_k existed. It would of course be much nicer if we could give an explicit formula – preferably not only for these constants but also for the density $p(d)$ and the layer sizes $\lambda_i(d)$ as functions of d, and, most importantly, the function $\mathrm{c}(d)$. This is possible, but the formulae involved turn out to be not very attractive.

Theorem 1.3.2

Set

$$\begin{aligned}
\phi^k &:= \frac{k-1}{k+1}\pi, \\
\phi(p) &:= \arccos\Big(\frac{\log p - \log(1-p)}{2\log p}\Big), \\
q^k &:= \phi^{-1}\big(\phi^k\big), \\
\lambda_i^k(p) &:= \frac{(-1)^{k-i}\sin i\phi - (-1)^i \sin(k-i+1)\phi + \sin(k+1)\phi}{\frac{-(-1)^k \sin\phi + \sin k\phi + \sin(k+1)\phi}{1+\cos\phi} + k\sin(k+1)\phi}, \\
&\quad \text{where } \phi = \phi(p), \\
\mathrm{d}(p) &:= \frac{1}{2} - \frac{1}{2}\sum_{i=1}^{k}(\lambda_i^k(p))^2 - (1-p)\sum_{i=1}^{k-1}\lambda_i^k(p)\cdot\lambda_{i+1}^k(p), \\
&\quad \text{for } p \in (q^k, q^{k+1}].
\end{aligned}$$

Then

$$\begin{aligned}
\mathrm{d} &: (q^k, q^{k+1}] \to (d_k, d'_{k+1}] \quad \textit{is strictly increasing}, \\
d'_k &= \lim_{p<q^k, p\to q^k} \mathrm{d}(p), \\
d_k &= \lim_{p>q^k, p\to q^k} \mathrm{d}(p), \\
p(d) &= \begin{cases} q^k, & \text{if } d \in (d'_k, d_k], \\ \mathrm{d}^{-1}(d), & \text{if } d \in (d_k, d'_{k+1}], \end{cases} \\
\mathrm{c}(d) &= \begin{cases} r\cdot d + [c(d'_k) - r\cdot d'_k], & \text{if } d \in (d'_k, d_k], \\ H(p)\sum_{i=1}^{k-1}\lambda_i^k(p)\lambda_{i+1}^k(p), & \text{if } d \in (d_k, d'_{k+1}], \end{cases} \\
&\quad \text{where } r = -\log q^k + \log(1-q^k) \text{ and } p = p(d).
\end{aligned}$$

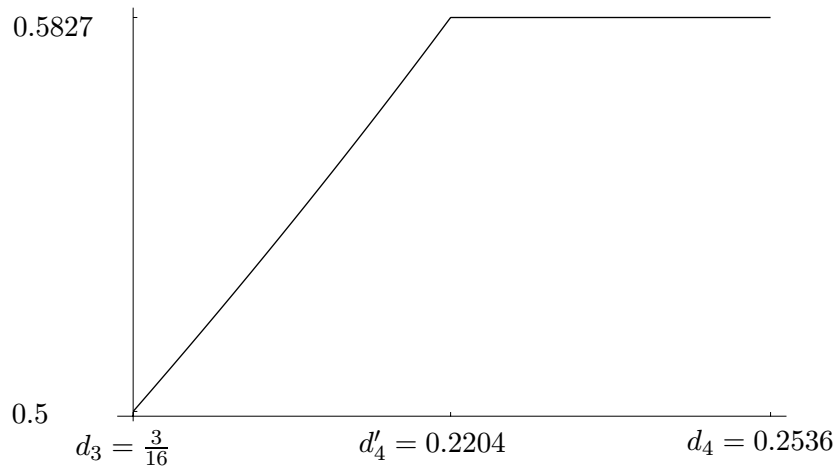

Figure 1.2: $p(d)$ in the range $[d_3, d_4]$.

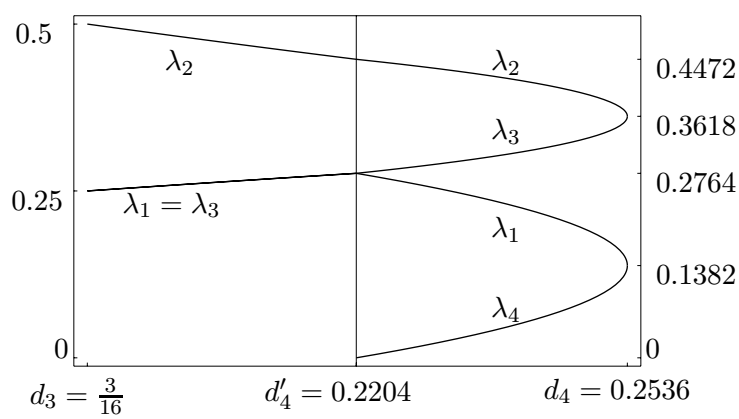

Figure 1.3: $\lambda_i(d)$ in the range $[d_3, d_4]$.

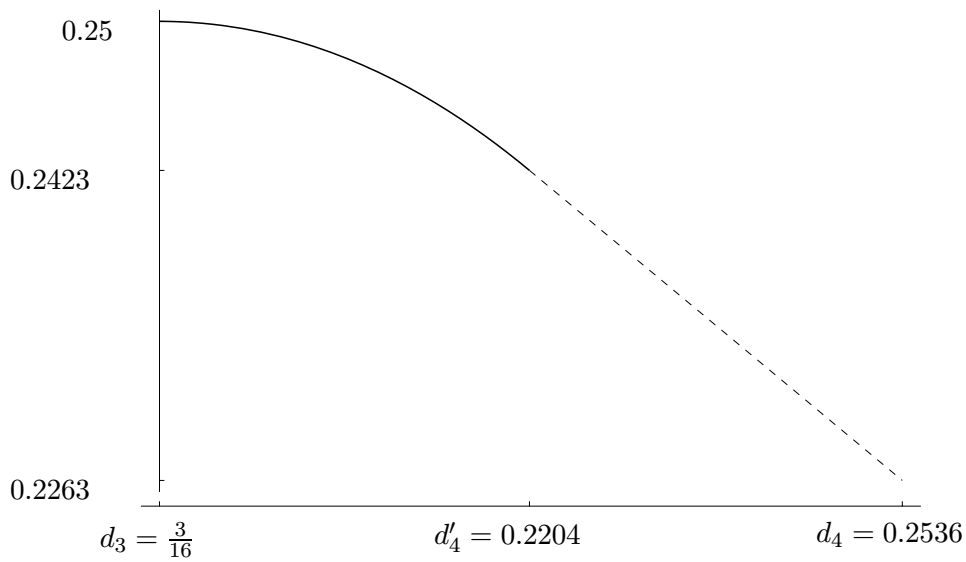

Figure 1.4: c(d) in the range $[d_3, d_4]$.

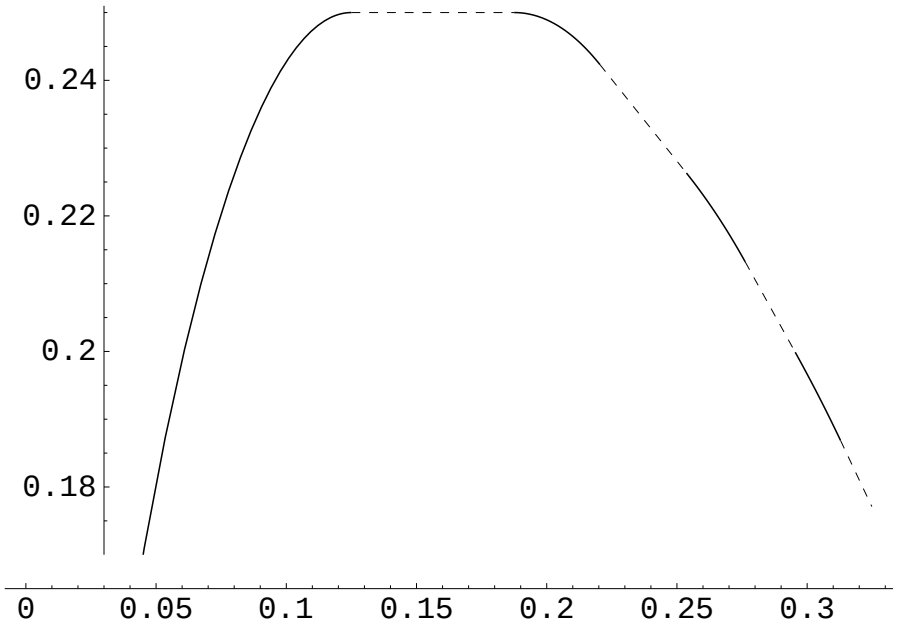

Figure 1.5: c(d) in the range $[0.05, 0.32]$.

Chapter 2

Partitionable posets

This chapter contains the first part of the proof of Theorem 1.3.1. The main idea of this part can be roughly expressed as follows: first we show that every poset is very close to one with a certain 'partitionable' structure. Then we prove in a second step that it suffices to consider the case where the partition classes are arranged in a 'linear' way. Let us first introduce a few definitions, since we can then describe the ideas more accurately before actually turning to the proofs.

Definitions. By a *k-configuration* Q we now mean a weighted poset with point set $\{x_1, \ldots, x_k\}$ where every point x_i carries weight λ_i (where $0 < \lambda_i < 1$ and $\sum_i \lambda_i = 1$) and every relation (x_i, x_j) carries weight $0 \leq p_{i,j} \leq 1$. Forced relations (x_i, x_j) must all have weight $p_{i,j} \equiv 1$.

We say that a poset $P = (X, P) \in \mathcal{P}_n$ has k-configuration Q, if there exists a partition of its point set $X = X_1 \cup \cdots \cup X_k$ into k antichains such that for $x \in X_i$ and $y \in X_j$ one can only have $x < y$ in P if $x_i < x_j$ in Q. On the other hand, if $x_i < x_j$ in Q then there must be exactly $p_{i,j}\,|X_i||X_j|$ comparable pairs $x < y$ with $x \in X_i$ and $y \in X_j$ in P. Furthermore we require that the partition classes satisfy $|X_i| = \lambda_i n$ for all i. Denote by $\mathcal{P}_{n,Q}$ the set of all posets in $\mathcal{P}_n$ that have configuration Q.

A poset $P = (X, P) \in \mathcal{P}_n$ will be called *k-partitionable*, if it has a k-configuration. Obviously, every poset is n-partitionable (in which case its configuration is just the poset itself), but we will be interested in partitionable posets with a constant number of classes. Notice that we have now regained some independence. If P is a poset with configuration Q, then for any cover relation $x_i <: x_j$ in Q, no matter how the relations between X_i and X_j in P are chosen, P remains a poset.

If the number of points in a k-configuration Q is clear or irrelevant, we will simply speak of a configuration. A configuration Q is called *linear*, if Q is a linear order. It is called *p-uniform*, if there exists a $0 \leq p \leq 1$ such that $p_{i,j} \equiv p$ for all cover relations (x_i, x_j) in Q. Comparing this with the terminology used in Section 1.3, a poset is a k-layer poset if it has

a p-uniform linear k-configuration, Q. In this case we will describe Q by $(\lambda_1, \ldots, \lambda_k; p)$, meaning that the points have the given weights and are related $x_1 <: \cdots <: x_k$.

The unique (up to isomorphism) *complete* poset $P \in \mathcal{P}_n$ *induced* by a configuration Q is obtained by letting $p_{i,j} \equiv 1$ for all relations (x_i, x_j) in Q.

Outline of the proofs. Using the terminology that we have just introduced, we now try to give an idea of what is to happen in this section. In Theorem 2.1.2 we show that every poset P is close to a partitionable poset P' that has the same number of relations. By that we mean that changing only very few relations in P suffices to obtain P'. The proof of this theorem is heavily based on Szemerédi's famous Regularity Lemma [60] which states that any large enough graph can be partitioned in such a way that the edges between the partition classes behave in a 'regular' way. We apply this lemma to the comparability graph G of P. Combining the transitivity and the regularity of G then shows that starting from the partition guaranteed by the lemma we only need to remove a few relations and add some others to turn P into the partitionable poset P'.

The closeness of every poset P to a partitionable poset P' as shown by Theorem 2.1.2 then implies that partitionable posets form something like a 'dense' subset of $\mathcal{P}_{n,d}$, and therefore we are able to prove that it suffices to enumerate partitionable posets.

Furthermore we subdivide the family of partitionable posets into smaller families which are defined by having the same configuration. It is easy to see that all posets with configuration Q have the same number of relations, which we denote by $\mathrm{d}(Q)n^2$, so these subfamilies really are contained in $\mathcal{P}_{n,d}$ (when $d = \mathrm{d}(Q)$). The number of posets in such a family, i.e. the number of posets with a given configuration Q, is readily determined (recall that one reason for introducing the concept of partitionable posets was independence which allows easy counting) and measured by a parameter called $\mathrm{c}(Q)$. We will sometimes refer to the parameters $\mathrm{d}(Q)$ and $\mathrm{c}(Q)$ as the d– and the c–value of Q. Since the number of different configurations is actually very small when compared to the number of posets with a given configuration, it is then not difficult to see that we can limit our attention to enumerating the subfamily with the highest possible $\mathrm{c}(Q)$. Thus we will arrive at the following theorem.

(Theorem 2.1.5) *For any $0 < d < \frac{1}{2}$ there exists a configuration Q such that $\mathcal{P}_{n,Q} \subseteq \mathcal{P}_{n,d}$ and*

$$|\mathcal{P}_{n,d}| = |\mathcal{P}_{n,Q}|\, 2^{o(n^2)} = 2^{\mathrm{c}(Q)\, n^2 + o(n^2)}, \qquad \textit{and thus} \qquad \mathrm{c}(d) \;=\; \mathrm{c}(Q).$$

Observe that a configuration Q satisfying the requirements of Theorem 2.1.5 is, by definition, d-significant. Hence the above theorem implies that Q is d-significant if and only if there is no other configuration Q' with $\mathrm{d}(Q') = \mathrm{d}(Q) = d$ and $\mathrm{c}(Q') > \mathrm{c}(Q)$. Thus we have reduced the original problem of determining $|\mathcal{P}_{n,d}|$ to finding this *extremal* configuration Q and computing $\mathrm{c}(Q)$.

The aim of the second section of this chapter is then to show that for every d there exists such a configuration which is p-uniform and linear.

(Theorem 2.2.11) *For every $0 < d < \frac{1}{2}$ there exists a d-significant p-uniform linear configuration.*

In other words, it suffices to count layer posets. To prove Theorem 2.2.11, we need to show that for every configuration Q there exists a p-uniform linear configuration Q' with $\mathrm{c}(Q') \geq \mathrm{c}(Q)$ – but at the same time, and this makes it hard, $\mathrm{d}(Q') = \mathrm{d}(Q)$ must hold.

Theorem 2.2.11 is proved by a series of lemmas. First we show that in fact every significant configuration Q must be p-uniform. For an arbitrary configuration Q we then construct a linear configuration Q' such that, roughly spoken, we can expect a rise in the c-value while loosing not too much in the d-value. On the other hand, by stretching a linear configuration, we also develop a method to increase the d-value, while controlling potential losses in the c-value. Combining these two moves with yet another modification proves the theorem.

Putting everything together, Chapter 2 shows that we only need to consider p-uniform linear configurations. More precisely it leaves us with the question: given d, find the linear configuration $Q^* = (\lambda_1, \ldots, \lambda_k; p)$ among all p-uniform linear configurations Q with $\mathrm{d}(Q) = d$ that achieves the highest possible $\mathrm{c}(Q)$. This will be addressed in Chapter 3.

2.1 Configurations

We start with a simple lemma which states that one can add or remove relations to a partitionable poset without forcing or destroying other relations and still maintain a partitionable poset.

Lemma 2.1.1 *For any two constants d, d' in the interval $(0, \frac{1}{2})$ and any $k \in \mathrm{N}$ there exists a $\bar{k} = \bar{k}(|d' - d|) \in \mathrm{N}$ such that the following holds. For every k-partitionable poset $P \in \mathcal{P}_{n,d}$ there exists a k'-partitionable poset $P' \in \mathcal{P}_{n,d'}$ with $k' \leq \bar{k}$. The new poset P' differs from P in exactly $|d' - d| \cdot n^2$ many relations.*

Proof. Let Q be the configuration of P. If $d' < d$, then first remove relations between X_i and X_j whenever $x_i <: x_j$ in Q. If in all such pairs no relations are left, then this will turn previously forced relations in Q into cover relations and so the process can continue until there are no relations at all.

If $d' > d$, then we will have to add relations and there are three ways to do so:

(i) Whenever $x_i <: x_j$ in Q simply add new relations between X_i and X_j. If this is not enough then

(ii) whenever $x_i \| x_j$ in Q, add new relations between X_i and X_j. Here some care is needed to avoid the forcing of other relations: let x_i be a point in Q that is not comparable

to all other points, and choose x_j to be a maximal point within $\mathrm{inc}(x_i)$. Again if this is not enough then

(iii) split the largest partition class X_i into two parts X_i^- and X_i^+, maintain all previous relations and add new relations $x < y$ where $x \in X_i^-$ and $y \in X_i^+$.

Repeating and combining these steps produces a k'-partitionable poset $P' \in \mathcal{P}_{n,d'}$ where d' is arbitrarily close to $\frac{1}{2}$. The constant k' depends only on d and d' – in fact it only depends on $|d' - d|$ – but not on n. 2

Theorem 2.1.2 *For every $\epsilon > 0$ there exist two constants k_0, n_0 such that for every poset $P \in \mathcal{P}_{n,d}$ with $n \geq n_0$ there is a k-partitionable poset $P' \in \mathcal{P}_{n,d}$ with $k \leq k_0$ that differs from P in at most ϵn^2 relations and in which each partition class has at least $n/(2k_0)$ points.*

For the proof of Theorem 2.1.2 we need Szemerédi's Regularity Lemma and some related definitions. Let $G = (V, E)$ be a graph and consider two disjoint subsets $A, B \subset V$. Denote by $E(A, B)$ the set of those edges in E that have one endpoint in A and one endpoint in B. Then the *density* $d(A, B)$ is defined as

$$d(A, B) = \frac{|E(A, B)|}{|A| \cdot |B|}.$$

For $\epsilon \in (0, 1)$ a pair A, B is called *ϵ-regular*, if for every $X \subseteq A$ and $Y \subseteq B$ satisfying

$$|X| > \epsilon|A|, \quad \text{and} \quad |Y| > \epsilon|B|$$

it is true that

$$|d(X, Y) - d(A, B)| < \epsilon.$$

To say it roughly, the Regularity Lemma [60] guarantees that for every graph one can find a partition of its vertex set into classes of almost the same size such that almost all pairs are regular. Natural modifications to the original proof easily give the following variant, for which we need a few more definitions:

Let $G_1, \ldots G_r$ be spanning subgraphs of a graph $G = (V, E)$. In this setting a partition $V = X_1 \cup \cdots \cup X_k$ is called *ϵ-regular*, if the classes X_i differ in size by at most 1 and all but at most ϵk^2 pairs are ϵ-regular for all G_i. Such a partition is said to *refine* another partition $V = V_1 \cup \cdots \cup V_{k'}$ if for each $1 \leq i \leq k$ there exists a $1 \leq j \leq k'$ so that $X_i \subseteq V_j$.

Theorem 2.1.3 (Szemerédi's Regularity Lemma) *For every $\tilde{\epsilon} > 0$ and $\ell, r \geq 1$ there exist two positive integers $\tilde{n}_0 = \tilde{n}_0(\tilde{\epsilon}, \ell, r)$ and $\tilde{k}_0 = \tilde{k}_0(\tilde{\epsilon}, \ell, r)$ such that the following is true. If $G = (X, E)$ is a graph with $|X| \geq \tilde{n}_0$ and $X = X_1' \cup \cdots \cup X_\ell'$ is a partition where the classes differ in size by at most one, and if $G_1, \ldots, G_r$ are spanning subgraphs of G, then there exists an $\tilde{\epsilon}$-regular partition $X = X_1 \cup \cdots \cup X_k$ with $\ell \leq k \leq \tilde{k}_0$ that refines the previous partition.* 2

Proof of Theorem 2.1.2. Let $G = (X, E)$ be the comparability digraph of P.

Set $\tilde{\epsilon} := \epsilon/12$. Choose an integer k' so that $1/k' < \tilde{\epsilon}$ and take an arbitrary partition $X = X'_1 \cup \cdots \cup X'_{k'}$ satisfying

$$\left\lfloor \frac{n}{k'} \right\rfloor \leq |X'_i| \leq \left\lceil \frac{n}{k'} \right\rceil \qquad \text{for all } \ 1 \leq i \leq k'.$$

In the following we will be somewhat sloppy by disregarding rounding.

Colour the edges of G in the following way. An edge (x, y) is coloured in *red*, if there exists a z, such that $x < z < y$. If not, then it is coloured in *blue* and in analogy to the situation in P we say that y *covers* x. Note that for every red edge (x, y) there must be a directed path $x, z_1, \ldots, z_k, y$ with $k \geq 1$ of blue edges.

Since we want to turn P into a partitionable poset, we will have to remove edges from G. Obviously red edges cannot be removed without destroying the transitivity. Thus by removing a family $\mathcal{F}$ of edges we *always* mean removing all blue edges in $\mathcal{F}$ and putting a spell on the red edges in $\mathcal{F}$: if later red edges in $\mathcal{F}$ turn blue, we will remove them as well. Notice that removing a blue edge results in a digraph which is the comparability digraph of a poset with one relation less.

We start by removing all edges of G that lie inside a class X'_i.
By the (ordered) pair (X'_i, X'_j) we denote the bipartite graph on the vertices $X'_i \cup X'_j$ that contains all edges that in G leave X'_i and enter X'_j. (Note that in (X'_i, X'_j) the edges are undirected.) Let

$$E_< := \bigcup_{i<j} E(X'_i, X'_j), \qquad E_> := \bigcup_{i>j} E(X'_i, X'_j),$$

and consider the two graphs $G_< := (X, E_<)$ and $G_> := (X, E_>)$.

Now apply Theorem 2.1.3 with parameters $\tilde{\epsilon}$, $\ell := k'$ and $r := 2$, a first partition $X = X'_1 \cup \cdots \cup X'_{k'}$ and the two spanning graphs $G_<$ and $G_>$. If we choose the constants n_0, k_0 in the statement of Theorem 2.1.2 so that $n_0/k' \geq \tilde{n}_0$ as well as both $k_0 \geq \tilde{k}_0$ and $k_0 \geq \bar{k}(\epsilon)$ (see Lemma 2.1.1) are ensured, we are then guaranteed a partition

$$X = X_1 \cup \cdots \cup X_k$$

with $k \leq k_0$ that refines $X = X'_1 \cup \cdots \cup X'_{k'}$ and has the property that for all but at most $\tilde{\epsilon}k^2$ pairs $i < j$ both (X_i, X_j) and (X_j, X_i) are $\tilde{\epsilon}$-regular.
Consider the following properties of an arbitrary pair A, B:

(i) all but at most $\tilde{\epsilon}|A|$ vertices in A have degree at least $2\tilde{\epsilon}|B|$, and analogously with the roles of A and B exchanged,

(ii) for every set $A' \subseteq A$ with $|A'| > \tilde{\epsilon}|A|$ the set of neighbours $\Gamma(A')$ must have cardinality at least $(1 - \tilde{\epsilon})|B|$, and analogously with the roles of A and B exchanged.
Any $\tilde{\epsilon}$-regular pair with density at least $3\tilde{\epsilon}$ satisfies property (i) – because if not then denote those vertices with degree less than $2\tilde{\epsilon}|B|$ by A' and $d(A', B) < 2\tilde{\epsilon}$ would contradict the

regularity. Any $\tilde{\epsilon}$-regular pair with density at least $\tilde{\epsilon}$ satisfies property (ii) – because if not then the pair $A', B \setminus \Gamma(A')$ would again contradict the regularity.

