Communauté UNIVERSITÉ Grenoble Alpes

THÈSE

Pour obtenir le grade de

DOCTEUR DE LA COMMUNAUTE UNIVERSITE GRENOBLE ALPES

Spécialité : **Informatique**

Arrêté ministériel : 25 mai 2016

Présentée par

Rémi de Joannis de Verclos

Thèse dirigée par **Frédéric Maffray** et
co-encadrée par **Louis Esperet** et **Jean-Sébastien Sereni**

préparée au sein du **Laboratoire G-SCOP**
dans **l'École Doctorale MSTII**

Applications des limites de structures combinatoires en géométrie et en théorie des graphes

Thèse soutenue publiquement le **20 juillet 2018**,
devant le jury composé de :

Monsieur Louis Esperet
Chargé de recherche CNRS, Laboratoire G-SCOP, Co-encadrant
Monsieur Daniel Kráľ
Professeur, University of Warwick, Examinateur
Monsieur Patrice Ossona de Mendez
Chargé de recherche CNRS, École des Hautes Études en Sciences Sociales, Rapporteur
Madame Myriam Preissmann
Directrice de recherche CNRS, Laboratoire G-SCOP, Présidente du Jury
Monsieur Jean-Sébastien Sereni
Chargé de recherche CNRS, Laboratoire ICUBE, Co-encadrant
Monsieur Asaf Shapira
Professeur, Université de de Tel-Aviv, Rapporteur
Monsieur Stéphan Thomassé
Professeur, ENS de Lyon, Examinateur

Abstract

This thesis is focused on problems related to the theory of combinatorial limits, a recent theory that opened links between different fields such as analysis, combinatorics, geometry and probability theory. The purpose of this thesis is to apply ideas coming from this framework to problems in extremal combinatorics.

In a first chapter we develop a theory of limits for *order types*, a geometrical object that encodes configuration of a set of points in the plane by the mean of the orientations of their triangles. The order type of a point set suffices to determine many of its properties, such as for instance the boundary of its convex hull. We show that the limit of a converging sequence of order types can be represented by random-free object analogous to a graphon. Further, we link this notion to the natural distributions of order types arising from the sampling of random points from some probability measure of the plane. We observe that in this mean, every probability measure gives rise to a limit of order types. We show that this map from probability measure on the plane to limit of order type is not surjective. Concerning its injectivity, we prove that if a measure has large enough support, for instance if its support contains an open ball, the limit of order types the measure generates suffices to essentially determine this measure.

A second chapter is focused on property testing. A tester is a randomized algorithm for distinguishing between objects satisfying a property from those that are at some distance at least ϵ from having it by means of the edition distance. This gives very efficient algorithms, and in particular algorithms whose complexity does not depend on the size of the input but only on the parameter ϵ. For graphs, it has been shown by Alon and Shapira that every hereditary property has such a tester. We contribute to the following question : which classes of graphs have a one-sided property tester with a number of queries that is a polynomial in $1/\epsilon$? We give a proof that the class of interval graphs has such a tester.

The theory of flag algebras is a framework introduced by Razborov closely related to dense limit of graphs, that gives a way to systematically derive bounds for parameters in extremal combinatorics. In a third chapter we present a program developed during my PhD. that implements this method. This program works as a library that can compute flag algebras, manipulate inequalities on densities and encode the optimization of some parameter in a semi-definite positive instance that can be given to a dedicated solver to obtain a bound on this parameter. This program is in particular used to obtain a new bound for the triangle case of the Caccetta-Häggkvist conjecture.

Résumé

La théorie des limites d'objets combinatoires est une récente théorie qui a permis de tisser des liens entre différents domaines tels que la combinatoire, l'analyse, la géométrie ou la théorie de la probabilité. Le but de cette thèse est d'appliquer des méthodes venant de cette théorie à des problèmes de combinatoire extrémale.

Un premier chapitre traite de limites d'objets géométriques appelés *types d'ordre*, un objet qui encode certaines propriétés des configurations d'ensembles de points du plan. Le type d'ordre d'un ensemble de points suffit à caractériser certaines des propriétés essentielles de cet ensemble, comme par exemple son enveloppe convexe. Je montre qu'une limite de types d'ordre peut être représentée par un objet analogue à un graphon à valeurs 0 ou 1. Les limites de type d'ordre sont à mettre en relation avec les méthodes d'échantillonnage d'ensembles de point du plan. En particulier, toute loi de probabilité sur le plan engendre une limite de types d'ordre qui correspond à l'échantillonnage de points suivant cette loi. Je montre d'une part que cette correspondance n'est pas surjective, c'est-à-dire qu'il existe des limites de types d'ordre ne venant pas de probabilité du plan, et j'étudie d'autre part son injectivité. Je montre que si le support d'une mesure de probabilité est suffisamment « gros », par exemple s'il contient une boule ouverte, alors la limite que cette mesure engendre suffit à caractériser cette mesure à une transformation projective près.

Un second chapitre traite de test de propriété. Un testeur de propriété est un algorithme aléatoire permettant de séparer les objets ayant une certaine propriété des objets à distance au moins ϵ de l'avoir, au sens de la distance d'édition. Ce domaine donne des algorithmes extrêmement rapides, et en particulier des algorithmes dont la complexité ne dépend pas de la taille de l'entrée mais seulement d'un paramètre de précision ϵ. Un résultat fondamental de ce domaine pour les graphes montré par Alon et Shapira est le suivant : toute classe de graphe héréditaire possède un tel testeur. Cette thèse contribue à la question suivante : quelles classes de graphes possèdent un testeur dont la complexité est un polynôme en $1/\epsilon$? Je montre qu'en particulier la classe des graphes d'intervalles possède un tel testeur.

La théorie des algèbres de drapeaux est un outil étroitement lié aux limites de graphes denses qui donne une méthode pour démontrer des bornes sur certains paramètres combinatoires à l'aide d'un ordinateur. Dans un troisième chapitre, je présente un programme écrit durant ma thèse qui implémente cette méthode. Ce programme fonctionne comme une bibliothèque pour calculer dans les algèbres de drapeaux, manipuler des inégalités sur les drapeaux ou encoder des problèmes d'optimisations par une instance de programme semi-défini positif qui peut ensuite être résolue par un solveur externe. Ce programme est en particulier utilisé pour obtenir un nouvelle borne pour le cas triangulaire de la conjecture de Caccetta-Häggkvist.

Contents

Introduction

Describing the structure of large combinatorial structures, such as those underlying the network of the internet, huge sets of data, or protein interaction graphs for instance, is an important challenge in combinatorics that has both a theoretical and practical interest. A theory of limits of combinatorial structure has been developed over the last decade and gives tools to tackle with such structures. The general idea of this theory is to study sequences of objects whose size tends to infinity (in some sense, this could be the mathematical concept corresponding to what a *very large* object is), and that share some asymptotic properties. This sequence is then condensed into some *limit object* that hopefully captures the essential properties of the sequence, while smoothing their local asperities.

The most studied case is the theory of limits of dense graphs. This theory has been developed by Borgs, Chayes, Lovász, Razborov, Sós, Szegedy and Vesztergombi. The density of a graph with k vertices in another graph is the probability that k random vertices induce this graph. A sequence of graphs $(G_n)_{n\in\mathbf{N}}$ *converges* if for every graph H, the density of H in G_n converges to some number $\ell(H)$ when n tends to infinity. The function $H \mapsto \ell(H)$ is the *limit* of the sequence $(G_n)_n$. This limit is a very abstract object that is completely different from a graph. As more "concrete" object to represent the limit object for converging sequences of dense graphs, the authors mentioned above introduced the notion of *graphons*, which is the contraction of *graph function*. A graphon is a symmetric measurable function $W : [0,1]^2 \to [0,1]$ that works as a (continuous) adjacency matrix and that is considered modulo reordering operations that are the continuous analogs of graph isomorphism. The *random sampling* of a graph of size n in a graphon W consists in the following process. Pick n random points $x_1, \ldots, x_n$ uniformly in $[0,1]$ and construct a graph whose vertices are the numbers $x_1, \ldots, x_n$ and put the edge from x_i to x_j with probability $W(x_i, x_j)$. The graphon $(G_n)_{n\in\mathbf{N}}$ represents the limit ℓ if for every graph H, a random sampling of a graph of size $|V(H)|$ in W gives H with probability exactly $\ell(H)$. The graphons make links between different fields of mathematics and in particular introduces analytic tools for graph theory.

Independently, Razborov developed in 2007 the theory of *Flag algebra*. This theory is based on the notion of density and apply to a large variety of combinatorial objects such as graphs, hypergraphs or permutations. In this theory, limits of are exactly the (positive) homomorphisms of a given algebra. Flag algebras provide an algorithmic method to prove bounds in extremal combinatorics with the help of a computer.

This thesis contains three chapters.

In a first chapter, we develop a theory of limits for *order types*. An order type is a combinatorial object coming from geometry that encodes the configuration of a set of points in the plane by the mean of the orientations of its triangles. The order type of a point set suffices to determine many of its geometrical properties, such as the boundary of its convex hull or which graphs can be drawn based on these points without crossing. Limits of order types can be defined analogously to limits of dense graphs, or equivalently as a particular case of as the limit of combinatorial structures, as for instance defined by Razborov. We show that the limit of a converging sequence of order types can be represented by a function $W : [0,1]^3 \to \{-1,1\}$ analogous to a graphon with values in $\{0,1\}$ (such a graphon is called *random-free*). Limits of order types have to be put in parallel with the following natural method for sampling configurations of n points on the plane: Fix a probability measure μ of $\mathbf{R}^2$, for instance the uniform measure on a square and sample n random points independently and according to μ. The distribution of order types given by such a process happens to define a limit of order types ℓ_μ. Such a limit is called a *geometric limit*. The main purpose of the chapter is to study the resulting mapping $\ell \mapsto \ell_\mu$ from probability to limits of order type. We show that this function is not surjective, that is there exists limits of order type that are not geometric. Concerning the injectivity, note that every measure μ whose support is in convex position generates the same limit of order types ℓ_C that assigns 1 to order types of points in convex position and 0 to the others. However, if two measures μ and μ' give the same limit $\ell_\mu = \ell_{\mu'}$ then they are equal up to a projective transformation whenever the support of these measures are large enough. We show two results in this direction. The first one show that this property holds if the support of μ contains an open ball. The second one shows that it holds if both μ and μ' have Hausdorff dimension strictly more than one.

A second chapter is focused on one-sided property testing of graphs. A tester is a randomized algorithm for distinguishing between objects satisfying a property from those that are at some distance at least ϵ from having it by mean of the edition distance. This gives very efficient algorithms, and in particular algorithms whose complexity does not depend on the size of the input but only on the parameter ϵ. For graphs, it has been shown by Alon and Shapira that every hereditary property has such a tester. However, the bound of the complexity of this tester, though it is independent from the size of the input, is a tower of exponential which makes it unusable in practice. We contribute to the following question: which classes of graphs have a one-sided property tester with a number of queries that is a polynomial in $1/\epsilon$? We give a proof that the class of interval graphs has such a tester.

A third chapter presents a program developed during my PhD that implements flag algebras. Flag algebras are a tool introduced in 2007 by Razborov closely related to limits of dense graphs, intended to solve problems of extremal combinatorics with the help of a computer. The theory of flag algebras applies to every combinatorial object with a notion of *induced substructure* such as for instance graphs, hypergraphs, permutations, or order types. The chosen object (the graph, the hypergraph, the permutation) is called a *flag*. The definition of flags is designed so that following notion of density exists: the *density* of F in G is the probability that a random induced substructure of G of same size as H is equal to H. The purpose of flag algebras are used to to show asymptotic results on these densities, in particular by encoding some properties of densities, such as relations coming from double-counting, in an algebra constituted of linear combinations of flags endowed with a product. The main feature of flag algebras is the semi-definite method, that consist in translating the search for the extremal value of a density using a costumed list of inequalities plus so-called Cauchy-Schwartz inequalities into semi-definite program. Such a problem can then be solved using an external solver, such as the solver `csdp`, to obtain a bound on the considered density. The program I wrote during my PhD can be used as a library that can compute flag algebras, manipulate inequalities on densities and encode the optimization of some parameter in a semi-definite positive instance that can then be given to a solver. In a joint work with Jean-Sébastien Sereni and Jan Volec, my program was used to give a new bound for the triangle case of the Caccetta-Häggkvist conjecture.

Chapter 1

Limits of order types

This section follows a joint work with Xavier Goaoc, Alfredo Hubbard, Jean-Sébastien Sereni and Jan Volec published in [21]. Part 1.5 has been written with Jean-Sébastien Sereni. Part 1.10.1 is a joint work with Xavier Goaoc and Alfredo Hubbard.

1.1 Chirotopes

The orientation of a triangle abc is *clockwise* if a is on the right of the line bc oriented from b to c. Similarly, the orientation of abc is *counter-clockwise* if a is on the left of the line bc oriented from b to c.

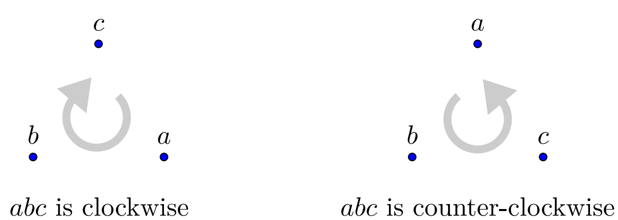


Figure 1.1 – Orientation of triangles.

Notice that this orientation depends on the order of the vertices of the triangle. Indeed, if abc is a clockwise triangle, then the triangle bac is counter-clockwise.

Algebraically, the orientation of a triangle can be computed using the determinant. For a vector $a = (x_a, y_a)$ of $\mathbf{R}^2$, let $a \oplus 1$ denote the vector $(x_a, y_a, 1) \in \mathbf{R}^3$. Consider three points a, b and c of $\mathbf{R}^2$. The value $\mathcal{A}(a, b, c) = \det(a \oplus 1, b \oplus 1, c \oplus 1)$ is positive whenever abc is a counter-clockwise triangle and negative whenever abc is a clockwise triangle. If a, b and c are aligned, then $\mathcal{A}(a, b, c) = 0$. Let sign : $\mathbf{R} \to \{-1, 0, 1\}$ be the sign function, that is the function that maps

$x \in \mathbf{R}$ to 1 if $x > 0$, to -1 if $x < 0$ and to 0 if $x = 0$. Given a subset $A \subseteq \mathbf{R}^2$ of points of the plane, the *chirotope* of A is the function

$$\chi_A : \left\{ \begin{array}{ccc} A^3 & \to & \{-1, 0, 1\} \\ (a, b, c) & \mapsto & \operatorname{sign}(\det(a \oplus 1, b \oplus 1, c \oplus 1)). \end{array} \right.$$

By the remark above, the chirotope of A encodes the orientation of every triangle of A.

Note that if the elements of the triple $(a, b, c) \in A^3$ are not pairwise distinct, for instance if $a = b$, then $\chi_A(a, b, c) = 0$. The chirotope χ_A is *degenerated* if $\chi_A(a, b, c) = 0$ for a triple $(a, b, c) \in A^3$ of pairwise distinct elements of A. From a geometrical perspective, the chirotope χ_A is degenerated if and only if A contains an aligned triple.

A chirotope $\chi_A : A^3 \to \{-1, 0, 1\}$ is anti-symmetrical, that is, for every $(a, b, c) \in A^3$ and every permutation σ of $\{a, b, c\}$, the value $\chi_A(\sigma(a), \sigma(b), \sigma(c))$ equals $\chi_A(a, b, c)$ times the signature of the permutation σ. In some cases, we would like to define a notion of triangle orientation on a set S that may not have prior natural geometrical structure. A *chirotope* then denotes any anti-symmetrical function χ from S^3 to $\{-1, 0, 1\}$. When there may be an ambiguity, we will call the function χ_A described above on a subset A of $\mathbf{R}^2$ the *standard chirotope* of A.

Deciding if such a function $\chi_A : A^3 \to \{0, 1\}$ is the standard chirotope of a finite subset of $\mathbf{R}^2$ is known to be NP-hard [39].

1.2 Order types

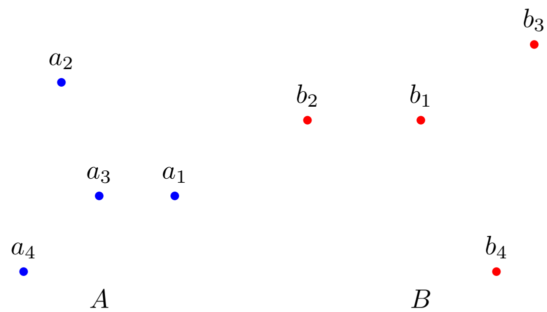
Similarly to graphs, two chirotopes $\chi_1 : A^3 \to \{-1, 0, 1\}$ and $\chi_2 : B^3 \to \{-1, 0, 1\}$ are *isomorphic* if there is an invertible map $\pi : A \to B$ such that $\chi_1(a, b, c) = \chi_2(\pi(a), \pi(b), \pi(c))$ for every $(a, b, c) \in A^3$. From a geometrical point of view, the standard chirotopes χ_A and χ_B with $A, B \subseteq \mathbf{R}^2$ are isomorphic if there is a bijection $\pi : A \to B$ that preserve the orientations of the triangles and the alignments.

An *order type* is an isomorphism equivalence class of finite chirotopes. The set of order types of size $n \in \mathbf{N}$ is $\mathcal{O}_n$. The set of all order types is $\mathcal{O}$. A *realization* of an order type ω is a (finite) point set $A \subseteq \mathbf{R}^2$ such that the standard chirotope χ_A belongs to the class ω. As a shortcut, *the order type of* $A \subseteq \mathbf{R}^2$ denotes the equivalence class of the standard chirotope χ_A.

We extend to order types any notion of a point set that does not depend on the choice of a realization. For instance, the *size* of an order type ω is the number of points in a set of points representing ω.

Similarly, an order type ω is degenerated if one representation (or equivalently, *every* representation) of ω is degenerated, that is, it contains an aligned triple. All order types we consider are always supposed to be non-degenerated.

Example 1.1. Consider the points of the plane $a_1 = (2, 0)$, $a_2 = (-1, 3)$, $a_3 = (0, 0)$, $a_1 = (-2, -2)$.



The triangles $a_2a_3a_4$, $a_1a_3a_2$, $a_1a_4a_3$ and $a_1a_4a_2$ are clockwise. Therefore the standard chirotope χ_A of the set $A = \{a_1, a_2, a_3, a_4\}$ is the function $\chi_A : A^3 \to \{0, 1\}$ with $\chi_A(a_2, a_3, a_4)$, $\chi_A(a_4, a_2, a_3)$, $\chi_A(a_3, a_4, a_2)$, $\chi_A(a_1, a_3, a_2)$, $\chi_A(a_2, a_1, a_3)$, $\chi_A(a_3, a_2, a_1)$, $\chi_A(a_1, a_4, a_3)$, $\chi_A(a_3, a_1, a_4)$, $\chi_A(a_4, a_3, a_1)$, $\chi_A(a_1, a_4, a_2)$, $\chi_A(a_2, a_1, a_4)$ and $\chi_A(a_4, a_2, a_1)$ equal to 1, with values on triples in reverse direction, namely $\chi_A(a_4, a_3, a_1)$, $\chi_A(a_3, a_2, a_4)$, $\chi_A(a_2, a_4, a_3)$, $\chi_A(a_2, a_3, a_1)$, etc. equal to -1. and with value zero if the same point is twice in the argument, i.e. $0 = \chi_A(a_1, a_1, a_1) = \chi_A(a_1, a_1, a_2) = \chi_A(a_1, a_2, a_1) = \chi_A(a_2, a_1, a_1) = \chi_A(a_2, a_2, a_3) = \dots$, and so on.

Set $B = \{b_1, b_2, b_3, b_4\}$ with $b_1 = (1, 3)$, $b_2 = (-2, 3)$, $b_3 = (4, 5)$ and $b_4 = (3, -1)$. The function

$$\pi : \left\{ \begin{array}{ccc} A & \to & B \\ a_1 & \mapsto & b_3 \\ a_2 & \mapsto & b_2 \\ a_3 & \mapsto & b_1 \\ a_4 & \mapsto & b_4 \end{array} \right.$$

is bijective and preserves the chirotopes, so χ_A and χ_B are equivalent, that is, they have same order type.

Example 1.2 (Small order types)**.** For each $n \in \mathbf{N}$, the number of order types of size n is finite. If $1 \leq n \leq 3$, there is a unique order type of size n. In Figure 1.2 is found a realization for each order type of size between 4 and 6. Order types of size 6 already give a hint on the complexity and variety of order types. There are 135 order types of size 7. The asymptotic number of different order types of size n is $2^{4n \log n(1+o(1))}$ [3].

Oswin Aichholzer made available a comprehensive list [1] of all the order types of size up to 10, based on his works with Aurenhammer and Krasser [1, 2]. The drawings of order types of size 6 in Figure 1.2 was made using this list.