The following third property is a strengthening of (i) and obviously not possessed by every regular pair:

(iii) all but at most $\tilde{\epsilon}|A|$ vertices in A have degree at least $(1-\tilde{\epsilon})|B|$, and analogously with the roles of A and B exchanged.

Now call a pair (X_i, X_j) *good*, if it has properties (i) and (ii), or if it has property (iii). Call it *bad* otherwise. Remove those edges in G that lie in a bad pair (X_i, X_j) or (X_j, X_i), where $i < j$. Denote by P' the poset obtained in this way. Since all $\tilde{\epsilon}$-regular pairs with density at least $3\tilde{\epsilon}$ are good, observe that up to now at most $5\tilde{\epsilon}n^2$ edges (that is, relations in P) have been removed, namely

at most $k'(\frac{n}{k'})^2 \leq \tilde{\epsilon}n^2$ edges inside the X'_i,
at most $k^2 \cdot 3\tilde{\epsilon} \cdot (\frac{n}{k})^2 \leq 3\tilde{\epsilon}n^2$ edges inside pairs (X_i, X_j) with density less than $3\tilde{\epsilon}$,
at most $\tilde{\epsilon}k^2(\frac{n}{k})^2 = \tilde{\epsilon}n^2$ edges inside irregular pairs (X_i, X_j).

Consider the digraph R with vertex set $\{X_1, \ldots, X_k\}$ and edges (X_i, X_j) if the pair (X_i, X_j) is good. We claim the following:

1. If $Y_1, Y_2, \ldots, Y_l$ is a dipath in R with $l \geq 3$, then (Y_1, Y_l) has property (iii).
2. R is acyclic.
3. All the edges left in G lie in good pairs.

These claims are easily verified as follows:

ad 1. Recall that G is still a comparability digraph, i.e. $(x, y) \in E$ and $(y, z) \in E$ imply that $(x, z) \in E$. We assume that (Y_1, Y_2) and (Y_2, Y_3) are good pairs and prove that this implies that (Y_1, Y_3) has property (iii). Then 1. follows by induction on l. Suppose (Y_1, Y_2) has property (iii), then all but at most $\tilde{\epsilon}|Y_1|$ vertices in Y_1 have degree at least $(1-\tilde{\epsilon})|Y_2|$, and it therefore suffices to check that all but at most $\tilde{\epsilon}|Y_3|$ vertices in Y_3 have degree more than $\tilde{\epsilon}|Y_2|$. This is guaranteed if (Y_2, Y_3) is good. Suppose on the other hand that (Y_1, Y_2) has property (i), then all but at most $\tilde{\epsilon}|Y_1|$ vertices in Y_1 have degree at least $2\tilde{\epsilon}|Y_2| > \tilde{\epsilon}|Y_2|$. No matter whether (Y_2, Y_3) now has property (ii) or (iii), we are done.

ad 2. By 1. we know that any cycle $Y_1, Y_2, \ldots, Y_{l-1}, Y_1$ implies that all but at most $\tilde{\epsilon}|Y_1|$ vertices in Y_1 have at least $(1-\tilde{\epsilon})|Y_1|$ neighbours in Y_1. This in turn implies a directed cycle $y_1, y_2, \ldots, y_{l-1}, y_1$ in G, which is impossible, since G is the comparability digraph of P'.

ad 3. This is obviously true for blue edges. For red edges it might seem a little surprising at first glance, since so far we never cared about removing red edges. Nevertheless it is true: consider a red edge (y_1, y_l) with $y_1 \in Y_1$ and $y_l \in Y_l$. Then there must be a path $y_1, y_2, \ldots, y_l$ of blue edges. Since blue edges can only be found in good pairs, there must be a directed path $Y_1, \ldots, Y_l$ in R. Since R is acyclic, we must have $Y_1 \neq Y_l$, and Claim 1 implies that (Y_1, Y_l) is also good.

By Claim 1 and 2, R is a comparability digraph and we denote by Q the corresponding poset with point set $\{x_1, \ldots, x_k\}$. Then Claim 1 and 3 assert that P' will have configuration Q if we complete all pairs (X_i, X_j) that satisfy property (iii) at the cost of at most

$$\binom{k}{2}\left(\tilde{\epsilon}\frac{n}{k}\cdot\frac{n}{k}+\tilde{\epsilon}\frac{n}{k}\cdot\frac{n}{k}\right) \;<\; \tilde{\epsilon}n^2$$

new relations (and then choose the weights λ_i and $p_{i,j}$ in Q accordingly). Note that inserting these new relations does not violate transitivity: any new edge (y, y') lies in a pair with property (iii), and if it together with another (new or old) edge (y', y'') requires the edge (y, y'') to exist, then since (y', y'') lies in a good pair, we know by Claim 1 that (y, y'') lies in a pair with property (iii), hence it either already exists or will be inserted anyway in the completion process.

In total we changed less than $6\tilde{\epsilon}n^2 = \frac{\epsilon}{2}n^2$ edges and the new poset now has configuration Q. In order to satisfy the requirements of Theorem 2.1.2 we have to make sure that it has the same number of relations as in the beginning, which means that we might have to add or remove at most $\frac{\epsilon}{2}n^2$ relations. This can be done as described by Lemma 2.1.1.

2

The above theorem marks the breakthrough on our way to Theorem 2.1.5; what remains to be done is now fairly straight forward. Denote by $\mathcal{P}^k_{n,d}$ the family of all k-partitionable posets from $\mathcal{P}_{n,d}$. Then Theorem 2.1.2 states that we can enumerate the set $\mathcal{P}_{n,d}$ in the following way: for every ϵ there exist two constants k_0, n_0 such that if $n \geq n_0$

$$\mathcal{P}_{n,d} = \bigcup_{k=1}^{k_0} \bigcup_{P\in\mathcal{P}^k_{n,d}} \Gamma_\epsilon(P),$$

where $\Gamma_\epsilon(P)$ denotes all those posets in $\mathcal{P}_{n,d}$ that can be constructed from P by changing at most ϵn^2 many relations: these are no more than

$$\binom{n^2}{\epsilon\cdot n^2} = 2^{(1+o(1))H(\epsilon)n^2},$$

where, as before, $H(\epsilon)$ denotes the entropy function. Thus the following corollary holds.

Corollary 2.1.4 *For every $\epsilon > 0$ there exists a constant k_0 such that*

$$|\mathcal{P}_{n,d}| = 2^{(1+o(1))H(\epsilon)n^2} \cdot \sum_{k=1}^{k_0} |\mathcal{P}^k_{n,d}|.$$

2

For a given configuration we would like to count the number of different posets that have this configuration. We therefore introduce some more definitions.

Let P be a k-partitionable poset with partition $X_1, \ldots, X_k$ and let Q be its configuration. Denote by $E(X_i, X_j)$ the set of relations between X_i and X_j and recall that $\lambda_i := |X_i|/n$ and $p_{i,j} := |E(X_i, X_j)|/|X_i| \cdot |X_j|$. When counting the number of posets with configuration Q it is clear that the degree of freedom we have lies in where we place the $p_{i,j}|X_i||X_j|$ relations between X_i and X_j when $x_i <: x_j$ in Q. Hence the number is equal to

$$\prod_{x_i <: x_j} \binom{\lambda_i \lambda_j n^2}{p_{i,j} \cdot \lambda_i \lambda_j n^2} = 2^{\sum_{i,j} H(p_{i,j})\lambda_i \lambda_j n^2 \, + o(n^2)}.$$

where the sum is taken over all pairs i, j with $x_i <: x_j$. (Actually we did have more freedom: since we are considering labeled posets we also had the choice of assigning points to classes. But this gives merely a factor of $O(n!) = O(2^{n \log n})$.) Let

$$\mathrm{c}(Q) := \sum_{x_i <: x_j} H(p_{i,j})\lambda_i \lambda_j,$$

and observe that Q determines the total number of relations in P. It must have dn^2 relations where

$$d = \mathrm{d}(Q) := \sum_{x_i \ll x_j} \lambda_i \lambda_j + \sum_{x_i <: x_j} p_{i,j} \lambda_i \lambda_j.$$

As mentioned before, we will refer to these parameters as the c-*value* and the d-*value* of the configuration Q.

Notice that since k is a constant and independent of n, there aren't actually all that many different k-configurations: there are at most $2^{k^2/2}$ posets, for each X_i there is a choice of at most n values to determine $|X_i|$. Finally for each pair (X_i, X_j) there is a choice of less than n^2 values to determine $|E(X_i, X_j)|$. Therefore in total there are certainly less than $2^{o(n^2)}$ different k-configurations. Hence if Q is a configuration with $\mathrm{d}(Q) = d$ and no other configuration has the same d-value and a higher c-value, then

$$|\mathcal{P}^k_{n,d}| = \sum_{Q'} 2^{c(Q')n^2} = 2^{c(Q) \cdot n^2 + o(n^2)},$$

where the sum is taken over all k-configurations Q' with $\mathrm{d}(Q') = d$. Together with Corollary 2.1.4 this now proves the following theorem that we had already announced at the beginning of this chapter.

Theorem 2.1.5 *For any $0 < d < \frac{1}{2}$ there exists a configuration Q such that $\mathcal{P}_{n,Q} \subseteq \mathcal{P}_{n,d}$ and*

$$|\mathcal{P}_{n,d}| = |\mathcal{P}_{n,Q}|\, 2^{o(n^2)} = 2^{\mathrm{c}(Q)\, n^2 + o(n^2)}, \qquad \textit{and thus} \qquad \mathrm{c}(d) \;=\; \mathrm{c}(Q).$$

2

We recall the following trivial but crucial observation: a configuration Q satisfying the requirements of Theorem 2.1.5 is, by definition, d-significant. Hence the above theorem implies that Q is d-significant if and only if there is no other configuration Q' with $\mathrm{d}(Q') = \mathrm{d}(Q) = d$ and $\mathrm{c}(Q') > \mathrm{c}(Q)$. Thus we have reduced the original problem of determining $|\mathcal{P}_{n,d}|$ to finding this *extremal* configuration Q and computing $\mathrm{c}(Q)$.

2.2 Linear configurations

The result of the previous section states that we can focus our attention on partitionable posets. More than that, it also says that in order to determine $\mathrm{c}(d)$ it suffices to compute the c-value of a d-significant configuration. In this section we prove that for each d we can always find a d-significant configuration with a particularly simple shape.

Lemma 2.2.1 *Any significant configuration Q must be p-uniform for some $p \in [0,1]$.*

Proof. Assume without loss of generality that we have two cover relations $x_1 <: x_3$ with weight p_1 and $x_2 <: x_4$ with weight p_2. We claim that $p_1 = p_2$. Let

$$m = p_1\lambda_1\lambda_3 + p_2\lambda_2\lambda_4.$$

Then let $m' := m/(\lambda_2\lambda_4)$ and obtain

$$m' - p_1\frac{\lambda_1\lambda_3}{\lambda_2\lambda_4} = p_2.$$

For the c-value of the configuration to be maximal, p_1 must be chosen in such a way that

$$H(p_1)\lambda_1\lambda_3 + H(p_2)\lambda_2\lambda_4 = H(p_1)\lambda_1\lambda_3 + H\Big(m' - p_1\frac{\lambda_1\lambda_3}{\lambda_2\lambda_4}\Big)\lambda_2\lambda_4 \tag{2.1}$$

is maximal.

Differentiating (2.1) with respect to p_1 and setting the derivative equal to zero we obtain

$$H'(p_1)\lambda_1\lambda_3 - H'(m' - p_1\frac{\lambda_1\lambda_3}{\lambda_2\lambda_4})\lambda_1\lambda_3 = 0.$$

Since H' is an injective function, this implies that

$$p_1 = m' - p_1\frac{\lambda_1\lambda_3}{\lambda_2\lambda_4} = p_2.$$

2

In other words, in a significant configuration Q all cover relations carry the same weight, which we will call the *density* of Q and denote by $p = \mathrm{p}(Q)$. We continue to show that for every d we can always find a d-significant configuration with a simple shape.

We first need some more definitions. For a configuration Q let

$$\sigma_1(Q) := \sum_{x_i <: x_j} \lambda_i \cdot \lambda_j \quad , \qquad \sigma_2(Q) := \sum_{x_i \ll x_j} \lambda_i \cdot \lambda_j,$$

$$\sigma_0(Q) := \sum_{i=1}^{k} (\lambda_i)^2 \quad , \qquad \sigma_\infty(Q) := \sum_{x_i \| x_j} \lambda_i \cdot \lambda_j.$$

Since $\sum_{i=1}^{k} \lambda_i = 1$ we have

$$2 \cdot (\sigma_1 + \sigma_2 + \sigma_\infty) + \sigma_0 = 1. \tag{2.2}$$

(In the definitions of the σ_r the index r roughly represents the distance between x_i and x_j in the cover graph of Q – if one considers 'distance 2' as 'finite distance of at least 2'.) Notice that using these definitions together with Lemma 2.2.1 we can now write

$$\mathrm{c}(Q) = H(p) \cdot \sigma_1(Q) \tag{2.3}$$

and

$$\mathrm{d}(Q) = p \cdot \sigma_1(Q) + \sigma_2(Q), \tag{2.4}$$

or equivalently

$$\mathrm{d}(Q) = \frac{1}{2} - \frac{1}{2}\sigma_0(Q) - (1-p)\sigma_1(Q) - \sigma_\infty(Q), \tag{2.5}$$

where $p = \mathrm{p}(Q)$.

Proposition 2.2.2 *For every configuration Q we must have*

$$\sigma_1(Q) \leq \frac{1}{4}.$$

Proof. Definitely the shortest way to prove this is by contradiction: if $\sigma_1(Q) > \frac{1}{4}$ then set $\mathrm{p}(Q) := p := \frac{1}{2}$ which implies $\mathrm{c}(Q) = H(p)\sigma_1(Q) > \frac{1}{4}$, thereby contradicting (1.11). But since this seems a bit like taking a sledgehammer to crack a nut, we also give an elementary proof here.

Consider the cover graph of Q with weight λ_i on the vertex x_i. It is obviously triangle-free. Denote by γ_i the sum of the weights of all neighbours of x_i. We propose the following process:

As long as there are two non-adjacent vertices x_i, x_j with positive weights λ_i, λ_j, take the vertex with smaller neighbourhood weight (say $\gamma_i \leq \gamma_j$) and shift its weight completely to the other vertex: $\lambda_i := 0$, $\lambda_j := \lambda_j + \lambda_i$.

We will check the following two observations. (1.) During this process $\sigma_1(Q)$ does not decrease and (2.) after the process there will be only two vertices with positive weight. Hence at the end $\sigma_1 \leq \frac{1}{4}$, which would prove the proposition.

To check (1.), simply observe that during one step of the process the loss in σ_1 is $\lambda_i \cdot \gamma_i$ and the win is $\lambda_i \cdot \gamma_j$; hence in total we don't lose anything.

As for (2.), since the graph is triangle–free, there are no cliques other than edges. In particular, let us restrict our attention to vertices with positive weight: if there are at least three of them, we will always find two non-adjacent ones. 2

The following proposition now determines $\mathrm{c}(d)$ and the structure of a d-significant configuration in the range $d \in (0, \frac{1}{8}]$ and thereby reproves the result by Kleitman and Rothschild as stated in (1.16).

Theorem 2.2.3 *For every $0 < d \leq \frac{1}{8}$ there exists a d-significant 2-configuration. It is (obviously) linear, the two layers have size $\frac{1}{2}$ and its density is $4d$. This implies that in this range*

$$\mathrm{c}(d) = \frac{1}{4} \cdot H(4d).$$

Proof. Consider an arbitrary d-significant configuration Q with density $p = \mathrm{p}(Q)$. We construct a 2-configuration Q' with point set $\{x_1', x_2'\}$ such that $\mathrm{d}(Q') = \mathrm{d}(Q)$ and $\mathrm{c}(Q') \geq \mathrm{c}(Q)$ by setting

$$x_1' <: x_2', \qquad \lambda_1' := \lambda_2' := \frac{1}{2},$$
$$p' := \mathrm{p}(Q') := 4d = 4p\sigma_1(Q) + 4\sigma_2(Q).$$

Obviously $\mathrm{d}(Q') = \frac{1}{4}p' = d = \mathrm{d}(Q)$ so it only remains to compare $\mathrm{c}(Q')$ and $\mathrm{c}(Q)$. Also observe that $d \leq \frac{1}{8}$ implies that $p' \leq \frac{1}{2}$, hence for any $p'' \leq p'$ we have $H(p') \geq H(p'')$. Furthermore $H(x)$ has the property that for any $0 < \alpha \leq 1$ we have that $H(\alpha \cdot x) \geq \alpha \cdot H(x)$. Equipped with these facts we abbreviate $\sigma_1 := \sigma_1(Q)$ and $\sigma_2 := \sigma_2(Q)$ and obtain

$$\begin{aligned} c(Q') &= H(p') \cdot \frac{1}{4} = H(4\sigma_2 + 4p\sigma_1) \cdot \frac{1}{4} \\ &\geq H(p \cdot 4\sigma_1) \cdot \frac{1}{4} \geq H(p)\sigma_1 \quad = \quad \mathrm{c}(Q), \end{aligned}$$

where we first applied the definition of p', then used $p' \leq \frac{1}{2}$, and finally relied on $4\sigma_1 \leq 1$ which is guaranteed by Proposition 2.2.2. 2

In the remaining part of this section we concentrate on the range $d \geq \frac{1}{8}$. Our aim is to prove that for every such d there exists a d-significant *linear* configuration. This will have to be done in several steps.

In the following lemma we choose to work with a poset P (with point set $[n]$ as usual) directly and not with its configuration Q. Denote by

$$\tilde{\sigma}_1(P) := \#\{i <: j \ \text{ where } \ i, j \in [n]\}, \qquad \tilde{\sigma}_2(P) := \#\{i \ll j \ \text{ where } \ i, j \in [n]\}$$
$$\tilde{\sigma}(P) := \tilde{\sigma}_1(P) + \tilde{\sigma}_2(P), \qquad \tilde{\sigma}_\infty(P) := \#\{i \| j \ \text{ where } \ i, j \in [n]\}.$$

Observe that trivially

$$\tilde{\sigma}(P) + \tilde{\sigma}_\infty(P) = \binom{n}{2}, \tag{2.6}$$

but more importantly, notice that for a partitionable poset P with configuration Q a pair $x_i \ll x_j$ contributes $\lambda_i\lambda_j$ to $\sigma_2(Q)$ and $|X_i||X_j| = \lambda_i\lambda_j n^2$ to $\tilde{\sigma}_2(P)$. Similarly a pair $x_i <: x_j$ contributes $\lambda_i\lambda_j$ to $\sigma_1(Q)$ and *at most* $|X_i||X_j| = \lambda_i\lambda_j n^2$ to $\tilde{\sigma}_1(P)$. Thus we have

$$\tilde{\sigma}_2(P) = \sigma_2(Q) \cdot n^2, \tag{2.7}$$

$$\tilde{\sigma}_1(P) \leq \sigma_1(Q) \cdot n^2,$$

and in the latter equality holds if and only if for all i, j with $x_i <: x_j$ in Q *all* cover relations between the two partition classes X_i and X_j exist in P. If this is the case then P is the *complete* partitionable poset *induced* by Q.

Lemma 2.2.4 *Suppose that P is a poset with height k. Then there exists a k-layer poset P' satisfying*

$$\tilde{\sigma}_1(P') \geq \tilde{\sigma}_1(P) \qquad \textit{and} \qquad \tilde{\sigma}(P') \geq \tilde{\sigma}(P).$$

Corollary 2.2.5 *For every k-configuration Q there exists a linear k'-configuration Q' with $k' \leq k$ such that*

$$\sigma_1(Q') \geq \sigma_1(Q) \qquad \textit{and} \qquad \sigma(Q') \geq \sigma(Q).$$

Proof of Corollary 2.2.5. Let P be the complete partitionable poset induced by Q. Apply Lemma 2.2.4 to P and obtain a k'-layer poset P'. Let Q' be the (linear) configuration of P'. Obviously

$$\sigma_1(Q') \geq \frac{\tilde{\sigma}_1(P')}{n^2} \geq \frac{\tilde{\sigma}_1(P)}{n^2} = \sigma_1(Q),$$

and

$$\sigma(Q') \geq \frac{\tilde{\sigma}(P')}{n^2} \geq \frac{\tilde{\sigma}(P)}{n^2} = \sigma(Q).$$

2

Proof of Lemma 2.2.4. Let $C_1, \ldots, C_k$ be a chain decomposition of P that is obtained by recursively removing maximum chains from P. Hence we have that $|C_1| \geq \cdots \geq |C_k|$ and furthermore that C_i is a maximal chain in $P - C_1 - \cdots - C_{i-1}$ for all $i \in [k]$. Denote by $c_i := |C_i|$. The underlying idea of the proof is to glue the chains C_i together again, but in such a way as to carefully control the parameters $\tilde{\sigma}$ and $\tilde{\sigma}_1$.

We first give bounds on $\tilde{\sigma}_1(P)$ and $\tilde{\sigma}_\infty(P)$. Within each chain C_i there can be at most $c_i - 1$ cover relations. Denote the number of cover relations between two chains C_i and C_j with $i < j$ by $\tilde{\sigma}_1(C_i, C_j)$. Hence

$$\tilde{\sigma}_1(P) \leq \sum_{i=1}^{k}(c_i - 1) + \sum_{i<j} \tilde{\sigma}_1(C_i, C_j), \tag{2.8}$$

and we claim that

$$\tilde{\sigma}_1(C_i, C_j) \leq \begin{cases} 2c_j - 2 & \text{, if } c_i = c_j \\ 2c_j - 1 & \text{, if } c_i = c_j + 1 \\ 2c_j & \text{, always.} \end{cases} \tag{2.9}$$

The best way to see this might be to view this as a bipartite graph with vertex sets C_i, C_j where an edge represents a cover relation. Since each point in one chain *can cover* at most one point and *can be covered* by at most one point from the other chain the graph has maximum degree at most 2. Let $C_i = [x_1, \ldots, x_{c_i}]$ and $C_j = [y_1, \ldots, y_{c_j}]$. Then x_1 cannot be covered by any element in C_j and x_{c_i} cannot cover any element in C_j (otherwise C_i would not be maximal) so they have degree at most one. Hence the sum of the degrees in C_i is bounded from above by $2c_i - 2$ (which settles the first case of the claim) and the sum of the degrees in C_j is bounded from above by $2c_j$ (which settles the third case). For the second case, where $c_i = c_j + 1$, the only possibility left to contradict our claim would be if $\tilde{\sigma}_1(C_i, C_j) = 2c_j = 2c_i - 2$, implying that all points in C_i *and* C_j have indeed degree 2 except for x_1 and x_{c_i}, which have degree 1. Now if x_1 did cover y_i for any $i > 1$, then the point y_1 could not be covered, hence $y_1 <: x_1$. Similarly y_1 can't cover any point other than x_2, for otherwise x_2 could not be covered. Thus $x_2 <: y_1$. But now $[x_1, y_1, x_2, x_3, \ldots, x_{c_i}]$ contradicts the maximality of C_i. This completes the proof of claim (2.9).

To give a lower bound on $\tilde{\sigma}_\infty(P)$ note that any incomparable pair in $C_i \cup C_j$ must obviously have one point in C_i and one point in C_j. Denote the number of such pairs by $\tilde{\sigma}_\infty(C_i, C_j)$. We claim that

$$\tilde{\sigma}_\infty(C_i, C_j) \geq c_j. \tag{2.10}$$

Suppose that a point $y \in C_j$ were comparable to all points $x \in C_i$. This would imply the existence of some t with $0 \leq t \leq c_i$ such that $x_t > y > x_{t+1}$. But then $[x_1, \ldots, x_t, y, x_{t+1}, \ldots, x_{c_i}]$ contradicts the maximality of C_i. Therefore every point $y \in C_j$ must be incomparable to at least one point $x \in C_i$, hence in total $\tilde{\sigma}_\infty(C_i, C_j) \geq c_j$, which proves claim (2.10).

If we now succeed in constructing a layer poset P' by taking the chains $C_1, \ldots, C_k$ as 'building blocks' (which means that, again, C_i is a maximum chain in $P' - C_1 - \cdots - C_{i-1}$)

and combining them in such a way that (2.8), (2.9), and (2.10) hold for P' with equality, then

$$\tilde{\sigma}_1(P') \geq \tilde{\sigma}_1(P)$$

and

$$\tilde{\sigma}_\infty(P') \leq \tilde{\sigma}_\infty(P).$$

By (2.6) it then follows immediately that

$$\tilde{\sigma}_1(P') \geq \tilde{\sigma}_1(P) \qquad \text{and} \qquad \tilde{\sigma}(P') \geq \tilde{\sigma}(P).$$

as required by the lemma.

We construct P' by letting the mid-points of all chains C_i form a single antichain. More precisely: re–number the points in the chains C_i so that

$$C_i = [x^i_{c_i-1}, \ldots, x^i_3, x^i_1, x^i_2, x^i_4, \ldots, x^i_{c_i}] \text{ if } c_i \text{ is even,}$$

$$C_i = [x^i_{c_i}, \ldots, x^i_3, x^i_1, x^i_2, x^i_4, \ldots, x^i_{c_i-1}] \text{ if } c_i \text{ is odd,}$$

so

$$[n] = \bigcup_{i=1}^{k} \bigcup_{j=1}^{c_i} \{x^i_j\}.$$

Now add the cover relations

$$\begin{array}{rcll}
\text{for all } i \in [k] \text{ with } c_i \geq 2 & : & x^i_2 <: x^j_1 & \text{for all } j \in [k] \\
\text{for all } i \in [k] \text{ and } s \geq 4 \text{ even with } c_i \geq s & : & x^i_s <: x^j_{s-2} & \text{whenever } c_j \geq s-2 \\
\text{for all } i \in [k] \text{ and } s \geq 1 \text{ odd with } c_i \geq s & : & x^i_s <: x^j_{s+2} & \text{whenever } c_j \geq s+2.
\end{array}$$

For an illustration of this construction see Figure 2.1, where only the cover relations involving x^3_1 are shown.

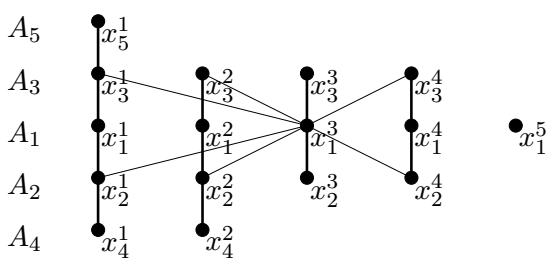

Figure 2.1: Construction of P'.

For $s = 1, \ldots, c_1$ the sets

$$A_s := \{x_s^i : \quad \text{for all } i \in [k] \text{ where } s \leq c_i\}$$

form antichains. They constitute the layers in P'. (An alternative description of the same construction is to say that P' is a c_1-layer poset with configuration $(\ldots, a_4, a_2, a_1, a_3, \ldots; 1)$ where $a_s := |A_s|/n$.)

Observe that for any two chains C_i and C_j with $i < j$ the only point in $C_i \cup C_j$ that is incomparable to $x_s^j \in C_j$ is x_s^i, hence inequality (2.10) holds with equality. Moreover it is easy to see that equality also holds in inequalities (2.8) and (2.9). 2

Starting from a configuration Q we will often build a new configuration Q', possibly with a different number of partition classes, different weights and relations. When doing so, we will usually refer to the original weights as λ_i, and to the new weights as $\lambda(x_i)$. Often we will shift weight ϵ from one point x_i to another point x_j, i.e. $\lambda(x_i) := \lambda(x_i) - \epsilon$ and $\lambda(x_j) := \lambda(x_j) + \epsilon$. Sometimes it will also be necessary to create a new point with some weight ϵ: this can be considered as shifting weight ϵ to a point which so far had weight zero.

If the old configuration Q and the new configuration Q' are clear from the context, then we will sometimes abbreviate $\sigma_r := \sigma_r(Q)$ and $\sigma'_r := \sigma_r(Q')$ for $r \in \{0, 1, 2\}$. From now on the case $r = \infty$ is irrelevant because we will be moving from one linear configuration to another, hence $\sigma_\infty = 0$ will always hold. As in the case of $\tilde{\sigma}$ define

$$\sigma(Q) := \sigma_1(Q) + \sigma_2(Q).$$

Therefore (2.2) now stands as

$$2\sigma(Q) + \sigma_0(Q) = 1. \tag{2.11}$$

Another trivial observation: if $\lambda_i \geq \lambda_j$ then shifting any weight $0 \leq \epsilon \leq \lambda_i - \lambda_j$ from x_i to x_j will not increase σ_0:

$$\sigma'_0 - \sigma_0 = (\lambda_i - \epsilon)^2 + (\lambda_j + \epsilon)^2 - \lambda_i^2 - \lambda_j^2 = -2\epsilon(\lambda_i - \lambda_j - \epsilon) \leq 0. \tag{2.12}$$

Lemma 2.2.6 *For every linear configuration Q and for every $0 < s \leq \sigma_1(Q)$ there exists a linear configuration Q' such that*

$$\sigma(Q') \geq \sigma(Q) \quad \textit{and} \qquad \sigma_1(Q') = s.$$

Proof. Assume w.l.o.g. that $s < \sigma_1(Q)$ for otherwise $Q' := Q$ does the job. Let $Q = [x_1, \ldots, x_k]$. We will shift weights several times, so denote by λ_i the original weights.

Let ϵ be such that $\epsilon < \lambda_i - 2\epsilon k$ for all $i \in [k]$. In a first round we add $2k$ new points $y_1, \ldots, y_{2k}$ with weight $\lambda(y_j) := \epsilon$ for all $j \in [2k]$ and reduce the weight of x_1 by $2\epsilon k$. Introduce the cover relations $x_1 <: y_1$ and $y_j <: y_{j+1}$ for all $j \in [2k-1]$ and add all forced relations induced by these.

Using observation (2.12) it is clear that $\sigma_0' \leq \sigma_0$. Hence by (2.11) $\sigma' \geq \sigma$.

In a second round, for all $i \in [k]$ consecutively, shift weight ϵ_i from x_i to y_{2i}, where $0 \leq \epsilon_i \leq \lambda(x_i) - \epsilon$. As before, use observation (2.12) to see that $\sigma_0' \leq \sigma_0$ and that thus by (2.11) $\sigma' \geq \sigma$.

So no matter how we choose ϵ and all the ϵ_i (provided they satisfy the above inequalities) the first assertion of the lemma is guaranteed.

Let Q' be the configuration created in this way. For a particular choice of ϵ make ϵ_i as large as possible, namely $\epsilon_1 := \lambda_1 - 2\epsilon k - \epsilon$ and $\epsilon_i := \lambda_i - \epsilon$ for $2 \leq i \leq k$. Then we have $\lambda(x_i) = \lambda(y_{2i-1}) = \epsilon$ for all $i \in [k]$, $\lambda(y_2) = \lambda_1 - 2\epsilon k$ and $\lambda(y_{2i}) = \lambda_i$ and hence an upper bound on $\sigma_1(Q')$ is given by

$$\sigma_1(Q') \leq \epsilon^2 \cdot k + \sum_{i=1}^{k} 2\epsilon \cdot \lambda_i \leq \epsilon \cdot (k+2).$$

This means that for a given s it is possible to choose ϵ sufficiently small so that the above 2-round-process can force $\sigma_1(Q')$ to become arbitrarily small. To ensure that the process produces $\sigma_1(Q') = s$, we choose ϵ so small that after the first round we still have $\sigma_1(Q') > s$ and $\epsilon(k+2) < s$. Then continuously increase the ϵ_i until at some point the second round must produce a Q' with $\sigma_1(Q') = s$. 2

Corollary 2.2.7 *For every configuration Q there exists a linear configuration Q'' satisfying*

$$\sigma_1(Q'') = \sigma_1(Q), \qquad \sigma_2(Q'') \geq \sigma_2(Q).$$

Proof. Apply Corollary 2.2.5 to Q and obtain a linear configuration Q' with

$$\sigma_1(Q') \geq \sigma_1(Q) \qquad \text{and} \qquad \sigma(Q') \geq \sigma(Q).$$

Now apply Lemma 2.2.6 to Q', setting $s := \sigma_1(Q) \leq \sigma_1(Q')$. We obtain a linear configuration Q'' with

$$\sigma_1(Q'') = s = \sigma_1(Q), \qquad \sigma(Q'') \geq \sigma(Q') \geq \sigma(Q),$$

and therefore

$$\sigma_2(Q'') \geq \sigma_2(Q),$$

as we were required to prove. 2

Lemma 2.2.8 *For every linear configuration Q and for every $0 \le s \le \sigma_2(Q)$ there exists a linear configuration Q' satisfying*

$$\sigma_1(Q') \ge \sigma_1(Q), \qquad \sigma_2(Q') = s.$$

Proof. Let $Q = [x_1, \ldots, x_k]$. Again denote by $\lambda_i := \lambda(x_i)$ the original weights. Start with x_1 and shift an increasing amount ϵ_1 of weight to x_3, until $\lambda(x_1) = 0$ and hence $\lambda(x_3) = \lambda_3 + \lambda_1$. Then move on to x_2, shifting weight ϵ_2 to x_4 until $\lambda(x_2) = 0$. Continue until the final step where weight ϵ_{k-2} is shifted from x_{k-2} to x_k.

Notice that whenever weight ϵ_i is shifted from x_i to x_{i+2}, we can be sure that x_i is the maximum of the chain and that the only point covered by x_i is x_{i+1}, which in turn also covers x_{i+2}. Hence $\sigma_1' \ge \sigma_1$ at all times.

Observe that if the process runs until the very end, the configuration Q' we are then left with has only two points x_{k-1} and x_k, and hence $\sigma_2(Q') = 0$. But since this process is continuous it must at one point produce a Q' with $\sigma_2(Q') = s$ for any $0 \le \sigma \le \sigma_2(Q)$. 2

Now we come back to the d- and c-value of a configuration Q. Recall that they are given by

$$\mathrm{d}(Q) = p \cdot \sigma_1(Q) + \sigma_2(Q), \qquad \mathrm{c}(Q) = H(p) \cdot \sigma_1(Q),$$

where $p = \mathrm{p}(Q)$.

Corollary 2.2.9 *For every linear configuration Q and for every d with $\frac{1}{8} \le d \le \mathrm{d}(Q)$ there exists a linear configuration Q' such that*

$$\mathrm{c}(Q') \ge \mathrm{c}(Q), \qquad \mathrm{d}(Q') = d.$$

Proof. Apply Lemma 2.2.8 to Q, with s slowly decreasing from $\sigma_2(Q)$ to 0. This means that on the one hand $\sigma_1(Q)$ does not decrease and thus $\mathrm{c}(Q)$ does not; and on the other hand it means that simultaneously $\sigma_2(Q)$ steadily approaches 0. Having arrived there, denote the new configuration by Q' and observe that $\mathrm{c}(Q') \ge \mathrm{c}(Q)$ and $\mathrm{d}(Q') = p\sigma_1(Q')$ where $p = \mathrm{p}(Q)$. If $p > 1/2$ then let now p approach $1/2$, thereby increasing $\mathrm{c}(Q') = H(p) \cdot \sigma_1(Q')$ and forcing $\mathrm{d}(Q')$ to approach $\frac{1}{2}\sigma_1(Q')$. Due to Proposition 2.2.2 we know that

$$\frac{1}{2} \cdot \sigma_1(Q') \le \frac{1}{8} \le d.$$

Hence this process must reach the point where $\mathrm{d}(Q') = d$ while maintaining at all times $\mathrm{c}(Q') \ge \mathrm{c}(Q)$. 2

Before closing this chapter, we quickly derive yet another corollary which will be of use in the next chapter.

Corollary 2.2.10 *For every significant linear configuration Q with $\mathrm{d}(Q) \geq \frac{1}{8}$ it must be true that*

$$\mathrm{p}(Q) \geq \frac{1}{2}.$$

Proof. Suppose to the contrary that there were such a Q with $p := \mathrm{p}(Q) < \frac{1}{2}$. Then increase p very slightly, thereby causing $H(p)$ to increase and thus both $\mathrm{d}(Q)$ and $\mathrm{c}(Q)$ must increase. Call the new configuration Q'' and set $d := \mathrm{d}(Q)$ so that

$$\frac{1}{8} \leq d < \mathrm{d}(Q'').$$

Applying Corollary 2.2.9 to Q'' and d, there must be a linear configuration Q' with

$$\mathrm{d}(Q') = d = \mathrm{d}(Q) \quad \text{and} \quad \mathrm{c}(Q') \geq \mathrm{c}(Q'') > \mathrm{c}(Q).$$

Hence Q cannot be significant. 2

Finally we summarize the results of this chapter in the following theorem.

Theorem 2.2.11 *For every $0 < d < \frac{1}{2}$ there exists a d-significant p-uniform linear configuration.*

Proof. For $d \leq \frac{1}{8}$ this has already been shown in Theorem 2.2.3. For $d > \frac{1}{8}$ consider an arbitrary d-significant configuration Q. By Corollary 2.2.7 there exists a linear configuration Q'' with

$$\sigma_1(Q'') = \sigma_1(Q) \ \text{ and } \ \sigma_2(Q'') \geq \sigma_2(Q).$$

Setting $\mathrm{p}(Q'') := \mathrm{p}(Q)$, this immediately implies that

$$c(Q'') = c(Q) \ \text{ and } \ d(Q'') \geq d(Q).$$

Now set $d := \mathrm{d}(Q)$ and apply Corollary 2.2.9 to Q''. Hence there must be a linear configuration Q' with

$$\mathrm{d}(Q') = d = \mathrm{d}(Q) \quad \text{and} \quad \mathrm{c}(Q') \geq \mathrm{c}(Q'') = \mathrm{c}(Q).$$

Thus if Q was d-significant so must be Q'. 2

Chapter 3

Layer posets

The aim of this chapter is to prove our main result as stated in Theorems 1.3.1 and 1.3.2. We first review the progress we have made in the previous chapter and then sketch the underlying ideas of this chapter.

Recall that by Theorem 2.1.5 for every d there exists a configuration Q which satisfies $\mathrm{c}(d) = \mathrm{c}(Q)$. By definition, this means that Q is d-significant (with $d = \mathrm{d}(Q)$). Since we know that there exists at least one such configuration, a sufficient and necessary condition for Q to be significant is that there is no configuration Q' such that

$$\mathrm{d}(Q') = \mathrm{d}(Q) \qquad \text{and} \qquad \mathrm{c}(Q') > \mathrm{c}(Q).$$

When searching for such a configuration, we can rely on Theorem 2.2.11, which states that for every d there exists a *linear p-uniform* d-significant configuration. Recall from Section 1.3 that such a configuration Q is fully described by the vector

$$Q := (\vec{\lambda}; p) := (\lambda_1, \ldots, \lambda_k;\ p),$$

and furthermore that (using (2.3) and (2.5)) in this case $\mathrm{c}(Q)$ and $\mathrm{d}(Q)$ are given by

$$\mathrm{c}(Q) = \mathrm{c}(\vec{\lambda}; p) = H(p) \cdot \sum_{i=1}^{k-1} \lambda_i \lambda_{i+1}$$

and

$$\mathrm{d}(Q) = \mathrm{d}(\vec{\lambda}; p) = \frac{1}{2} - \frac{1}{2} \sum_{i=1}^{k} (\lambda_i)^2 - (1-p) \cdot \sum_{i=1}^{k-1} \lambda_i \lambda_{i+1}.$$

Moreover notice that we can restrict our attention to $d > \frac{3}{16}$ because in the range $(0, \frac{3}{16}]$ the function $\mathrm{c}(d)$ is determined by (1.14) and Theorem 2.2.3. Finally observe that, due to Corollary 2.2.10, we can therefore assume that $\mathrm{p}(Q) \geq \frac{1}{2}$. Hence a solution of the following

constrained extremum problem must be a d-significant configuration, and it is exactly this configuration that we will be seeking.

$$\boxed{\begin{array}{lll} \text{Choose } k,\ \lambda_1,\dots,\lambda_k,\ \text{and } p & \text{such as to maximize} & H(p)\sum_{i=1}^{k-1}\lambda_i\lambda_{i+1}, \\ & \text{subject to} & \frac{1}{2}-\frac{1}{2}\sum_{i=1}^{k}\lambda_i^2-(1-p)\sum_{i=1}^{k-1}\lambda_i\lambda_{i+1}=d, \\ & & \sum_{i=1}^{k}\lambda_i=1, \\ & & \lambda_i>0 \qquad \text{for all } i=1,\dots,k,\ \text{and} \\ & & \frac{1}{2}\le p\le 1. \end{array}} \tag{3.1}$$

Denote by $Q^*=(\vec{\lambda};p)$ a solution of the above maximization problem. In the following we will derive necessary conditions that $(\vec{\lambda};p)$ must satisfy. (Obviously the idea here is that at the end we will actually be able to deduce the values $\lambda_1,\dots,\lambda_k,p$ from these conditions.)

From the theory of Lagrange multipliers we know that for any local extremum $z=(z_1,\dots,z_m)$ of a function $f:\mathrm{R}^m\to\mathrm{R}$ under the side constraints $g_i(z)=0$ for $i\in[\ell]$ there must exist real numbers t_i such that

$$\frac{\partial f}{\partial z_i}(z)=t_1\cdot\frac{\partial g_1}{\partial z_i}(z)+\cdots+t_\ell\cdot\frac{\partial g_\ell}{\partial z_i}(z)$$

for all $i\in[m]$. Hence in our case, where $m=k+1$ and $\ell=2$, there must be real numbers r,u such that

$$\frac{\partial}{\partial p}:\qquad H'(p)\sum_{i=1}^{k-1}\lambda_i\lambda_{i+1}+\sum_{i=1}^{k-1}\lambda_i\lambda_{i+1}\cdot r=0, \tag{3.2}$$

$$\frac{\partial}{\partial \lambda_i}:\qquad H(p)\lambda_{i+1}+H(p)\lambda_{i-1}+\Big(-\lambda_i-(\lambda_{i-1}+\lambda_{i+1})(1-p)\Big)\cdot r+u=0, \tag{3.3}$$

for all $i\in[k]$, where we set $\lambda_0=\lambda_{k+1}=0$ for convenience.

From (3.2) we obtain immediately that p and r must satisfy $H'(p)=-r$. Using this, (3.3) simplifies to

$$\lambda_iH'(p)+\lambda_{i-1}\Big(H(p)+(1-p)H'(p)\Big)+\lambda_{i+1}\Big(H(p)+(1-p)H'(p)\Big)=-u. \tag{3.4}$$

Setting

$$a(p):=-H'(p)=\log(p)-\log(1-p)$$

and

$$b(p) := -H(p) - (1-p)H'(p) = \log p,$$

it is easy to see that (3.4) corresponds to a system of linear equations

$$T^k(p) \cdot \vec{\lambda} = \vec{u}, \qquad \text{where} \tag{3.5}$$

$$T^k(p) = \begin{pmatrix} a(p) & b(p) & 0 & 0 & \dots & 0 & 0 \\ b(p) & a(p) & b(p) & 0 & \dots & 0 & 0 \\ 0 & b(p) & a(p) & b(p) & \dots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & 0 & 0 & \dots & b(p) & a(p) \end{pmatrix}, \qquad \vec{\lambda} = \begin{pmatrix} \lambda_1 \\ \lambda_2 \\ \vdots \\ \vdots \\ \lambda_k \end{pmatrix}, \qquad \vec{u} = \begin{pmatrix} u \\ u \\ \vdots \\ \vdots \\ u \end{pmatrix}.$$

Outline of the proofs. Before the technical details become overwhelming, let us pause here to re–evaluate our position and to sketch the line of thought in what is to come.

Let us suppose for a moment that k and p are such that $T^k(p)$ is non–singular. Then the system (3.5) has a unique solution $\vec{\lambda}$. In other words, for such values k and p there exists only one configuration of the form $(\lambda_1, \dots, \lambda_k; p)$ that can possibly be the solution Q^* of (3.1) – call this configuration the (k,p)-*candidate* and denote it by $Q_{k,p}$. It is determined in Proposition 3.1.4. Since (3.5) is (only) a necessary condition for a configuration to be the solution of (3.1), it is clear that on the one hand Q^* must be a (k,p)-candidate for some k, p, but also on the other hand that not every candidate is significant, since there might well be another candidate $Q_{k',p'}$ which has the same d-value but a higher c-value.

So we are faced with the problem of comparing the c-values of candidates that have the same d-value. As the computations will show, it is relatively easy to determine $Q_{k,p}$ and hence $\mathrm{d}(Q_{k,p})$ as a function of k and p. However here we really need the inverse, i.e. for a given d we must first determine those candidates with d-value d and then compare their c-values. After a glance at the formulae needed to describe $Q_{k,p}$, this seems very hard indeed.

We therefore propose a different approach. We will prove that there is a sequence of constants $0 < q^3 < q^4 < \dots < q^*$ so that for each k and all $p \notin [q^k, q^{k+1}]$ the respective (k,p)-candidate $Q_{k,p}$ cannot be significant. We will prove this by comparing $Q_{k,p}$ to a different configuration Q', that is obtained by slightly modifying $Q_{k,p}$, and that achieves the same d-value but a higher c-value. (Note that usually Q' will not be a (p', k')-candidate, so the idea is not to elect the best possible candidate but to simply 'dismiss' all the others.)

We have tried to illustrate this approach in Figure 3.1. Here we took $Q_{3,p}$ and $Q_{4,p}$ and plotted their d-value against their c-value. The dashed parts of the curves are those to be dismissed, while the solid part represents the part $p \in [q^3, q^4]$ and $p \in [q^4, q^5]$ respectively.

Let us say that as p increases from q^k to q^{k+1}, the function $\mathrm{d}(Q_{k,p})$ increases from d_k to d'_{k+1}. What we will then have to show is that the intervals $[d_k, d'_{k+1}]$ don't overlap for

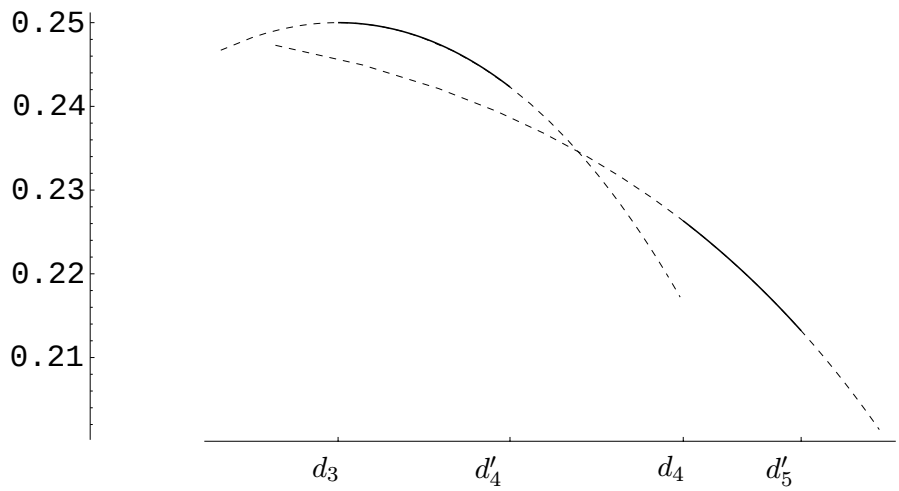

Figure 3.1: $Q_{3,p}$ and $Q_{4,p'}$ where $p \in [0.45, 0.68]$ and $p' \in [0.52, 0.65]$.

different k and k', so that for $d \in [d_k, d'_{k+1}]$ the configuration $Q_{k,p}$ (where $p \in [q^k, q^{k+1}]$) is the only candidate left and must therefore be significant.