Example 1.3 (Points in convex position)**.** A set of points P is in *convex position* if for every point $x \in P$, the convex hull of $P \setminus \{x\}$ is different from the convex hull of P. This notion only depends of the order type of P. This allows us to say that an order type ω is in *convex position* if a (or equivalently, every) realization of ω is in convex position. For each size $n \geq 3$, there is exactly one order type

[1] http://www.ist.tugraz.at/aichholzer/research/rp/triangulations/ordertypes/

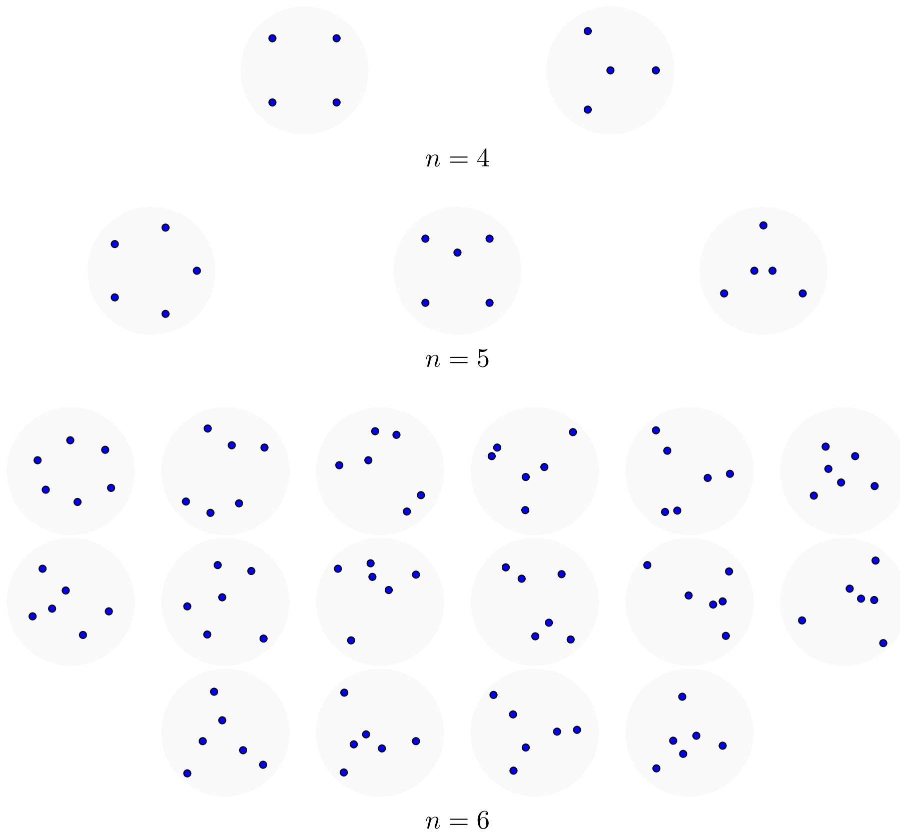


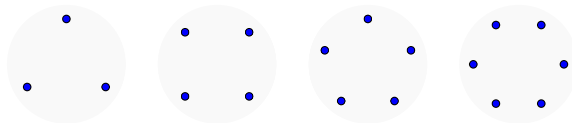
Figure 1.2

in convex position. Let C_n denote this order type. Figure 1.3 represents C_n for $n \in \{3, \ldots, 6\}$.

The set of families of points that realize a fixed order type ω is the *realization space* of ω. A theorem of Mnëv [39] proves that realization spaces can be arbitrarily complex in the sense that for every semi-algebraic space is homotope to the realization space of an order type. The most simple non-trivial case of this theorem is the existence of an order type ω and two point sets realizing ω such that one cannot be continuously deformed into the other while preserving the order type ω for every intermediate point set.

1.3 Sampling order types

A method to generate an order type of size n is the following. Take n (independent) random points, uniformly in the unit square. As every order type ω has



C_3 C_4 C_5 C_6

Figure 1.3 – Order types in convex position.

a representation in the unit square and such a representation is stable under a small perturbation, i.e. the realization space of ω is an open set, every realizable order type of size n will be generated with a positive probability.

This construction can be done using any probability measure μ on $\mathbf{R}^2$. For such a probability measure μ, let $\mathbf{H}(n, \mu)$ be the random order type given by a tuple of n μ^n-random points.

Definition 1.1. Let $\ell_\mu : \mathcal{O} \to \mathbf{R}$ be the function that assigns to an order type ω the probability that $\mathbf{H}(n, \mu) = \omega$.

A necessary and sufficient condition to ensure that $\mathbf{H}(n, \mu)$ is almost surely not degenerated is that μ charges no line, i.e. every line of $\mathbf{R}^2$ has zero μ-measure. In the following, we consider only probability measures that charge no line.

If ω_1 and ω_2 are order types of respective size n_1 and n_2, the *density* $p(\omega_1, \omega_2)$ of ω_1 in ω_2 is the probability on a random set P_1 chosen uniformly among the $\binom{n_2}{n_1}$ subsets of a set of points S_2 with order type ω_2 that P_1 has order type ω_1. This value does not depend on the choice of S_2. Limits of order types are then defined as follows.

Definition 1.2. A function $\ell : \mathcal{O} \to [0, 1]$ is a *limit of order types* if there is a sequence of order types $(\omega_n)_{n \in \mathbf{N}}$ such that $|\omega_n| \to \infty$ and $\lim_{n\to\infty} p(\omega, \omega_n) = \ell(\omega)$ for every $\omega \in \mathcal{O}$.

This definition of limits is a particular case of a general definition of limits of combinatorial structures. The general framework is discussed in Chapter 3.

If μ is a probability measure on $\mathbf{R}^2$ that charges no line, the function ℓ_μ is a limit of order types. To see this, we apply a general criterion given by Theorem 3.3 in [40] (See also [33]): Let $n_1, n_2 \in \mathbf{N}$ and $\omega_1 \in \mathcal{O}_{n_1}$, $\omega_2 \in \mathcal{O}_{n_2}$ and $\omega \in \mathcal{O}_{n_1+n_2}$ be order types and choose a point set S with order type ω. Let $p(\omega_1, \omega_2; \omega)$ be the probability, for a partition $S = S_1 \cup S_2$ chosen uniformly at random among the $\binom{n_1+n_2}{n_1}$ partitions with $|S_1| = n_1$ and $|S_2| = n_2$, that S_1 has order type ω_1 and S_2 has order type ω_2. Note that $p(\omega_1, \omega_2; \omega)$ does not depend on the choice of S. Then the criterion is the following. A function $\ell : \mathcal{O} \to \mathbf{R}^+$ is a limit of order types if and only if for every $n \in \mathbf{N}$,

$$\sum_{\omega \in \mathcal{O}_n} \ell(\omega) = 1 \tag{1.1}$$

and for every $n_1, n_2 \in \mathbf{N}$, $\omega_1 \in \mathcal{O}_{n_1}$ and $\omega_2 \in \mathcal{O}_{n_2}$, the following holds.

$$\ell(\omega_1)\ell(\omega_2) = \sum_{\omega \in \mathcal{O}_{n_1+n_2}} p(\omega_1, \omega_2; \omega)\ell(\omega). \tag{1.2}$$

This condition is equivalent to the fact that ℓ is an algebra homomorphism of from the flag algebra, defined in Chapter 3, to $\mathbf{R}$.

Equality (1.2) for $\ell = \ell_\mu$ is the formula of total probability applied to the possible order types of a μ^n-random set.

Let us show Equality (1.2) for ℓ_μ. Consider a $\mu^{n_1+n_2}$-random subset S of $\mathbf{R}^2$. Let $S = S_1 \cup S_2$ be a random partition of S chosen uniformly among those partitions that satisfy $|S_1| = n_1$ and $|S_2| = n_2$. For three order types $\omega_1 \in \mathcal{O}_{n_1}$, $\omega_2 \in \mathcal{O}_{n_2}$ and $\omega \in \mathcal{O}_{n_1+n_2}$, it follows from the definitions that the conditional property $\mathbf{P}(S_1 \in \omega_1 \text{ and } S_2 \in \omega_2 | S \in \omega)$ is equal to the split density $p(\omega_1, \omega_2; \omega)$ and the probability that $S \in \omega$ is $\ell_\mu(\omega)$. As a consequence, the right side of the equation expresses the probability $\mathbf{P}(S_1 \in \omega_1 \text{ and } S_2 \in \omega_2)$ using conditioning on the outcome of the order type of S. Since S_1 and S_2 are independent random sets with respective distribution μ^{n_1} and μ^{n_2}, the right side is equal to $\mathbf{P}(S_1 \in \omega_1)\mathbf{P}(S_2 \in \omega_2) = \ell_\mu(\omega_1)\ell_\mu(\omega_2)$, which proves (1.2).

Definition 1.3. A limit of order types ℓ for which there exists μ such that $\ell = \ell_\mu$ is a *geometric limit*.

Example 1.4 (Limit of order types in convex position)**.** Recall for $n \in \mathbf{N}$, the order type $C_n \in \mathcal{O}_n$ is the order type in convex position, defined in Example 1.3.

It is easy to see that the density $p(\omega, C_n)$ of any $\omega \in \mathcal{O}_m$ with $m \leq n$ is 1 if $\omega = C_m$ and 0 otherwise. As a consequence, the sequence of order types $(C_n)_{n\in\mathbf{N}}$ converges, and the limit of this sequence is the function

$$\ell_\diamond : \begin{cases} \mathcal{O} & \to & \mathbf{R}^+ \\ C_n & \mapsto & 1 \\ \omega & \mapsto & 0 \quad \text{if } \omega \neq C_n \text{ where } n = |\omega| \end{cases} \tag{1.3}$$

If ν is a probability measure on R^n endowed with the Borel algebra, then the *support* $\operatorname{supp} \nu$ of ν is the collection of all points $x \in R^n$ such that every open neighborhood of x (with respect to the Euclidean topology) has positive measure. It is known (and standard) that for such a measure every set disjoint from the support has Lebesgue measure 0.

Example 1.5. Let $\mu_\circ$ be the uniform probability measure on the unit circle S_1, that is the probability given by the 1-dimensional Lebesgue measure. Since every subset of S_1 is in convex position, the limit $\ell_{\mu_\circ}$ equals the limit of order types in convex position $\ell_\diamond$ of example 1.4. The limit $\ell_\diamond$ is therefore a geometric limit.

More generally, if $\mathcal{C}$ is a strictly convex loop in $\mathbf{R}^2$ (that is, $\mathcal{C}$ is equal to its convex hull) and μ is any probability measure whose support is contained in $\mathcal{C}$, and that charges no line, then it holds that $\ell_\mu = \ell_\diamond$. This shows that the limit $\ell_\diamond$ has a wide family of representations as a geometric limit.

1.4 Geometric limits

Consider an oracle that takes input a size n and generates an order type of size n chosen in some random way. For $\omega \in \mathcal{O}$, write $\ell(\omega)$ the probability for this oracle to output ω on the input $n = |\omega|$. If $n_1, n_2 \in \mathbf{N}$, we consider two experiments. On one side, let ω be a random order type of size n generated by the above procedure and let ω_1 and ω_2 be the order types generated by splitting the vertices of ω into two random parts of respective size n_1 and n_2, chosen uniformly among the $\binom{n_1+n_2}{n_1}$ ways to do it. On the other side let ω_1' and ω_2' be two order types of respective size n_1 and n_2 directly generated by two independent runs of our procedure. A natural hypothesis on a random generator is to ask that the random variables (ω_1, ω_2) and (ω_1', ω_2') are identically distributed. In any way, ℓ is a limit of order types if and only if this hypothesis holds for every $n_1, n_2 \in \mathbf{N}$.

Another natural question is to know whether every such process, and further any limit, can be expressed as a geometric limit. The following theorem provides a negative answer to this question.

Theorem 1.1. *There is a limit ℓ of order types such that $\ell \neq \ell_\mu$ for every probability measure μ on the plane.*

We construct such a limit in Example 1.6.

Example 1.6. Given a positive real r, let $\mathcal{C}_r$ be the circle of radius r centered on the origin. Let ν_r be the uniform probability distribution on the disk $\mathcal{C}_r$, that is proportional to the length measure. For $\epsilon > 0$, define the probability measure μ_ϵ as the mixture $\mu_\epsilon = \frac{1}{2}\nu_1 + \frac{1}{2}\nu_\epsilon$. Note that $\operatorname{supp}\mu_\epsilon = \mathcal{C}_1 \cup \mathcal{C}_\epsilon$.

By the compacity of the set of limits of order types, there is a sequence of positive numbers $(\epsilon_n)_{n\in\mathbf{N}}$ that tends to 0 such that the sequence (of limits) $\ell_{\mu_{\epsilon_n}}$ tends to a limit $\ell_\odot$. By somehow making $\ell_\odot$ explicit, we will actually show that this limit is unique and further that ℓ_{μ_ϵ} tends to $\ell_\odot$ when ϵ tends to 0.

The limit $\ell_\odot$ appears to have no geometric representation and will therefore serve as counter-example in the proof of Theorem 1.1. Before discussing this theorem, we show an other way to represent $\ell_\odot$.

The idea is to construct a probability space by gluing a circle of "infinitesimally small" radius in $\mathbf{R}^2$.

More formally, let $\mathbf{R}(X)$ denote the set of algebraic fractions with real coefficients, that is the set of quotients of the form P/Q, where P and Q are polynomials with real coefficients and $Q \neq 0$. This set $\mathbf{R}(X)$ is endowed by a field structure. We will represent $\ell_\odot$ using a measure on a non-standard plane $\mathbf{R}(X)^2$. In this space, X shall be thought as a positive number smaller than all positive (standard) reals.

As in Example 1.6, let $\mathcal{C}_1$ be the unit disk of $\mathbf{R}^2$ centered on the origin, that is $D = \left\{ (a,b) \;\middle|\; a^2 + b^2 = 1, a \in \mathbf{R} \text{ and } b \in \mathbf{R} \right\}$ and let ν_1 be the uniform probability measure on $\mathcal{C}_1$ (i.e. the one-dimensional Lebesgue measure, which corresponds to the length). Note that $\mathcal{C}_1$ is a subset of $\mathbf{R}(X)^2$ and further ν_1 induces a probability measure on $\mathbf{R}(X)^2$.

Let $\mathcal{C}_X \subseteq \mathbf{R}(X)^2$ be the *circle of radius* X centered in the origin, that is the set $\{ (aX, bX) \mid a^2 + b^2 = 1 \text{ and } (a,b) \in \mathbf{R}^2 \}$. Let ν_X be the probability measure on C_X defined by $\nu_X(A) = \nu_1^1(\{ (a,b) \mid (aX, bX) \in A \})$ whenever $A \subseteq \mathbf{R}^2$. Similarly to Example 1.6, let μ_X be the probability measure on $\mathbf{R}(X)$ defined by $\mu_X = \frac{1}{2}\nu_1 + \frac{1}{2}\nu_2^X$.

We now define a chirotope $\chi : (\mathbf{R}(X)^2)^3 \to \{-1, 0, 1\}$ on the plane $\mathbf{R}(X)^2$ that extend the standard chirotope of $\mathbf{R}$. For three points $a = (A_1, A_2)$, $b = (B_1, B_2)$ and $c = (C_1, C_2)$ of $\mathbf{R}(X)^2$, the determinant $\mathcal{A}(a, b, c)$ of the matrix

$$\begin{pmatrix} 1 & A_1 & A_2 \\ 1 & B_1 & B_2 \\ 1 & C_1 & C_2 \end{pmatrix}$$

is an algebraic fraction.

The *sign* of an algebraic fraction $R = P/Q$ is the sign of $R(\epsilon)$ when $\epsilon > 0$ is small enough, that is

$$\operatorname{sign} R := \lim_{\substack{\epsilon \to 0 \\ \epsilon > 0}} \operatorname{sign} R(\epsilon).$$

This limit exists because the sign of $R(\epsilon)$ is eventually constant when ϵ tends to 0. Indeed, this is clear if $R = 0$, and otherwise it suffices to take ϵ smaller than every positive roots of P and Q, which are finitely many. Note that if $R = P$ is a non-zero polynomial, then the sign of R is the sign of the first (i.e. with the lower power) non-zero coefficient of P.

This makes it possible to define the orientation of a triangle $abc \in (\mathbf{R}^2(X))^3$ as

$$\chi(a, b, c) := \operatorname{sign}(\mathcal{A}(a, b, c)).$$

By the property stated above, the orientation of abc in $\mathbf{R}^2(X)$ is the orientation of the triangle $a(\epsilon)b(\epsilon)c(\epsilon)$ of the standard plane $\mathbf{R}^2$ when ϵ is small enough. Moreover, this definition is consistent on $(\mathbf{R}^2)^3$ with the definition of the chirotope on $\mathbf{R}^2$. If $A = \{(P_1, Q_1), \ldots, (P_n, Q_n)\}$ is a finite subset of $\mathbf{R}(X)^2$ then the chirotope χ restricted to A^3 is a realizable order type. It is for instance realizable as a standard chirotope χ_B in the plane (i.e. with $B \subseteq \mathbf{R}^2$), with $B = \{(P_1(\epsilon), Q_1(\epsilon)), \ldots, (P_n(\epsilon), Q_n(\epsilon))\}$, where $\epsilon > 0$ is small enough so that $\operatorname{sign} P_i(\epsilon) = \operatorname{sign} P_i$ and $\operatorname{sign} Q_i(\epsilon) = \operatorname{sign} Q_i$ for $i \in \{1, \ldots, n\}$.

We now deduce an explicit definition of $\ell_\odot$ from the triple $(\mu_X, \mathcal{C}_1 \cup \mathcal{C}_X, \chi)$, constituted by the measure μ_X, its support $\mathcal{C}_1 \cup \mathcal{C}_X$ and the chirotope χ on $\mathcal{C}_1 \cup \mathcal{C}_X \subseteq (R(X)^2)^3$. For an order type $\omega \in \mathcal{O}$, write $\ell_{\mu_X}(\omega)$ the probability that a $(\mu_X)^{|\omega|}$-random set $S \subseteq \mathbf{R}(X)^2$ induces an order type equivalent to ω. We claim that $\forall \omega \in \mathcal{O}, \ell_{\mu_X}(\omega) = \ell_\odot(\omega)$.

This general construction, where we use a chirotope-like function and a probability measure in a set different from $\mathbf{R}^2$ will be generalized as a kernel in the next section.

A heuristic proof of Theorem 1.1. We aim to show that $\ell = \ell_\odot$ is a counter-example. Suppose otherwise that $\ell_\odot = \ell_\mu$ for some probability measure μ on $\mathbf{R}^2$.

(1) We would like to assume that the support of μ is made of two parts $\mathcal{C}_1'$ and $\mathcal{C}_X'$ that mimic the behaviors of $\mathcal{C}_1$ and $\mathcal{C}_X$ respectively. In particular, a triangle of $\mathcal{C}_1$ that contains the origin should correspond to a triangle of $\mathcal{C}_1'$ that contains $\mathcal{C}_X'$.
(2) More precisely, we would like that $\mathcal{C}_1'$ and $\mathcal{C}_X'$ have the same topological properties as $\mathcal{C}_1$ and $\mathcal{C}_X$, in particular that $\mathcal{C}_1$ is a loop around $\mathcal{C}_X$.
Assuming this, we now construct triangles of $\mathcal{C}_1'$ of arbitrarily small area that contains $\mathcal{C}_X'$, from which we deduce that $\mathcal{C}_X'$ is concentrated on a line, which yields a contradiction. □

The theory exposed in the following chapter will make it possible to properly prove Theorem 1.1. Concerning assumption (1), one may wonder if measures realizing $\ell_\odot$ may look completely different from the measure μ_X, while generating the same order types densities.

Proposition 1.8 is an adaptation of a result from the theory of dense graphs that answers this question by giving a bijection ρ between (almost all) the *points* of $\operatorname{supp}\mu$ and $\operatorname{supp}\mu_X$ that preserves the chirotopes almost everywhere. This justifies (an adapted version of) hypothesis (1). An application of Corollary 1.23 is then used to justify that there is a homeomorphism from $\mathcal{C}_1 \setminus N_1$ to $\mathcal{C}_2 \setminus N_2$, where N_1 and N_2 are countable set. In particular, $\mathcal{C}_1'$ may not be a loop, so hypothesis (2) is not exact. However, the points of discontinuity are countable, which is enough to make the argument work.

A complete proof of Theorem 1.1 using Corollary 1.23 is given in section 1.8.

1.5 Kernels

In this section, we adapt the notion of kernels to three-variables functions. This notion generalizes the notion of geometric limit by first allowing probability measures of any space J instead of probability measures on $\mathbf{R}^2$ and secondly by using any real-valued function of J^3 instead of the chirotope of the plane. The main result of this section is Proposition 1.8, which shows the equivalence for two kernels to have same sampling distribution and to be isomorphic. These results are direct adaptations of similar results on two-dimensional kernels and graphon. This part is written to justify as fast as possible the results we need for kernels. See the book of Lázló Lovász [32] for (a lot) more details on two-dimensional kernels.

If (J_1, μ_1) and (J_2, μ_2) are two probability spaces, a map $\rho\colon J_1 \to J_2$ is *measure preserving* if for every μ_2-measurable set $A' \subseteq J_2$, the set $A = \rho^{-1}(A')$ is μ_1-measurable and $\mu_1(A) = \mu_2(A')$. A *kernel* is a triple (J, μ, W) where (J, μ) is a probability space and W is a measurable map from J^3 to $\mathbf{R}$. Two points x and x' in J are *twins for* (J, μ, W) if $W(x, y, z) = W(x', y, z)$ for μ^2-almost every pair $(y, z) \in J^2$. The kernel (J, μ, W) is *twin-free* if it admits no twins. For the sake of readability, we may write only W to denote kernel (J, μ, W) when there is no ambiguity about the set and the measure involved.

If (J,μ,W) is a kernel and $\phi : J_0 \to J$ is a measure-preserving map, let W^ϕ be the function from J_0^3 to $\mathbf{R}$ defined by $W^\phi(x,y,z) = W(\phi(x),\phi(y),\phi(z))$ for every $(x,y,z) \in J_0^3$.

The L_1-*norm* of a kernel (J,μ,W) is the value

$$\|W\|_1 = \int_{J^3} |W(x,y,z)| \,\mathrm{d}\mu(x)\,\mathrm{d}\mu(y)\,\mathrm{d}\mu(z)$$

If (J_1,μ_1,W_1) and (J_2,μ_2,W_2) are two kernels, the *distance* between them is

$$\mathrm{d}(W_1,W_2) = \inf \left\| W_1^{\phi_1} - W_2^{\phi_2} \right\|_1$$

where the infimum is taken over every choice of a probability space (J,μ) and a pair of measure-preserving maps $\phi_i\colon J \to J_i$ for $i \in \{1,2\}$. In this last definition, we may assume that μ is a coupling measure on $J = J_1 \times J_2$ and that ϕ_i is the natural projection of J on J_i for $i \in \{1,2\}$. The L_∞-*norm* of (J,μ,W) is the value $\|W\|_\infty = \sup_{J^3} |W|$.