As we saw in Figure 3.1, not only do the intervals $[d_3, d'_4]$ and $[d_4, d'_5]$ not overlap, but there are actually gaps between them. Of course this cannot mean that we here have a total lack of significant configurations. Indeed, recall that we had assumed $T^k(p)$ to be non–singular. However there are values of p and k where this is not true. Here there is no unique solution to (3.5), so we can let one of the components of $\vec{\lambda}$ have some arbitrary value x, neglect one of the equations and solve the remaining system, which then always turns out to be non–singular. Denote the solution of this system by $Q_{k,p}(x)$. It is determined im Proposition 3.1.5. For some of these singular couples k, p (namely when $p = q^k$) the candidates $Q_{k,p}(x)$ turn out to be significant, at least for small x. Moreover, as we slowly increase x one can observe that in these cases $Q_{k,p}(x)$ beautifully builds a 'linear bridge' accross the gaps noticed above. This is illustrated by the bold line in Figure 3.2.

We have subdivided the remaining part of this chapter into four sections. In Section 3.1 we determine explicit representations of the candidates $Q_{k,p}$ and $Q_{k,p}(x)$. As described above, Section 3.2 then shows that many of them cannot be significant. In Section 3.3 we then prove some facts about the behaviour of the remaining candidates and complete the proof of Theorems 1.3.1 and 1.3.2. In the course of these three sections we will need a few lemmas and propositions whose proofs are technical in nature and do not provide any new insight. So whenever a lemma or a proposition is stated without proof, the latter can be found in Section 3.4.

We introduce one last definition. Observe that the candidates $Q_{k,p}$ and $Q_{k,p}(x)$ as

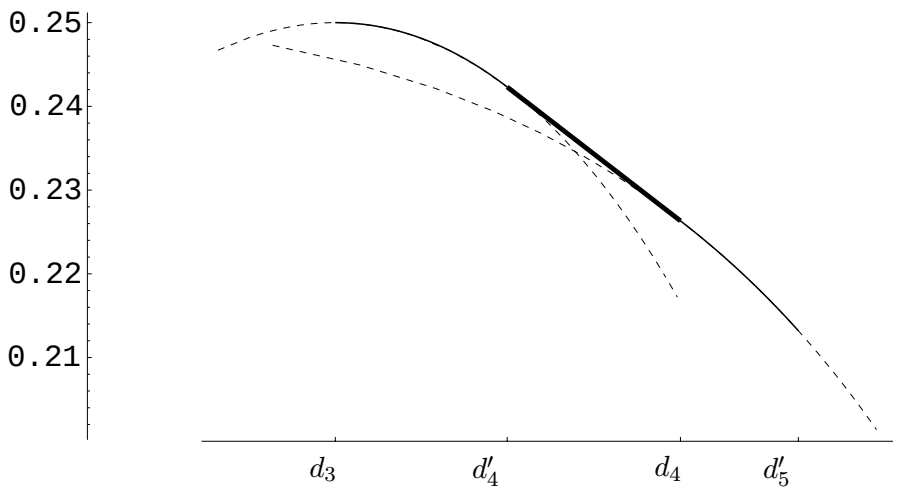

Figure 3.2: $Q_{3,p}$, $Q_{4,p'}$ and $Q_{4,q^4}(x)$ where $q^4 = 0.5827$ and $x \in [0, 0.1382]$.

computed in (3.5) and (3.10) respectively don't automatically satisfy the condition $0 < \lambda_i$ for all $i \in [k]$. If they do, we will call them *feasible*.

3.1 Determining candidates

At first it might seem as if (3.5) wouldn't actually produce a unique candidate $Q_{k,p}$ (even if $T^k(p)$ is non–singular), because one is allowed to choose u arbitrarily. However solutions $\vec{\lambda}$ and $\vec{\lambda}'$ of (3.5) are linear dependent and therefore there is only one solution that simultaneously satisfies $\sum \lambda_i = 1$. Hence for every fixed k and p where $T^k(p)$ is non-singular, $Q_{k,p}$ is given by $(\vec{\lambda}; p)$ with

$$\vec{\lambda}(p) := \frac{\vec{\xi}(p)}{|\vec{\xi}(p)|} \quad \text{and} \quad \vec{\xi}(p) := \left[T^k(p)\right]^{-1} \cdot \vec{1}. \tag{3.6}$$

Here $\vec{1}$ is the vector whose components are all 1, and $|\vec{v}|$ denotes the sum of the components. (This somewhat abuses the notion of a norm, since it could happen that $|\vec{v}| < 0$.)

The next two lemmas compute the determinant and inverse of a tridiagonal $k \times k$ matrix T^k with a on the main diagonal and b on the two subdiagonals for arbitrary real numbers a and b.

Lemma 3.1.1

$$t_k := \det(T^k) = \begin{cases} \frac{1}{D}\left[\left(\frac{a+D}{2}\right)^{k+1} - \left(\frac{a-D}{2}\right)^{k+1}\right] \textit{ where } D = \sqrt{a^2 - 4b^2} & , \textit{ if } \frac{a}{2b} \notin (-1,1) \\ b^k \sin(k+1)\phi / \sin\phi \textit{ where } \phi = \arccos\frac{a}{2b} & , \textit{ if } \frac{a}{2b} \in (-1,1). \end{cases}$$

Corollary 3.1.2 *$t_k = 0$ if and only if $a^2 = 4b^2$ or $\frac{a}{2b} = \cos\frac{j\pi}{k+1}$ for some $j \in \{1, \dots, k\}$. In particular, for $-1 < a/2b < 1$ it follows from $t_k = 0$ that $t_{k-1} \neq 0$.* 2

Lemma 3.1.3 *The inverse $M^k = (T^k)^{-1}$ is given by (m_{ij}) where*

$$m_{ij} := \begin{cases} (-1)^{i+j} b^{i-j} \frac{t_{j-1} t_{k-i}}{t_k} & , \text{if } j < i \\ (-1)^{i+j} b^{j-i} \frac{t_{i-1} t_{k-j}}{t_k} & , \text{if } j \geq i. \end{cases}$$

Before we use this to compute $\vec{\lambda}$, observe that in the case of our particular matrix $T^k(p)$ the function $\frac{a(p)}{2b(p)}$ is strictly decreasing for $\frac{1}{2} \leq p \leq 1$ and ranges from 0 to $-\infty$, attaining value -1 at $p = 0.682328... =: q^*$. Hence for $\frac{1}{2} \leq p \leq q^*$ the function $\phi(p) := \arccos\frac{a(p)}{2b(p)}$ is strictly increasing and ranges from $\pi/2$ to π. We will make use of the one-to-one relation between p and $\phi(p)$ and sometimes express functions in terms of $\phi = \phi(p)$ instead of p. Hopefully this will make things easier to read, as the two preceding lemmas seem to indicate.

Now Lemmas 3.1.1 and 3.1.3 enable us to give an explicit formula for $Q_{k,p}$ as suggested in (3.6).

Proposition 3.1.4 *Denote by $M^k(p) = (m^k(p)_{ij})$ the inverse of $T^k(p)$. For all k and all p with $\pi/2 < \phi < \pi$ and $\phi \neq \frac{j\pi}{k+1}$ for all $j \in \{1, \dots, k\}$, we must have*

$$Q_{k,p} = \Big(\lambda_1^k(p), \dots, \lambda_k^k(p);\ p\Big),$$

where

$$\lambda_i^k(p) := \frac{\xi_i^k(p)}{|\vec{\xi}^k(p)|} = \frac{(-1)^{k-i}\sin i\phi - (-1)^i \sin(k-i+1)\phi + \sin(k+1)\phi}{\frac{-(-1)^k \sin\phi + \sin k\phi + \sin(k+1)\phi}{1+\cos\phi} + k\sin(k+1)\phi},$$

$$\xi_i^k(p) := \sum_{j=1}^k m_{i,j}^k(p) = \frac{(-1)^{k-i}\sin i\phi - (-1)^i \sin(k-i+1)\phi + \sin(k+1)\phi}{2b(p)(1+\cos\phi)\sin(k+1)\phi}.$$

In particular, $Q_{k,p}$ is symmetric, i.e. for all $i \in [k]$ we have

$$\lambda_i^k(p) \ = \ \lambda_{k-i+1}^k(p).$$

Proof. In view of (3.6) it suffices to check the explicit formulae; namely

$$\sum_{j=1}^k m_{i,j}^k(p) \ = \ \frac{(-1)^{k-i}\sin i\phi - (-1)^i \sin(k-i+1)\phi + \sin(k+1)\phi}{2b(p)(1+\cos\phi)\sin(k+1)\phi} \tag{3.7}$$

and

$$|\vec{\xi}^k(p)| = \frac{\frac{-(-1)^k \sin\phi + \sin k\phi + \sin(k+1)\phi}{1+\cos\phi} + k\sin(k+1)\phi}{2b(p)(1+\cos\phi)\sin(k+1)\phi}. \tag{3.8}$$

This very technical task is deferred to Section 3.4. The symmetry of the $\lambda_i^k(p)$ is indeed obvious from the definition. 2

Now we turn to the candidates $Q_{k,p}(x)$, i.e. to those pairs k, p where

$$\frac{a(p)}{2b(p)} = \cos\frac{j\pi}{k+1} \text{ for some } j \in \{1, \ldots, k\}. \tag{3.9}$$

Here $T^k(p)$ is singular and there is no unique solution to the system of equations (3.5), which we now restate as

$$T^k(p) \cdot \vec{\ell} = \vec{u}.$$

Hence we may choose one ℓ_i-value arbitrarily, say $\ell_k = x$, and neglect the last equation of the system. The requirement for the remaining ℓ_i is that there exists a u' such that

$$\begin{pmatrix} a(p) & b(p) & 0 & 0 & \ldots & 0 & 0 & 0 & 0 \\ b(p) & a(p) & b(p) & 0 & \ldots & 0 & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & \ldots & b(p) & a(p) & b(p) & 0 \\ 0 & 0 & 0 & 0 & \ldots & 0 & b(p) & a(p) & b(p) \end{pmatrix} \cdot \begin{pmatrix} \ell_1 \\ \ell_2 \\ \vdots \\ \vdots \\ \ell_{k-1} \\ x \end{pmatrix} = \begin{pmatrix} u' \\ \vdots \\ \vdots \\ \vdots \\ u' \end{pmatrix}$$

which is equivalent to

$$T^{k-1}(p) \cdot \begin{pmatrix} \ell_1 \\ \vdots \\ \ell_{k-2} \\ \ell_{k-1} \end{pmatrix} = \begin{pmatrix} u' \\ \vdots \\ u' \\ u' - b(p) \cdot x \end{pmatrix}, \tag{3.10}$$

whose solution is uniquely determined since $\det(T_{k-1}) \neq 0$, as we saw in Corollary 3.1.2. Not surprisingly, the solution of (3.10) is related to that of (3.5).

Proposition 3.1.5 *For all $x \in (0,1)$, all k and all p with $\phi = \frac{j\pi}{k+1}$ for some $j \in \{1, \ldots, k\}$*

$$Q_{k,p}(x) = \Big(\ell_1^k(p,x), \ldots, \ell_k^k(p,x); p\Big),$$

where

$$\begin{aligned}
\ell_i^k(p,x) &:= \lambda_i^{k-1}(p) + x\cdot\mu_i^{k-1}(p) \quad \textit{for } i=1,\dots,k-1,\\
\ell_k^k(p,x) &:= x,\\
\mu_i^{k-1}(p) &:= -\lambda_i^{k-1}(p)\Big(1-b(p)\cdot\xi_{k-1}^{k-1}(p)\Big) - b(p)m_{i,k-1}^{k-1} \quad \textit{for } i=1,\dots,k-1,\\
\mu_k^{k-1}(p) &:= 1.
\end{aligned}$$

Proof. For $i=1,\dots,k-1$ denote by $\vec{r}_i$ the ith row of $T^{k-1}(p)$ and by $\vec{m}$ the $(k-1)$st column of $M^{k-1}(p)$, hence $\vec{r}_i\cdot\vec{m}=\delta_{i,k-1}$. Since $\vec{\lambda}^{k-1}(p)$ is a solution of $T_{k-1}(p)\cdot\vec{\lambda}^{k-1}(p)=\vec{u}$ for an appropriate value u, we know that $\vec{r}_i\cdot\vec{\lambda}^{k-1}(p)=u$ for all $1\le i\le k-1$. Now let $s:=1-b(p)\cdot\xi_{k-1}^{k-1}(p)$, hence

$$\vec{r}_i\cdot\begin{pmatrix}\ell_1^k(p,x)\\ \vdots\\ \ell_{k-1}^k(p,x)\end{pmatrix} = u + x\cdot\vec{r}_i\cdot\begin{pmatrix}\mu_1^{k-1}(p)\\ \vdots\\ \mu_{k-1}^{k-1}(p)\end{pmatrix}$$

$$= u - x\cdot s\cdot\vec{r}_i\cdot\begin{pmatrix}\lambda_1^{k-1}(p)\\ \vdots\\ \lambda_{k-1}^{k-1}(p)\end{pmatrix} - x\cdot b(p)\cdot\vec{r}_i\cdot\vec{m} = u - s\cdot x\cdot u - x\cdot b(p)\cdot\delta_{i,k-1}.$$

Setting $u':=u-s\cdot x\cdot u$, this satisfies (3.10). Furthermore

$$\sum_{i=1}^{k-1}\mu_i^{k-1}(p) = -s\sum_{i=1}^{k-1}\lambda_i^{k-1}(p) - b(p)\cdot\sum_{i=1}^{k-1}m_{i,k-1}^{k-1}(p) = -1,$$

because

$$s = 1-b(p)\cdot\xi_{k-1}^{k-1}(p), \quad \sum_{i=1}^{k-1}\lambda_i^{k-1}(p)=1 \quad \text{and} \quad \sum_{i=1}^{k-1}m_{i,k-1}^{k-1}(p)=\xi_{k-1}^{k-1}(p).$$

Therefore

$$|\vec{\ell}| = \sum_{i=1}^{k}\ell_i^k(p,x) = \sum_{i=1}^{k-1}\lambda_i^{k-1}(p) + x\cdot\sum_{i=1}^{k-1}\mu_i^{k-1}(p) + x = 1.$$

2

Let us introduce some definitions for those p where $T^k(p)$ is singular. Set $\phi_j^k := \frac{k-j}{k+1}\pi$ and define q_j^k by $\phi(q_j^k)=\phi_j^k$. We shall see that ϕ_1^k plays a special role, therefore let $\phi^k:=\phi_1^k$ and $q^k=q_1^k$. We will often use the following trivial but helpful observation:

$$\begin{aligned}
\sin(k-i)\phi_j^k &= \sin[(k+1)\tfrac{k-j}{k+1}\pi - (i+1)\phi_j^k]\\
&= (-1)^{k-j}\sin[-(i+1)\phi_j^k] \;=\; -(-1)^{k-j}\sin(i+1)\phi_j^k. \qquad (3.11)
\end{aligned}$$

In order to obtain an explicit formula for $\vec{\ell}$ as determined in Proposition 3.1.5, we will need the following technical lemma.

Lemma 3.1.6

$$m_{i,k-1}^{k-1}(q_j^k) = (-1)^{i+j}\frac{\sin i\phi_j^k}{b(q_j^k)\sin\phi_j^k}$$

$$\xi_{k-1}^{k-1}(q_j^k) = \begin{cases} \frac{1}{b(q_j^k)} & , \textit{ if } j \textit{ is odd} \\ \frac{1}{b(q_j^k)} - \frac{1}{b(q_j^k)(1+\cos\phi_j^k)} & , \textit{ if } j \textit{ is even.} \end{cases}$$

$$\mu_i^{k-1}(q_j^k) = \begin{cases} (-1)^i\frac{\sin i\phi_j^k}{\sin\phi_j^k} & , \textit{ for all } 1 \le i \le k \quad \textit{if } j \textit{ is odd} \\ -(-1)^i\frac{\sin i\phi_j^k}{\sin\phi_j^k} - \frac{\lambda_i^{k-1}(q_j^k)}{1+\cos\phi_j^k} & , \textit{ for all } 1 \le i \le k-1 \quad \textit{if } j \textit{ is even.} \end{cases}$$

For $1 \le i \le k-1$ we have

$$\ell_i^k(q_j^k, x) = \begin{cases} \lambda_i^{k-1}(q_j^k) - x\cdot b(q_j^k)\cdot m_{i,k-1}^{k-1}(q_j^k) & , \textit{ if } j \textit{ is odd} \\ (1-\frac{x}{1+\cos\phi_j^k})\lambda_i^{k-1}(q_j^k) - x\cdot b(q_j^k)\cdot m_{i,k-1}^{k-1}(q_j^k) & , \textit{ if } j \textit{ is even.} \end{cases}$$

For an odd number j the terms $\mu_i^{k-1}(q_j^k)$ have some nice properties that we will use later, especially in the case where $j = 1$. Notice that $\mu_i^{k-1}(\phi_j^k)$ is the same as $\mu_i^{k-1}(q_j^k)$. We will frequently switch between the two expressions.

Lemma 3.1.7 *Now let $j \in [k]$ be an odd number. Then*

$$(i) \qquad \mu_{k+1-i}^{k-1}(\phi_j^k) = -\mu_i^{k-1}(\phi_j^k)$$

$$(ii) \qquad \sum_{i=1}^{k}\mu_i^{k-1}(\phi_j^k) = 0$$

$$(iii) \qquad \sum_{i=1}^{k}\mu_i^{k-1}(\phi_j^k)^2 = \frac{k+1}{2\sin^2(\phi_j^k)}$$

$$(iv) \qquad \sum_{i=1}^{k-1}\mu_i^{k-1}(\phi_j^k)\cdot\mu_{i+1}^{k-1}(\phi_j^k) = -\frac{(k+1)\cos(\phi_j^k)}{2\sin^2(\phi_j^k)}.$$

Let $j' \neq j$ be another odd number in $[k]$. Then

$$(v) \qquad \sum_{i=1}^{k}\mu_i^{k-1}(\phi_j^k)\cdot\mu_i^{k-1}(\phi_{j'}^k) = 0$$

$$(vi) \qquad \sum_{i=1}^{k-1}\mu_i^{k-1}(\phi_j^k)\cdot\mu_{i+1}^{k-1}(\phi_{j'}^k) = 0.$$

3.2 Dismissing candidates

In this section we follow the approach outlined in the beginning of the chapter. Namely we shall use the explicit representation just derived for $Q_{k,p}$ and show that for $p \notin [q^k, q^{k+1}]$ the candidate $Q_{k,p}$ cannot be significant. We then proceed analogously in the case of candidates $Q_{k,p}(x)$. The following lemma is our main tool in this enterprise.

Lemma 3.2.1 *Consider an arbitrary configuration $(\lambda_1, \ldots, \lambda_r; p)$ where $p < 1$ is a fixed constant. Let $(\nu_1, \ldots, \nu_{r'})$ be a vector with $|\vec{\nu}| = 0$, $r \leq r'$ and $\nu_i \geq 0$ for all $r < i \leq r'$. Set $\lambda_{r+1} = \cdots = \lambda_{r'} = 0$ and define*

$$\alpha_1 := 2\sum_{i=1}^{r'} \lambda_i \nu_i \quad \textit{and} \quad \beta_1 := \sum_{i=1}^{r'} \nu_i^2,$$

$$\alpha_2 := \sum_{i=1}^{r'-1} (\nu_i \lambda_{i+1} + \nu_{i+1}\lambda_i) \quad \textit{and} \quad \beta_2 := \sum_{i=1}^{r'-1} \nu_i \nu_{i+1}.$$

If

$$-b(p) \cdot \alpha_2 - \frac{a(p)}{2} \cdot \alpha_1 \ > \ 0, \tag{3.12}$$

or

$$\alpha_1 = \alpha_2 = 0 \ \textit{and} \ -b(p) \cdot \beta_2 - \frac{a(p)}{2} \cdot \beta_1 > 0, \tag{3.13}$$

then $(\lambda_1, \ldots, \lambda_r; p)$ cannot be significant.

Proof. Suppose to the contrary that $(\lambda_1, \ldots, \lambda_r; p)$ were significant. Define

$$\lambda_i' := \lambda_i + \epsilon \cdot \nu_i \quad \text{for all } i = 1, \ldots, r',$$

$$\gamma_1 := \alpha_1 + \epsilon \cdot \beta_1, \qquad \gamma_2 := \alpha_2 + \epsilon \cdot \beta_2,$$

$$p' := p + \epsilon \cdot q, \text{ where } q := \frac{(1-p)\gamma_2 + \frac{\gamma_1}{2}}{\sum_{i=1}^{r-1} \lambda_i \lambda_{i+1} \ + \epsilon\gamma_2}.$$

We claim that $\mathrm{d}(\lambda_1, \ldots, \lambda_r; p) = \mathrm{d}(\lambda_1', \ldots, \lambda_r'; p')$, which is verified as follows.

$$\begin{aligned} \sum_{i=1}^{r'} (\lambda_i')^2 - \sum_{i=1}^{r} (\lambda_i)^2 \ &= \ \sum_{i=1}^{r'} \left(\lambda_i^2 + \epsilon^2 \nu_i^2 + 2\epsilon\lambda_i\nu_i - \lambda_i^2\right) \\ &= \ \epsilon \sum_{i=1}^{r'} \left(\epsilon\nu_i^2 + 2\lambda_i\nu_i\right) \ = \ \epsilon\gamma_1, \end{aligned}$$

$$\begin{aligned}\sum_{i=1}^{r'-1}\lambda'_i\lambda'_{i+1}-\sum_{i=1}^{r-1}\lambda_i\lambda_{i+1} &= \sum_{i=1}^{r'-1}((\lambda_i+\epsilon\nu_i)(\lambda_{i+1}+\epsilon\nu_{i+1})-\lambda_i\lambda_{i+1})\\ &= \epsilon\sum_{i=1}^{r'-1}(\epsilon\nu_i\nu_{i+1}+\nu_i\lambda_{i+1}+\nu_{i+1}\lambda_i) \;=\; \epsilon\gamma_2.\end{aligned}$$

$$\mathrm{d}(\lambda'_1,\ldots,\lambda'_{r'};p') = \frac{1}{2}-\frac{1}{2}\sum_{i=1}^{r'}(\lambda'_i)^2-(1-p-\epsilon q)\cdot\sum_{i=1}^{r'-1}(\lambda'_i\lambda'_{i+1})$$

$$= \frac{1}{2}-\frac{1}{2}\sum_{i=1}^{r}\lambda_i^2-\frac{\epsilon\gamma_1}{2}-(1-p)(\sum_{i=1}^{r-1}\lambda_i\lambda_{i+1}+\epsilon\gamma_2)+\epsilon q(\sum_{i=1}^{r-1}\lambda_i\lambda_{i+1}\;+\epsilon\gamma_2)$$

$$= \mathrm{d}(\lambda_1,\ldots,\lambda_r;p)-\frac{\epsilon\gamma_1}{2}-(1-p)\epsilon\gamma_2+\epsilon\Big((1-p)\gamma_2+\frac{\gamma_1}{2}\Big) \;=\mathrm{d}(\lambda_1,\ldots,\lambda_r;p).$$

Thus we have shown that the configurations $(\vec{\lambda};p)$ and $(\vec{\lambda}';p')$ have the same d-value. Furthermore, if ϵ is sufficiently small and $(\vec{\lambda};p)$ is a feasible configuration, then so must be $(\vec{\lambda}';p')$. Moreover we have $|\vec{\lambda}'|=|\vec{\lambda}|+\epsilon|\vec{\nu}|=|\vec{\lambda}|=1$. Hence if we now prove that for all small enough ϵ the c-value of $(\vec{\lambda}';p')$ is greater than that of $(\vec{\lambda};p)$, the latter cannot be significant.