If (J,μ,W) is a kernel, the *neighborhood pseudo-distance of* (J,μ,W) is the function $d_W\colon J \times J \to \mathbf{R}^+$ given by

$$\begin{aligned} \forall (x,x') \in J^2, \qquad d_W(x,x') &= \|W(x,\cdot,\cdot) - W(x',\cdot,\cdot)\|_1 \\ &= \int_{J^2} |W(x,y,z) - W(x',y,z)| \,\mathrm{d}\mu(y)\,\mathrm{d}\mu(z). \end{aligned}$$

It is straightforward to check that d_W is reflexive and satisfies the triangular inequality. The measure μ has *full support* if every nonempty open set of (J,d_W) has positive μ-measure. The kernel (J,μ,W) is *pure* if the three following properties are verified.

1. The map d_W is a metric on J;
2. the metric space (J,d_W) is complete and separable; and
3. the measure μ has full support.

The aforementioned properties of d_W imply that the first condition boils down to asking that $x = x'$ whenever $d_W(x,x') = 0$ for every $(x,x') \in J^2$, that is, (J,W,μ) is twin-free.

If n is an integer, let $[n]$ denote the set $\{1,\ldots,n\}$. Given a probability space (J,μ), a *step function* on (J,μ) is a kernel (J,μ,U) for which there is a partition $\bigcup_{i=1}^m P_i$ of J such that the function U is constant on $P_i \times P_j \times P_k$ for every $(i,j,k) \in [m]^3$. Since the set $U^{-1}(\{x\})$ is measurable for every $x \in \mathbf{R}$, we may always assume that P_i is measurable for $i \in [m]$. For a kernel (J,μ,W), a positive integer n and an n-tuple $S = (x_1,\ldots,x_n) \in J^n$, the *function induced by S on W* is the function $\mathbf{H}(W,S)$ from $[n]^3$ to $\mathbf{R}$ that maps (i,j,k) to $W(x_i,x_j,x_k)$ for every (i,j,k) in $[n]^3$. Note that if μ_n^u is the uniform probability measure on $[n]$ (i.e. such that $\mu_n^u(A) = |A|/n$ for every $A \subseteq [n]$), then $([n],\mu_n^u,\mathbf{H}(W,S))$ is a kernel. Moreover, since $[n]$ is a finite set, $\mathbf{H}(W,S)$ is also a step function on the probability space $([n], \mu_n^u)$ using the partition of $[n]$ into n singletons.

For a positive integer n we define the random step function $\mathbf{H}(W, n)$ as the function $\mathbf{H}(W, S)$, where S is a μ^n-random tuple. If H is a function from $[n]^3$ to $\mathbf{R}$ the *density* of H in W is the probability $t(H, W)$ that $\mathbf{H}(W, n) = H$.

Example 1.7. The triple $(\mu_X, \mathcal{C}_1 \cup \mathcal{C}_X, \chi)$ defined in Section 1.4 defines a kernel with values in $\{-1, 0, 1\}$. Recall that μ_X is a probability measure that is uniform on a circle $\mathcal{C}_1$ with radius 1 and uniform in a circle $\mathcal{C}_X$ with same center and *infinitely smaller* radius; and such that $\mu_X(\mathcal{C}_1) = \frac{1}{2} = \mu_X(\mathcal{C}_X)$.

For the sake of readability, set $J := \mathcal{C}_1 \cup \mathcal{C}_X$. Let us compute the neighborhood distance d_χ of this kernel. Note that if $a \in J$, then $\chi(a, y, z) \in \{1, -1\}$ for μ_X^2-almost every $(y, z) \in J^2$. It follows that

$$d_\chi(a, b) = 2 \cdot \left| \left\{ (y, z) \in J^2 \;\middle|\; \chi(a, y, z) \neq \chi(b, y, z) \right\} \right|$$

for every $a, b \in J$. Geometrically speaking, $d_\chi(a, b)$ is therefore (twice) the probability for a μ_X^2-random pair $(y, z) \in J^2$ that the line through y and z separates a and b.

Take $a, b \in J$ and let θ be the absolute value of the angle between $\vec{Oa}$ and $\vec{Ob}$ taken in $[-\pi, \pi]$ (so $\theta \in [0, \pi]$) where O stands for the origin $(0, 0)$. If x and y are two points of a same circle $\mathcal{C} \in \{\mathcal{C}_1, \mathcal{C}_X\}$, let $\operatorname{arc}(x, y) \subseteq \mathcal{C}$ denote the smallest path from x to y in $\mathcal{C}$, which is an arc of $\mathcal{C}$. In the following, we consider a μ_X^2-random pair (y, z).

If $a, b \in \mathcal{C}_1$, note that for every pair $(y, z) \in J$ of disjoint points of the line through y and z cross $\mathcal{C}_1$ in exactly two points, so by symmetry of J, the average number of crossings of the line through y and z with $\operatorname{arc}(a, b)$ is $2\frac{\theta}{2\pi} = \frac{\theta}{\pi}$. The line through y and z crosses $\operatorname{arc}(a, b)$ twice if and only if $y, z \in \operatorname{arc}(a, b)$, which happens with probability $(\mu_X(\operatorname{arc}(a, b)))^2 = \frac{\theta^2}{16\pi^2}$. It follows that

$$d_\chi(a, b) = 2\left(\frac{\theta}{\pi} - \frac{\theta^2}{16\pi^2}\right) = \frac{\theta}{\pi}\left(2 - \frac{\theta}{8\pi}\right).$$

Assume $a, b \in \mathcal{C}_X$. The line through y and z crosses $\mathcal{C}_X$ in two points (μ_X^2-almost) if and only (y, z) is not in $\mathcal{C}_1^2$, which happens with probability $\frac{3}{4}$. By the symmetry of J, the expected number of crossings with $\operatorname{arc}(a, b)$ is $2 \cdot \frac{3}{4} \cdot \frac{\theta}{2\pi} = \frac{3\theta}{4\pi}$. The line through y and z crosses $\operatorname{arc}(a, b)$ twice in the following cases: $(y, z) \in \operatorname{arc}(a, b)^2$; or one of y and z is in $\operatorname{arc}(a, b)$ and the other one is in the half of $\mathcal{C}_1$ above the line from a to b. This happens with probability $(\frac{\theta}{4\pi})^2 + 2 \cdot \frac{\theta}{4\pi} \cdot \frac{1}{4} = \frac{\theta}{4\pi}(\frac{\theta}{4\pi} + \frac{1}{2})$. It follows that

$$d_\chi(a, b) = 2\left(\frac{3\theta}{4\pi} - \frac{\theta}{4\pi}\left(\frac{\theta}{4\pi} + \frac{1}{2}\right)\right) = \frac{\theta}{\pi}\left(\frac{11}{8} - \frac{\theta}{8\pi}\right).$$

Assume $a \in \mathcal{C}_1$ and $b \in \mathcal{C}_X$. Assuming that (y, z), which happens with probability $\frac{1}{4}$, the line through $y \in \mathcal{C}_1$ and $z \in \mathcal{C}_1$ either separates a and b or separates $-a$ and b. By the symmetry of $\mathcal{C}_1$, it follows that the line through y and z separates a and b with probability at least $\frac{1}{4} \cdot \frac{1}{2} = \frac{1}{8}$. Consequently,

$$d_\chi(a, b) \geq 2 \cdot \frac{1}{8} = \frac{1}{4}.$$

It follows that the topology induced by d_χ on J is the natural topology of two disjoint circles of the euclidean plane. In particular, J is a compact set for d_χ. It follows that the kernel (μ_x, J, χ) is twin-free, that J is complete. Consequently, (μ_x, J, χ) is a pure kernel.

1.5.1 Weak isomorphism

The purpose of this section is to prove the following statement.

Theorem 1.2. *If (J_1, μ_1, W_1) and (J_2, μ_2, W_2) are two kernels such that $t(H, W_1) = t(H, W_2)$ for every function H, then $\mathrm{d}(W_1, W_2) = 0$.*

We start with a lemma to approximate the symmetric function of a kernel by a step function on the same probability space. A function $V \colon J^3 \to \mathbf{R}$ is *simple* if it is measurable and if its image $V(J^3)$ is finite. A simple function V can thus be expressed by a finite sum $V = \sum_{i=1}^{k} a_i \mathbf{1}_{A_i}$, where a_i is a real number and $\mathbf{1}_{A_i}$ the indicator function of a measurable set A_i for each $i \in [k]$. It is a classic property of measure theory that for every positive ε, every measurable bounded function W can be approximated by a simple function V in such a way that $\|W - V\|_\infty \leq \varepsilon$.

Lemma 1.3. *If ε is a positive number and (J, μ, W) is a kernel with $\|W\|_\infty < \infty$, then there is a step function V on (J, μ) such that $\|W - V\|_1 \leq \varepsilon$.*

Proof. The *symmetric difference* of two sets A and B is the set $A \triangle B$ of elements that are in either of A and B and not in their intersection. We need the following property coming from measure theory.

If (J, μ) is a probability space and X is a non-empty family of measurable subsets of J such that

- X generates the σ-algebra of measurable sets; and
- X is stable under finite unions and complementary operations,

then for every $\varepsilon > 0$ and every measurable set A there is $B \in X$ such that $\mu(A \triangle B) \leq \varepsilon$, where $A \triangle B$ stands for the symmetric difference of A and B, that is $A \triangle B = (A \setminus B) \cup (B \setminus A)$.

One can prove the statement above by showing that the family of sets A for which there is indeed such an element $B \in X$ contains X and is stable by taking complements and countable unions.

The result is true if $A \in X$, since it then suffices to take $B = A$. Assume that A satisfies that $\mu(A \triangle B) \leq \varepsilon$ for some $B \in X$, then the complementary $\overline{A}$ of A (in J) also satisfies $\mu(\overline{A} \triangle \overline{B}) = \mu(A \triangle B) \leq \varepsilon$ and $\overline{B}$ belongs to X since X is stable under complementary. Let $(A_i)_{i\in\mathbf{N}}$ be a countable family of measurable subsets satisfying the property and let us prove that the property holds for the set $A = \bigcup_{i\in\mathbf{N}} A_i$. For each $i \in \mathbf{N}$, we know that there is a set $B_i \in X$ such that $\mu(A_i \triangle B_i) \le \varepsilon/2^i$. Taking $S_k = \bigcup_{i=1}^{k} B_i$ for $k \in \mathbf{N}$ and $S_\infty = \bigcup_{i\in\mathbf{N}} B_i$, we have $\mu(A \triangle S_\infty) \le \sum_{i\in\mathbf{N}} \mu(A_i \triangle B_i) \le 2\varepsilon$. Note that $(S_k)_{k\in\mathbf{N}}$ is an increasing sequence of sets of X whose union is S_∞. Since μ is a probability measure, $\mu(S_\infty)$ is finite. It follows that the real number sequence $(\mu(S_k))_{k\in\mathbf{N}}$ tends to $\mu(S_\infty)$ as k tends to infinity. Let $k \in \mathbf{N}$ be an index such that $\mu(S_\infty) - \mu(S_k) \le \varepsilon$. Then $\mu(S_\infty \triangle S_k) = \mu(S_\infty) - \mu(S_k) \le \varepsilon$ because $S_k \subseteq S_\infty$. It thus follows that $\mu(A \triangle S_k) \le \mu(A \triangle S_\infty) + \mu(S_\infty \triangle S_k) \le 3\varepsilon$. This finishes the proof of the asserted property.

Let $\mathcal{B}$ be the set of boxes of the form $P_1 \times P_2 \times P_3$, where P_1, P_2 and P_3 are measurable subsets of J and let X be the set of finite unions of elements of $\mathcal{B}$. By the definition of the product of probability spaces, the σ-algebra of the probability space (J^3, μ^3) is generated by the family $\mathcal{B}$, and further by X. Moreover, the family X is trivially stable by finite unions and it is not difficult to check that X is stable by complementary as well. As a consequence, X satisfies the property above. Also note that if B is an element of X then $\mathbf{1}_B$ is a step function on (J, μ).

We are now ready to prove the statement of the lemma. Fix $\varepsilon > 0$ and let $V = \sum_{i=1}^{k} a_i \mathbf{1}_{A_i}$ be a simple function on J such that $\|W - V\|_\infty \le \varepsilon/2$. By the property above, for each $i \in [k]$, there is a set $B_i \in X$ such that $\mu(A_i \triangle B_i) \le \frac{\varepsilon}{2k|a_i|}$. The function $U = \sum_{i=1}^{k} a_i \mathbf{1}_{B_i}$ is a step function as a linear combination of step functions. Moreover,

$$\|U - V\|_1 \le \sum_{i=1}^{k} |a_i| \cdot \|\mathbf{1}_{A_i} - \mathbf{1}_{B_i}\|_1 = \sum_{i=1}^{k} |a_i| \cdot \mu(A_i \triangle B_i) \le \frac{\varepsilon}{2}.$$

Consequently, $\|W - U\|_1 \le \|W - V\|_1 + \|V - U\|_1 \le \varepsilon$, which finishes the proof. □

We now bound the difference between the norm of the symmetric function W of a kernel and the expected value of the norm of the random step function $\mathbf{H}(W, n)$.

Lemma 1.4. *Let (J, μ, W) be a kernel and n a positive integer, then*

$$|\mathbf{E}(\|\mathbf{H}(W, n)\|_1) - \|W\|_1| \le \frac{3}{n} \|W\|_\infty .$$

Proof. Let $S = (x_1, \ldots, x_n)$ be a μ^n-random tuple of J^n. The expected value E of $\|\mathbf{H}(W, S)\|_1$ is equal to

$$\int_{J^n} \frac{1}{n^3} \sum_{1\le i,j,k\le n} |W(x_i, x_j, x_k)| \, \mathrm{d}\mu^n(S) = \frac{1}{n^3} \sum_{1\le i,j,k\le n} I_{i,j,k},$$

where $I_{i,j,k}$ is the integral $\int_{J^n} |W(x_i, x_j, x_k)| \,\mathrm{d}\mu^n(S)$. First note that $0 \le I_{i,j,k} \le \|W\|_\infty$ for every $(i,j,k) \in [n]^3$. If moreover i, j and k are pairwise different, then $I_{i,j,k} = \|W\|_1$. The number triples of $[n]^3$ whose elements are pairwise different is greater than $n^3 - 3n^2$. It follows that $(n^3 - 3n^2)\|W\|_1 \le n^3 E \le (n^3 - 3n^2)\|W\|_1 + 3n^2 \|W\|_\infty$, and further $-\frac{3}{n}\|W\|_1 \le E - \|W\|_1 \le \frac{3}{n}\|W\|_\infty$, which yields the conclusion since $\|W\|_1 \le \|W\|_\infty$. ☐

An *atom* of a measure μ is a μ-measurable set A with $\mu(A) > 0$ and such that every μ-measurable subset of A has μ-measure $\mu(A)$ or 0. All our applications concern Borel measures. In this case, an atom is precisely a singleton with positive measure up to removing a set of measure zero. A measure is *atom-free* if it admits no atom. In the next proof we need the following technical result from measure theory.

Lemma 1.5 (Sikorski [42]). *Let μ be an atom-free measure on a set J and let $A \subseteq J$ be a measurable set with finite μ-measure. Then for every every non-negative number $x \le \mu(A)$, there is a measurable subset $B \subseteq A$ with $\mu(B) = x$.*

Proof. If y is a real number and $B_1, B_2 \subseteq J_0$ are measurable sets satisfying $\mu(B_1) \le y \le \mu(B_2)$, define

$$\alpha(B_1, B_2, y) := \sup\{\, \mu(C) \mid \mu(C) \le y, B_1 \subseteq C \subseteq B_2 \,\}$$

where only measurable sets C are considered.

We first prove that for every measurable B and y with $y \le \mu(B) < \infty$, there exists a measurable set $C \subseteq B$ satisfying $\alpha(C, B, y) = \mu(C)$. To see this, fix a sequence $(\epsilon_n)_{n\in\mathbf{N}}$ of positive numbers that tends to 0. Then define an increasing sequence of sets $(C_n)_{n\in\mathbf{N}}$ satisfying $\mu(C_n) \le y$ for every $n \in \mathbf{N}$ as follows. Set $C_0 = \varnothing$ and for $i \ge 1$ choose a measurable set C_i with $C_{i-1} \subseteq C_i \subseteq B$ with $\mu(C_i) \ge \alpha(C_{i-1}, B, y) - \epsilon_i$ which exists by the definition of α. Set $C = \bigcup_{n\in\mathbf{N}} C_n$ and note that $\alpha(C, B, y) \le \alpha(C_n, B, y)$ since $C_n \subseteq C$ for every $n \in \mathbf{N}$. It follows that

$$\mu(C_n) \ge \alpha(C_{n-1}, B, y) - \epsilon_n \ge \alpha(C, B, y) - \epsilon_n$$

for every $n \ge 1$. Letting n tends to infinity gives $\mu(C) \ge \alpha(C, B, y)$. This upper bound of the supremum $\alpha(C, B, y)$ is in particular reached by C, so $\mu(C) = \alpha(C, B, y)$. This proves the claimed property.

By the property above, there exists $C_1 \subseteq A$ with $\alpha(C_1, A, x) = \mu(C_1) \le x$. A second application of this property gives a measurable set $C_2 \subseteq A \setminus C_1$ with $\alpha(C_2, A \setminus C_1, \mu(A) - x) = \mu(C_2) \le \mu(A) - x$.

We claim that the set $S := A \setminus (C_1 \cup C_2)$ is an atom unless it has measure 0. Indeed assume otherwise that S has a measurable subset T with $0 < \mu(T) < \mu(S)$. Then since $\alpha(C_1, A, x) = \mu(C_1)$, we have $\mu(C_1 \cup T) > x$. Similarly, since $\alpha(C_2, A \setminus C_1, \mu(A) - x) = \mu(C_2)$, we have $\mu(C_2 \cup T \setminus S) > \mu(A) - x$. This yields the following contradiction $\mu(A) = \mu(C_2 \cup T \setminus S) + \mu(C_1 \cup T) > \mu(A)$.

Since μ is atom-free, it follows that $\mu(T) = 0$, i.e. $C_1 \cup C_2 = A$. Further, $x \ge \mu(C_1) = \mu(A \setminus C_2) \ge x$, so $\mu(C_1) = x$. ☐

We need a last lemma before turning to the demonstration of Theorem 1.2.

Lemma 1.6. *Let (J, μ, W) be a kernel such that $\|W\|_\infty < \infty$ and consider the random kernel $([n], \mu_n^u, \mathbf{H}(W, n))$ for $n \geq 1$. The random sequence $\mathrm{d}(W, \mathbf{H}(W, n))$ tends almost surely to 0 as n goes to infinity.*

Proof. As a first step, we prove that the result holds if W is a step function. In this case, we fix a partition $P_1, \dots, P_m$ of J such that the function W is constant on $P_i \times P_j \times P_k$ whenever $(i, j, k) \in [m]^3$.

Fix a positive integer n and let $S = (x_1, \dots, x_n)$ be an n-tuple of J^n. For each $i \in [m]$, we define Q_i to be the set of indices $j \in [n]$ such that $x_j \in P_i$ and we let n_i be the size of Q_i. We assert that

$$\mathrm{d}(W, \mathbf{H}(W, S)) \leq 3 \sum_{i=1}^{m} \left| \mu(P_i) - \frac{n_i}{n} \right| \cdot \|W\|_\infty .$$

To prove this assertion, we first construct an atom-free measurable space (J_0, μ_0) and a measure-preserving map ϕ_1 from J_0 to J. To this end, it is enough to take $J_0 = J \times [0, 1]$; the measure μ_0 equal to the product measure of μ and the Lebesgue measure on $[0, 1]$; and ϕ_1 equal to the natural projection of $J_0 = J \times [0, 1]$ to J, that is, $\phi_1(x, y) = x$ for every $(x, y) \in J \times [0, 1]$. (Note that if μ is atom-free then we may take $(J_0, \mu_0) = (J, \mu)$ and ϕ_1 equal to the identity instead.)

As consequence of Lemma 1.5, if $x_1, \dots, x_\ell$, is a sequence of non-negative numbers such that $\sum_{i=1}^{\ell} x_i = \mu_0(A)$, then A admits a partition $A = \cup_{i=1}^{\ell} A'_i$ with $\mu_0(A'_i) = x_i$ for $i \in [\ell]$.

For every $i \in [m]$, we set $a_i = \min\{\mu(P_i), \mu_n^u(Q_i)\}$. Since $\mu_0(\phi_1^{-1}(P_i)) = \mu(P_i) \geq a_i$, there is a subset $R_i \subseteq \phi_1^{-1}(P_i)$ such that $\mu(R_i) = a_i$. The set $J'_0 = J_0 \setminus \bigcup_{i=1}^{m} R_i$ has measure $1 - \sum_{i=1}^{m} a_i = \sum_{i=1}^{m} (\mu_n^u(Q_i) - a_i)$. It follows that there is a partition of J'_0 into m measurable subsets $T_1, \dots, T_m$ such that $\mu_0(T_i) = \mu_n^u(Q_i) - a_i$ for every $i \in [m]$.

We now construct a measure-preserving function ϕ_2 such that $\phi_2(R_i \cup T_i) = Q_i$ for each $i \in [m]$. Recall that $Q_1, \dots, Q_m$ is a partition of $[n]$. We know that $\mu_0(R_i \cup T_i) = \mu_n^u(Q_i) = n_i/n$ for each $i \in [m]$, so there is a partition $\bigcup_{j \in Q_i} B_j^i$ of $R_i \cup T_i$ into n_i parts each of μ_0-measure $1/n$, which are indexed by the elements of Q_i. If $j \in Q_i$, then we set $\phi_2(y) = j$ for every $y \in B_j^i$. Doing this for each $i \in [m]$ naturally defines a function ϕ_2 from J_0 to $[n]$. (Indeed, for every $x \in J_0$ there is a unique pair $(i, j) \in [m] \times [n]$ such that $x \in B_j^i$, and we set $\phi_2(x) = j$.) For convenience, we note that for each $j \in [n]$, there is exactly one index $i \in [m]$ such that B_j^i is defined, henceforth we abbreviate B_j^i as B_j.

The function ϕ_2 satisfies that $\phi_2(R_i \cup T_i) = Q_i$ for $i \in [m]$. Moreover, it is measure preserving since $\mu_0(\phi_2^{-1}(A)) = \sum_{j \in A} \mu_0(B_j) = |A|/n = \mu_n^u(A)$ for every $A \subseteq [n]$. To deduce our assertion, it remains to prove that

$$\left\| W^{\phi_1} - \mathbf{H}(W, S)^{\phi_2} \right\|_1 \leq \sum_{i=1}^{m} \left| \mu(P_i) - \mu_n^u(Q_i) \right| .$$

If a belongs to R_i for some $i \in [m]$, then by construction $\phi_1(a) \in P_i$ and moreover $\phi_2(a) \in Q_i$, that is, $x_{\phi_2(a)} \in P_i$. Consequently, for every $(i,j,k) \in [m]^3$, the functions W^{ϕ_1} and $\mathbf{H}(W,S)^{\phi_2}$ are equal (and constant) on $P_i \times P_j \times P_k$.

It follows that W^{ϕ_1} and $\mathbf{H}(W,S)^{\phi_2}$ are equal everywhere except on $(J_0' \times J_0 \times J_0) \cup (J_0 \times J_0' \times J_0) \cup (J_0 \times J_0 \times J_0')$, which has μ_0^3-measure at most $3\mu_0(J_0') = 3\sum_{i=1}^m (\mu_n^u(Q_i) - a_i) \le 3\sum_{i=1}^m |\mu(P_i) - \mu_n^u(Q_i)|$. This concludes the proof of our assertion.

Let S be a μ^n-random tuple. For each $i \in [m]$, the parameter n_i follows a binomial law of parameter $\mu(P_i)$. Fix $\varepsilon > 0$. By Hoeffding's inequality, the probability that $\left|\mu(P_i) - \frac{n_i}{n}\right| \ge \varepsilon$ is at most $2e^{-2\varepsilon^2 n}$. Thus the union bound yields that $\mathrm{d}(W, \mathbf{H}(W,S)) \le 3m\varepsilon \left\|W\right\|_\infty$ with probability at least $1 - 2me^{-2\varepsilon^2 n}$. This proves that $\mathrm{d}(W, \mathbf{H}(W,n))$ tends almost surely to 0 when n goes to infinity.

We now prove the statement in the general case. By Lemma 1.3, there is a step function V on (J,μ) such that $\left\|W - V\right\|_1 \le \varepsilon$. For every positive integer n, one can couple the random variables $\mathbf{H}(V,n)$ and $\mathbf{H}(W,n)$ such that

$$\mathbf{E}(\left\|\mathbf{H}(V,n) - \mathbf{H}(W,n)\right\|_1) \le \frac{3}{n} \left\|V - W\right\|_\infty . \tag{1.4}$$

Indeed, let $S = (x_1, \ldots, x_n)$ be a single μ^n-random tuple of J^n, and note that $\mathbf{H}(V,S) - \mathbf{H}(W,S) = \mathbf{H}(V - W, S)$. Lemma 1.4 applied to $V - W$ yields Equation (1.4).

Since V is a step function, we know from the first part of the proof that as n goes to infinity, the random sequence $\mathrm{d}(V, \mathbf{H}(V,n))_n$ tends almost surely to 0. Let N be large enough to ensure that $\mathbf{E}(\mathrm{d}(V, \mathbf{H}(V,n))) \le \varepsilon$ and that $\frac{3}{n}\left\|V - W\right\|_\infty \le \varepsilon$ for every $n \ge N$.

It follows that for every $n \ge N$,

$$\begin{aligned}
\mathbf{E}(\mathrm{d}(W, \mathbf{H}(W,n))) &\le \mathbf{E}(\mathrm{d}(W,V) + \mathrm{d}(V, \mathbf{H}(V,n)) + \mathrm{d}(\mathbf{H}(V,n), \mathbf{H}(W,n))) \\
&\le \mathrm{d}(W,V) + \mathbf{E}(\mathrm{d}(V, \mathbf{H}(V,n))) + \mathbf{E}(\left\|\mathbf{H}(V,n) - \mathbf{H}(W,n)\right\|_1) \\
&\le \left\|W - V\right\|_1 + \mathbf{E}(\mathrm{d}(V, \mathbf{H}(V,n))) + \frac{3}{n}\left\|V - W\right\|_\infty \\
&\le 3\varepsilon.
\end{aligned}$$

□

We are now ready to establish Theorem 1.2.

Proof of Theorem 1.2. The assumption implies that for every positive integer n, the random variables $\mathbf{H}(W_1,n)$ and $\mathbf{H}(W_2,n)$ have the same distribution. Consequently, if H_n is a random function equivalent to both $\mathbf{H}(W_1,n)$ and $\mathbf{H}(W_2,n)$, then by lemma 1.6 both $\mathrm{d}(W_1,H_n)$ and $\mathrm{d}(W_2,H_n)$ tend almost surely to 0. It therefore follows from the triangular inequality $\mathrm{d}(W_1,W_2) \le \mathrm{d}(W_1,H_n) + \mathrm{d}(H_n,W_2)$ that $\mathrm{d}(W_1,W_2)$ equals 0. □

1.5.2 Kernel isomorphisms

Theorem 1.2 shows that if two kernels are equal in terms of densities, then they are equal for the distance d. The purpose of this section is to show that in this case, there is an isomorphism between (almost all) the vertices of these kernels that preserves the characteristic of the kernels. This is established by Proposition 1.8.

Lemma 1.7. *If (J_1, μ_1, W_1) and (J_2, μ_2, W_2) are two kernels on Borel probability spaces such that $\mathrm{d}(W_1, W_2) = 0$, then there exist a Borel probability space (J, μ) and for each $i \in \{1, 2\}$ a measure-preserving map $\varphi_i \colon J \to J_i$ such that*

$$W_2^{\varphi_2}(x, y, z) = W_1^{\varphi_1}(x, y, z)$$

for μ-almost every $(x, y, z) \in J^3$.

Proof. We only sketch the argument, as it is a straightforward extension of results already published.

We may first assume that $J_1 = [0, 1] = J_2$ and $\mu_1 = \mu_2$ is the Lebesgue measure Λ_1 of $[0, 1]$. This follows from the fact that for $i \in \{1, 2\}$ there is a measure-preserving function ψ_i from $([0, 1], \Lambda_1)$ to (J_i, μ_i) [30, Theorem A.9]. Indeed, in this case it holds that $\mathrm{d}(W_i^{\psi_i}, W_i) = 0$ for $i \in \{1, 2\}$ by the definition of d, and further that $\mathrm{d}(W_1^{\psi_1}, W_2^{\psi_2}) = 0$. Assume then that the Lemma holds for the kernels $([0, 1], \Lambda_1, W_1^{\psi_1})$ and $([0, 1], \Lambda_1, W_2^{\psi_2})$ and yields a probability space (J, μ) and measure-preserving functions $\varphi_1, \varphi_2 \colon J \to [0, 1]$ such that $(W_1^{\psi_1})^{\varphi_1} = (W_2^{\psi_2})^{\varphi_2}$ μ^3-almost everywhere. Then the Lemma holds for W_1 and W_2 with the same space (J, μ) and the functions $\varphi_1' = \psi_1 \circ \varphi_1$ and $\varphi_2' = \psi_2 \circ \varphi_2$.

Now a direct adaptation of the proof of Theorem 8.13 in the book by Lovász [32, p. 136] gives the lemma. We just give an outline of the argument. We aim to show the theorem for $J = J_1 \times J_2$, and ϕ_i being the projection of J on J_i for $i \in \{1, 2\}$. We know that $\mathrm{d}(W_1, W_2) = \inf_\mu \left\| W_1^{\phi_1} - W_2^{\phi_2} \right\|_1^\mu$, where μ ranges over all coupling measures of $J = J_1 \times J_2$. It suffices to show that this last infimum is in fact a minimum to deduce the theorem. As we assumed that $J_1 = [0, 1] = J_2$, the space of coupling measures is compact in the weak topology. Consequently, it is enough to show that the function $\mu \mapsto \left\| W_1^{\phi_1} - W_2^{\phi_2} \right\|_1^\mu$ is *lower semicontinuous*, i.e., if $(\mu_n)_n$ weakly converges to μ then

$$\liminf_n \left\| W_1^{\phi_1} - W_2^{\phi_2} \right\|_1^{\mu_n} \geq \left\| W_1^{\phi_1} - W_2^{\phi_2} \right\|_1^{\mu}.$$

This last inequality is Inequality (8.21) on p. 137 of *loc. cit.* and the proof follows as in the book. □

Proposition 1.8. *Let (J_1, μ_1, W_1) and (J_2, μ_2, W_2) be two twin-free kernels on Borel probability spaces such that $t(H, W_1) = t(H, W_2)$ for every function H. There exist two sets $N_1 \subseteq J_1$ and $N_2 \subseteq J_2$ with $\mu_1(N_1) = 0 = \mu_2(N_2)$ and an invertible and measure-preserving map $\rho \colon J_1 \setminus N_1 \to J_2 \setminus N_2$ such that*

1. *ρ^{-1} is measure preserving; and*
2. *W_1 and W_2^ρ are equal μ_1^3-almost everywhere.*

Proof. Theorem 1.2 ensures that Lemma 1.7 applies: let (J, μ) and φ_1, φ_2 be the probability space and the applications given by this lemma, respectively.

By Theorem A.9 in [30], we may assume that $(J, \mu) = ([0, 1], \Lambda_1)$, up to composing φ_1 and φ_2 by a measure preserving map from $[0, 1]$ to J. By Lemma 8.9 in [30] there is a measure preserving function $\psi : J_1 \times [0, 1] \to J$ such that $\varphi_1(\psi(x, t)) = x$ for $\mu_1 \times \Lambda_1$ every $(x, t) \in J_1 \times [0, 1]$.

Since $W_1^{\varphi_1} = W_2^{\varphi_2}$ almost everywhere and ψ is measure preserving, it follows that

$$\begin{aligned} W_2^{\varphi_2}(\psi(x, t_1), \psi(y, t_2), \psi(z, t_3)) &= W_1^{\varphi_1}(\psi(x, t_1), \psi(y, t_2), \psi(z, t_3)) \\ &= W_1(x, y, z) \end{aligned} \tag{1.5}$$

for $\mu_1^3 \times \Lambda_1^3$-almost every $(x, y, z, t_1, t_2, t_3) \in J_1^3 \times [0, 1]^3$. Further,

$$W_2^{\varphi_2}(\psi(x, t_1), \psi(y, t_2), \psi(z, t_3)) = W_1(x, y, z) = W_2^{\varphi_2}(\psi(x, t_1'), \psi(y, t_2), \psi(z, t_3))$$

for almost every $(x, y, z, t_1, t_1', t_2, t_3)$. Since $\varphi_2 \circ \psi$ is measure preserving, this can be rewritten as

$$W_2(\varphi_2 \circ \psi(x, t_1), y', z') = W_2(\varphi_2 \circ \psi(x, t_1'), y', z')$$

for almost every $(x, y', z', t_1, t_1') \in J_1 \times J_2^2 \times [0, 1]^2$. Since W_2 is twin-free, it follows that $\psi(x, t_1) = \psi(x, t_1')$ for almost every $(x, t_1, t_1') \in J_1^2 \times [0, 1]$. Consequently, we can fix t_0 such that $\varphi_2 \circ \psi(x, t_0) = \varphi_2 \circ \psi(x, t)$ for $\mu_1 \times \Lambda_1$-almost every (x, t). Define $\rho_0(x) = \varphi_2(\psi(x, t_0))$, then $\rho_0(x) = \varphi_2(\psi(x, t))$ for almost every (x, t). As a first consequence, $\mu_1({\rho_0}^{-1}(A)) = \mu_2(\varphi_2^{-1}(A)) \cdot \Lambda_1([0, 1]) = \mu_2(A)$ for every μ_2-measurable set A, so ρ_0 is measure preserving. As a second consequence, (1.5) can be rewritten as

$$W_2^{\rho_0}(x, y, z) = W_2^{\varphi_2}(\psi(x, t_0), \psi(y, t_0), \psi(z, t_0)) = W_1(x, y, z)$$

for μ_1^3-almost every (x, y, z).

Let $N_1 \subseteq J_1$ be a nullset such that for every $x \in J_1 \setminus N_1$, $W_2^{\rho_0}(x, y, z) = W_1(x, y, z)$ holds for μ_1^2-almost every (y, z). Set $N_2 = J_2 \setminus \rho_0(J_1 \setminus N_1)$ and let ρ be the restriction of ρ_0 from $J_1 \setminus N_1$ to $J_2 \setminus N_2$.

We now prove that ρ is injective. Suppose that $\rho(x) = \rho(x')$, for some $x, x' \in J_1 \setminus N_1$. Then in particular

$$W_2^\rho(x, y, z) = W_2(\rho(x), \rho(y), \rho(z)) = W_2(\rho(x'), \rho(y), \rho(z)) = W_2^\rho(x', y, z)$$

for every $(x, y) \in J_2^2$. Since $x, x' \notin N_1$, it follows that

$$W_1(x, y, z) = W_2^\rho(x, y, z) = W_2^\rho(x', y, z) = W_1(x', y, z)$$

for μ_1^2-almost every (y, z). Since W_1 is twin-free, we conclude that $x = x'$, which proves that ρ is injective.

This further allows us to apply Theorem A.7 in [30] applies, that shows that the inverse of ρ is measure preserving. This concludes the proof of the Lemma. □

Corollary 1.9. *We may assume that the map ρ of Proposition 1.8 is an isometry from $(J_1 \setminus N_1, d_{W_1})$ to $(J_2 \setminus N_2, d_{W_2})$. Moreover,*

- *if μ_1 has full support and (J_2, d_{W_2}) is complete, then we may assume that $N_1 = \varnothing$; and*
- *conversely, if (J_1, d_{W_1}) is complete and μ_2 has full support, then we may assume that $N_2 = \varnothing$.*

In particular, if (J_1, μ_1, W_1) and (J_2, μ_2, W_2) are pure kernels, then $N_1 = \varnothing = N_2$.

Proof. Let $\rho\colon J_1 \setminus N_1 \to J_2 \setminus N_2$ be the map given by Proposition 1.8 applied to (J_1, μ_1, W_1) and (J_2, μ_2, W_2). We first prove that we may restrict ρ to a set D_1 of μ_1-measure 1 such that if we fix any $x \in D_1$, then $W(x, y, z)$ equals $W^\rho(x, y, z)$ for μ_1^2-almost every pair $(y, z) \in J^2$. For $x \in J_1$, we define $I(x) \subseteq J_1^2$ to be the set of pairs (y, z) such that $W(x, y, z) \neq W^\rho(x, y, z)$. Further, let A be the set composed of each $x \in J_1$ such that $\mu_1^2(I(x)) > 0$.

We assert that $\mu_1(A) = 0$. To prove this, we set $A_\varepsilon = \left\{ x \in J_1 \;\middle|\; \mu_1^2(I(x)) > \varepsilon \right\}$ and we notice that $A = \bigcup_n A_{\varepsilon_n}$ where the union is taken over a decreasing sequence $(\varepsilon_n)_{n\in\mathbf{N}}$ that tends to 0. As the union is countable, it suffices to prove that $\mu_1(A_\varepsilon) = 0$ for every $\varepsilon > 0$ to conclude that $\mu_1(A) = 0$. Fixing $\varepsilon > 0$, it follows from the definitions that $W(x, y, z) \neq W^\rho(x, y, z)$ for every triple in $\{ (x, y, z) \mid x \in A_\varepsilon \text{ and } (y, z) \in I(x) \}$, which is a set of μ_1^3-measure at least $\varepsilon \cdot \mu_1(A_\varepsilon)$. Because of Property 2 of Proposition 1.8, the previous statement implies that $\varepsilon \cdot \mu_1(A_\varepsilon) = 0$, hence $\mu_1(A_\varepsilon) = 0$.

We define $D_1 = J_1 \setminus N_1 \setminus A$ and $D_2 = \rho(D_1) = J_2 \setminus N_2 \setminus \rho(A)$. We know that $\mu_1(D_1) = 1$ and the equality $\mu_2(D_2) = 1$ follows from the fact that ρ^{-1} is measure preserving. The restriction $\rho_{|D_1}\colon D_1 \to D_2$ of ρ to D_1 is an isometry between the metric spaces (D_1, d_{W_1}) and (D_2, d_{W_2}). Indeed, fixing $(x, x') \in D_1^2$, we know from the construction of D_1 that for μ^2-almost every pairs $(y, z) \in J_1^2$ we have $W_1(x, y, z) = W_2^\rho(x, y, z)$ and $W_1(x', y, z) = W_2^\rho(x', y, z)$. So in particular $W_1(x, y, z) - W_1(x', y, z) = W_2^\rho(x, y, z) - W_2^\rho(x', y, z)$. Consequently,

$$\begin{aligned}
d_{W_1}(x, x') &= \int_{J_1^2} |W_1(x, y, z) - W_1(x', y, z)| \,\mathrm{d}\mu_1^2(y, z) \\
&= \int_{J_1^2} |W_2^\rho(x, y, z) - W_2^\rho(x', y, z)| \,\mathrm{d}\mu_1^2(y, z) \\
&= \int_{J_2^2} |W_2(\rho(x), y', z') - W_2(\rho(x'), y', z')| \,\mathrm{d}\mu_2^2(y', z') \\
&= d_{W_2}(\rho(x), \rho(x')).
\end{aligned}$$

This proves that $\rho_{|D_1}$ is an isometry.

Now we assume that μ_1 has full support and (J_2, d_{W_2}) is complete. In this case, $\rho_{|D_1}$ extends by continuity to an injective map $\tilde{\rho}$ on J_1. To prove this, it suffices to show that $\rho_{|D_1}$ is absolutely continuous and D_1 is dense in (J_1, d_{W_1}).

The absolute continuity follows from the fact that $\rho_{|D_1}$ is an isometry. The set D_1 is dense in J_1 because every open set included in $J_1 \setminus D_1$ is an open nullset, and hence is empty as μ_1 has full support. By continuity of d_{W_1} and d_{W_2} towards themselves the extension $\tilde{\rho}$ is an isometry.

To prove the second item, it suffices to apply the previous proof to the inverse $(\rho_{|D_1})^{-1}$ of $\rho_{|D_1}$, where the roles played by (J_1, μ_1, W_1) and (J_2, μ_2, W_2) are inverted. □

1.6 Order types on the sphere

Why the sphere?

In the following, it will be convenient to manipulate pure kernels, in particular in order to apply Corollary 1.9. Formally, it is always possible to transform a kernel into an equivalent kernel that is pure by completing the neighborhood distance space and taking the quotient by the twin relation (see for instance the construction in the book of Lovász [32, Section 13.3.1. Purifying kernels]). However, in the case of a kernel (D, μ, χ_D), where μ is a probability measure of $\mathbf{R}^2$ with support D, the completion can be done in a geometric way by embedding $\mathbf{R}^2$ into (a hemisphere of) the two-dimensional sphere S_2, which is compact. This completion also boils down to add a *point at infinity* for each direction to which a sequence of points of D tends. In the following, we shall consider chirotopes on the whole sphere.

Consider the chirotope $\chi : (x, y, z) \mapsto \operatorname{sign}\det(x, y, z)$ defined on the space $\mathbf{R}^3 \setminus \{0\}$. With this definition, the standard chirotope $\chi_{\mathbf{R}^2}$ of $\mathbf{R}^2$ is written as $\chi_{\mathbf{R}^2}(a, b, c) = \chi(a \oplus 1, b \oplus 1, c \oplus 1)$ for every a, b and c in $\mathbf{R}^2$. As the sign of the determinant $\det(a, b, c)$ does note change when one of a, b and c is multiplied by a positive number, the vectors $x \in \mathbf{R}^3 \setminus \{0\}$ and λx are indistinguishable for χ whenever $\lambda > 0$. Therefore consider the quotient E of $\mathbf{R}^3 \setminus \{0\}$ by the corresponding equivalence relation: x is equivalent to y if $y = \lambda x$ for some positive real number λ. By the remark stated above, the chirotope χ of $\mathbf{R}^3 \setminus \{0\}$ is properly defined on E. A possible set of representatives for the quotient E is the unit sphere $S_2 \subseteq \mathbf{R}^3$. The plane $\mathbf{R}^2 \times \{1\}$ represents the same elements of E as an open hemisphere of S_2, composed of all points (x, y, z) of S_2 with $z > 0$. We write $h_{z>0}$ this open hemisphere.

More explicitly, the plane can be mapped to $h_{z>0}$ in chirotope-preserving way by the radial projection

$$\pi_{S_2} : \left\{ \begin{array}{rcl} \mathbf{R}^2 & \to & S_2 \\ a & \mapsto & \frac{a \oplus 1}{\|a \oplus 1\|_2} \end{array} \right.$$

We shall use this function to push a probability measures on $\mathbf{R}^2$ to the sphere.

Spherical transformations

It can be proved that the functions from $\mathbf{R}^2$ to $\mathbf{R}^2$ that preserves the chirotope of $\mathbf{R}^2$, i.e. the functions that preserves alignments and the orientations of the triangles, are exactly the affine transformations with positive determinant. A projective transform f of $\mathbf{R}^2$ also preserves the chirotope on a half-plane D_f bounded by the line sent to infinity.

Note also that if $f \in \mathrm{GL}^+(\mathbf{R}^3)$ is a linear mapping with positive determinant, then f preserves the chirotope of $\mathbf{R}^3 \setminus \{0\}$ since $\det(f(a), f(b), f(c)) = \det(f) \cdot \det(a, b, c)$ and $\det(f) > 0$. This function induces a map from the set of representatives S_2 to itself that maps $x \in S_2$ to $\frac{f(x)}{\|f(x)\|_2}$. This last map from S_2 to S_2 is called a *spherical transform*. It can be proved that spherical transforms are exactly the functions that preserve the chirotope of S_2.