We will use the Taylor expansion of $H(p)$ to compute the value of $H(p')$: there exists a $0\le\eta'\le 1$ such that

$$H(p')=H(p+\epsilon q)=H(p)+\epsilon qH'(p)+\frac{\epsilon^2q^2}{2}H''(p+\eta'\epsilon q),$$

therefore there exists a positive constant η, such that for all sufficiently small ϵ we have $H(p')\ge H(p)+\epsilon qH'(p)-\eta\epsilon^2q^2$. This implies

$$\mathrm{c}(\lambda'_1,\ldots,\lambda'_{r'};p')-\mathrm{c}(\lambda_1,\ldots,\lambda_r;p)=H(p')\sum_{i=1}^{r'-1}(\lambda'_i\lambda'_{i+1}) \;-\; H(p)\sum_{i=1}^{r-1}(\lambda_i\lambda_{i+1})$$

$$\ge (H(p)+\epsilon qH'(p)-\eta\epsilon^2q^2)(\sum_{i=1}^{r-1}\lambda_i\lambda_{i+1}+\epsilon\gamma_2) \;-\; H(p)\sum_{i=1}^{r-1}\lambda_i\lambda_{i+1}$$

$$=\epsilon\gamma_2H(p)+q(\epsilon H'(p)-\eta\epsilon^2q)(\sum_{i=1}^{r-1}\lambda_i\lambda_{i+1}+\epsilon\gamma_2)$$

$$=\epsilon\gamma_2H(p)+(\epsilon H'(p)-\eta\epsilon^2q)((1-p)\gamma_2+\frac{\gamma_1}{2})$$

$$=\epsilon\gamma_2\big(H(p)+(1-p)H'(p)\big)+\epsilon H'(p)\frac{\gamma_1}{2}-\eta\epsilon^2q((1-p)\gamma_2+\frac{\gamma_1}{2}).$$

$$= \epsilon\Big(- \gamma_2 b(p) - a(p)\frac{\gamma_1}{2} - \epsilon \cdot \eta q((1-p)\gamma_2 + \frac{\gamma_1}{2})\Big).$$

Now either of the two conditions (3.12) or (3.13) immediately guarantees that we can choose ϵ so small that the last expression certainly becomes positive, which completes the proof of this lemma. 2

Corollary 3.2.2 *No configuration* $(\lambda_1, \ldots, \lambda_k, p)$ *with*

$$b(p) \cdot \lambda_{k-2} + a(p) \cdot \lambda_{k-1} > 0$$

is significant.

Proof. We would like to apply Lemma 3.2.1 with $r := k$, $r' := k+1$ and

$$\nu_{k-1} := -1, \ \nu_{k+1} := 1, \quad \text{and} \quad \nu_i := 0 \ \text{ for } i = 1, \ldots, k-2, k.$$

Then

$$\alpha_1 = -2\lambda_{k-1} \quad \text{and} \quad \alpha_2 = -\lambda_{k-2} - \lambda_k + \lambda_k = -\lambda_{k-2},$$

hence

$$-b(p)\alpha_2 - \frac{a(p)}{2}\alpha_1 = b(p)\lambda_{k-2} + a(p)\lambda_{k-1} > 0,$$

which means that (3.12) is satisfied. Hence Lemma 3.2.1 can be applied and $(\lambda_1, \ldots, \lambda_k, p)$ is not significant. 2

Proposition 3.2.3 *For* $p < q^k$ *no configuration* $(\lambda_1, \ldots, \lambda_k; p)$ *that is symmetric (i.e. one that satisfies* $\lambda_i = \lambda_{k+1-i}$ *for* $i = 1, \ldots, k$*) can be significant.*
In particular, for $p < q^k$ *the configuration* $Q_{k,p} = (\lambda_1^k(p), \ldots, \lambda_k^k(p); p)$ *is not significant.*

Proof. We apply Lemma 3.2.1 with $r := r' := k$ and $\nu_i := \mu_i^{k-1}(q^k)$ and claim that (3.13) holds.

First we use the properties of ν_i as guaranteed by Lemma 3.1.7 to see that indeed $\alpha_1 = \alpha_2 = 0$:

$$\alpha_1 = 2\sum_{i=1}^{k} \lambda_i \nu_i = 0$$

because of $\lambda_i = \lambda_{k-i+1}$ and $\nu_i = -\nu_{k-i+1}$. Similarly

$$\alpha_2 = \sum_{i=1}^{k-1} (\nu_i \lambda_{i+1} + \nu_{i+1}\lambda_i) = 0,$$

since $\nu_{i+1}\lambda_i = \nu_{i+1}\lambda_{k-i+1} = -\nu_{k-(i+1)+1}\lambda_{k-i+1} = -\nu_{k-i}\lambda_{k-i+1}$.

Again using Lemma 3.1.7 now observe that

$$-\frac{\beta_2}{\beta_1} = -\frac{\sum_{i=1}^{k-1} \nu_i \nu_{i+1}}{\sum_{i=1}^{k} \nu_i^2} = \cos \phi^k.$$

Since $\beta_1 > 0$ and $b(p) < 0$, it now follows that

$$\begin{aligned} -b(p)\beta_2 - \frac{a(p)}{2}\beta_1 > 0 \quad &\Leftrightarrow \quad \frac{a(p)}{2b(p)} > -\frac{\beta_2}{\beta_1} \\ &\Leftrightarrow \quad \cos\phi > \cos\phi^k, \end{aligned}$$

which is true, because $p < q_k$ is equivalent to $\phi < \phi^k \leq \pi$. 2

Proposition 3.2.4 *For $p > q^{k+1}$ the configuration $Q_{k,p} = (\lambda_1^k(p), \ldots, \lambda_k^k(p); p)$ is not significant.*

Proof. We would like to apply Corollary 3.2.2 and therefore need to show that for a fixed $p > q^{k+1}$

$$b(p)\lambda_{k-2}^k(p) + a(p)\lambda_{k-1}^k(p) > 0. \tag{3.14}$$

First we consider the case $q^{k+1} < p < q^*$. Since $a(p) > 0$ for $p > \frac{1}{2}$ and since we can assume that $\lambda_{k-1}^k(p) > 0$ (otherwise $Q_{k,p}$ wouldn't even be feasible), (3.14) is equivalent to

$$1 > -\frac{b(p)}{a(p)} \frac{\lambda_{k-2}^k(p)}{\lambda_{k-1}^k(p)}.$$

Using the explicit formula in Proposition 3.1.4 (observe that the denominator is the same for all λ_i and therefore cancels) we have to show that

$$1 > \frac{\sin(k-2)\phi - (-1)^k \sin 3\phi + \sin(k+1)\phi}{-2\cos\phi(-\sin(k-1)\phi + (-1)^k \sin 2\phi + \sin(k+1)\phi)} \quad \text{for } \frac{k}{k+2}\pi < \phi < \pi. \tag{3.15}$$

We defer this hardly interesting computation to Section 3.4.

Now consider the case $p \geq q^*$. Here $a(p)/2b(p) \leq -1$ and by Lemma 3.1.1

$$t_i(p) = \det(T^i(p)) > 0. \tag{3.16}$$

Recall that

$$\lambda_i^k(p) = \frac{\xi_i^k(p)}{|\vec{\xi}^k(p)|} \quad \text{and} \quad T^k(p) \cdot \vec{\xi}^k = \vec{1},$$

which implies that $b(p)\xi^k_{k-2}(p) + a(p)\xi^k_{k-1}(p) + b(p)\xi^k_k(p) = 1$ and therefore

$$b(p)\lambda^k_{k-2}(p) + a(p)\lambda^k_{k-1}(p) = \frac{1 - b(p)\xi^k_k(p)}{|\vec{\xi}^k(p)|}. \tag{3.17}$$

By Lemma 3.1.3 together with the fact that $t_0 = 1$ it follows that

$$\xi^k_k(p) = \sum_{j=1}^{k} m^k_{k,j}(p) = \frac{(-b(p))^k}{t_k(p)} \sum_{j=1}^{k} (-1)^j b(p)^{-j} t_{j-1}(p).$$

We use this to prove that the recursion

$$\xi^{i+1}_{i+1}(p) \cdot t_{i+1}(p) = t_i(p) - b(p) \cdot \xi^i_i(p) \cdot t_i(p) \tag{3.18}$$

holds for all $i \in \mathrm{N}$, which is easily verified as follows:

$$\begin{aligned} t_i(p) - b(p)\xi^i_i(p)t_i(p) &= t_i(p) + (-b(p))^{i+1} \sum_{j=1}^{i} (-1)^j b(p)^{-j} t_{j-1}(p) \\ &= (-b(p))^{i+1} \sum_{j=1}^{i+1} (-1)^j b(p)^{-j} t_{j-1}(p) \\ &= \xi^{i+1}_{i+1}(p) t_{i+1}(p) \ . \end{aligned}$$

Now coming back to our aim (3.14), we use (3.17) to obtain

$$b(p)\lambda^k_{k-2}(p) + a(p)\lambda^k_{k-1}(p) = \frac{t_k(p) - b(p)\xi^k_k(p)t_k(p)}{|\vec{\xi}^k(p)|t_k(p)} = \frac{\xi^{k+1}_{k+1}(p)t_{k+1}(p)}{|\vec{\xi}^k(p)|t_k(p)}.$$

We claim that this term is always positive. First observe that by (3.16), $t_i(p) > 0$ for all $i \in \mathrm{N}$. Furthermore this implies that the matrix $T^k(p)$ is positive definite, hence $M^k(p) = (T^k(p))^{-1}$ is also positive definite, thus $|\vec{\xi}^k(p)| = \vec{1}^T \cdot M^k(p) \cdot \vec{1}$ is positive. Hence if the configuration is feasible we must have $\xi^1_1(p) > 0$, thus the recursion (3.18) and the fact that $b(p) < 0$ imply that $\xi^i_i(p) > 0$ for all $i \in [k+1]$ and therefore $\xi^{k+1}_{k+1}(p) \cdot t_{k+1}(p)$ must also be positive. 2

Now we move on to dismissing candidates $Q_{k,p}(x)$, which means dealing with those special cases, where $p = q^k_j$ for some integer $0 \le j \le k$. Recall that here

$$Q_{k,q^k_j}(x) = \Big(\ell^k_1(q^k_j, x), \ldots, \ell^k_k(q^k_j, x); q^k_j\Big),$$

as stated in Proposition 3.1.5. Note that the layer sizes $\lambda_i^k(p)$ are not defined for $p = q_j^k$, since here the matrix $T^k(p)$ is singular. Nevertheless we need $\lambda_i^k(q_j^k)$ when working with $\ell_i^k(q_j^k, x)$. This motivates the following definition:

$$\lambda_i^k(q_j^k) := \lim_{p \to q_j^k,\ p \geq q_j^k} \lambda_i^k(p).$$

The following proposition states some very useful facts about $\lambda_i^k(q_j^k)$ and $\ell_i^k(q_j^k, x)$.

Proposition 3.2.5 *Let $j \in [k]$ be an odd number and denote by $x^* := \lambda_k^k(q_j^k)$. Then*

$$\begin{aligned}
&(i) && \lambda_i^{k-1}(q_j^k) = \frac{\sin\phi_j^k - (-1)^i(\sin i\phi_j^k + \sin(i+1)\phi_j^k)}{(k+1)\sin\phi_j^k} \\
&(ii) && \lambda_i^k(q_j^k) = \frac{1 + (-1)^{i+1}\cos i\phi_j^k}{k+1} \\
&(iii) && \lambda_{k-1}^{k-1}(q_j^k) = 2\lambda_k^k(q_j^k) \\
&(iv) && \lambda_{i-1}^{k-1}(q_j^k) + \lambda_i^{k-1}(q_j^k) = 2\lambda_i^k(q_j^k) \\
&(v) && \ell_i^k(q_j^k, x^*) = \lambda_i^k(q_j^k) \\
&(vi) && \ell_i^k(q_j^k, x^* + x) = \ell_{k-i+1}^k(q_j^k, x^* - x).
\end{aligned}$$

This proposition has several important consequences. First, (v) states that at the point $x = x^*$ the configurations $Q_{k,p}(x)$ and $Q_{k,p}$ 'meet'. Moreover, it is equivalent to

$$\lambda_i^k(q_j^k) - x^* \cdot \mu_i^{k-1}(q_j^k) = \lambda_i^{k-1}(q_j^k),$$

which implies that

$$\ell_i^k(q_j^k, x) = \lambda_i^{k-1}(q_j^k) + x \cdot \mu_i^{k-1}(q_j^k) = \lambda_i^k(q_j^k) + (x - x^*) \cdot \mu_i^{k-1}(q_j^k). \tag{3.19}$$

This will turn out to be useful, since it will allow us to use the symmetry of $\vec{\lambda}$ to partly cover for the lack of symmetry in $\vec{\ell}$.

Also, by (vi), we know that

$$\ell_1^k(q_j^k, 2x^* + x) = \ell_k^k(q_j^k, -x) = -x,$$

hence the configuration $Q_{k,q_j^k}(x)$ is only feasible for $x \in [0, 2x^*]$, and therefore, because of the symmetry expressed in (vi), we only need to consider $x \in (0, x^*)$. Finally (iii) and (iv) prove statement (vi) of our main theorem 1.3.1.

Because of the different structure of $Q_{k,q_j^k}(x)$ for j odd and even, the dismissal of $Q_{k,q_j^k}(x)$ will be done separately.

Proposition 3.2.6 *For any even number j with $0 \le j \le k$ and any $x \in (0,1)$, the configuration $Q_{k,q_j^k}(x) = (\ell_1^k(q_j^k,x),\dots,\ell_k^k(q_j^k,x);q_j^k)$ is not significant.*

Proof. Let $p := q_j^k$ and $\phi = \phi(p)$. Since we can assume that $p \ge \frac{1}{2}$, we also know that $\phi \ge \pi/2$. Recall from Lemma 3.1.6 that for $i \in [k-1]$,

$$\ell_i^k(p,x) = \lambda_i^{k-1}(p) + x \cdot \mu_i^{k-1}(p) = (1 - \frac{x}{1+\cos\phi})\lambda_i^{k-1}(\phi) - x \cdot b(p) \cdot m_{i,k-1}^{k-1}(p)$$

and $\ell_k^k(p,x) = x$. We consider two different cases, depending on whether

$$(1 - \frac{x}{1+\cos\phi})\frac{1}{|\vec{\xi}^{k-1}(p)|} \tag{3.20}$$

is positive or not.

Suppose that it is positive. Then we would like to apply Corollary 3.2.2 and therefore need to show that

$$b(p)\ell_{k-2}^k(p) + a(p)\ell_{k-1}^k(p) > 0. \tag{3.21}$$

Similarly to the proof of Proposition 3.2.4 we deduce from $T^{k-1}(p)\cdot\vec{\xi}^{k-1} = \vec{1}$ and $M^{k-1}(p)\cdot T^{k-1}(p) = id$ that

$$b(p)\xi_{k-2}^{k-1}(p) + a(p)\xi_{k-1}^{k-1}(p) = 1, \qquad b(p)m_{k-1,k-2}^{k-1}(p) + a(p)m_{k-1,k-1}^{k-1}(p) = 1.$$

Hence condition (3.21) is equivalent to

$$(1 - \frac{x}{1+\cos\phi})\frac{1}{|\vec{\xi}^{k-1}(p)|} - x \cdot b(p) > 0,$$

which is true since (3.20) is positive and $b(p)$ is negative.

Now suppose that (3.20) is negative. In this case we would like to apply Lemma 3.2.1 with $r := k$, $r' := k+1$, $\lambda_i := \ell_i^k(p,x)$ and

$$\nu_k := -1,\ \nu_{k+1} := 1, \quad \text{and} \quad \nu_i := 0 \ \text{ for } i = 1,\dots,k-1.$$

Then

$$\alpha_1 = -2\ell_k^k(p,x) \ \text{ and } \ \alpha_2 = -\ell_{k-1}^k(p,x) + \ell_k^k(p,x),$$

hence the condition $-b(p)\alpha_2 - \frac{a(p)}{2}\alpha_1 > 0$ now reduces to

$$\begin{aligned} 0 &\overset{!}{<} b(p)\cdot\ell_{k-1}^k(p,x) - b(p)\cdot\ell_k^k(p,x) + a(p)\cdot\ell_k^k(p,x) \\ &= b(p)\cdot(1 - \frac{x}{1+\cos\phi})\cdot\lambda_{k-1}^{k-1}(\phi) - b(p)^2 x\cdot m_{k-1,k-1}^{k-1}(p) - b(p)\cdot x + a(p)\cdot x \\ &= b(p)\cdot\xi_{k-1}^{k-1}(p)\cdot(1 - \frac{x}{1+\cos\phi})\frac{1}{|\vec{\xi}^{k-1}(p)|} - x\cdot b(p), \end{aligned}$$

where we used the fact (see Lemma 3.1.6) that

$$m_{k-1,k-1}^{k-1}(p) = (-1)^{k-1}\frac{\sin(k-1)\phi}{b(p)\sin\phi} = \frac{2\cos\phi}{b(p)} = \frac{a(p)}{b(p)^2}.$$

Since (3.20) is negative it therefore suffices to show that $\xi_{k-1}^{k-1}(p)$ is non–negative. Thanks to Lemma 3.1.6 we know that

$$\xi_{k-1}^{k-1}(p) = \frac{1}{b(p)} - \frac{1}{b(p)(1+\cos\phi)} = \frac{\cos\phi}{b(p)(1+\cos\phi)},$$

which is non–negative because $\pi \geq \phi \geq \pi/2$. 2

Proposition 3.2.7 *For any odd number j with $1 < j \leq k$ and any $x \in (0,1)$, the configuration $Q_{k,q_j^k}(x) = (\ell_1^k(q_j^k,x),\ldots,\ell_k^k(q_j^k,x);q_j^k)$ is not significant.*

Proof. We take the same approach as in the proof of Proposition 3.2.3: since $j > 1$ we have that $q_j^k < q^k$, but unfortunately the configuration $(\ell_1^k(q_j^k,x),\ldots,\ell_k^k(q_j^k,x);q_j^k)$ is not symmetric. However, recall from (3.19) that it can be written as

$$\ell_i^k(q_j^k,x) = \lambda_i^k(q_j^k) + (x-x^*)\cdot\mu_i^{k-1}(q_j^k)$$

where the $\lambda_i^k(q_j^k)$ are symmetric. Now we apply Lemma 3.2.1 with

$$r := r' := k, \qquad \nu_i := \mu_i^{k-1}(q^k) \qquad \lambda_i := \ell_i^k(q_j^k,x)$$

and claim that (3.13) holds.

Again we use the properties of ν_i as guaranteed by Lemma 3.1.7 to see that indeed $\alpha_1 = \alpha_2 = 0$:

$$\alpha_1 = 2\sum_{i=1}^{k}(\lambda_i^k(q_j^k) + (x-x^*)\cdot\mu_i^{k-1}(q_j^k))\nu_i = 0$$

because

$$\lambda_i^k(q_j^k) = \lambda_{k-i+1}^k(q_j^k), \quad \nu_i = -\nu_{k-i+1}, \quad \sum_{i=1}^{k}\mu_i^{k-1}(q_j^k)\cdot\nu_i = 0.$$

Similarly

$$\alpha_2 = \sum_{i=1}^{k-1}\Big(\nu_i[\lambda_{i+1}^k(q_j^k) + (x-x^*)\cdot\mu_{i+1}^{k-1}(q_j^k)] + \nu_{i+1}[\lambda_i^k(q_j^k) + (x-x^*)\cdot\mu_i^{k-1}(q_j^k)]\Big) = 0,$$

since $\nu_{i+1}\lambda_i^k(q_j^k) = \nu_{i+1}\lambda_{k-i+1}^k(q_j^k) = -\nu_{k-(i+1)+1}\lambda_{k-i+1}^k(q_j^k) = -\nu_{k-i}\lambda_{k-i+1}^k(q_j^k)$ and

$$\sum_{i=1}^{k} \nu_i \cdot \mu_{i+1}^{k-1}(q_j^k) = \sum_{i=1}^{k} \nu_{i+1} \cdot \mu_i^{k-1}(q_j^k) = 0$$

according to the last statement in Lemma 3.1.7.

The rest of the proof is identical to that of Proposition 3.2.3, since the ν_i are chosen in exactly the same way and because $q_j^k < q^k$. 2

Thus we have dismissed all candidates $Q_{k,q_j^k}(x)$ other than $Q_{k,q^k}(x)$ in the range $x \in [0, x^*]$.

3.3 Linearity and monotonicity

In this section we complete the proof of Theorems 1.3.1 and 1.3.2. We start with the statement (vii) of Theorem 1.3.1, which we prove here in a slightly more general form.