The case of the plane described above then appears as a particular case of the spherical setting: a projective transform f of $\mathbf{R}^2$ is equal on the domain D_f on which it preserves chirotopes to a function $f_0 = \pi_{S_2}^{-1} \circ g \circ \pi_{S_2}$, where g is a spherical transform and the domain D_f is exactly the set of points on which f_0 is properly defined, i.e. $D_f = \left\{ \pi_{S_2}^{-1}(x) \;\middle|\; x \in h_{z>0} \text{ and } g(x) \in h_{z>0} \right\}$. Moreover, f is a (positive) affine transform of $\mathbf{R}^2$ exactly in the case where $D_f = h_{z>0}$.

Spherical geometry

We have to be careful while defining the lines through a to b. Indeed, if a and $b = -a$ are antipodal points of S_2, there exists infinitely many such lines since every third point c is aligned (that is, on a same great circle) with a and b. If $a \in S_2$ and $b \in S_2$ are different non-antipodal points, let $h(a, b)$ be the (unique!) line (i.e. great circle) that contains both a and b.

Further, we let $h^+(a, b)$ be the hemisphere of all points $x \in S_2$ such that the orientation of (a, b, x) is counter-clockwise, and $h^-(a, b)$ the hemisphere of all points $x \in S_2$ such that the orientation of (a, b, x) is clockwise.

We deliberately use the same notations for the plane. If $(a, b) \in \mathbf{R}^2$, let $h(a, b)$ to be the line through a and b. Similarly, we let $h^+(a, b) \subseteq \mathbf{R}^2$ be the half-space of all points $x \in \mathbf{R}^2$ such that (a, b, x) is counter-clockwise, and $h^-(a, b)$ the half-space of all points $x \in \mathbf{R}^2$ such that (a, b, x) is clockwise.

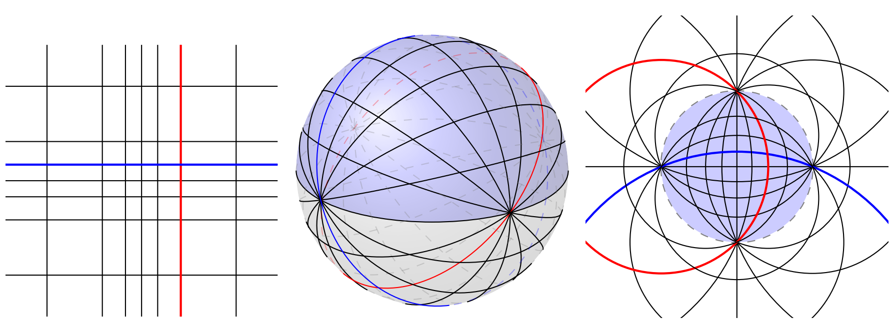


Figure 1.4 – The same grid. Left: On the plane corresponding to $h_{z>0}$. Middle: On the sphere. Right: Stereographic projection of the sphere.

The geometry of S_2 is naturally called *spherical geometry*. In spherical geometry, the lines are the great circles, that is the circle of S_2 with same center as the sphere. The lines are also the geodesics. It is well known that many properties differ from spherical to plane geometry. In particular, spherical geometry does not satisfy the parallel postulate of Euclid (there are no parallel lines in the sphere). Regarding the orientation of the triangles, however, S_2 is an extension of $\mathbf{R}^2$.

Two points of the a and b of the sphere $S_2$are *antipodal* if $b = -a$. Note that if a and b are antipodal, then every (spherical) line containing a contains b. If a and b are non-antipodal points of S_2, the (spherical) line $h(a, b)$ is the circle centered on the origin (of $\mathbf{R}^3$) that passes through a and b. The *(spherical) segment* $[a, b]$ is the shortest of the two arcs of $h(a, b)$ going from a to b. These notations are the same as for the plane on purpose because they are consistent when identifying the open hemisphere $h_{z>0}$ with the plane $\mathbf{R}^2$ via π_{S_2}. Indeed, if $a, b \in h_{z>0}$ then $\pi_{S_2}^{-1}(h \cap h(a, b))$ equals the line $h(\pi_{S_2}^{-1}(a), \pi_{S_2}^{-1}(b))$. Moreover, $[a, b] \subseteq h_{z>0}$ and $\pi_{S_2}^{-1}([a, b]) = [\pi_{S_2}^{-1}(a), \pi_{S_2}^{-1}(b)]$.

As a consequence, one can reason on the plane $\mathbf{R}^2$ for every geometric reasoning on the hemisphere h. To illustrate reasoning on the sphere on this plane sheet of paper, we use a stereographic projection that maps S_2 to $\mathbf{R}^2$ by the function $(x, y, z) \in S_2 \mapsto (\frac{x}{z+1}, \frac{y}{z+1})$. See Figure 1.4.

Spherical chirotope

Let a, b and c be three points of the sphere $S_2 \subseteq \mathbf{R}^3$. The *orientation* of a triangle abc in the sphere S_2 is given by $\chi(a, b, c) = \operatorname{sign}(\det(a, b, c))$. Analogously to the plane, the triangle abc is *clockwise* if $\chi(a, b, c) = 1$ and *counter-clockwise* if $\chi(a, b, c) = -1$. The three points a, b and c are aligned if $\chi(a, b, c) = 0$. This defines a chirotope on the sphere S_2.

This chirotope on the open hemisphere $h_{z>0}$ is consistent with the standard

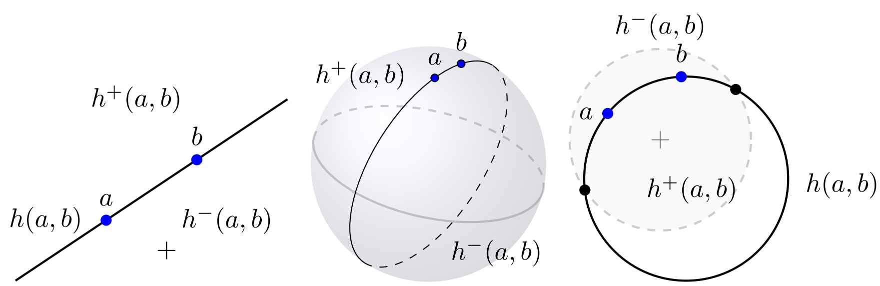


Figure 1.5 – A line on the sphere. Left: Radial projection of $h(a, b) \cap h_{z>0}$ on the plane. Middle: The whole spherical line $h(a, b)$. Right: Stereographic projection.

chirotope of $\mathbf{R}^2$, using the homeomorphism π_{S_2}. Indeed, it is easy to check that

$$\begin{aligned}\operatorname{sign}\det(a\oplus 1, b\oplus 1, c\oplus 1) &= \operatorname{sign}\det\left(\frac{a\oplus 1}{\|a\oplus 1\|_2}, \frac{b\oplus 1}{\|b\oplus 1\|_2}, \frac{c\oplus 1}{\|c\oplus 1\|_2}\right)\\ &= \operatorname{sign}\det(\pi_{S_2}(a), \pi_{S_2}(b), \pi_{S_2}(c))\end{aligned}$$

for every $(a, b, c) \in (\mathbf{R}^2)^3$.

As a consequence, every order type on the plane can be realized in the sphere. On the contrary, there are order types in the sphere that cannot be realized in the plane. The smallest example is the order type of size 4 represented in particular by the vertices of a regular tetrahedron, so for instance by the points $\frac{1}{\sqrt{3}}(1,1,1)$, $\frac{1}{\sqrt{3}}(1,-1,-1)$, $\frac{1}{\sqrt{3}}(-1,1,-1)$and $\frac{1}{\sqrt{3}}(-1,-1,1)$. See Figure 1.6.

Figure 1.6 – The regular tetrahedron on the sphere.

Recall that a line of the sphere is a great circle. We say that a probability measure ν on S_2 does not charge lines if for each such line $L \subseteq S_2$, it holds that $\nu(L) = 0$. A probability measure on S_2 that does not charge lines generates a limit of (spherical) order types similarly to measure on the plane. If ν is such a measure that does not charge lines, let ℓ_ν be the function that assigns to every order type ω realizable on the sphere the pro probability that a $\mu^{|\omega|}$-random set of points has order type ω.

1.7 Kernel isomorphism for spherical (or plane) measures

The purpose of this section is to specialize Proposition 1.8 (or rather its corollary 1.9) to kernels defined by measure on the sphere. If two probability measures ν_1 and ν_2 satisfy $\ell_{\nu_1} = \ell_{\nu_2}$, an application of Proposition 1.8 already gives a measure preserving bijection $\rho : \operatorname{supp}\nu_1 \setminus \mathcal{N}_1 \to \operatorname{supp}\nu_2 \setminus \mathcal{N}_2$, that preserves the chirotope *almost* everywhere and where both $\mathcal{N}_1$ and $\mathcal{N}_2$ are sets of zero measure (for ν_1 and ν_2, respectively).

In this section, we use the geometrical properties of these particular kernels to construct such a function ρ that moreover preserves the chirotope *everywhere* it is defined and we take the sets $\mathcal{N}_1$ and $\mathcal{N}_2$ as small as possible. This is achieved by Theorem 1.22.

Note that since a probability measures on $\mathbf{R}^2$ is a special cases case of measures of S_2 for our concerns, every result in this section involving a measure μ

on S_2 also holds when μ is measure on the plane. The reader can also think that the proves are done for measures on the plane as the arguments remain the same.

1.7.1 Twins in measures of S_2 (and further, of $\mathbf{R}^2$)

We give an example that shows that we need to remove a countable number of twins.

Example 1.8. We construct a probability measure on $\mathbf{R}^2$ whose support is illustrated on Figure 1.7. First, we construct a countable set of consecutive arcs of S_1. The choices for the bound of these arcs are arbitrary. For $i \in \mathbf{N}$, define

$$a_i = \frac{1}{2^i} \quad \text{and} \quad b_i = \frac{3}{2} \cdot \frac{1}{2^i}.$$

Note that $\sum_{i\in\mathbf{N}} |b_i - a_i| = \sum_{i\in\mathbf{N}} \frac{1}{2^{i+1}} = 1$. For each $i \in \mathbf{N}$, let p_i and q_i be the points of the unit circle S_1 with polar coordinates $(1, a_i)$ and $(1, b_i)$ respectively, and let $A_i = \{ (\cos\theta, \sin\theta) \mid a_i \leq \theta \leq b_i \}$ be the arc of the circle S_1 from p_i to q_i. Let μ_1 be the probability measure with support $D := \bigcup_{i\in\mathbf{N}} A_i$ that is uniform on this support (i.e. μ_1 is equal to the length measure Λ_1 on D), so that $\mu_1(A_i) = \Lambda_1(A_i) = \frac{1}{2^{i+1}}$.

Recall the convex limit $\ell_\diamond$ defined in Example 1.4. Since $D \subseteq S_1$ is a convex curve, we have $\ell_{\mu_1} = \ell_\diamond$. Moreover, the points p_i and q_{i+1} are twins for this measure for every $i \in \mathbf{N}$. Indeed, if b and c are two distinct points of $D \backslash \{p_i, q_{i+1}\}$ then p_i and q_{i+1} are on the same side of the line $h(b, c)$ because $h(b, c)$ crosses the circle S_1 in b and c, which are not between p_i and q_{i+1} in S_1. Consequently, $\chi(p_i, b, c) = \chi(q_{i+1}, b, c)$ for μ_1^2-almost every $(b, c) \in D^2$, i.e. p_i and q_{i+1} are twins.

As a consequence, μ_1 has a countable family of twins $\{ \{p_i, q_{i+1}\} \mid i \in \mathbf{N} \}$. Moreover, we know that $\ell_{\mu_1} = \ell_\diamond$ is also realized by the uniform measure μ_2 on the whole circle S_1. Now, if there is a set $\mathcal{N}_1 \subseteq D$ and a measure preserving bijection $\rho : D \backslash \mathcal{N}_1 \to \operatorname{supp} \mu_2$ that preserves the chirotope, then ρ in particular preserves twins. Since μ_2 admit no twins, the set $D \setminus \mathcal{N}_1$ contains no pair of twins of μ_1. It follows that $\mathcal{N}_1$ contains at least one element of $\{p_i, q_{i+1}\}$ for each $i \in \mathbf{N}$, so $\mathcal{N}_1$ is at least countable.

In the rest of this section, we shall prove that twins of spherical measures are rather exceptional. Proposition 1.15 shows a strong necessary condition for the existence of twins and Lemma 1.14 shows that in any case there are at most countably many of them.

Given three points a, b and c of S_2 (or $\mathbf{R}^2$), the *wedge* $w(a, b, c)$ is the set

$$w(a, b, c) = \{ x \in S_2 \mid \chi(a, c, x) \neq \chi(b, c, x),\ \chi(a, c, x) \neq 0 \text{ and } \chi(b, c, x) \neq 0 \}.$$

In words, $w(a, b, c)$ is the set of points x such that acx and bcx are triangles with different orientations. If a, b and c are in general position, i.e. they are not aligned, $w(a, b, c)$ is equal to $(h^+(a, c) \triangle h^+(b, c)) \setminus (h(a, c) \cup h(b, c))$. See

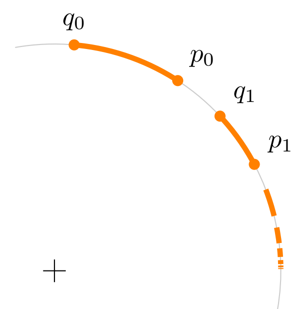


Figure 1.7 – The support of μ_1 in Example 1.8

Figure 1.8 for an illustration. In the degenerated cases, that is when a, b and c are aligned, the following holds. If $a = b$, then $w(a,b,c) = \varnothing$. If $b = -a$ and $c \notin \{a, b\}$ then $w(a,b,c) = S_2 \setminus h(a,c)$. If $b \notin \{a, -a\}$ and c is in one of the open segments $]a,b[$, and $]-a,-b[$, then $w(a,b,c) = S_2 \setminus h(a,b)$. In the remaining case, i.e. if $c \in [a,-b] \cup [-a,b]$, then $w(a,b,c) = \varnothing$.

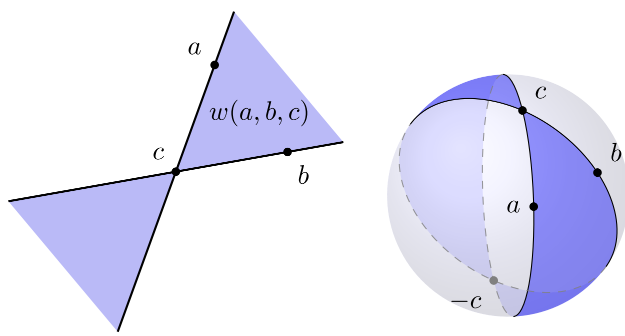


Figure 1.8 – The wedge $w(a,b,c)$.

The next Lemma exhibit a relation between twins and wedges. This relation is also a characterization of twins in a measure that charges no line.

Lemma 1.10. *Let μ be a measure on S_2 or $\mathbf{R}^2$ and let a and b be two twins for the standard chirotope of μ. For every $c \in \operatorname{supp} \mu$, the intersection of $\operatorname{supp} \mu$ and $w(a,b,c)$ is empty.*

Proof. Let $c \in \operatorname{supp}(\mu)$. If $c' \in \operatorname{supp}(\mu) \cap w(a,b,c)$, then none of $\chi(a,c,c')$ and $\chi(b,c,c')$ is 0 and, moreover, $\chi(a,c,c') \neq \chi(b,c,c')$. By the continuity of the determinant, the set

$$X = \{\ (d,d') \mid \chi(a,d,d') = \chi(a,c,c') \text{ and } \chi(b,d,d') = \chi(b,c,c')\ \}$$

is open. Since (c,c') belongs to the support of μ^2 and is also in X, it follows that $\mu^2(X) > 0$. This contradicts the hypothesis that a and b are twins for χ. $\square$

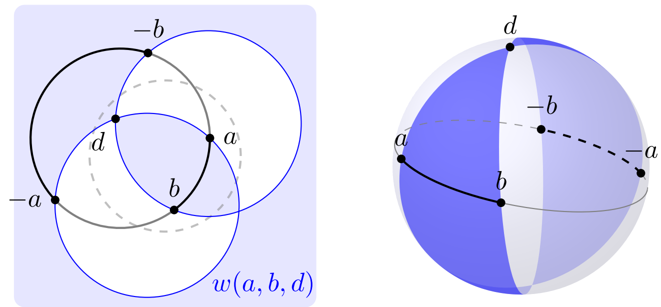


Figure 1.9 – $w(a, b, d) \cap h(a, b) =]a, b[\cup]-a, -b[$

Lemma 1.11. *Let μ be a measure on S_2 such that no line contains* $\operatorname{supp} \mu$*. Let a and b be twins for the standard chirotope of μ. Then a and b are not antipodal.*

Proof. Assume otherwise that there is $a \in \operatorname{supp} \mu$ with $-a \in \operatorname{supp} \mu$ such that a and $-a$ are twins. If $\operatorname{supp} \mu = \{a, -a\}$, then $\operatorname{supp} \mu$ is in particular contained in a line, so we assume there is a point c in $\operatorname{supp} \mu$ that is different from a and $-a$. By Lemma 1.10, we know that $w(a, -a, c) \cap \operatorname{supp} \mu = \varnothing$. Since in this special case $w(a, -a, c) = S_2 \setminus h(a, c)$, it follows that $w(a, -a, c)$ is contained in the line $h(a, c)$, which contradicts the hypothesis. □

Lemma 1.12. *Let μ be a probability measure of S_2 such that no line contains* $\operatorname{supp}(\mu)$*. Let W be the standard chirotope on* $\operatorname{supp}(\mu)$*. Every $x \in \operatorname{supp} \mu$ has at most one twin for the kernel* $(\operatorname{supp} \mu, \mu, W)$*.*

Proof. Suppose on the contrary that a, b and c are three distinct points in $\operatorname{supp} \mu$ such that (a, b), (a, c) and (b, c) are twins for $(\operatorname{supp} \mu, \mu, W)$. By Lemma 1.11, we know that $c \notin \{-a, -b\}$.

Assume first that a, b and c are aligned. As no line contains $\operatorname{supp} \mu$, there is a fourth point $d \in \operatorname{supp} \mu \setminus h(a, b)$. Note that the wedge $w(a, b, d)$ contains the open segments $]a, b[$ and $]-a, -b[$. See Figure 1.10 for an illustration of this. Since a and b are twins, it follows from Lemma 1.10 that $c \notin [a, b]$ and $c \notin [-a, -b]$. By the symmetry of the roles played by a, b and c, we may also assume that $b \notin [a, c]$ and $a \notin [b, c]$. The (spherical) line $h(a, b)$ is the disjoint union of the (spherical) segments $[a, b]$, $]b, -a[$, $[-a, -b]$ and $]-b, a[$. We know that $c \notin [a, b] \cup [-a, -b]$. Moreover, if $c \in]-b, a[$ then $a \in [b, c]$ and, similarly, if $c \in]b, -a[$ then $b \in [a, c]$. In each case, we get a contradiction.

We now assume that a, b and c are not aligned. Applying Lemma 1.10 to each of $w(a, b, c)$, $w(b, c, a)$ and $w(c, a, b)$ yields that $\operatorname{supp} \mu$ is contained in $h(a, b) \cup h(a, c) \cup h(b, c)$, as illustrated in Figure 1.10. Note that at this point the proof is finished if we further assume that μ charges no line.

Observe that none of a, b and c can be an atom. Suppose indeed that a is an atom. As $b \in \operatorname{supp} \mu$, there exists a neighborhood N_b of b with positive measure

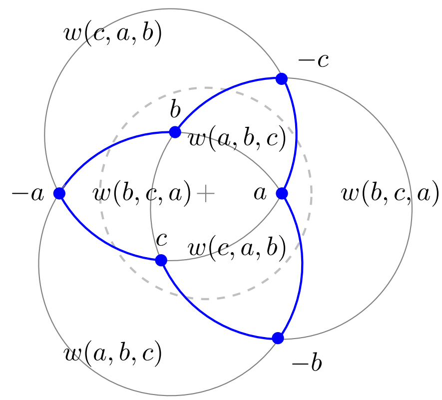


Figure 1.10 – $S_2 \setminus (w(a,b,c) \cup w(b,c,a) \cup w(c,a,b))$ is equal to the union of the segments $[a,-c]$, $[-c,b]$, $[b,-a]$, $[-a,c]$, $[c,-b]$ and $[-b,a]$.

that is disjoint from $h(a,c)$. For every $b' \in N_b$, it holds that $W(a,b',a) = 0 \neq W(c,b',a)$, which yields a contradiction since $\mu^2(\{a\} \times N_b) > 0$ and a and c are twins.

As a result, $\operatorname{supp}\mu$ contains points that are not atoms, which implies that $\operatorname{supp}\mu$ is infinite. Consequently, one of $h(a,b)$, $h(a,c)$ and $h(b,c)$ contains an infinite number of points in $\operatorname{supp}\mu$. Without loss of generality, we may thus assume that $h(a,b)$ contains two distinct points x and y in $\operatorname{supp}\mu \setminus \{a,b\}$. Let N_x and N_y be neighborhoods of x and y that do not intersect $h(a,c) \cup h(b,c)$, respectively. We know that $\mu(N_x) > 0$ and, further, $\mu(N_x \cap h(a,b)) > 0$. Likewise, $\mu(N_y \cap h(a,b)) > 0$. For every $x' \in N_x \cap h(a,b)$ and every $y' \in N_y \cap h(a,b)$, it holds that $W(a,x',y') \neq 0 = W(c,x',y')$, which contradicts that a and c are twins. $\square$

The two following results prove that the set of twins of a measure on the plane and the sphere respectively is countable. We chose to first show it on the plane and then derive the result on the sphere from the result on the plane.

Proposition 1.13. *Let μ be a probability measure on $\mathbf{R}^2$ such that no line contains* $\operatorname{supp}(\mu)$. *Let W be the standard chirotope on* $\operatorname{supp}(\mu)$. *The set T of twins of the kernel* $(\operatorname{supp}(\mu), \mu, W)$ *is countable. Furthermore, if μ is atom-free then $\mu(T) = 0$.*

Proof. The second part of the statement directly follows from the first. For convenience, we set $D = \operatorname{supp}(\mu)$ and for every set $X \subseteq \mathbf{R}^2$ we define X_D to be $X \cap D$. We first prove that for every line L, the set of points on L_D that have a twin is countable. Assume that L_D is not countable and let $a, b \in D$ be two twins such that $a \in L_D$. We first deal with the case where $D \subseteq L \cup h(a,b)$. Since no line contains D, we know that $h(a,b) \neq L$. Let x and y be two distinct

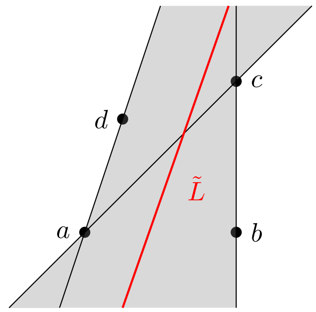


Figure 1.11 – The set $R = w(a,b,c) \cup w(c,d,a) \cup h(a,c) \cup h(a,d) \cup h(b,c)$ is figured by the shaded part of the plane. The line $\tilde{L}$ is contained in R.

points of $L_D \setminus \{a\}$, which exist as L_D is not finite. By the definition of the support, every neighborhood of x in $\mathbf{R}^2$ has positive μ-measure, and similarly for neighborhoods of y. Let N_x and N_y be two disjoint neighborhoods of x and y that do not intersect $h(a,b)$, respectively. We know that both $\mu(N_x \cap D)$ and $\mu(N_y \cap D)$ are positive. For every $x' \in N_x \cap D$ and $y' \in N_y \cap D$, we have $W(x', y', b) \neq 0 = W(x', y', a)$, which contradicts that a and b are twins.

Suppose now that D is not contained in $L \cup h(a,b)$, so there exists a third point $c \in D \setminus (L \cup h(a,b))$. Furthermore, the set $L \setminus D$ is open in the line L, hence it is the union of a countable family F of open intervals of L that are pairwise disjoint. (This is obtained by taking the collection of maximal open intervals contained in $L \setminus D$. There are countably many intervals, as is seen by realizing that every set X of disjoint open intervals of $\mathbf{R}$ is countable: one can indeed choose a rational number in each element of X, thereby defining an injection of X in $\mathbf{Q}$.) It suffices to prove that the point a must be a bound of an interval in F. The wedge $w(a,b,c)$ is open and intersects L. Moreover, by Lemma 1.10, the intersection of $w(a,b,c)$ and D is empty, which implies that $w(a,b,c) \cap (L \setminus D)$ is an open interval I, of which a is a bound. As $a \in D$, the unique interval in F that contains I also admits a as a bound. This finishes the proof of our assertion.

Now consider two twins a and b. By what precedes, we may assume that the set of twins is not contained in the line $h(a,b)$. Therefore there exists a point $c \in D \setminus h(a,b)$ that has a twin $d \in D \setminus h(a,b)$. Applying Lemma 1.10 to (a,b,c) and to (c,d,a), we infer that $w(a,b,c) \cup w(c,d,a)$ does not intersect D. There exists a vector $\vec{u}$ with rational coordinates such that the line L colinear to $\vec{u}$ through c is contained in $w(a,b,c) \cup \{c\}$. Moreover, we can ensure that $\vec{u}$ is not colinear to $\vec{cd}$. Therefore, there is $\varepsilon_0 > 0$ such that for any $\varepsilon \in (0, \varepsilon_0)$, the translated of L by $\varepsilon\vec{cd}$ is contained in $R = w(a,b,c) \cup w(c,d,a) \cup h(a,c) \cup h(a,d) \cup h(b,c)$. Therefore, there exists such a translated $\tilde{L}$ of L that goes through a point with rational coordinates. Figure 1.11 provides an illustration.

Moreover, if x and y are two twins in D that do not belong to a same connected component of $\mathbf{R}^2 \setminus \tilde{L}$, then necessarily they belong to $h(a,c) \cup h(a,d) \cup$

$h(b, c)$. To establish this, observe that if x and y do not belong to a line $h(s, t)$ with $s, t \in D$, then x and y must actually belong to a same connected component of $\mathbf{R}^2 \setminus h(s, t)$, for otherwise s would belong to $w(x, y, t)$ thereby contradicting Lemma 1.10. Applying this for $(s, t) \in \{(a, c), (a, d), (b, c)\}$ yields that x and y belong to a same connected component of $\mathbf{R}^2 \setminus (h(a, c) \cup h(a, d) \cup h(b, c))$. Hence x and y belongs to a same connected component of $\mathbf{R}^2 \setminus R$ because none of them belongs to $w(a, b, c) \cup w(c, d, a)$. In particular, x and y belong to a same connected component of $\mathbf{R}^2 \setminus \tilde{L}$. Thus whenever two twins are separated by $\tilde{L}$, one of them belongs to $h(a, c) \cup h(a, d) \cup h(b, c)$. We saw that the number of points on $D \cap (h(a, c) \cup h(a, d) \cup h(b, c))$ that have a twin is countable. Moreover, Lemma 1.12 ensures that every point in D has at most one twin. Therefore, we infer that $\tilde{L}$ separates a countable number of twins (including a and b).

As a result, to each pair (a, b) of twins in D we can associate a line with a rational equation that separates a and b such that each such line is associated to countably many pairs of twins. As there is a countable number of such lines, we deduce that the set of twins in D is countable. This concludes the proof. □

Let us extend Proposition 1.13 to measures on the sphere.

Lemma 1.14. *Let μ be a measure on S_2 such that no line contains* $\operatorname{supp} \mu$*. The set* $\operatorname{Twin}(\mu)$ *of twins of μ is countable.*

Proof. Lemma 1.11 shows that if a and b are twins for μ then a and b are not antipodal. It follows that there is an open hemisphere of S_2 that contains both a and b.

Moreover, the sphere S_2 can be covered by a finite number of open hemispheres such that every pair of non-antipodal points is covered by one of these hemispheres. To see this, note that $(S_2)^2 \subseteq \bigcup_h h^2$, where the union is taken over all open hemispheres. Since $(S_2)^2$ is compact and every h^2 is open, it is possible to extract a finite covering of $(S_2)^2$.

Consequently, it suffices to show that every open hemisphere h of S_2 contains countably many pairs of twins to deduce that there are countably many pairs of twins, and further deduce the lemma.

If $\mu(h) = 0$, then h does not intersect $\operatorname{supp} \mu$, so we assume $\mu(h) > 0$. Consider the measure μ_h defined on h by $\mu_h(A) = \frac{\mu(A)}{\mu(h)}$ for every measurable set $A \subseteq h$. It follows from Proposition 1.13 that the number of twins of μ_h is countable. Note that, moreover, a pair (a, b) of twins of μ with $a \in h$ and $b \in h$ is a pair of twins of μ_h, so there are countably many such pairs. □

Hausdorff dimension

Let E be a metric space. The *diameter* $\operatorname{diam} A$ of a subset $A \subseteq E$ is the largest distance between two points of A, that is $\operatorname{diam} A = \sup_{x,y \in E^2} d(x, y)$. For each $s \geq 0$ define the *Hausdorff measure of dimension s* as

$$\mathcal{H}^s(A) = \lim_{\epsilon \to 0} \inf_S \sum_{A \in S} (\operatorname{diam} A)^s$$

where the infimum is taken over all families S of subsets of E of diameter at most ϵ that cover E, i.e. such that $E \subseteq \bigcup_{A \in S} A$.

For integer values of s, the measure $\mathcal{H}^s$ coincides with the Lebesgue measure Λ_s of dimension s. One can prove that there is a value $\alpha \in [0, +\infty]$ such that $\mathcal{H}_s(A) = 0$ whenever $s > \alpha$ and $\mathcal{H}_s(A) = +\infty$ whenever $s < \alpha$. This number α is called the *Hausdorff dimension of* E and is written $\dim_H E$. Hausdorff dimensions need not be integers.

A function f from a metric space (E, d_E) to a metric space (F, d_F) is *Lipchitz continuous* if there exists a constant $c > 0$ such that

$$d_E(f(x), f(y)) \leq c \cdot d_F(x, y)$$

for every $(x, y) \in E^2$. The function f is *locally Lipchitz continuous* if every $x \in E$ has neighborhood on which f is Lipchitz continuous. If f is locally Lipchitz continuous, then $\dim_H f(E) \leq \dim_H E$. For more details on the Hausdorff dimension, see for instance [43, chapter 7].

Hausdorff dimension of a set with a twin

The existence of a twin in a measure of S_2 implies very strong constraints on its support. The next result shows that the support of a measure with a twin is a (at most) one-dimensional object.

Proposition 1.15. *Let μ be a probability measure on S_2 with standard chirotope χ. If the kernel* $(\operatorname{supp} \mu, \mu, \chi)$ *has a twin, then* $\operatorname{supp} \mu$ *has Hausdorff dimension at most* 1.

Proof. Let a and b be two twins for μ. We know from Lemma 1.11 that a and b are not antipodal. Up to apply a spherical transformation, we may assume that $a = \frac{1}{\sqrt{2}}(1, 1, 0)$ and $b = \frac{1}{\sqrt{2}}(-1, 1, 0)$.

We first show that $\operatorname{supp} \mu$ intersects $h^+(a, b) = h_{z>0} = \{ (x, y, z) \in S_2 \mid z > 0 \}$ in a curve of Hausdorff dimension at most 1.

Consider the projection $\pi_{S_2}^{-1}$ from $h^+(a, b)$ to the plane. Geometrically, the points a and b (that are in the the adherence of $h^+(a, b)$) are "mapped to the points at infinity" in the directions $(1, 1)$ and $(-1, 1)$ respectively, that is, the intersection of the lines containing a (resp. b) with $h_{z>0}$ are mapped by $\pi_{S_2}^{-1}$ to parallel lines of $\mathbf{R}^2$ with direction $(1, 1)$ (resp. $(-1, 1)$). This is illustrated by Figure 1.12. Since a and b are twins, we know by Lemma 1.10 that $w(a, b, c)$ does not intersect $\operatorname{supp} \mu$. For a point $c \in h^+(a, b)$ with coordinates (x_c, y_c) in the plane, the intersection of $w(a, b, c)$ with $h^+(a, b)$ is therefore the wedge $w(c + (1, 1), c + (-1, 1), c)$. It follows that every pair of (x, y) in the support of μ in $\mathbf{R}^2$ satisfies

$$|y - y_c| \leq |x - x_c|. \tag{1.6}$$

Let A be the set of abscissas x such that there is $a \in \operatorname{supp} \mu \cap h^+(a, b)$ with abscissa x. It follows from (1.6) that this element a is unique, let f be the function defined on D be that maps such a pair to the corresponding $a \in$

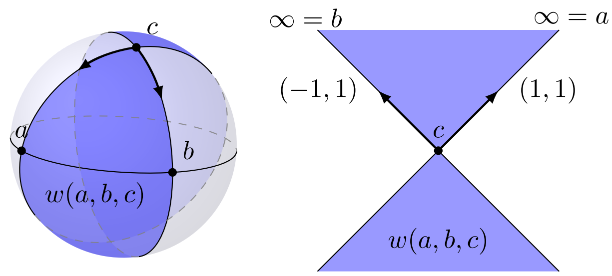


Figure 1.12 – The wedge $w(a, b, c)$. Left: On the sphere. Right: The upper hemisphere projected on the plane.

$\operatorname{supp}\mu \cap h^+(a, b)$ with abscissa x. By Equation (1.6), the function f is Lipchitz continuous. It follows from the Lipschitz continuity of f that the Hausdorff dimension of $f(A)$ is at most that of D, which is at most 1. Consequently, the smoothness of π_{S_2} yields that $\operatorname{supp}\mu \cap h^+(a, b) = g(f(A))$ also has Hausdorff dimension at most 1.

One similarly shows that $\operatorname{supp}\mu \cap h^-(a, b)$ has Hausdorff dimension at most 1. As $\operatorname{supp}\mu \cap h(a, b)$ also has Hausdorff dimension at most 1, we infer that the Hausdorff dimension of $\operatorname{supp}\mu$ is at most 1. □

1.7.2 Pseudo-aligned triples

We present in Example 1.9 two probability measures μ_1 and μ_2 with bounded respective supports and no twins that represent the same limit $\ell_{\mu_1} = \ell_{\mu_2}$ and such that there is no measure preserving function $\rho : \operatorname{supp}\mu_1 \to \operatorname{supp}\mu_2$ that preserves the chirotopes on *every* triple of $\operatorname{supp}\mu_1$.

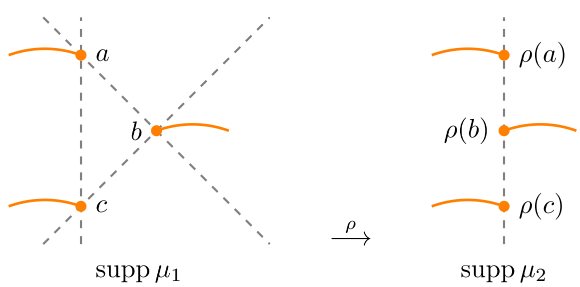


Figure 1.13

Example 1.9. The measures we describe are illustrated on Figure 1.13. We first construct a convex curve I from $(0, 0)$ to $(1, 0)$ that is flat enough. The purpose of this curve is to build a segment-like piece of measure while being consistent with our choice to consider measures that charge no line (so that the generated order types are non-degenerated). If the reader does not insist on this constraint, she or he can think that I is the interval $[0,1] \times \{0\}$. Let I be the arc from $(0,0)$ to $(1,0)$ of center $(\frac{1}{2}, -2)$ and suitable radius, that is $\frac{\sqrt{17}}{2}$. Define the points $a = (0,1)$, $b = (1,0)$ and $c = (0,-1)$. Define the sets $I_a = a - I$, $I_b = b + I$ and $I_c = c - I$. Let μ_1 be the probability measure on $\mathbf{R}^2$ satisfying $\mu_1(I_a) = \mu_1(I_b) = \mu_1(I_c) = \frac{1}{3}$ that is uniform on each I_x with $x \in \{a,b,c\}$ that is, proportional to the Lebesgue measure Λ_1 on these sets.

Similarly, define $b' = (0,0)$ and $I'_b = b' + I$ and let μ_2 be the probability measure with $\mu_2(I_a) = \mu_2(I'_b) = \mu_2(I_c) = \frac{1}{3}$ that is uniform on these three arcs. It can be checked that I is flat enough so that I_b, I'_b and I_c (resp. I_a, I_b and I'_b) lie below (resp. above) any line spanned by two points of I_a (resp. I_c) and any line spanned by two points of I_b or I'_b are below I_a and above I_c. Note also that (a,b,c) is the only aligned triple of $\operatorname{supp}\mu_2$.

Note that $\operatorname{supp}\mu_1 = I_a \cup I_b \cup I_c$ and $\operatorname{supp}\mu_2 = I_a \cup I'_b \cup I_c$ and that μ_1 is equal to μ_2 on $I_a \cup I_c$. Now, let $\rho : \operatorname{supp}\mu_1 \to \operatorname{supp}\mu_2$ be the function such that $\rho(x) = x$ for $x \in I_a \cup I_c$ and such that $\rho(b+x) = b' + x$ for $b+x \in I_b$ with $x \in I$. It is clear from the construction that $\mu_1 = \mu_2 \circ \rho$. Moreover, one may check that if χ is the standard chirotope of $\mathbf{R}^2$, then for $(x,y,z) \in \operatorname{supp}\mu_1^3$, we have $\chi(x,y,z) = \chi^\rho(x,y,z)$ unless $\{x,y,z\} = \{a,b,c\}$. Since $\operatorname{supp}\mu_1$ contains no aligned triple, there is no bijection between $\operatorname{supp}\mu_1$ and $\operatorname{supp}\mu_2$ that preserves the chirotope everywhere.

It is possible to make a fractal construction where this situation appears on a countable number of points.

The next definition captures the properties of the triple (a,b,c) in Example 1.9. Let $a,b,c \in (S_2)^3$ be three points that are pairwise non-antipodal. The open *pseudo-hemisphere* defined by (a,b,c) is $p(a,b,c) = h^+(a,b) \cap h^+(b,c) \cap h^+(a,c)$.

Observe that if b belongs to the (spherical) segment $[a,c]$, then $p(a,b,c)$ is the hemisphere $h^+(a,b)$. In this case, $h^-(a,b) = p(c,b,a)$. If a, b and c are aligned points, and none of them belongs to the segment define by the two others, then $p(a,-b,c) = h^+(a,b)$. This situation is specific to the sphere, it happens for instance when $a = (1,0,0)$, $b = (-\frac{1}{2}, -\frac{\sqrt{3}}{2}, 0)$ and $c = (-\frac{1}{2}, \frac{\sqrt{3}}{2}, 0)$.

Let $\overline{p}(a,b,c)$ denote the closure of $p(a,b,c)$. Pseudo-hemispheres are illustrated in Figure 1.14.

Definition 1.4. Let D be a subset of S_2. A triple $(a,b,c) \in D^3$ is *pseudo-aligned towards* D if it is not aligned and one of the following properties holds:

- $\overline{p}(a,b,c) \cup \overline{p}(c,b,a)$ covers D; or
- $\overline{p}(a,-b,c) \cup \overline{p}(c,-b,a)$ covers D.

The existence of pseudo-aligned triples put some contraints in D, for instance in this case D is not connected. Note also that a point that belongs a pseudo-aligned triple of D is on are on the boundary of D (See Figure 1.14).

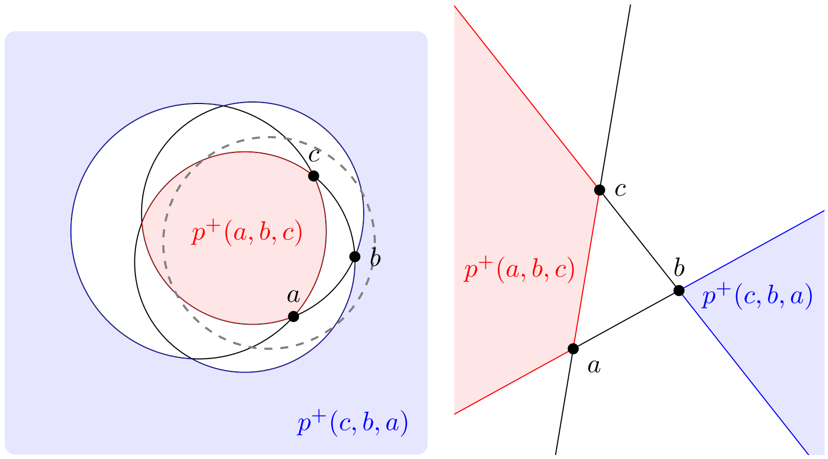


Figure 1.14 – Pseudo-hemispheres. Left: stereographic projection of the sphere. Right: intersection with the plane.

In the following, we shall only use the following proposition to control the quantity of pseudo-aligned.

Proposition 1.16. *The set of pseudo-aligned triples of a set $D \subseteq S_2$ is countable.*

Proof. Let P be the set of points $x \in D$ contained in a pseudo-aligned triple of D. Equivalently to the proposition, we prove that P is countable.

More precisely, we prove that the set P is extremal in D in the sense that for every $x \in P$ there is a non-empty open set $U_x \subseteq S_2$ such that for every $y \in U_x$, the point x is the unique point of D such that $\|x - y\|_2 = \inf_{x' \in D} \|x' - y\|_2$.

In this case, the sets U_x and U_y are disjoint whenever x and y are different element of D. As family of disjoint open sets of a separable space, the family $\{ U_x \mid x \in P \}$ is countable, so P is countable.

It remains to show the existence of such a U_x for every x contained in a pseudo-aligned triple of D. Let (a, b, c) be a pseudo aligne triple, so D is a subset of $A := \overline{p}(a, b', c) \cup \overline{p}(c, b', a)$ with $b' = \pm b$. Take $\epsilon_a > 0$ small enough to ensure that $A \cap B(a, \epsilon_a) \subseteq \overline{h}^+(a, b') \cap \overline{h}^+(a, c)$. It then suffices to take $U_a = h^-(a, b') \cap h^-(a, c) \cap B(a, \frac{\epsilon_a}{2})$. It is easy to check that for every $x \in U_a$, the point of D that is the closest to x is in $B(x, \frac{\epsilon_a}{2}) \subseteq B(a, \epsilon_a)$ and thus is a. Symetrically, take $U_c = h^-(b', c) \cap h^-(a, c) \cap B(a, \frac{\epsilon_c}{2})$ where $\epsilon_c > 0$ is chosen such that $A \cap B(c, \epsilon_c) \subseteq \overline{h}^+(b', c) \cap \overline{h}^+(a, c)$. For b', the same construction works with $U_b = h^-(c, b) \cap h^-(b, c) \cap B(b, \frac{\epsilon_b}{2})$ (resp. $U_b = h^-(a, -b) \cap h^-(-b, c) \cap B(b, \frac{\epsilon_b}{2})$) if $b' = b$ (resp. $b' = -b$) where ϵ_b is chosen such that $A \cap B(b, \epsilon_b) \subseteq \overline{h}^+(c, b) \cap \overline{h}^+(b, a)$ (resp. $A \cap B(b, \epsilon_b) \subseteq \overline{h}^+(a, -b) \cap \overline{h}^+(-b, c)$). This concludes the proof of the propostion.

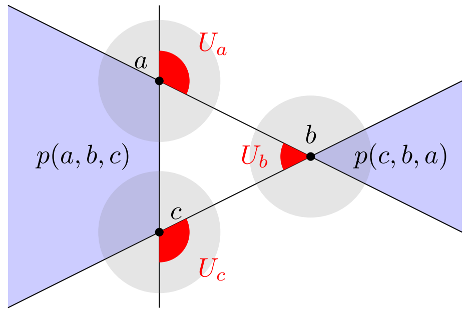


Figure 1.15 – Construction of U_a, U_b and U_c.