Proposition 3.3.1 *For an odd number $j \in [k]$ consider a configuration $(\lambda_1, \ldots, \lambda_r; q_j^k)$ with $r \leq k$. Set $\lambda_{r+1} = \cdots = \lambda_k = 0$ and $\mu_i := \mu_i^{k-1}(q_j^k)$ for $i = 1, \ldots, k$. There exists a constant u such that for all x*

$$\mathrm{c}(\lambda_1 + x\mu_1, \ldots, \lambda_k + x\mu_k; q_j^k) \;=\; -a(q_j^k) \;\cdot\; \mathrm{d}(\lambda_1 + x\mu_1, \ldots, \lambda_k + x\mu_k; q_j^k) \;+ u\ .$$

Proof. Let $p := q_j^k$. We need to show that

$$H(p)\sum_{i=1}^{k-1}(\lambda_i + x\mu_i)(\lambda_{i+1} + x\mu_{i+1}) =$$
$$-a(p)\cdot\Big(\frac{1}{2} - \frac{1}{2}\sum_{i=1}^{k}(\lambda_i + x\mu_i)^2 - (1-p)\sum_{i=1}^{k-1}(\lambda_i + x\mu_i)(\lambda_{i+1} + x\mu_{i+1})\Big) \;+ u.$$

Both sides are polynomials in x of degree 2. We compare their coefficients:

$$x^0: \quad H(p)\sum_{i=1}^{k-1}\lambda_i\lambda_{i+1} = -a(p)\Big(\frac{1}{2} - \frac{1}{2}\sum_{i=1}^{k}\lambda_i^2 - (1-p)\sum_{i=1}^{k-1}\lambda_i\lambda_{i+1}\Big) + u \tag{3.22}$$

$$x^1: \quad H(p)\sum_{i=1}^{k-1}(\lambda_i\mu_{i+1} + \lambda_{i+1}\mu_i) =$$

$$-a(p)\Big(-\frac{1}{2}\sum_{i=1}^{k}2\lambda_i\mu_i - (1-p)\sum_{i=1}^{k-1}(\lambda_i\mu_{i+1} + \lambda_{i+1}\mu_i)\Big) \tag{3.23}$$

$$x^2: \quad H(p)\sum_{i=1}^{k-1}\mu_i\mu_{i+1} = -a(p)\Big(-\frac{1}{2}\sum_{i=1}^{k}\mu_i^2 - (1-p)\sum_{i=1}^{k-1}\mu_i\mu_{i+1}\Big) \tag{3.24}$$

Obviously u can be chosen so that (3.22) holds, so we only need to check (3.23) and (3.24). Comparing the coefficients of λ_i in (3.23) results in

$$H(p)(\mu_{i+1} + \mu_{i-1}) = -a(p)(-\mu_i - (1-p)(\mu_{i+1} + \mu_{i-1})).$$

Using the facts $H(p) = (1-p)a(p) - b(p)$ and $\frac{a(p)}{2b(p)} = \cos\phi$ this is equivalent to

$$\mu_{i+1} + \mu_{i-1} = -2\cos\phi \cdot \mu_i,$$

which is true because by Lemma 3.1.6

$$\begin{aligned}\mu_{i-1} + \mu_{i+1} &= (-1)^{i-1}\cdot\frac{\sin(i-1)\phi + \sin(i+1)\phi}{\sin\phi} = (-1)^{i-1}\cdot\frac{2\sin i\phi\cdot\cos\phi}{\sin\phi}\\ &= -2\cos\phi\cdot\mu_i.\end{aligned}$$

Applying the same facts as above, (3.24) is equivalent to

$$-\sum_{i=1}^{k-1}\mu_i\cdot\mu_{i+1} = \cos\phi\sum_{i=1}^{k}\mu_i^2,$$

which can be seen to be true using Lemma 3.1.7. 2

To complete the proof of the main result there is now only one piece missing: we still need to prove that as x increases $\mathrm{d}(Q_{k,q^k}(x))$ increases, and that as p increases $\mathrm{d}(Q_{k,p})$ increases. This will be done in the following two propositions. The proof of the latter is very technical and hence postponed to Section 3.4. For the proof of the first, recall that in the case of $q_j^k = q^k$ we have $x^* = \lambda_k^k(q^k)$.

Proposition 3.3.2 *As x increases from 0 to x^*, the function*

$$\mathrm{d}(Q_{k,q^k}(x)) = \mathrm{d}(\ell_1^k(q^k,x),\ldots,\ell_k^k(q^k,x),q^k)$$

is strictly increasing.

Proof. Let us abbreviate $\ell_i(x) := \ell_i^k(q^k,x)$, $\lambda_i := \lambda_i^k(q^k)$ and $\mu_i := \mu_i^{k-1}(q^k)$. Then recall that by (3.19) we can write

$$\ell_i(x) = \lambda_i - x^* \cdot \mu_i + x \cdot \mu_i,$$

(where we will make use of the fact that the λ_i are symmetric) and that the function under consideration is given by

$$\mathrm{d}(\ell_1^k(q^k,x),\ldots,\ell_k^k(q^k,x),q^k) = \frac{1}{2} - \frac{1}{2}\sum_{i=1}^{k}\ell_i(x)^2 - (1-q^k)\sum_{i=1}^{k-1}\ell_i(x)\ell_{i+1}(x). \tag{3.25}$$

We claim that the derivatives of the two sums involved are negative. Since $\lambda_i = \lambda_{k-i+1}$ and $\mu_i = -\mu_{k-i+1}$ we can use that

$$\sum_{i=1}^{k}\lambda_i\mu_i = 0 = \sum_{i=1}^{k-1}(\lambda_i\mu_{i+1} + \lambda_{i+1}\mu_i).$$

Thus the derivative of the first sum is given by

$$\sum_{i=1}^{k}\left(2[\lambda_i - x^*\mu_i]\mu_i + 2x\cdot\mu_i^2\right) = 2(x-x^*)\sum_{i=1}^{k}\mu_i^2,$$

which is easily seen to be negative, because $x < x^*$.

Similarly, the derivative of the second sum in (3.25) is

$$\sum_{i=1}^{k-1}\left([\lambda_i - x^*\mu_i]\mu_{i+1} + [\lambda_{i+1} - x^*\mu_{i+1}]\mu_i + 2x\cdot\mu_i\mu_{i+1}\right)$$

$$= 2(x-x^*)\sum_{i=1}^{k-1}\mu_i\mu_{i+1},$$

which is negative by Lemma 3.1.7. 2

Proposition 3.3.3 *As p increases from q^k to q^{k+1}, the function*

$$\mathrm{d}(Q_{k,p}) = \mathrm{d}(\lambda_1^k(p), \ldots, \lambda_k^k(p), p)$$

is strictly increasing, while

$$\mathrm{c}(Q_{k,p}) = \mathrm{c}(\lambda_1^k(p), \ldots, \lambda_k^k(p), p)$$

is strictly decreasing.

We close this section with the proof of the main theorems. The only thing we have to do is to assemble all the information gathered so far. While doing so, we will refer to previous theorems, lemmas or propositions by simply mentioning their reference numbers in brackets.

Proof of Theorems 1.3.1 and 1.3.2. We know that $\mathrm{c}(d) = \mathrm{c}(Q)$ if Q is a d-significant configuration [2.1.5]. We have determined a family $\mathcal{C}$ of candidates $Q_{k,p}$ and $Q_{k,p}(x)$, of which we know [2.2.1, 2.2.10, 2.2.11] that for every d there exists a d-significant configuration in $\mathcal{C}$. We have defined constants q^k and x^* (depending on k) and have proved [Section 3.2] that $Q_{k,p}$ can only be significant if $p \in [q^k, q^{k+1}]$, and that $Q_{k,p}(x)$ can only be significant if $p = q^k$ and $x \in [0, x^*]$. Let us denote the subfamily formed by these remaining configurations by $\mathcal{C}'$. We have seen that the configurations 'meet' at the beginning (by definition of $Q_{k,p}(x)$) and end (by [3.2.5]) of their intervals:

$$Q_{k-1,q^k} = Q_{k,q^k}(0) \quad \text{and} \quad Q_{k,q^k}(x^*) = Q_{k,q^k}.$$

We have shown [3.3.2,3.3.3] that $\mathrm{d}(Q_{k,p})$ is strictly increasing as a function in p, and that $\mathrm{d}(Q_{k,q^k}(x))$ is strictly increasing as a function in x. We may therefore deduce that the infinite sequence of constants defined by

$$d'_k := \mathrm{d}(Q_{k-1,q^k}) \quad \text{and} \quad d_k := \mathrm{d}(Q_{k,q^k})$$

does indeed satisfy

$$d_2 < d'_3 < d_3 < d'_4 < d_4 < \cdots < \frac{1}{2}$$

which in turn implies that for every d there exists only one configuration Q^* in $\mathcal{C}'$ with $\mathrm{d}(Q^*) = d$. We can therefore conclude that it must be d-significant. Now it only remains to check that in the above intervals the configurations and their parameters [3.1.4, 3.1.5, 3.3.1] are as stated in the theorems. This completes the proof of our main result as stated in Theorems 1.3.1 and 1.3.2. 2

3.4 Technical computations

This section contains the proofs of those lemmas and propositions that we stated in the previous sections of this chapter but didn't prove there. All of them are fairly technical and hence probably not very easy to read. We will very often use trigonometric identities (without refering to them explicitly), such as $\sin(x \pm y) = \sin x \cos y \pm \cos x \sin y$ or $\sin x \sin y = \frac{1}{2}(\cos(x-y) - \cos(x+y))$ for example.

Proof of Lemma 3.1.1. By expanding the determinant of T^k by the first row we obtain the recursion

$$t_k = at_{k-1} - b^2 t_{k-2}. \tag{3.26}$$

In both cases the claimed identity for t_k is then easily verified by induction on k. (In the second case use $\cos 2x = 2\cos^2 x - 1$ to verify the case $\ell = 2$ and $2\sin\frac{x+y}{2}\cos\frac{x-y}{2} = \sin x + \sin y$ for the induction step.) 2

Proof of Lemma 3.1.3. Let us first deal with the case $2 \le i, j \le k-1$. Here we need to show that

$$bm_{i,j-1} + am_{i,j} + bm_{i,j+1} = \delta_{i,j}.$$

If indeed $i = j$ then applying the definition of the $m^k_{i,j}$ and using the recursion (3.26) from the preceding proof

$$\begin{aligned} bm_{i,i-1} + am_{i,i} + bm_{i,i+1} &= -b^2\frac{t_{i-2}t_{k-i}}{t_k} + a\frac{t_{i-1}t_{k-i}}{t_k} - b^2\frac{t_{i-1}t_{k-i-1}}{t_k} \\ &= \frac{1}{t_k}\Big(-b^2 t_{i-2}t_{k-i} + t_{i-1}t_{k-i+1}\Big). \end{aligned}$$

It therefore suffices to show that for all $i = 2, \dots, k$

$$t_k = -b^2 t_{i-2}t_{k-i} + t_{i-1}t_{k-i+1},$$

which we prove by induction on i. Since $t_0 = 1$ and $t_1 = a$, the case $i = 2$ is identical to (3.26). For the induction step we need to show that that

$$-b^2 t_{i-2}t_{k-i} + t_{i-1}t_{k-i+1} = t_k \overset{!}{=} -b^2 t_{i-1}t_{k-i-1} + t_i t_{k-i},$$

which is equivalent to

$$t_{i-1} \cdot (t_{k-i+1} + b^2 t_{k-i-1}) = t_{k-i}(t_i + b^2 t_{i-2}).$$

Once again using (3.26), both sides of the last equality can be seen to be equal to $a \cdot t_{i-1} \cdot t_{k-i}$.

Now suppose that $i < j$, the case $i > j$ can be dealt with analogously. Then

$$bm_{i,j-1} + am_{i,j} + bm_{i,j+1} \quad = \quad (-1)^{i+j} b^{j-i} \frac{t_{i-1}}{t_k} \Big(- t_{k+1-j} + at_{k-j} - b^2 t_{k-j-1} \Big) \ = 0,$$

as required.

To deal with those cases where $\{i, j\} \cap \{1, k\} \neq \emptyset$, simply define $t_{-1} := 0$, and observe that this agrees with the recursion (3.26). Now we can extend the definition $m^k_{i,j}$ to cover those cases where $\{i, j\} \cap \{0, k+1\} \neq \emptyset$ – here $m^k_{i,j} = 0$ and the same proof as above applies. 2

In the following we prove a number of trigonometric identities that will later turn out to be useful.

Lemma 3.4.1

$$(i) \quad \sum_{j=1}^{k} (-1)^j \sin j\theta = \frac{-\sin\theta + (-1)^k \sin k\theta + (-1)^k \sin(k+1)\theta}{2 + 2\cos\theta},$$

$$(ii) \quad \sum_{j=1}^{k} \cos(j \cdot \theta + \vartheta) = \frac{-\cos\vartheta + \cos(\theta + \vartheta) + \cos(k \cdot \theta + \vartheta) - \cos((k+1)\theta + \vartheta)}{2 - 2\cos\theta}.$$

Proof. We shall use the fact that

$$\sin x = \frac{1}{2i}(e^{ix} - e^{-ix}) \qquad \text{and} \qquad \cos x = \frac{1}{2}(e^{ix} + e^{-ix}).$$

$$\begin{aligned}
\sum_{j=1}^{k} (-1)^j \sin j\theta &= \frac{1}{2i} \sum_{j=1}^{k} (-1)^j (e^{ij\theta} - e^{-ij\theta}) \\
&= -\frac{1}{2i} [\frac{e^{i\theta}(1 - (-1)^k e^{ik\theta})}{1 + e^{i\theta}} - \frac{e^{-i\theta}(1 - (-1)^k e^{-ik\theta})}{1 + e^{-i\theta}}] \\
&= -\frac{1}{2i} \frac{e^{i\theta} + 1 - (-1)^k e^{i(k+1)\theta} - (-1)^k e^{ik\theta} - e^{-i\theta} - 1 + (-1)^k e^{-i(k+1)\theta} + (-1)^k e^{-ik\theta}}{(1 + e^{i\theta})(1 + e^{-i\theta})} \\
&= \frac{-\sin\theta + (-1)^k \sin k\theta + (-1)^k \sin(k+1)\theta}{2 + 2\cos\theta}.
\end{aligned}$$

$$\begin{aligned}
\sum_{j=1}^{k}\cos(j\theta+\vartheta) &= \frac{1}{2}\sum_{j=1}^{k}(e^{i(j\theta+\vartheta)}+e^{-i(j\theta+\vartheta)}) \\
&= \frac{1}{2}[\frac{e^{i(\theta+\vartheta)}(1-e^{ik\theta})}{1-e^{i\theta}}+\frac{e^{-i(\theta+\vartheta)}(1-e^{-ik\theta})}{1-e^{-i\theta}}] \\
&= \frac{1}{2}\frac{e^{i(\theta+\vartheta)}[1-e^{-i\theta}-e^{ik\theta}+e^{i(k-1)\theta}]}{(1-e^{i\theta})(1-e^{-i\theta})} \\
&\quad +\frac{1}{2}\frac{e^{-i(\theta+\vartheta)}[1-e^{i\theta}-e^{-ik\theta}+e^{-i(k-1)\theta}]}{(1-e^{i\theta})(1-e^{-i\theta})} \\
&= \frac{\cos(\theta+\vartheta)-\cos\vartheta-\cos[(k+1)\theta+\vartheta]+\cos[k\theta+\vartheta]}{2-2\cos\theta}.
\end{aligned}$$

2

Lemma 3.4.2

$$(i)\quad \sum_{j=1}^{k}(\sin j\theta)^2=\frac{1}{2}k+\frac{1}{4}-\frac{\sin(2k+1)\theta}{4\sin\theta},$$

$$(ii)\quad \sum_{j=1}^{k}\sin j\theta\cdot\sin(j+1)\theta=\frac{1}{2}(k+1)\cos\theta-\frac{\sin(2k+2)\theta}{4\sin\theta}.$$

Proof.

$$\begin{aligned}
\sum_{j=1}^{k}(\sin j\theta)^2 &= -\frac{1}{4}\sum_{j=1}^{k}(e^{2ij\theta}+e^{-2ij\theta}-2) \\
&= -\frac{1}{4}[\frac{e^{2i\theta}-e^{2i(k+1)\theta}}{1-e^{2i\theta}}+\frac{e^{-2i\theta}-e^{-2i(k+1)\theta}}{1-e^{-2i\theta}}]+\frac{1}{2}k \\
&= -\frac{e^{2i\theta}-1-e^{2i(k+1)\theta}+e^{2ik\theta}+e^{-2i\theta}-1-e^{-2i(k+1)\theta}+e^{-2ik\theta}}{4(1-e^{2i\theta})(1-e^{-2i\theta})}+\frac{1}{2}k \\
&= -\frac{2\cos 2\theta-2-2\cos 2(k+1)\theta+2\cos 2k\theta}{4(2-2\cos 2\theta)}+\frac{1}{2}k \\
&= -\frac{\cos 2\theta-1-\cos 2(k+1)\theta+\cos 2k\theta}{4-4\cos 2\theta}+\frac{1}{2}k \\
&= \frac{1}{4}+\frac{\cos 2(k+1)\theta-\cos 2k\theta}{8(\sin\theta)^2}+\frac{1}{2}k \\
&= \frac{1}{4}+\frac{-2\sin(2k+1)\theta\sin\theta}{8(\sin\theta)^2}+\frac{1}{2}k.
\end{aligned}$$

$$
\begin{aligned}
\sum_{j=1}^{k} \sin j\theta \cdot \sin(j+1)\theta &= -\frac{1}{4}\sum_{j=1}^{k}(e^{i(2j+1)\theta} - e^{i\theta} - e^{-i\theta} + e^{-i(2j+1)\theta}) \\
&= \frac{1}{2}k\cos\theta - \frac{1}{4}[\frac{e^{3i\theta} - e^{i(2k+3)\theta}}{1-e^{2i\theta}} + \frac{e^{-3i\theta} - e^{-i(2k+3)\theta}}{1-e^{-2i\theta}}] \\
&= \frac{1}{2}k\cos\theta - \frac{e^{3i\theta} - e^{i\theta} - e^{i(2k+3)\theta} + e^{i(2k+1)\theta}}{4(1-e^{2i\theta})(1-e^{-2i\theta})} \\
&\quad - \frac{e^{-3i\theta} - e^{-i\theta} - e^{-i(2k+3)\theta} + e^{-i(2k+1)\theta}}{4(1-e^{2i\theta})(1-e^{-2i\theta})} \\
&= \frac{1}{2}k\cos\theta - \frac{\cos 3\theta - \cos\theta - \cos(2k+3)\theta + \cos(2k+1)\theta}{4(1-\cos 2\theta)} \\
&= \frac{1}{2}k\cos\theta - \frac{-2\sin 2\theta\sin\theta + 2\sin(2k+2)\theta\sin\theta}{8(\sin\theta)^2} \\
&= \frac{1}{2}k\cos\theta + \frac{2\sin\theta\cos\theta - \sin(2k+2)\theta}{4\sin\theta}.
\end{aligned}
$$

2

Proof of (3.7) and (3.8) in the proof of Proposition 3.1.4.

$$
\begin{aligned}
\xi_i^k(p) = \sum_{j=1}^{k} m^k(p)_{ij} &= \sum_{j=1}^{i} m^k(p)_{ij} + \sum_{j=i+1}^{k} m^k(p)_{ij} \\
&= \frac{(-1)^i}{b(p)\sin\phi\sin(k+1)\phi} \cdot [\sin(k-i+1)\phi \sum_{j=1}^{i}(-1)^j \sin j\phi \\
&\qquad + \sin i\phi \underbrace{\sum_{j=i+1}^{k}(-1)^j \sin(k-j+1)\phi}_{=(-1)^{k+1}\sum_{j=1}^{k-i}(-1)^j \sin j\phi}] \\
&= \frac{(-1)^i}{b(p)\sin\phi\sin(k+1)\phi} \cdot [\sin(k-i+1)\phi \frac{-\sin\phi + (-1)^i \sin i\phi + (-1)^i \sin(i+1)\phi}{2+2\cos\phi} \\
&\qquad -(-1)^k \sin i\phi \frac{-\sin\phi + (-1)^{k-i}\sin(k-i)\phi + (-1)^{k-i}\sin(k-i+1)\phi}{2+2\cos\phi}] \\
&= \frac{-(-1)^i \sin(k-i+1)\phi\sin\phi + \sin(k-i+1)\phi\sin(i+1)\phi}{b(p)\sin\phi\sin(k+1)\phi\cdot(2+2\cos\phi)} \\
&\quad + \frac{(-1)^{k-i}\sin\phi\sin i\phi - \sin i\phi\sin(k-i)\phi}{b(p)\sin\phi\sin(k+1)\phi\cdot(2+2\cos\phi)}.
\end{aligned}
$$

Finally, observe that

$$\begin{aligned}
&\sin(k-i+1)\phi\sin(i+1)\phi - \sin i\phi\sin(k-i)\phi \\
&\quad = \sin(k-i+1)\phi\cdot[\sin i\phi\cos\phi + \cos i\phi\sin\phi] \\
&\qquad - \sin i\phi\cdot[\sin(k-i+1)\phi\cos\phi - \cos(k-i+1)\phi\sin\phi] \\
&\quad = \sin\phi\cdot[\sin(k-i+1)\phi\cos i\phi + \cos(k-i+1)\phi\sin i\phi] \\
&\quad = \sin\phi\sin(k+1)\phi,
\end{aligned}$$

which proves (3.7). As to the proof of (3.8),

$$|\vec{\xi}(p)| = \sum_{i=1}^{k}\xi_i^k(p) = \frac{\sum_{i=1}^{k}(-1)^{k-i}\sin i\phi - (-1)^i\sin(k-i+1)\phi + \sin(k+1)\phi}{2b(p)(1+\cos\phi)\sin(k+1)\phi}.$$

Consider the numerator:

$$\begin{aligned}
< N > &= (-1)^k\sum_{i=1}^{k}(-1)^i\sin i\phi - \underbrace{\sum_{i=1}^{k}(-1)^i\sin(k-i+1)\phi}_{=(-1)^{k+1}\mathrm{P}_{i=1}^{k}(-1)^i\sin i\phi} + k\sin(k+1)\phi \\
&= 2(-1)^k\sum_{i=1}^{k}(-1)^i\sin i\phi + k\sin(k+1)\phi \\
&= 2(-1)^k\frac{-\sin\phi + (-1)^k\sin k\phi + (-1)^k\sin(k+1)\phi}{2+2\cos\phi} + k\sin(k+1)\phi \\
&= \frac{-(-1)^k\sin\phi + \sin k\phi + \sin(k+1)\phi}{1+\cos\phi} + k\sin(k+1)\phi,
\end{aligned}$$

which proves (3.8). 2

Proof of Lemma 3.1.6. We have to compute the value of the functions $m_{i,k-1}^{k-1}$, ξ_{k-1}^{k-1}, μ_i^{k-1} and ℓ_i^k at the point q_j^k. To improve readability, we omit the term q_j^k in our calculations and let $\phi := \phi_j^k$. Then

$$\begin{aligned}
m_{i,k-1}^{k-1} &= (-1)^{i+k-1}b^{k-1-i}\frac{t_{i-1}t_0}{t_{k-1}} = (-1)^{i+k-1}b^{-1}\frac{\sin i\phi}{\sin k\phi} \\
&\overset{(3.11)}{=} (-1)^{i+k-1}b^{-1}\frac{\sin i\phi}{-(-1)^{k-j}\sin\phi} \\
&= (-1)^{i+j}\frac{\sin i\phi}{b\sin\phi}.
\end{aligned}$$

To compute ξ_{k-1}^{k-1}, we use the formula in Proposition 3.1.4 to see that

$$\begin{aligned}\xi_{k-1}^{k-1} &= \frac{-(-1)^{k-j}\sin 2\phi - (-1)^{k-1}\sin\phi - (-1)^{k-j}\sin\phi}{-(-1)^{k-j}2b(1+\cos\phi)\sin\phi} \\ &= \begin{cases} \frac{2\cos\phi\sin\phi+2\sin\phi}{2b(\cos\phi\sin\phi+\sin\phi)} = \frac{1}{b} & \text{, if } j \text{ odd,} \\ \frac{2\cos\phi\sin\phi}{2b(1+\cos\phi)\sin\phi} = \frac{1}{b} - \frac{1}{b(1+\cos\phi)} & \text{, if } j \text{ even.} \end{cases}\end{aligned}$$

Finally, first use (3.11) to see that (for j odd) the formula for μ_k^{k-1} is indeed equal to one and then observe that the above results show that the expressions for μ_i^{k-1} and ℓ_i^k follow directly from their definitions in Proposition 3.1.5. 2

Proof of Lemma 3.1.7. We will use the explicit formula

$$\mu_i^{k-1}(q_j^k) = (-1)^i \frac{\sin i\phi_j^k}{\sin\phi_j^k}$$

from Lemma 3.1.6 together with the observation (3.11) and the abbreviations $\phi := \phi_j^k$, $\mu_i := \mu_i^{k-1}(q_j^k)$, $\phi' := \phi_{j'}^k$, $\mu_i' := \mu_i^{k-1}(q_{j'}^k)$. First

$$\mu_{k+1-i} = (-1)^{k+1-i}\frac{\sin(k-(i-1))\phi}{\sin\phi} = -(-1)^{k+1-i+k-j}\frac{\sin i\phi}{\sin\phi} = -\mu_i,$$

which proves (i). This immediately implies (ii). To prove (iii)-(vi), we apply Lemma 3.4.2. Hence

$$\begin{aligned}\sum_{i=1}^k \mu_i^2 &= \frac{1}{\sin^2\phi}\sum_{i=1}^k(\sin i\phi)^2 = \frac{1}{\sin^2\phi}(\frac{k}{2}+\frac{1}{4}-\frac{\sin(2k+1)\phi}{4\sin\phi}) \\ &\overset{(3.11)}{=} \frac{1}{\sin^2\phi}(\frac{k}{2}+\frac{1}{4}+\frac{1}{4}) = \frac{k+1}{2\sin^2\phi}, \quad \text{and} \\ \sum_{i=1}^{k-1}\mu_i\mu_{i+1} &= \frac{-1}{\sin^2\phi}\sum_{i=1}^k \sin i\phi\sin(i+1)\phi = \frac{-1}{\sin^2\phi}(\frac{k+1}{2}\cos\phi - \frac{\sin(2k+2)\phi}{4\sin\phi}) \\ &\overset{(3.11)}{=} -\frac{(k+1)\cos\phi}{2\sin^2\phi},\end{aligned}$$

which proves (iii) and (iv). To prove (v) and (vi), observe that

$$\begin{aligned}\sum_{i=1}^k \mu_i'\mu_i &= \frac{1}{\sin\phi'\sin\phi}\sum_{i=1}^k \sin i\phi'\sin i\phi \\ &= \frac{1}{\sin\phi'\sin\phi}\left(\sum_{i=1}^k\cos i(\phi'-\phi) - \sum_{i=1}^k\cos i(\phi'+\phi)\right).\end{aligned}$$

We claim that the last two sums are identical. By Lemma 3.4.1

$$\sum_{i=1}^{k} \cos i(\phi' \pm \phi) = \frac{-1 + \cos(\phi' \pm \phi) + \cos k(\phi' \pm \phi) - \cos(k+1)(\phi' \pm \phi)}{2 - 2\cos(\phi' \pm \phi)}.$$

Notice that

$$\phi' - \phi = \frac{j - j'}{k+1}\pi, \qquad \phi' + \phi = \frac{2k - j' - j}{k+1}\pi,$$

As j and j' are odd, we know that $(k+1)(\phi' \pm \phi)$ is an even multiple of π and hence $\cos(k+1)(\phi' \pm \phi) = 1$ and $\cos k(\phi' \pm \phi) = \cos((k+1)(\phi' \pm \phi) - (\phi' \pm \phi)) = \cos(\phi' \pm \phi)$. Therefore

$$\sum_{i=1}^{k} \cos i(\phi' \pm \phi) = \frac{-2 + 2\cos(\phi' \pm \phi)}{2 - 2\cos(\phi' \pm \phi)} = -1,$$

which completes the proof of (v). Similarly for the proof of (vi)

$$\sum_{i=1}^{k-1} \mu'_{i+1}\mu_i = \sum_{i=1}^{k} \mu'_{i+1}\mu_i = \frac{-1}{\sin\phi' \sin\phi}\left(\sum_{i=1}^{k} \cos[i(\phi' - \phi) + \phi'] - \sum_{i=1}^{k} \cos[i(\phi' + \phi) + \phi']\right),$$

$$\begin{aligned}\sum_{i=1}^{k} \cos[i(\phi' \pm \phi) + \phi'] &= \frac{-\cos\phi' + \cos((\phi' \pm \phi) + \phi') + \cos(-(\phi' \pm \phi) + \phi') - \cos\phi'}{2 - 2\cos(\phi' \pm \phi)} \\ &= \frac{-2\cos\phi' + 2\cos(\phi' \pm \phi)\cos\phi'}{2 - 2\cos(\phi' \pm \phi)} = -\cos\phi'.\end{aligned}$$

2

Proof of (3.15) in the proof of Proposition 3.2.4. Recall that

$$(1 - \frac{2}{k+2})\pi < \phi < \pi$$

and hence

$$(r-1)\cdot\pi < r\cdot\phi < r\cdot\pi \qquad \text{for } r \in (0, \frac{k+2}{2}]. \tag{3.27}$$

Observe that

$$\frac{\sin(k-2)\phi - (-1)^k \sin 3\phi + \sin(k+1)\phi}{-2\cos\phi(-\sin(k-1)\phi + (-1)^k \sin 2\phi + \sin(k+1)\phi)}$$

$$= \frac{(-1)^{k+1}\sin 3\phi + \sin(k-2)\phi + \sin(k+1)\phi}{(-1)^{k+1}(\sin 3\phi + \sin\phi) + \sin(k-2)\phi - \sin(k+2)\phi} \quad =: \frac{A(k,\phi)}{B(k,\phi)}.$$

In order to prove (3.15), we need to show that $B(k,\phi) > A(k,\phi)$ or $B(n,\phi) < A(n,\phi)$, depending on whether $B(n,\phi)$ is positive or negative. This is done in the following two claims.

Claim 1: If $\phi^{k+1} < \phi < \pi$, then $B(k,\phi)$ is positive for odd k and negative for even k.

Claim 2: If $\phi^{k+1} < \phi < \pi$, then $B(k,\phi) - A(k,\phi)$ is positive for odd k and negative for even k.

Proof of Claim 1: Rearranging the terms involved in $B(k,\phi)$, we get

$$\begin{aligned} B(k,\phi) &= (-1)^{k+1} 2\sin 2\phi \cos\phi - 2\cos k\phi \sin 2\phi \\ &= 2\sin 2\phi \Big((-1)^{k+1}\cos\phi - \cos k\phi \Big) \\ &= \begin{cases} 2\sin 2\phi \cdot 2\sin\frac{k+1}{2}\phi \sin\frac{k-1}{2}\phi & , k \text{ odd} \\ -2\sin 2\phi \cdot 2\cos\frac{k+1}{2}\phi \cos\frac{k-1}{2}\phi & , k \text{ even.} \end{cases} \end{aligned}$$

Now with the help of (3.27)

$$\pi < 2\phi < 2\pi, \quad (\frac{k-1}{2} - 1)\pi < \frac{k-1}{2}\phi < \frac{k-1}{2}\pi < \frac{k+1}{2}\phi < \frac{k+1}{2}\pi$$

and thereby Claim 1 follows, because $\sin 2\phi$ is negative, and each of the pairs $\sin\frac{k-1}{2}\phi$, $\sin\frac{k+1}{2}\phi$ and $\cos\frac{k-1}{2}\phi$, $\cos\frac{k+1}{2}\phi$ has opposite signs.

Proof of Claim 2:

$$\begin{aligned} B(k,\phi) - A(k,\phi) &= -\sin(k+2)\phi - \sin(k+1)\phi + (-1)^{k+1}\sin\phi \\ &= -2\sin(k+\frac{3}{2})\phi\cos\frac{1}{2}\phi + 2(-1)^{k+1}\sin\frac{1}{2}\phi\cos\frac{1}{2}\phi \\ &= \begin{cases} 2\cos\frac{1}{2}\phi \cdot (\sin\frac{1}{2}\phi - \sin(k+\frac{3}{2})\phi) & , k \text{ odd} \\ 2\cos\frac{1}{2}\phi \cdot (-\sin\frac{1}{2}\phi - \sin(k+\frac{3}{2})\phi) & , k \text{ even} \end{cases} \\ &= \begin{cases} -4\cos\frac{1}{2}\phi\cos\frac{k+2}{2}\phi\sin\frac{k+1}{2}\phi & , k \text{ odd} \\ -4\cos\frac{1}{2}\phi\sin\frac{k+2}{2}\phi\cos\frac{k+1}{2}\phi & , k \text{ even.} \end{cases} \end{aligned}$$

Moreover, again with the help of (3.27)

$$-\frac{1}{2}\pi < \frac{1}{2}\phi < \frac{1}{2}\pi, \quad \frac{k-1}{2}\pi < \frac{k+1}{2}\phi < \frac{k+1}{2}\pi, \quad \frac{k}{2}\pi < \frac{k+2}{2}\phi < \frac{k+2}{2}\pi$$

and thereby Claim 2 follows, because $\cos\frac{1}{2}\phi$ is positive, $\cos\frac{k+2}{2}\phi$ and $\sin\frac{k+1}{2}\phi$ have opposite signs and $\sin\frac{k+2}{2}\phi$ and $\cos\frac{k+1}{2}\phi$ have the same sign. 2

Proof of Proposition 3.2.5. We use the shorthand $\phi := \phi_j^k$, $\lambda_i^r := \lambda_i^r(q_j^k)$ and $\mu_i^r := \mu_i^r(q_j^k)$ for $r = k-1, k$. Moreover, observation (3.11) and the analogue for cos, namely

$$\cos(k-i)\phi_j^k = (-1)^{k-j}\cos(i+1)\phi_j^k, \tag{3.28}$$

will prove to be useful. Recall that, by Proposition 3.1.4, $\lambda_i^k = \frac{\eta_i^k(\phi)}{\psi^k(\phi)}$, where

$$\eta_i^k(\phi) := (-1)^{k-i}\sin i\phi - (-1)^i \sin(k-i+1)\phi + \sin(k+1)\phi,$$

$$\psi^k(\phi) := \frac{-(-1)^k \sin\phi + \sin k\phi + \sin(k+1)\phi}{1+\cos\phi} + k\sin(k+1)\phi.$$

To prove part (i), we repeatedly use (3.11) together with the fact that j is odd.

$$\begin{aligned}
\eta_i^{k-1}(\phi) &= (-1)^{k-1-i}\sin i\phi - (-1)^{k+i}\sin(i+1)\phi + (-1)^k \sin\phi \\
\psi^{k-1}(\phi) &= \frac{-(-1)^{k-1}\sin\phi + (-1)^k \sin 2\phi + (-1)^k \sin\phi}{1+\cos\phi} + (k-1)(-1)^k \sin\phi \\
&= (-1)^k\Big(\frac{2\sin\phi + 2\sin\phi\cos\phi}{1+\cos\phi} + (k-1)\sin\phi\Big) \\
&= (-1)^k(k+1)\sin\phi.
\end{aligned}$$

For the proof of (ii), notice that $\eta_i^k(\phi) = 0 = \psi^k(\phi)$. We therefore use L'Hôpital's rule to determine λ_i^k and differentiate $\eta_i^k(\phi)$ and $\psi^k(\phi)$ with respect to ϕ:

$$\begin{aligned}
(\eta_i^k)'(\phi) &= (-1)^{k-i} i\cos i\phi - (-1)^i(k-i+1)\cos(k-i+1)\phi + (k+1)\cos(k+1)\phi \\
&= (-1)^{k-i} i\cos i\phi + (-1)^{k+i}(k-i+1)\cos i\phi - (k+1)(-1)^k \\
&= (-1)^{k-1}(k+1)\Big(1 + (-1)^{i+1}\cos i\phi\Big).
\end{aligned}$$

Similarly,

$$\begin{aligned}
(\psi^k)'(\phi) &= \frac{\Big(-(-1)^k\cos\phi + k\cos k\phi + (k+1)\cos(k+1)\phi\Big)(1+\cos\phi)}{(1+\cos\phi)^2} \\
&\quad - \frac{\Big(-(-1)^k\sin\phi + \sin k\phi + \sin(k+1)\phi\Big)(-\sin\phi)}{(1+\cos\phi)^2} + k(k+1)\cos(k+1)\phi \\
&= \frac{\Big(-(-1)^k\cos\phi + k(-1)^{k-1}\cos\phi + (k+1)(-1)^{k-1}\Big)}{1+\cos\phi} + k(k+1)(-1)^{k-1} \\
&= \frac{(-1)^{k-1}\Big((k+1)\cos\phi + (k+1)\Big)}{1+\cos\phi} + k(k+1)(-1)^{k-1} \quad = (-1)^{k-1}(k+1)^2.
\end{aligned}$$

To prove (iii), simply observe that by (i) and (ii)

$$\lambda_{k-1}^{k-1} = \frac{\sin\phi + \sin 2\phi + \sin\phi}{(k+1)\sin\phi} = 2\cdot\frac{\sin\phi(1+\cos\phi)}{(k+1)\sin\phi} = 2\lambda_k^k.$$

Before we proceed to the proof of (iv)-(vi), we claim that

$$\lambda_i^{k-1} - \lambda_{i-1}^{k-1} \quad = \quad -2x^* \cdot \mu_i^{k-1} \qquad \text{for } i \geq 2, \tag{3.29}$$

which is immediately verified as follows:

$$\begin{aligned}
\lambda_i^{k-1} - \lambda_{i-1}^{k-1} &= \frac{-(-1)^i(\sin i\phi + \sin(i+1)\phi) + (-1)^{i-1}(\sin(i-1)\phi + \sin i\phi)}{(k+1)\sin\phi} \\
&= (-1)^{i-1}\frac{2\sin i\phi + \sin(i+1)\phi + \sin(i-1)\phi}{(k+1)\sin\phi} \\
&= (-1)^{i-1}\frac{2\sin i\phi + 2\sin i\phi\cos\phi}{(k+1)\sin\phi} \\
&= \frac{2(-1)^{i-1}\sin i\phi(1+\cos\phi)}{(k+1)\sin\phi} \quad = \quad -2\lambda_k^k\mu_i^{k-1} \quad = \quad -2x^*\mu_i^{k-1},
\end{aligned}$$

which proves (3.29). On the other hand,

$$\begin{aligned}
\lambda_i^{k-1} + \lambda_{i-1}^{k-1} &= \frac{2\sin\phi - (-1)^i(\sin i\phi + \sin(i+1)\phi) - (-1)^{i-1}(\sin(i-1)\phi + \sin i\phi)}{(k+1)\sin\phi} \\
&= \frac{2\sin\phi + (-1)^{i-1}(\sin(i+1)\phi) - \sin(i-1)\phi)}{(k+1)\sin\phi} \\
&= \frac{2\sin\phi + 2(-1)^{i-1}\cos i\phi\sin\phi}{(k+1)\sin\phi} \quad = \quad 2\frac{1+(-1)^{i-1}\cos i\phi}{k+1} \quad = \quad 2\lambda_i^k,
\end{aligned}$$

which proves (iv). The case $i = k$ of (v) follows from the definition of x^*. For $i < k$ we use (3.29) and (iv):

$$\begin{aligned}
\ell_i^k(x^*) &= \lambda_i^{k-1} + x^* \cdot \mu_i^{k-1} \quad = \quad \lambda_i^{k-1} - \frac{\lambda_i^{k-1} - \lambda_{i-1}^{k-1}}{2} \\
&= \frac{\lambda_i^{k-1} + \lambda_{i-1}^{k-1}}{2} \quad = \quad \lambda_i^k.
\end{aligned}$$

For the case $i = k$ of (vi) use the fact that (by Lemma 3.1.6) $\mu_1^{k-1} = -1$ together with (iii) to see

$$\ell_1^k(x^* - x) = \lambda_1^{k-1} + (x^* - x)\mu_1^{k-1} = \lambda_{k-1}^{k-1} - (x^* - x) = 2\lambda_k^k - (x^* - x) = x^* + x = \ell_k^k(x^* + x).$$

For $i < k$ we have by Lemma 3.1.7

$$\begin{aligned}
\ell_{k-i+1}^k(x^* - x) &= \lambda_{k-i+1}^{k-1} + (x^* - x) \cdot \mu_{k-i+1}^{k-1} \;=\; \lambda_{i-1}^{k-1} - (x^* - x) \cdot \mu_i^{k-1} \\
&= \lambda_{i-1}^{k-1} + \frac{\lambda_i^{k-1} - \lambda_{i-1}^{k-1}}{2} + x \cdot \mu_i^{k-1} \;=\; \lambda_i^{k-1} - \frac{\lambda_i^{k-1} - \lambda_{i-1}^{k-1}}{2} + x \cdot \mu_i^{k-1} \\
&= \lambda_i^{k-1} + x^* \cdot \mu_i^{k-1} + x \cdot \mu_i^{k-1} \quad = \quad \ell_i^k(x^* + x). \qquad 2
\end{aligned}$$

We now approach the proof of Proposition 3.3.3, for which we need two auxiliary lemmas. First recall again that

$$\lambda_j^k(p) = \frac{\eta_j^k(p)}{\psi^k(p)},$$

where

$$\eta_j^k(p) = (-1)^{k-j}\sin j\phi - (-1)^j \sin(k-j+1)\phi + \sin(k+1)\phi.$$

Lemma 3.4.3

$$\eta_j^k(p) - \eta_{j+1}^k(p) = \begin{cases} 4\cdot(-1)^j\cdot\cos\frac{k+1}{2}\phi\cdot\cos\frac{1}{2}\phi\cdot\sin\frac{2k-4j}{4}\phi & , \textit{ if } k \textit{ is even} \\ 4\cdot(-1)^j\cdot\sin\frac{k+1}{2}\phi\cdot\cos\frac{1}{2}\phi\cdot\cos\frac{2k-4j}{4}\phi & , \textit{ if } k \textit{ is odd.} \end{cases}$$

Proof. Assume that k is even. Then

$$\begin{aligned} &\eta_j^k(p) - \eta_{j+1}^k(p) \\ &= (-1)^j\sin j\phi - (-1)^j\sin(k-j+1)\phi - [-(-1)^j\sin(j+1)\phi + (-1)^j\sin(k-j)\phi] \\ &= (-1)^j\cdot 2\cdot\cos\frac{k+1}{2}\phi\cdot\sin(-\frac{k-2j+1}{2}\phi) - (-1)^j\cdot 2\cdot\cos\frac{k+1}{2}\phi\cdot\sin\frac{k-2j-1}{2}\phi \\ &= -(-1)^j\cdot 2\cdot\cos\frac{k+1}{2}\phi\cdot[\sin\frac{k-2j+1}{2}\phi + \sin\frac{k-2j-1}{2}\phi] \\ &= -(-1)^j\cdot 2\cdot\cos\frac{k+1}{2}\phi\cdot 2\cdot\sin\frac{2k-4j}{4}\phi\cdot\cos\frac{1}{2}\phi. \end{aligned}$$

The proof for k odd is almost identical. 2

Lemma 3.4.4 *Let $\phi^k < \phi < \pi$. Then for all integers i, j with $0 \le i < j < \frac{k}{2}$*

$$(i) \qquad (-1)^i\sin(\frac{k}{2}-i)\phi \begin{cases} <0 & , \textit{ if } k\equiv 0(4) \\ >0 & , \textit{ if } k\equiv 2(4) \end{cases}$$

$$(ii) \qquad (-1)^i\cos(\frac{k}{2}-i)\phi \begin{cases} >0 & , \textit{ if } k\equiv 1(4) \\ <0 & , \textit{ if } k\equiv 3(4) \end{cases}$$

$$(iii) \qquad (-1)^i\cos(\frac{k}{2}-i)\phi - (-1)^j\cos(\frac{k}{2}-j)\phi \begin{cases} <0 & , \textit{ if } k\equiv 0(4) \\ >0 & , \textit{ if } k\equiv 2(4) \end{cases}$$

$$(iv) \qquad (-1)^i\sin(\frac{k}{2}-i)\phi - (-1)^j\sin(\frac{k}{2}-j)\phi \begin{cases} <0 & , \textit{ if } k\equiv 1(4) \\ >0 & , \textit{ if } k\equiv 3(4) \end{cases}$$

$$(v) \qquad (-1)^k\cos(k-2i-1)\phi - \cos\phi > 0.$$

Proof. Since $\frac{k-1}{k+1}\pi < \phi < \pi$ we can write

$$\phi = (1 - \frac{2\delta}{k+1})\pi, \quad \text{for some } \delta \in (0,1).$$

Hence, for all $0 \le i < \frac{k}{2}$:

$$(\frac{k}{2} - i)\phi = (\frac{k}{2} - i)\pi - \delta \cdot \underbrace{\frac{k-2i}{k+1}}_{\in (0,1)} \pi.$$

Now use the fact that $(-1)^i \sin x = \sin(x + i\pi)$ and $(-1)^i \cos x = \cos(x + i\pi)$ to see that

$$(-1)^i \sin(\frac{k}{2} - i)\phi = (-1)^i \sin\left((\frac{k}{2} - i)\pi - \delta\frac{k-2i}{k+1}\pi\right) = \sin\left(\frac{k}{2}\pi - \delta\frac{k-2i}{k+1}\pi\right)$$

and analogously $(-1)^i \cos(\frac{k}{2} - i)\phi = \cos(\frac{k}{2}\pi - \delta\frac{k-2i}{k+1}\pi)$. This already suffices to establish (i) and (ii). For (iii) and (iv) apply the same argument again together with the obvious fact that $\delta\frac{k-2i}{k+1} > \delta\frac{k-2j}{k+1}$ since $i < j$.

To prove (v), observe that

$$(-1)^k \cos(k-2i-1)\phi - \cos\phi \quad = \quad -\cos\left(2\delta\frac{k-2i-1}{k+1}\pi\right) + \cos\left(\frac{2\delta}{k+1}\pi\right).$$

But since $2\delta\frac{k-2i-1}{k+1}\pi > \frac{2\delta}{k+1}\pi$ and $2\delta\frac{k-2i-1}{k+1}\pi < 2\pi - \frac{2\delta}{k+1}\pi$, this expression is positive. 2

Proof of Proposition 3.3.3. We will prove that for $q^k < p < q^*$ the functions

$$\mathrm{c}(Q_{k,p}) = H(p)\sum_{i=1}^{k-1} \lambda_i^k(p)\lambda_{i+1}^k(p)$$

$$1 - 2 \cdot \mathrm{d}(Q_{k,p}) = \sum_{i=1}^{k} (\lambda_i^k(p))^2 + 2(1-p)\sum_{i=1}^{k-1} \lambda_i^k(p)\lambda_{i+1}^k(p)$$

are strictly decreasing. In the following we will abbreviate $\lambda_i^k(p)$ by λ_i, and set $\lambda_0 = \lambda_{k+1} = 0$ for convenience. It obviously suffices to show that both of the involved sums are strictly decreasing. Denote the sums by A and B respectively. Their derivatives are

$$A' = 2\sum_{i=1}^{k} \lambda_i\lambda_i', \quad B' = \sum_{i=1}^{k-1}(\lambda_i'\lambda_{i+1} + \lambda_i\lambda_{i+1}') = \sum_{i=1}^{k}(\lambda_{i-1} + \lambda_{i+1})\lambda_i'.$$

Notice that $\sum_{i=1}^{k} \lambda_i' = 0$ since $\sum_{i=1}^{k} \lambda_i = 1$. Recall that by the original Lagrange conditions as stated in (3.5) we know that there exists a u such that for all i

$$b(p) \cdot \lambda_{i-1} + a(p) \cdot \lambda_i + b(p) \cdot \lambda_{i+1} = u.$$

Combining these two facts results in

$$\sum_{i=1}^{k} b(p) \cdot (\lambda_{i-1} + \lambda_{i+1})\lambda_i' + a(p)\lambda_i\lambda_i' = 0,$$

which is the same as $b(p)B' + a(p)A'/2 = 0$. Since $b(p)$ is negative and $a(p)$ is positive, B' and A' must have the same sign. We introduce a third sum

$$C := \sum_{i=0}^{k} (\lambda_{i+1} - \lambda_i)^2 \ = 2(A - B).$$

Since

$$C' = 2A' - 2B' = A' \cdot (2 + \frac{a(p)}{b(p)})$$

and $\frac{a(p)}{b(p)} > -2$ for $p < q^*$, C' has the same sign as A' and B'. Our aim is now to show that

$$C' = 2 \ \sum_{i=1}^{k} (\lambda_{i+1} - \lambda_i)(\lambda_{i+1}' - \lambda_i')$$

is indeed negative.