□

1.7.3 Geometric purification

Fix a probability measure ν on the sphere S_2 that charges no line. Let D be the support of ν and let $W := \chi_D$ be the standard chirotope on D, and consider the kernel (D, ν, W). We aim to define a kernel equivalent to (D, ν, W) that is pure. Let Twin(D) be the set of twins of (D, ν, W) and let $\sim$ be the twin equivalence relation, defined by $x \sim y$ if and only if x and y are twins.

We now define the *quotient kernel* of (D, ν, W) by the twin relation $\sim$. Let $\tilde{D} = D/\sim$ be the set of equivalence classes of D by the relation $\sim$. For $x \in D$, we write $\overline{x} \in \tilde{D}$ the equivalence class of x, that is, the set of twins of x (including x). Let $\tilde{\pi} : D \to \tilde{D}$ be the natural projection that maps $x \in D$ to $\overline{x}$ (so in particular if x has no twin then $\tilde{\pi}(x) = \{x\}$). The probability measure ν naturally induces a quotient measure ν' on $\tilde{D}$ defined by $\tilde{\nu}(A) = \nu(\tilde{\pi}^{-1}(A))$ for every measurable subset A of $\tilde{D}$. It follows from this definition that $\tilde{\pi}$ is a measure preserving function from (D, ν) to $(\tilde{D}, \tilde{\nu})$. It remains to define a chirotope $\tilde{W}$ on $\tilde{D}$. The property we require for this function is that for $\overline{a}, \overline{b}, \overline{c} \in \tilde{D}^3$ the value $\tilde{W}(\overline{a}, \overline{b}, \overline{c})$ belongs to the set $\left\{ W(a', b', c') \,\middle|\, a' \in \overline{a}, b' \in \overline{b}, c' \in \overline{c} \right\}$. To fix things and to ensure that $\tilde{W}$ is measurable, define $\tilde{W}(\overline{a}, \overline{b}, \overline{c}) := \min_{(a',b',c') \in \overline{a} \times \overline{b} \times \overline{c}} W(a', b', c')$. This choice is of course non-canonical, a max function or every other measurable criterion would do. Note in particular that if a, b and c have no twin, then $\tilde{W}(\overline{a}, \overline{b}, \overline{c}) = W(a, b, c)$. The *quotient kernel* of (D, ν, W) is $(\tilde{D}, \tilde{\nu}, \tilde{W})$.

By the remark above, we have $\tilde{W}^{\tilde{\pi}} = W$ on $(D \setminus \text{Twin}(\nu))^3$. Moreover, the set Twin$(\nu)$ is countable by Lemma 1.14, so $\nu(\text{Twin}(\nu)) = 0$. As a consequence, (D, ν, W) and $(\tilde{D}, \tilde{\nu}, \tilde{W})$ are equivalent kernels. In particular, the random chirotope $\mathbf{H}(\tilde{W}, n)$ is almost surely realizable on the sphere.

Note that we have two topologies for D: the topology given by the neighborhood distance d_W of the kernel W and the Euclidean topology of S_2. It will be very usefull in the following that these topologies happen to be equivalent outside of the set of twins, as shown by the next lemmas.

Lemma 1.17. *The pseudo-distance d_W is continuous toward the euclidean distance $\|\cdot\|_2$. That is if $\|x_n - x\|_2 \to 0$, then $d_W(x_n, x) \to 0$.*

Proof. We take a sequence $(x_n)_{n\in\mathbf{N}} \in \operatorname{supp}\nu^{\mathbf{N}}$ that converges to a limit x in the sense that $\|x_n - x\|_2$ tends to 0 and we show that $d_W(x_n, x)$ tends to 0. For every pair $(y, z) \in \operatorname{supp}\nu^2$ such that x, y and z are not aligned, the set of points $x' \in \mathbf{R}^2$ satisfying the equation $W(x, y, z) = W(x', y, z)$ is an open hemisphere containing x. Consequently, $W(x_n, y, z)$ is eventually equal to $W(x, y, z)$. It follows that the sequence $f_n(y, z) = W(x, y, z) - W(x_n, y, z)$ tends to 0 for ν^2-almost every pair $(y, z) \in \operatorname{supp}\nu^2$. Indeed, this last statement is true if x, y and z are not aligned and the set of pairs (y, z) aligned to x is a ν^2-nullset because ν does not charge lines. We conclude by an application of the dominated convergence theorem that the sequence $d_W(x_n, x) = \int_{\operatorname{supp}\nu^2} f_n(y, z)\,\mathrm{d}\nu(y, z)$ tends to 0 when n goes to infinity. □

Lemma 1.18. *For every $(x, y) \in D^2$,*

$$d_W(x, y) = d_{\tilde{W}}(\overline{x}, \overline{y}).$$

Proof. We first prove that for $a \in S_2$, the equality $W(a, b, c) = \tilde{W}(\overline{a}, \overline{b}, \overline{c})$ holds for ν^2-almost every $(b, c) \in D^2$. Since $\nu(\mathrm{Twin}(\nu)) = 0$, it is enough to prove this equality for ν^2-almost all $(b, c) \in (D \setminus \mathrm{Twin}(\nu))^2$.

For $a \in S_2$ and $(b, c) \in (D \setminus \mathrm{Twin}(\nu))^2$, we know from the definition of $\tilde{W}$ that $\tilde{W}(\overline{a}, \overline{b}, \overline{c}) = W(a_0, b_0, c_0)$ for some $a_0 \in \overline{a}$, $b_0 \in \overline{b}$ and $c_0 \in \overline{c}$. Further, if b and c have no twin, we have $\tilde{W}(\overline{a}, \overline{b}, \overline{c}) = W(a_0, b, c)$ because $\overline{b} = \{b\}$ and $\overline{c} = \{b\}$. Note that the point $a_0 \in \overline{a}$ is either equal to a or a twin of a. In both cases, the equality $W(a', b, c) = W(a, b, c)$ is satisfied by ν^2-almost every $(b, c) \in D^2$. This proves the claimed property.

We deduce that $d_W(x, y) = d_{\tilde{W}}(\overline{x}, \overline{y})$ for every $(x, y) \in \operatorname{supp}\nu^2$ using the following computation.

$$\begin{aligned}
d_W(x, y) &= \int_{D^2} |W(x, b, c) - W(y, b, c)| d\nu^2(b, c) \\
&= \int_{D^2} |\tilde{W}(\overline{x}, \overline{b}, \overline{c}) - \tilde{W}(\overline{y}, \overline{b}, \overline{c})| d\nu^2(b, c) \quad \text{by the property proved above.} \\
&= \int_{\tilde{D}^2} |\tilde{W}(\overline{x}, u, v) - \tilde{W}(\overline{y}, u, v)| d\tilde{\nu}^2(u, v) \quad \text{as } \tilde{\pi} : x \mapsto \overline{x} \text{ is measure preserving.} \\
&= d_{\tilde{W}}(\overline{x}, \overline{y}).
\end{aligned}$$

□

Lemma 1.19. *The function $\tilde{\pi} : x \mapsto \overline{x}$ is continuous from the metric space $(D, \|.\|_2)$ to $(\tilde{D}, d_{\tilde{W}})$.*

Proof. Let us first check that $d_{\tilde{W}}$ is a metric on $\tilde{W}$. Indeed, as mentioned after the definition of a pure kernel, this boils down to checking that $\tilde{W}$ has no twin. Assume otherwise that $\overline{x} \in \tilde{D}$ and $\overline{y} \in \tilde{D}$ are twins of $\tilde{W}$. It follows that $d_{\tilde{W}}(\overline{x},\overline{y}) = 0$ and by the Lemma 1.18, $d_W(x,y) = d_{\tilde{W}}(\overline{x},\overline{y}) = 0$. Further, x and y are twins of W, so $\overline{x} = \overline{y}$.

Let us prove the continuity of the projection function $\tilde{\pi}$. Let $(x_n)_{n\in\mathbf{N}} \in D^{\mathbf{N}}$ be a sequence that converges to a limit $x \in D$ in the sense that $\|x_n - x\|_2 \underset{n\to\infty}{\longrightarrow} 0$. By Lemma 1.17, the sequence $d_W(x_n, x)$ tends to 0 and by Lemma 1.18 we have $d_W(x_n, x) = d_{\tilde{W}}(\pi(x_n), \pi(x))$ for every $n \in \mathbf{N}$. Consequently,

$$d_{\tilde{W}}(\tilde{\pi}(x_n), \tilde{\pi}(x)) \underset{n\to\infty}{\longrightarrow} 0,$$

which indeed proves that $\tilde{\pi} : (D, \|.\|_2) \to (\tilde{D}, d_{\tilde{W}})$ is continuous. □

Lemma 1.20. *The metrics d_W and $\|.\|_2$ are equivalent on $\operatorname{supp}\nu \setminus \operatorname{Twin}(\nu)$, that is they induce the same topology.*

Proof. We already know from Lemma 1.17 that d_W is continuous toward $\|.\|_2$, so it suffices to prove that if $d_W(x_n, x)$ tends to 0 for some sequence $(x_n)_{n\in\mathbf{N}} \in (\operatorname{supp}\nu \setminus \operatorname{Twin}(\nu))^{\mathbf{N}}$ and $x \in \operatorname{supp}\nu \setminus \operatorname{Twin}(\nu)$ then $\|x_n - x\|$ tends to 0 when n goes to infinity.

We apply a classical compacity argument. As a close subset of the compact S_2, the set $\operatorname{supp}\nu$ is compact with respect to $\|.\|_2$. Assume for the sake of contradiction that there is a sequence $(d_W(x_n, x))_{n\in\mathbf{N}}$ that tends to 0 while $(\|x_n - x\|)_{n\in\mathbf{N}}$ does not. Since $\operatorname{supp}\nu$ is compact, it is possible to extract a subsequence $(x_{\phi(n)})_{n\in\mathbf{N}}$ that converges with respect to $\|.\|_2$ to a limit $y \in \operatorname{supp}\nu$ different from x. Lemma 1.17 yields that $d_W(x_{\phi(n)}, y) \to 0$, and further $d_W(x, y) = 0$. Since x has no twin, it follows that $x = y$, which yields a contradiction. □

Proposition 1.21. *The kernel $(\tilde{D}, \tilde{\nu}, \tilde{W})$ is pure.*

Proof. We already know that $(\tilde{D}, \tilde{\nu}, \tilde{W})$ has no twins. Let us show that the metric space $(\tilde{D}, d_{\tilde{W}})$ is compact, hence complete. By Lemma 1.19, the projection $\tilde{\pi}$ is a continuous function from $(D, \|\cdot\|_2)$ to $(\tilde{D}, d_{\tilde{W}})$. Moreover, D is compact for $\|.\|_2$, as a closed subset of the compact S_2. It follows that $\tilde{D} = \tilde{\pi}(D)$ is compact with respect to $d_{\tilde{W}}$.

It remains to check that $\tilde{\nu}$ has full support. Let U be an open subset of $\tilde{D}$ with $\tilde{\nu}(U) = 0$. Then $\tilde{\pi}^{-1}(U)$ is an open of D, and further $\nu(\tilde{\pi}^{-1}(U)) = \tilde{\nu}(U) = 0$ because $\tilde{\pi}$ is measure-preserving. By the definition of $D = \operatorname{supp}\nu$, the measure ν has full support on D. Consequently, $\tilde{\pi}^{-1}(U) = \varnothing$ and further $U = \varnothing$. This finishes to prove that $(\tilde{D}, \tilde{\nu}, \tilde{W})$ is pure. □

1.7.4 Kernel isomorphisms on the sphere

We are now ready to prove the main theorem of Section 1.7. Recall that pseudo-aligned points are defined in Definition 1.4. See also Property 1.16.

Theorem 1.22. *Let ν_1 and ν_2 be probability measures on S_2 that charge no line, and such that $\ell_{\nu_1} = \ell_{\nu_2}$. There are countable sets $\mathcal{N}_1 \subseteq \operatorname{supp}\nu_1$ and $\mathcal{N}_2 \subseteq$*

$\operatorname{supp}\nu_2$, *and a measure-preserving homeomorphism* $\rho : \operatorname{supp}\nu_1 \setminus \mathcal{N}_1 \to \operatorname{supp}\nu_2 \setminus \mathcal{N}_2$ *that preserves the chirotope, that is such that* $\chi(x, y, y) = \chi(\rho(x), \rho(y), \rho(z))$ *for every* $(x, y, z) \in (\operatorname{supp}\nu_1 \setminus \mathcal{N}_1)^3$.

Proof. For $i \in \{1, 2\}$, let D_i be the support of ν_i and W_i be the standard chirotope of D_i. Recall that (D_i, ν_i, W_i) is a kernel. Let $(\tilde{D}_i, \tilde{\nu}_i, \tilde{W}_i)$ be the quotient kernel of (D_i, ν_i, W_i) as defined in Section 1.7.3. (Recall that $\tilde{D}_i$ is the quotient of D_i by the twin relation). We know from Proposition 1.21 that the kernel $(\tilde{D}_i, \tilde{\nu}_i, \tilde{W}_i)$ is pure. Hence, Corollary 1.9 applies to $\tilde{W}_1$ and $\tilde{W}_2$: there is an invertible map $\tilde{\rho} : \tilde{D}_1 \to \tilde{D}_2$ such that

- $\tilde{\rho}$ and $\tilde{\rho}^{-1}$ are measure preserving;
- $\tilde{W}_1$ and $\tilde{W}_2^{\tilde{\rho}}$ are equal $(\tilde{\nu}_1)^3$-almost everywhere; and
- $\tilde{\rho}$ is an isometry from $(\tilde{D}_1, d_{\tilde{W}_1})$ to $(\tilde{D}_2, d_{\tilde{W}_2})$.

We now remove the twins from $\operatorname{supp}\nu_1$ and $\operatorname{supp}\nu_2$. Let $\mathcal{N}_1^{\text{twins}}$ be the set of points $a \in \operatorname{supp}\nu_1$ such that a has a twin for ν_1 or the image $\tilde{\rho}(\overline{a})$ of the twin class $\overline{a}$ of a contains a twin for ν_2. Let similarly define $\mathcal{N}_2^{\text{twins}}$ as the set of points $b \in \operatorname{supp}\nu_2$ such that b has a twin or $\tilde{\rho}^{-1}(\overline{b})$ contains twins for ν_1. It follows from Lemma 1.14 that $\mathcal{N}_1^{\text{twins}}$ and $\mathcal{N}_2^{\text{twins}}$ are countable.

It follows from the definitions that the image of of twin class of $\mathcal{N}_1^{\text{twins}}$ by $\tilde{\rho}$ is the set of twin classes of $\mathcal{N}_1^{\text{twins}}$. Let us define a new domain $E_i = \operatorname{supp}\nu_i \setminus \mathcal{N}_i^{\text{twins}}$ for $i \in \{1, 2\}$. Since $\mathcal{N}_i^{\text{twins}})$ is countable, $\nu_i(E_i) = \nu_i(D_i) = 1$. Note that for $a \in E_i$, the twin class of a is the singleton $\{a\}$. The function $\tilde{\rho}$ therefore induces a bijection ρ between E_1 and E_2 that satisfies $\tilde{\rho}(\{a\}) = \{\rho(a)\}$.

This function ρ inherit the following properties from $\tilde{\rho}$: ρ is measure preserving with respect to the probabilities ν_1 and $W_1 = W_2^{\rho}$ μ_1^3-almost everywhere on E_1^3. We know that $\tilde{\rho}$ is an isometry from the metric space $(\tilde{D}_1, d_{\tilde{W}_1})$ to the metric space $(\tilde{D}_2, d_{\tilde{W}_2})$. By Lemma 1.18, ρ is thereofore an isometry, and further an homeomorphism, from (E_1, d_{W_1}) to (E_2, d_{W_2}). By Lemma 1.20, the metric d_{W_i} is equivalent to the euclidean metric of the sphere on E_i for $i \in \{1, 2\}$. It follows that $\rho : E_1 \to E_2$ is a homeomorphism for the Euclidean topology of S_2.

If $i \in \{1, 2\}$, let $T_i^+ \subseteq D_i^3$ be the set of triples (a, b, c) that form a counterclockwise triangle, that is such that $W_i(a, b, c) = 1$. Similarly let $T_i^- \subseteq D_i^3$ be the set of triples (a, b, c) with $W_i(a, b, c) = -1$.

The sets T_i^+ and T_i^- are open in E_i^3 for the euclidean topology. Indeed, T_i^+ (resp. T_i^-) is the reverse image of the open set $[0, \infty[$ (resp. $]-\infty, 0]$) by the determinant $\det : E_i \subseteq \mathbf{R}^3 \to \mathbf{R}$, which is continuous.

The measure ν_i has full support on $D_i \subseteq \operatorname{supp}\nu_i$, hence $(\nu_i)^3$ has full support on D_i^3. Let us show that

$$\rho^3(T_1^+) \cap T_2^- = \varnothing. \tag{1.7}$$

To see this, it suffices to show that the set $U := \rho^3(T_1^+) \cap T_2^-$ is empty. This set U is open in E_2^2 because it is the intersection of the open set T_2^- with $\rho^3(T_1^+)$, which is open as the image of an open set by the homeomorphism ρ^3. Moreover, since $W_1 = W_2^\rho$ ν_2^3-almost everywhere, we have $\nu_2^3(U) = 0$. Further, ν_2 has full support on D_2, so $U = \varnothing$.

By symmetry of the roles played by T_i^+ and T_i^- for $i \in \{1,2\}$, it follows similarly that $\rho^3(T_1^-) \cap T_2^+ = \varnothing$. In particular, for every $(a,b,c) \in S_1^3$, it holds that $W_2^\rho(a,b,c) \geq 0$ if $W_1(a,b,c) = 1$ and $W_2^\rho(a,b,c) \leq 0$ if $W_1(a,b,c) = -1$.

We now prove that if the image by ρ^3 of an aligned triple is not aligned, then this image corresponds, up to reordering, to a pseudo-aligned triple of E_2. We first claim that for a, b and c in D_1 and a sign $s \in \{-1,1\}$, it holds that

$$\rho(p(a,sb,c)) \subseteq \overline{p}(\rho(a), s \cdot \rho(b), \rho(c)).$$

Indeed, the pseudo hemisphere $p(a,sb,c)$ is defined as the set of points x such that $s \cdot \chi(a,b,x) = -1$, $s \cdot \chi(b,a,x) = -1$ and $\chi(a,c,x) = -1$, so if $x \in p(a,sb,c) \cap D_1$ then $s \cdot \chi^\rho(a,b,x) \leq 0$, $s \cdot \chi^\rho(b,a,x) \leq 0$ and $\chi^\rho(a,c,x) \leq 0$, i.e. $\rho(c)$ belongs to $\overline{p}(\rho(a), s \cdot \rho(b), \rho(c))$.

Let $(a,b,c) \in D_1^3$ be an triple and assume $W_2^\rho(a,b,c) \neq 0$. Up to reordering a, b and c, we can choose $s \in \{-1,1\}$ such that $p(a,sb,c) = h^+(a,c)$ and $p(c,sb,a) = h^+(c,a)$, and further

$$E_1 \setminus h(a,b) \subseteq p(a,sb,c) \cup p(c,sb,a).$$

By the property above,

$$\rho(E_1 \setminus h(a,b)) \subseteq \overline{p}(\rho(a), s\rho(b), \rho(c)) \cup \overline{p}(\rho(c), s\rho(b), \rho(a)).$$

Since the set on the right side of the above inclusion is closed, it suffices to show that $E_2 = \rho(E_1)$ is in the closure of $\rho(E_1 \setminus h(a,b))$ to conclude that (a,b,c) is pseudo-aligned toward E_2. To see this, recall that $E_1 \setminus h(a,b)$ is constructed as the support of ν_1 minus a countable set and the line $h(a,b)$, which are both sets of ν_1-measure 0, so $\nu_1(E_1 \setminus h(a,b)) = 1$. Further, $\nu_2(\rho(E_1 \setminus h(a,b))) = 1$ and by the definition of the support, the adherence of $\rho(E_1 \setminus h(a,b))$ indeed contains $\operatorname{supp} \mu_2 \supseteq E_2$.

By the symmetry of the roles played by E_1 and E_2, ρ and ρ^{-1}, it also holds that if $W_1(a,b,c) \neq 0$ then $W_2^\rho(a,b,c) \neq 0$ unless (a,b,c) correspond to a pseudo-algined triple of E_1 (up to reordering).

It remains to remove the pseudo-aligned triples of E_1 and E_2. For $i \in \{1,2\}$, let $\mathcal{N}_i^{\text{pseudo}}$ be the set of points that belong to a pseudo-aligned triple of E_i. By Proposition 1.16, $\mathcal{N}_i^{\text{pseudo}}$ is countable.

Set $\mathcal{N}_1 := \mathcal{N}_1^{\text{twins}} \cup \mathcal{N}_1^{\text{pseudo}} \cup \rho^{-1}(\mathcal{N}_2^{\text{pseudo}})$ and $\mathcal{N}_2 := \mathcal{N}_2^{\text{twins}} \cup \mathcal{N}_2^{\text{pseudo}} \cup \rho(\mathcal{N}_1^{\text{pseudo}})$. The sets $\mathcal{N}_1$ and $\mathcal{N}_2$ are countable, and $\rho(\operatorname{supp} \nu_1 \setminus \mathcal{N}_1) = \operatorname{supp} \nu_2 \setminus \mathcal{N}_2$. By the properties proved above, $W(a,b,c) = W^\rho(a,b,c)$ whenever $(a,b,c) \in (\operatorname{supp} \nu_1 \setminus \mathcal{N}_1)^3$. This finishes the proof of Theorem 1.22. $\square$

1.7.5 Bijection on the plane

Using the bijection between the plane and an open hemisphere, we prove a version of Theorem 1.22 for measures on the plane. In this theorem, the sets $\mathcal{N}_1$ and $\mathcal{N}_2$ may contain a line, which corresponds to the line mapped to infinity.