Recall that

$$\lambda_j = \lambda_j^k(p) = \frac{\eta_j^k(p)}{\psi^k(p)}, \quad \eta_j^k(p) = (-1)^{k-j} \sin j\phi - (-1)^j \sin(k - j + 1)\phi + \sin(k + 1)\phi,$$

and note that the definition also allows for $\eta_0^k(p) = \eta_{k+1}^k(p) = 0$. By Lemma 3.4.3

$$\eta_j^k(p) - \eta_{j+1}^k(p) = \begin{cases} 4 \cdot (-1)^j \cdot \cos \frac{k+1}{2}\phi \cdot \cos \frac{1}{2}\phi \cdot \sin(\frac{k}{2} - j)\phi & , \text{ if } k \text{ is even} \\ 4 \cdot (-1)^j \cdot \sin \frac{k+1}{2}\phi \cdot \cos \frac{1}{2}\phi \cdot \cos(\frac{k}{2} - j)\phi & , \text{ if } k \text{ is odd.} \end{cases}$$

For k even this implies that

$$\frac{\lambda_i - \lambda_{i+1}}{\lambda_0 - \lambda_1} = (-1)^i \frac{\sin(\frac{k}{2} - i)\phi}{\sin \frac{k}{2}\phi}$$

and hence, as $\lambda_0 = 0$, also

$$\lambda_{i+1} = \lambda_i + (-1)^i \frac{\sin(\frac{k}{2} - i)\phi}{\sin \frac{k}{2}\phi} \cdot \lambda_1.$$

For k odd we analogously obtain

$$\lambda_{i+1} = \lambda_i + (-1)^i \frac{\cos(\frac{k}{2} - i)\phi}{\cos \frac{k}{2}\phi} \cdot \lambda_1.$$

Moreover, induction on i gives

$$\lambda_i = (1 + \frac{\sum_{j=1}^{i-1} y_j}{y_0}) \cdot \lambda_1, \qquad \text{where } y_j := \begin{cases} (-1)^j \sin(\frac{k}{2} - j)\phi & , \text{ for } k \text{ even,} \\ (-1)^j \cos(\frac{k}{2} - j)\phi & , \text{ for } k \text{ odd.} \end{cases} \tag{3.30}$$

Observe that by Lemma 3.4.4 we know that for a given k all

$$y_i \text{ with } 0 \leq i < \tfrac{k}{2} \text{ have the same sign.} \tag{3.31}$$

Setting $i := k + 1$ in (3.30), we deduce from $\lambda_{k+1} = 0$ that

$$\sum_{j=1}^{k} y_j = -y_0.$$

Taking $\sum_{i=1}^{k+1} \lambda_i = 1$ and replacing λ_i by (3.30) gives

$$1 = (k+1)\lambda_1 + \frac{\lambda_1}{y_0} \sum_{j=1}^{k} (k + 1 - j) y_j = -\frac{\lambda_1}{y_0} \sum_{j=1}^{k} j y_j,$$

and hence

$$\lambda_1 = -\frac{y_0}{\sum_{j=1}^{k} j y_j}. \tag{3.32}$$

Furthermore

$$\begin{aligned} \lambda_1' &= -\frac{y_0'}{\sum_{j=1}^{k} j y_j} + \frac{y_0 \cdot \sum_{j=1}^{k} j y_j'}{(\sum_{j=1}^{k} j y_j)^2} \\ &= \lambda_1 \cdot (\frac{y_0'}{y_0} - \frac{\sum_{j=1}^{k} j y_j'}{\sum_{j=1}^{k} j y_j}). \end{aligned} \tag{3.33}$$

Now determine λ_i' by differentiating (3.30)

$$\lambda_i' = (1 + \frac{\sum_{j=1}^{i-1} y_j}{y_0}) \cdot \lambda_1' + (\frac{\sum_{j=1}^{i-1} y_j'}{y_0} - \frac{y_0' \sum_{j=1}^{i-1} y_j}{(y_0)^2}) \cdot \lambda_1.$$

Hence,

$$\lambda_{i+1}' - \lambda_i' = \frac{\lambda_1'}{y_0} \cdot y_i + \frac{\lambda_1}{y_0} \cdot (y_i' - \frac{y_0'}{y_0} \cdot y_i).$$

As $\lambda_{i+1} - \lambda_i = \frac{\lambda_1}{y_0} \cdot y_i$, this implies

$$\begin{aligned}
C' &= \sum_{i=0}^{k} (\lambda_{i+1} - \lambda_i)(\lambda_{i+1}' - \lambda_i') \\
&= \sum_{i=0}^{k} \left(\frac{\lambda_1 \cdot \lambda_1'}{(y_0)^2} \cdot y_i^2 + \frac{\lambda_1 \cdot \lambda_1}{(y_0)^2} \cdot (y_i y_i' - \frac{y_0'}{y_0} \cdot y_i^2) \right) \\
&\overset{(3.33)}{=} \frac{(\lambda_1)^2}{(y_0)^2} \cdot \sum_{i=0}^{k} \left(y_i y_i' - y_i^2 \cdot \frac{\sum_{j=1}^{k} j y_j'}{\sum_{j=1}^{k} j y_j} \right) \\
&= \frac{(\lambda_1)^2}{(y_0)^2 \cdot \sum_{j=1}^{k} j y_j} \cdot \sum_{i=0}^{k} \sum_{j=1}^{k} (y_i y_i' j y_j - y_i^2 j y_j') \\
&\overset{(3.32)}{=} -\frac{(\lambda_1)^3}{(y_0)^3} \cdot \sum_{i=0}^{k} \sum_{j=1}^{k} (y_i y_i' j y_j - y_i^2 j y_j') \\
&= -\frac{(\lambda_1)^3}{(y_0)^3} \cdot \sum_{i=0}^{k} \sum_{j=0}^{k} (y_i y_i' j y_j - y_i^2 j y_j')
\end{aligned}$$

Since we want to prove that $C' < 0$ and since we know that $\lambda_1 > 0$, our aim is to show that the above double sum has the same sign as y_0. To do this we first simplify this sum by grouping some terms, namely the pairs $i, k-i$ and $j, k-j$.

Observe that for k even

$$y_{k-i} = (-1)^{k-i} \cdot \sin(\frac{k}{2} - (k-i))\phi = (-1)^i \sin(-(\frac{k}{2} - i))\phi = -y_i$$

hence $y_{k-i}' = -y_i'$ and analogously for k odd. Therefore

$$j y_j + (k-j) y_{k-j} = (2j-k) y_j, \qquad j y_j' + (k-j) y_{k-j}' = (2j-k) y_j'.$$

That is, using that for k even $y_{\frac{k}{2}} = y'_{\frac{k}{2}} = 0$,

$$\begin{aligned}
\sum_{i=0}^{k}\sum_{j=0}^{k}(y_i y'_i j y_j - y_i^2 j y'_j) &= 2\sum_{0\le i<\frac{k}{2}}\ \sum_{0\le j<\frac{k}{2}}(y_i y'_i(2j-k)y_j - y_i^2(2j-k)y'_j) \\
&= 2\sum_{0\le i<j<\frac{k}{2}}(y_i y'_i(2j-k)y_j - y_i^2(2j-k)y'_j + y_j y'_j(2i-k)y_i - y_j^2(2i-k)y'_i) \\
&= 2\sum_{0\le i<j<\frac{k}{2}}[y'_i y_j(y_i(2j-k) - y_j(2i-k)) - y_i y'_j((2j-k)y_i - y_j(2i-k))] \\
&= 2\sum_{0\le i<j<\frac{k}{2}}(y'_i y_j - y_i y'_j)\cdot(y_i(2j-k) - y_j(2i-k)).
\end{aligned}$$

So, we have reduced the problem to the two factors $y'_i y_j - y_i y'_j$ and $y_i(2j-k) - y_j(2i-k)$ for $i<j$. We claim that the factor $y_i(2j-k) - y_j(2i-k)$ has the same sign as y_0. Consider its derivative: for k even

$$\begin{aligned}
&(y_i(2j-k) - y_j(2i-k))' \\
&= (2j-k)(-1)^i(\frac{k}{2}-i)\cos(\frac{k}{2}-i)\phi - (2i-k)(-1)^j(\frac{k}{2}-j)\cos(\frac{k}{2}-j)\phi \\
&= -2(\frac{k}{2}-j)(\frac{k}{2}-i)\cdot[(-1)^i\cos(\frac{k}{2}-i)\phi - (-1)^j\cos(\frac{k}{2}-j)\phi].
\end{aligned}$$

Analogously for k odd:

$$(y_i(2j-k) - y_j(2i-k))' = 2(\frac{k}{2}-j)(\frac{k}{2}-i)\cdot[(-1)^i\sin(\frac{k}{2}-i)\phi - (-1)^j\sin(\frac{k}{2}-j)\phi].$$

Now apply Lemma 3.4.4 (iii) and (iv) to see that these expressions are positive if $k \equiv 0, 3(4)$ and negative otherwise. As the value of $y_i(2j-k) - y_j(2i-k)$ is zero at $\phi = \pi$ the value has to be negative for all $\phi \in (\frac{k-1}{k+1}\pi, \pi)$ if $k \equiv 0, 3(4)$ and positive otherwise. By Lemma 3.4.4 (i) and (ii) we know that y_0 behaves identically.

Now consider $y'_i y_j - y_i y'_j$: we claim that it is positive for all k and all $i<j$. Since this expression is exactly the numerator of the derivative of y_i/y_j, it suffices to show that the latter is positive. Moreover observe that

$$\begin{aligned}
\left(\frac{y_i}{y_j}\right)' &= \left(\frac{y_i}{y_{i+1}}\frac{y_{i+1}}{y_{i+2}}\cdots\frac{y_{j-1}}{y_j}\right)' \\
&= \left(\frac{y_i}{y_{i+1}}\right)'\frac{y_{i+1}}{y_{i+2}}\cdots\frac{y_{j-1}}{y_j} + \frac{y_i}{y_{i+1}}\left(\frac{y_{i+1}}{y_{i+2}}\right)'\cdots\frac{y_{j-1}}{y_j} + \cdots + \frac{y_i}{y_{i+1}}\frac{y_{i+1}}{y_{i+2}}\cdots\left(\frac{y_{j-1}}{y_j}\right)'.
\end{aligned}$$

Due to (3.31), the terms y_i/y_{i+1} are all positive, and it therefore suffices to show that $(y_i/y_{i+1})' > 0$ for all $0 \le i < \frac{k}{2} - 1$. We claim that

$$y_i' y_{i+1} - y_i y_{i+1}' = -(-1)^k \frac{1}{2} \sin(k-2i-1)\phi + \frac{1}{2}(k-2i-1)\sin\phi, \tag{3.34}$$

which can be checked as follows. For k even:

$$\begin{aligned}
& y_i' y_{i+1} - y_i y_{i+1}' \\
&= -(\frac{k}{2} - i)\cos(\frac{k}{2} - i)\phi \sin(\frac{k}{2} - i - 1)\phi + (\frac{k}{2} - i - 1)\cos(\frac{k}{2} - i - 1)\phi \sin(\frac{k}{2} - i)\phi \\
&= -(\frac{k}{2} - i)\frac{1}{2}[\sin(k-2i-1)\phi + \sin(-\phi)] + (\frac{k}{2} - i - 1)\frac{1}{2}[\sin(k-2i-1)\phi + \sin\phi] \\
&= -\frac{1}{2}\sin(k-2i-1)\phi + \frac{1}{2}(k-2i-1)\sin\phi.
\end{aligned}$$

For k odd:

$$\begin{aligned}
& y_i' y_{i+1} - y_i y_{i+1}' \\
&= (\frac{k}{2} - i)\sin(\frac{k}{2} - i)\phi \cos(\frac{k}{2} - i - 1)\phi - (\frac{k}{2} - i - 1)\sin(\frac{k}{2} - i - 1)\phi \cos(\frac{k}{2} - i)\phi \\
&= (\frac{k}{2} - i)\frac{1}{2}[\sin\phi + \sin((k-2i-1)\phi)] - (\frac{k}{2} - i - 1)\frac{1}{2}[\sin(-\phi) + \sin(k-2i-1)\phi] \\
&= \frac{1}{2}\sin(k-2i-1)\phi + \frac{1}{2}(k-2i-1)\sin\phi.
\end{aligned}$$

At the point $\phi = \pi$ (3.34) has value zero. Again we consider the derivative

$$(y_i' y_{i+1} - y_i y_{i+1}')' = -\frac{1}{2}(k-2i-1)[(-1)^k \cos(k-2i-1)\phi - \cos\phi].$$

By Lemma 3.4.4 (v) this is negative and therefore $y_i' y_{i+1} - y_i y_{i+1}'$ must be positive for $\phi < \pi$. 2

Bibliography

[1] M. Ajtai, J. Komlós, and E. Szemerédi, Sorting in $C \log n$ parallel steps, *Combinatorica* 3, 1–19, 1983.

[2] M.H. Albert and A.M. Frieze, Random graph orders, *Order* 6, 19–30, 1989.

[3] V.E. Alekseev, On the entropy values of hereditary classes of graphs, *Discrete Appl. Mathematics* 3, 191–199, 1993.

[4] G. Birkhoff, *Lattice Theory*, Colloquium Publications 25, Amer. Math. Soc., Providence, 1967.

[5] B. Bollobás, *Random Graphs*, Academic Press, London, 1985.

[6] B. Bollobás, Hereditary properties of graphs: asymptotic enumeration, global structure, and colouring, *Documenta Mathematica*, Volume 3, ICM 1998, 333–342, 1998. Academic Press, London, 1985.

[7] B. Bollobás and G. Brightwell, Graphs whose every transitive orientation contains almost every relation, *Israel J. Math.* 59, 112-128, 1987.

[8] B. Bollobás and A. Thomason, Hereditary and monotone properties of graphs, in *The mathematics of Paul Erdős II* (R.L. Graham and J. Nešetřil, eds.), Algorithms and Combinatorics, Vol. 14, Springer–Verlag, 70–78, 1997.

[9] B. Bollobás and P. Winkler, The longest chain among random points in euclidean space, *Proc. Amer. Math. Soc.* 103, 347–353, 1988.

[10] K.H. Borgwardt, Some distribution independent results about the asymptotic order of the average number of pivot steps in the simplex algorithm, *Mathematics of Operations Research* Vol. 7, 441–462, 1982.

[11] G. Brightwell, Models of random partial orders, in *Surveys in Combinatorics* (K. Walker, ed.), Cambridge University Press, 53-83, 1993.

[12] G. Brightwell, H.J. Prömel, and A. Steger, The average number of linear extensions of a partial order, *J. Combinat. Theory (A)* 73, 193–206, 1996.

[13] F. Chung and R.L. Graham, *Erdős on Graphs: His Legacy of Unsolved Problems*, A K Peters, Wellesley, 1998.

[14] D. Dhar, Entropy and phase transitions in partially ordered sets, *J. Math. Phys.* 19(8), 1711–1713, 1978.

[15] D. Dhar, Asymptotic enumeration of partially ordered sets, *Pacific Journal of Mathematics* 90 (2), 299–305, 1980.

[16] R.P. Dilworth, A decomposition theorem for partially ordered sets, *Ann. Math.* 51, 161–165, 1950.

[17] B. Dushnik and E.W. Miller, Partially ordered sets, *Amer. J. Math.* 63, 600–610, 1941.

[18] M.E. Dyer and A.M. Frieze, The solution of some NP-hard problems in polynomial expected time, *J. Algorithms* 10, 451–489, 1989.

[19] M.H. El-Zahar, Enumeration of ordered sets, in *Algorithms and Order* (I. Rival, ed.) NATO ASI Series, Kluwer Academic publishers, Dordrecht, 327–352, 1989.

[20] P. Erdős, Some remarks on the theory of graphs, *Bull. Austral. Math. Soc.* 53, 292–294, 1947.

[21] P. Erdős, P. Frankl, and V. Rödl, The asymptotic enumeration of graphs not containing a fixed subgraph and a problem for hypergraphs having no exponent, *Graphs and Combinatorics* 2, 113–121, 1986.

[22] P. Erdős, H. Kierstead, and W.T. Trotter, The dimension of random ordered sets, *Random Structures and Algorithms* 2, 253–275, 1991.

[23] P. Erdős, D.J. Kleitman, and B.L. Rothschild, Asymptotic enumeration of K_n-free graphs, in *International Colloquium on Combinatorial theory, Atti dei Convegni Lincei* 17, Vol. 2, Rome, 19–27, 1976.

[24] P. Erdős and A. Rényi, On random graphs, I, *Publ. Math. Debrecen* 6, 290–297, 1959.

[25] P. Erdős and A. Rényi, On the evolution of random graphs, *Publ. Math. Inst. Hungar. Acad. Sci.* 5, 17–61, 1961.

[26] P. Erdős and A. Rényi, On the evolution of random graphs, *Bull. Inst. Internat. Statist.* 38(4), 343–347, 1961.

[27] P. Erdős and A. Rényi, On the strength of connectedness of a random graph, *Acad. Sci. Hungar.* 12, 261–267, 1961.

[28] P. Erdős and A. Rényi, On random matrices, *Magyar Tud. Akad. Mat. Kutató Int. Közl.* 8, 455–461, 1964.

[29] P. Erdős and A. Rényi, On random matrices, II, *Studia Sci. Math. Hungar.* 3, 459–464, 1968.

[30] P. Erdős and M. Simonovits, A limit theorem in graph theory, *Studia Sci. Math. Hung.* 1, 51–57, 1966.

[31] P. Erdős and A.H. Stone, On the structure of linear graphs, *Bull. Amer. Math. Soc.* 5, 1089-1091, 1946.

[32] Z. Füredi and J. Kahn, On the dimension of ordered sets of bounded degree, *Order* 3, 17–20, 1986.

[33] P. Hanlon, Counting interval graphs, *Trans. Amer. Math. Soc.* 272, 383–426, 1982.

[34] M. Karoński: Random Graphs, in *Handbook of Combinatorics*, Volume 1, 351–380. North-Holland, Amsterdam, 1995

[35] R.M. Karp, The probabilistic analysis of combinatorial optimization algorithms, *Proc. International Congress of Mathematicians*, Warzawa 1989, 1601–1609, 1984.

[36] D.A. Klarner, The number of graded partially ordered sets, *J. Combin. Theory* 6, 12–19, 1969.

[37] V. Klee and G.J. Minty, How good is the simplex algorithm?, in *Inequalities – III*, Academic Press, 159–175, 1972.

[38] D.J. Kleitman and B.L. Rothschild, The number of finite topologies, *Proc. Amer. Math. Soc.* 25, 276–282, 1970.

[39] D.J. Kleitman and B.L. Rothschild, Asymptotic enumeration of partial orders on a finite set, *Trans. Amer. Math. Soc.* 205, 205–220, 1975.

[40] D.J. Kleitman and B.L. Rothschild, A phase transition on partial orders, *Physica* 96A, 254–259, 1979.

[41] D.J. Kleitman and D.B. Wilson, On the number of graphs which lack small cycles, *preprint*, 1997.

[42] D.J. Kleitman and K.J. Winston, The asymptotic number of lattices, *Ann. Disc. Math.* 6, 243–249, 1980.

[43] W. Klotz and L. Lucht, Endliche Verbände, *Reine Angew. Math.* 247, 58–68, 1971.

[44] Ph.G. Kolaitis, H.J. Prömel, and B.L. Rothschild, K_{l+1}-free graphs: asymptotic structure and a 0–1 law, *Trans. Amer. Math. Soc.* 303, 637–671, 1987.

[45] B. Kreuter, *Extremale und Asymptotische Graphentheorie für verbotene bipartite Untergraphen*, Diplomarbeit, Universität Bonn, 1994.

[46] L. Kučera, Graphs with small chromatic number are easy to color, *Inform. Process. Lett.* 30, 233–236, 1989.

[47] E. Lamken and B.L. Rothschild, The number of odd–cycle–free graphs, in *Finite and Infinite Sets* (A. Hajnal, L. Lovász, and V.T. Sós, eds.), Colloq. Math. Soc. János Bolyai 37, North–Holland, Amsterdam, 547–553, 1984.

[48] B.F. Logan and L.A. Schepp, A variational problem for Young tableaux, *Advances in Mathematics* 26, 206–222, 1977.

[49] T. Łuczak, On triangle–free random graphs, *preprint*, 1998.

[50] R. Möhring, Almost all comparability graphs are UPO, *Discrete Math.* 50, 63–70, 1984.

[51] H.J. Prömel, Counting unlabeled structures, *J. Combinat. Theory (A)* 44, 83–93, 1987.

[52] H.J. Prömel and A. Steger, Excluding induced subgraphs II: extremal graphs, *Discrete Appl. Mathematics* 44, 283–294, 1993.

[53] H.J. Prömel and A. Steger, Excluding induced subgraphs III: a general asymptotic, *Random Structures and Algorithms* 3, 19–31, 1992.

[54] H.J. Prömel and A. Steger, The asymptotic number of graphs not containing a fixed color–critical subgraph, *Combinatorica* 12, 463–473, 1992.

[55] H.J. Prömel and A. Steger, Coloring clique-free graphs in linear expected time, *Random Structures and Algorithms* 3, 375–402, 1992.

[56] H.J. Prömel and A. Steger, On the asymptotic structure of sparse triangle-free graphs, *Journal of Graph Theory* 21, 137–151, 1996.

[57] E.R. Scheinerman and J. Zito, On the size of hereditary classes of graphs, *J. Combinat. Theory (B)*, 16–39, 1994.

[58] R. Stanley, Enumeration of posets generated by disjoint unions and ordinal sums, *Proc. Amer. Math. Soc.* 45(2), 295–299, 1974.

[59] A. Steger, *Asymptotic properties of H-free graphs*, Habilitationsschrift, Universität Bonn, 1993.

[60] E. Szemerédi, Regular partitions of graphs, in *Problèmes en Combinatoire et Théorie des Graphes*, Proc. Colloque Inter. CNRS (J.-C. Bermond, J.-C. Fournier, M. Las Vergnas, D. Sotteau, eds.), CNRS, Paris, 399–401, 1978.

[61] W.T. Trotter, *Combinatorics and Partially Ordered Sets: Dimension Theory.* Johns Hopkins University Press, Baltimore and London, 1992.

[62] W.T. Trotter, Applications of the probabilistic method to partially ordered sets, in *The mathematics of Paul Erdős II* (R.L. Graham and J. Nešetřil, eds.), Algorithms and Combinatorics, Vol. 14, Springer–Verlag, 214–228, 1997.

[63] P. Turán, Egy gráfelméleti szélőértékfeladatról, *Mat. Fiz. Lapok* 48, 436–452, 1941.

[64] P. Turán, On the theory of graphs, *Colloq. Math.* 3, 19–30, 1954.

[65] A.M. Veršik and S.V. Kerov, Asymptotics of the Plancherel measure of the symmetric group and the limiting form of Young tableaux, *Dokl. Akad. Nauk. SSR* 233, 1024–1028, 1977.

[66] H.S. Wilf, Backtrack: an $O(1)$ expected time algorithm for the graph coloring problem, *Inform. Process. Lett.* 18, 119–121, 1984.

[67] P. Winkler, Random structures and 0-1 laws, in *Finite and Infinite Combinatorics of Sets and Logic* (N.W. Sauer, R.E. Woodrow, and B. Sands, eds.) NATO ASI Series, Kluwer Academic publishers, Dordrecht, 1991.