Corollary 1.23. *Let μ_1 and μ_2 be two probability measures on $\mathbf{R}^2$ that charge no line and such that $\ell_{\mu_1} = \ell_{\mu_2}$. For each $i \in \{1,2\}$, there is a set $\mathcal{N}_i \subseteq \operatorname{supp}\mu_i$ that is the union of a countable union of points and at most one line, and a measure-preserving homeomorphism $\rho : \operatorname{supp}\mu_1 \setminus \mathcal{N}_1 \to \operatorname{supp}\mu_2 \setminus \mathcal{N}_2$ that preserves the chirotopes.*

Proof. Recall that the projection π_{S_2} is a chirotope-preserving bijection from $\mathbf{R}^2$ to $h_{z>0}$. Theorem 1.22 applies to the spherical probability measures $\nu_1 = \mu_1 \circ \pi_{S_2}^{-1}$ and $\nu_2 = \mu_2 \circ \pi_{S_2}^{-1}$ and gives two countables sets $\mathcal{N}_1^0 \subseteq \operatorname{supp}\nu_1$ and $\mathcal{N}_2^0 \subseteq \operatorname{supp}\nu_2$ and a (ν_1,ν_2)-measure-preserving homeomorphism $\rho : \operatorname{supp}\nu_1 \setminus \mathcal{N}_1^0 \to \operatorname{supp}\mu_2 \setminus \mathcal{N}_2^0$ that preserves the chirotopes.

Note that for $i \in \{1,2\}$, the support of ν_i may not be included in $h_{z>0}$ because the spherical line $h_\infty = \{ (x,y,z) \in S_2 \mid z = 0 \}$ is in the adherence of $h_{z>0}$. However, it holds that $\operatorname{supp}\nu_i \subseteq h_{z>0} \cup h_\infty$.

It follows that $\pi_{S_2}^{-1} \circ \rho \circ \pi_{S_2}$ is a measure-preserving and chirotope-preserving homeomorphism on the set on which it is well-defined to its image, that is from $\operatorname{supp}\mu_1 \setminus \mathcal{N}_1$ to $\operatorname{supp}\mu_2 \setminus \mathcal{N}_2$ where

$$\mathcal{N}_1 := \pi_{S_2}^{-1}(\mathcal{N}_1^0) \cup \pi_{S_2}^{-1}(\rho^{-1}(h_\infty \cap \operatorname{supp}\nu_2))$$

and

$$\mathcal{N}_2 := \pi_{S_2}^{-1}(\mathcal{N}_2^0) \cup \pi_{S_2}^{-1}(\rho(h_\infty \cap \operatorname{supp}\nu_1)).$$

It remains to show that for $i \in \{1,2\}$ the set $\mathcal{N}_i$ is the union of a countable set and at most one line. To see this, it suffices to note that $\pi_{S_2}^{-1}(\mathcal{N}_2^0)$ and $\pi_{S_2}(\mathcal{N}_1^0)$ are countable because $\mathcal{N}_2^0$ and $\mathcal{N}_1^0$ are countable and that $\pi_{S_2}^{-1}(\rho^{-1}(h_\infty \cap \operatorname{supp}\nu_2))$ (resp. $\pi_{S_2}^{-1}(\rho(h_\infty \cap \operatorname{supp}\nu_1)))$ is contained in a line since h_∞ is a line and both $\pi_{S_2}^{-1}$ and ρ^{-1} (resp. ρ) preserve the chirotope. $\square$

1.8 Non representable limit

We have enough material to prove Theorem 1.1.

Proof of Theorem 1.1. Let us prove that there is no measure on the plane representing the limit $\ell_\odot$ introduced in Example 1.6. Assume for the sake of contradiction that $\ell_\odot = \ell_\mu$ for some probability measure μ, that is, $\ell_\odot$ is generated by the kernel $(\operatorname{supp}\mu, \mu, W)$, where W is the standard chirotope of $\operatorname{supp}\mu \subseteq \mathbf{R}^2$. Recall that $\ell_\odot$ is the distribution given by the kernel $(\mathcal{C}_1 \cup \mathcal{C}_X, \mu_X, \chi)$ described in Section 1.4.

Let $(\tilde{D}, \tilde{\mu}, \tilde{W})$ be the quotient kernel of (D, μ, W) with $D := \operatorname{supp}\mu$, as describded in Section 1.7.3. We know that $(\tilde{W}, \tilde{\mu}, \operatorname{supp}\tilde{\mu})$ is pure by Proposition 1.21 and that $(\mu_X, \mathcal{C}_1 \cup \mathcal{C}_X, \chi)$ is also pure by Example 1.7. Since these kernels generate the same limit $\ell_\odot$, Corollary 1.9 applies and gives an isometry $\tilde{\rho}$ from $(\mathcal{C}_1 \cup \mathcal{C}_X, d_\chi)$ to $(\tilde{D}, d_{\tilde{W}})$ that preserves the chirotopes on μ_X^3-almost every triples. We have seen in Example 1.7 that the topology d_χ is the natural topology of an union of cycles on $\mathcal{C}_1 \cup \mathcal{C}_X$ and it follows from Lemma 1.20 that $d_{\tilde{W}}$ is equivalent to the euclidean metric on $D \setminus \mathrm{Twin}(\mu)$. Recall that $\mathrm{Twin}(\mu)$ is countable by Proposition 1.13 and further $\mu_X(N_1) = 0$. Setting $\mathcal{N}_1 := \tilde{\rho}^{-1}(\mathrm{Twin}(\mu))$, it follows that $\tilde{\rho}$ induces a continuous function ρ from $\mathcal{C}_1 \cup \mathcal{C}_X \setminus \mathcal{N}_1$, with the metric of the union of circles, to $D \setminus \mathrm{Twin}(\mu)$, that moreover is (μ_X, μ)-measure preserving and such that $W^\rho = \chi$.

Let $a \in \mathcal{C}_1$ be a random point of $\mathcal{C}_1$ and write $(1, \theta_a)$ the polar coordinates of a, where $\theta_a \in [0, 2\pi[$, then set $b := -a$. For every $n \in \mathbf{N}$, let a_n be a random point of the arc of points with polar coordinate $(1, \theta)$ for some $\theta \in]\theta_a, \theta_a + \frac{1}{n}[$. Let c be a random point of the half of $\mathcal{C}_X$ with polar coordinates (X, θ) with $\theta \in]\theta_a, \theta_a + \pi[$, so that c is inside the triangle aba_n for every $n \in \mathbf{N}$. See Figure 1.16. Set $S = \{a, b, c\} \cup \{ a_n \mid n \in \mathbf{N} \}$.

Since $\mu_X(\mathcal{N}_1) = 0$ and $W^\rho = \chi$ on μ_x^3-almost every triple, the following events happen with probability 1:

- $S \subseteq \mathcal{C}_1 \setminus \mathcal{N}_1$; and
- ρ preserves the chirotopes on S.

Fix a, $(a_n)_{n\in\mathbf{N}}$ and c such that these events happen. It follows that $\rho(c)$ is inside every triangle $\rho(a)\rho(b)\rho(a_n)$. Note that the sequence $(a_n)_{n\in\mathbf{N}}$ tends to a. Since ρ is continuous, it follows that $(\rho(a_n))_{n\in\mathbf{N}}$ tends to $\rho(a)$. Moreover, $\rho(b) \neq \rho(a)$, so the intersection I of the interior of the triangle $\rho(a)\rho(b)\rho(a_n)$ for every $n \in \mathbf{N}$ is empty. This contradicts the fact that $\rho(b) \in I$ and finishes the proof.

□

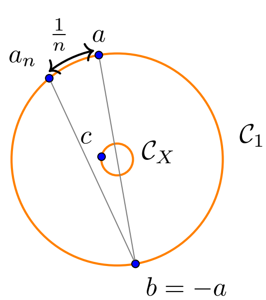


Figure 1.16 – Points of S.

The proof of Theorem 1.1 would be simpler if we could apply Corollary1.23 to deduce that the function ρ preserves chirotopes *everywhere*, so that it suffices to avoid N_1 when choosing a, b, c and $(a_n)_{n\in\mathbf{N}}$ instead of considering random objects. Unfortunately, Corollary 1.23 cannot directly apply to μ and μ_X since μ_X is not a probability measure of $\mathbf{R}^2$. However, all the arguments of the proof of 1.23 applies (up to some tedious checks) and leads to the same conclusion.

1.9 Existence of a limit kernel and regularity lemma

1.9.1 Semi-algebraic relations

A set $A \subseteq \mathbf{R}^n$ is *semi-algebraic* if it can be described as the set of points $(x_1, \ldots, x_n) \in \mathbf{R}^n$ that satisfy a finite number of given polynomial inequalities, that is, inequalities of the form $P(x_1, \ldots, x_n) \geq 0$ or $P(x_1, \ldots, x_n) > 0$, where P is a multi-variate polynomial. This can be reformulated as follows: the set A is semi-algebraic if there are polynomials $P_1, \dots, P_k \in \mathbf{R}[X_1, \dots, X_n]$ and a k-ary Boolean function Φ such that $(x_1, \dots, x_n) \in \mathbf{R}_n$ belongs to A if and only if it satisfies

$$\Phi\left(P_1(x_1, \dots, x_n) > 0, \dots, P_k(x_1, \dots, x_n) > 0\right). \tag{1.8}$$

Let $E \subseteq P^m$ be an m-ary relation on a set P. This set P may not be a subset of $\mathbf{R}^n$. Then E is a *semi-algebraic relation* if there is a function g from P to $\mathbf{R}^n$, for some n and a semi-algebraic set $A \subseteq (\mathbf{R}^n)^m$ such that

$$(x_1, \dots, x_m) \in E \Leftrightarrow (g(x_1), \dots, g(x_m)) \in A.$$

The relation E has *complexity* $k+d+n+m$ if the set A is described by a relation of type (1.8) with$k$ n-variate polynomials and such that $\max_{i=1}^{k} \deg P_i \leq d$ for every $i \in \{1, \dots, k\}$.

This notion generalizes to functions into finite sets as follows. A function f from P^k to a finite set F is *semi-algebraic* if $f^{-1}(\{e\}) \subseteq P^k$ is a semi-algebraic relation for each $e \in F$. That is, if there is a map $g : P \to \mathbf{R}^n$ and a family $((A_e)_{e \in F}) \subseteq (\mathbf{R}^n)^F$ of semi-algebraic sets such that

$$f(x_1, \dots, x_k) = e \quad \text{if and only if} \quad (g(x_1), \dots, g(x_k)) \in A_e$$

The *complexity* of the function f is then the maximum complexity among the sets $f^{-1}(\{e\})$ with $e \in F$. In the literature, a function from $\mathbf{R}^n$ to $\mathbf{R}^m$ is called *semi-algebraic* if its graph is a semi-algebraic subset of $\mathbf{R}^n \times \mathbf{R}^m$. This is different from the notion above since in our definition there is no structure on P. If, however, P happens to be a subset of $\mathbf{R}^p$ and f is a function from $P^k = \mathbf{R}^{pk}$ to a finite subset of $\mathbf{R}$ and the graph of f is a semi-algebraic subset of $\mathbf{R}^{mk} \times \mathbf{R}$, then in particular f is semi-algebraic with respect to the definition above.

Example 1.10 (Interval graphs)**.** Let $G = (V, E)$ be an interval graph, the edge relation $E \subseteq V^2$ is a 2-ary semi-algebraic relation. Indeed, let $([a_v, b_v])_{v \in V}$ be a representation of G as an interval graph. Note that for $(u, v) \in V^2$, the intervals $[a_u, b_u]$ and $[a_v, b_v]$ intersects if and only if $a_u \leq b_v$ and $a_v \leq b_u$. As a consequence,

$$E = \left\{ \, (u, v) \in V^2 \;\middle|\; (a_u, b_u, a_v, b_v) \in A \, \right\}.$$

where A is the semi-algebraic set

$$A = \left\{ \, (a_1, b_1, a_2, b_2) \in (\mathbf{R}^2)^2 \;\middle|\; b_1 - a_2 \geq 0 \text{ and } b_2 - a_1 \geq 0 \, \right\}.$$

Further, the complexity of E is $2 + 1 + 2 + 2 = 7$.

Observation. A chirotope χ on the sphere (so in particular on the plane) is a 3-ary semi-algebraic function. Indeed, the sets

$$\chi^{-1}(-1) = \left\{ \, (x, y, z) \in (S_2)^3 \;\middle|\; \det(x, y, z) < 0 \, \right\},$$

$$\chi^{-1}(0) = \left\{ \, (x, y, z) \in (S_2)^3 \;\middle|\; \det(x, y, z) = 0 \, \right\}, \text{ and}$$

$$\chi^{-1}(1) = \left\{ \, (x, y, z) \in (S_2)^3 \;\middle|\; \det(x, y, z) > 0 \, \right\}$$

are semi-algebraic since the determinant is a polynomial function.

1.9.2 Semi-algebraic regularity lemma

Jacob Fox, János Pach and Andrew Suk proved a strong regularity lemma for semi-algebraic structures. Before stating this result, we need some vocabulary.

Let $E \subseteq P^m$ be an m-ary relation on a set P. For $i \in [m]$, let P_i be a subset of P. The m-tuple $(P_1, \dots, P_m)$ is *homogeneous* if $P_1 \times \cdots \times P_k$ is a subset of either E or $P^m \setminus E$.

Theorem 1.24 (Regularity Lemma, Theorem 1.3. in [18])**.** *For every positive integer N there exists a constant $c = c(N) > 0$ with the following property. Let $0 < \epsilon < \frac{1}{2}$ and let E be a semi-algebraic relation with complexity at most N. Then P has an equitable partition $P = P_1 \cup \cdots \cup P_m$ into $m \leq (1/\epsilon)^c$ parts whose size differ by at most 1 and such that all but an ϵ-fraction of the k-tuples of parts (P_1) is homogeneous for E.*

1.9.3 Limits of semi-algebraic relations

A *box approximation* $((P_i)_{i=1}^M, S, f)$ of a kernel (J, μ, W) is a measurable partition $(P_i)_{i=1}^M$ of J along with a set $S \subseteq \{ P_{i_1} \times P_{i_2} \times P_{i_3} \mid 1 \leq i_1, i_2, i_3 \leq M \}$ and a function $f : S \to \mathbf{R}$ that satisfies the property that $W(a, b, c) = f(B)$ for every $B \in S$ and $(a, b, c) \in B$. The *error* of this approximation is the value $\sum_{B \notin S} \mu^3(B)$, where the sum is taken on the boxes of the form $P_{i_1} \times P_{i_2} \times P_{i_3}$ that are not in S. The *number of parts* of $((P_i)_{i=1}^M, S, f)$ is M. The box approximation $((P_i)_{i=1}^M, S, f)$ is *equitable* if $\mu(P_i) = \mu(P_j)$ for every $(i, j) \in \{1, \dots, M\}^2$. If $([0, 1[, \Lambda_1, W)$ is a kernel on $[0, 1[$, a box approximation $((P_i)_{i=1}^M, S, f)$ of W is *polished* if it is equitable and each part of the partition is an interval of the form $[a, b[$ with $0 \leq a < b < 1$. Note that if $((P_i)_{i=1}^M, S, f)$ is a polished box approximation then $P_i = [\frac{i-1}{M}, \frac{i}{M}[$ up to a reordering.

For a decreasing function $M : \mathbf{R}_*^+ \to \mathbf{N}$, a kernel (J, μ, W) is *M-regular* if for every $\epsilon > 0$, this kernel has a box approximation with $M(\epsilon)$ parts and error at most ϵ.

Every kernel (J, μ, W) is equivalent to a kernel $([0, 1], \Lambda_1, W')$ on an interval in the sense that $\mathrm{d}(W, W') = 0$. In the case where (J, μ) is a Borel space, this is direct consequence of Theorem A.9 in [29], which gives a measure preserving bijection between the unit interval and every Borel probability space. A proof of this fact in the general case for kernels of two variables for instance appears in the proof of Theorem 7.1 in [29] given by Svante Janson. The argument extends directly to functions of three variables, and, for that matter, to any finite number of variables. In this section, we therefore focus on kernels on intervals.

Let P be a set of size n and g be a function from P^k to a finite set F. Let μ_P^u be the uniform probability measure on P. We associate to g by a graphon $([0, 1[, \Lambda_1, W_g)$, by $W_g(x, y, z) = f(i, j, k)$ whenever $x \in [\frac{i-1}{n}, \frac{i}{n}[$, $y \in [\frac{j-1}{n}, \frac{j}{n}[$ and $z \in [\frac{k-1}{n}, \frac{k}{n}[$.

Proposition 1.25. *If for $i \in \{1, 2\}$, $([0, 1[, \Lambda_1, W_i)$ is a kernel that has a box approximation with M_i parts and error ϵ_i, then the kernel $([0, 1[, \Lambda_1, W_1 + W_2)$*

has a box approximation with M_1M_2 *parts and error* $\epsilon_1+\epsilon_2$.

Proof. For $i \in \{0,1\}$, let $((P^i_j)_{j=1}^{M_i}, S_i, f_i)$ be a box approximation of $([0,1[, \Lambda_1, W_i)$ with error ϵ_i. For $j_1 \in [M_1]$ and $j_2 \in [M_2]$, define $P_{(j_1,j_2)} := P^1_{j_1} \cap P^2_{j_2}$. This yields a partition of $[0,1[$ with M_1M_2 parts. Let S be the set of boxes $P_{(j_1,j_2)} \times P_{(k_1,k_2)} \times P_{(\ell_1,\ell_2)}$ such that $P_{j_1} \times P_{k_1} \times P_{\ell_1}$ is in S_1 and $P_{j_2} \times P_{k_2} \times P_{\ell_2}$ is in S_2. Note that every box $B \in S$ is contained is a box $B_i \in S_i$, so W_i is constant on B for $i \in \{1,2\}$, and hence W_1+W_2 is constant on B. It remains to define the function $f : S \to \mathbf{R}$ for every $B \in S$ as the (unique) value of W_1+W_2 in B. This defines a box approximation $((P_{(j_1,j_2)})_{j_1 \in [M_1] j_2 \in [N_2]}, S, f)$ of $([0,1[, \Lambda_1, W_1+W_2)$

Let us prove the upper bound on the error of this box approximation. It follows from the definition of S that $\bigcup_{B \in S} B = \left(\bigcup_{B_1 \in S_1} B_1\right) \cap \left(\bigcup_{B_2 \in S_2} B_2\right)$. Consequently,

$$\bigcup_{B \notin S} B = \left(\bigcup_{B_1 \notin S_1} B_1\right) \cup \left(\bigcup_{B_2 \notin S_2} B_2\right).$$

It follows that $\sum_{B \notin S} \Lambda_1^3(B) \leq \sum_{B_1 \notin S_1} \Lambda_1^3(B_1) + \sum_{B_2 \notin S_2} \Lambda_1^3(B_2) \leq \epsilon_1+\epsilon_2$. This conclude the proof of the lemma. □

Theorem 1.24 implies the following statement.

Lemma 1.26 (Regularity lemma, kernel version)**.** *For every set* P*, if* g *is a semi-algebraic function of complexity* k *from* P^3 *to a finite set* F*, the kernel* $([0,1[, \Lambda_1, W_g)$ *is* M*-regular, for some function* M *depending only on* k *and* $|F|$.

Proof. We first prove this result in the case where g is the indicator function $\mathbf{1}_E$ of a semi-algebraic set $E \subseteq P^3$ with complexity k. Let $P = P_1, \dots, P_m$ be the partition given by Theorem 1.24 applied to E. Let $S_0 \subseteq [m]^3$ be the set of triples $(i,j,k) \in [m]^3$ such that $P_i \times P_j \times P_k$ is homogeneous for E. We know that $|S_0| \geq (1-\epsilon)m^3$ and $m \leq \epsilon^{-c(k)}$, where $c(k)$ is the constant coming from Theorem 1.24. Note that the triple $((Q_i)_{i=1}^m, S, f)$, where $Q_i = \bigcup_{j \in P_i} [\frac{j-1}{n}, \frac{j}{n}[$, $S = \{\ P_i \times P_j \times P_k \mid (i,j,k) \in S_0\ \}$ and the value of f on $Q_i \times Q_j \times Q_k \in S$ is the value of $\mathbf{1}_E$ on $P_i \times P_j \times P_k$, is a box approximation of $([0,1[, \Lambda_1, W_{\mathbf{1}_E})$ with at most $m \leq \epsilon^{-c(k)}$ parts. Since $\Lambda_1(Q_i \times Q_j \times Q_k) = \frac{|P_i|}{m}\frac{|P_j|}{m}\frac{|P_k|}{m} \leq \frac{(m+1)^3}{m^3} \leq 2^3$ for every $(i,j,k) \in [m^3]$, the error of $((Q_i)_{i=1}^m, S, f)$ is at most 8ϵ. This suffices to finish the proof in the case $g = \mathbf{1}_E$.

For the general case, we write

$$g = \sum_{e \in F} \mathbf{1}_{g^{-1}(e)}.$$

It directly follows from the first part and Proposition 1.25 that g has a box-approximation with $\epsilon^{-c(k)|F|}$ parts and error $8|F|\epsilon$. This suffices to prove the theorem for the function $M(\epsilon') = \left(\frac{8|F|}{\epsilon}\right)^{c(k)|F|}$. □

The box approximation $((P'_i)_{i=1}^{M'}, S', f')$ is a *refinement* of the box approximation $((P_i)_{i=1}^{M}, S, f)$ if

- for every $i \in [M']$ there is $j \in [M]$ such that $P'_i \subseteq P_j$;
- for every $(i, j, k) \in [M']^3$ and $B \in S$ such that $P_i \times P_j \times P_k \subseteq B$, it holds that $P_i \times P_j \times P_k \in S'$; and
- $f(B) = f'(B')$ whenever $B' \subseteq B$ for $B \in S$ and $B' \in S'$.

The next lemma allows us to refine a (polished) box approximation. It is formulated for kernels on $[0, 1[$, since every kernel is equivalent to a kernel on an interval. We use the right open interval $[0, 1[$, so it can be partitioned into intervals of the same form $[a, b[$.

Lemma 1.27. *Let $M : \mathbf{R}_+^* \to \mathbf{N}$ be a decreasing function, $m \in \mathbf{N}$, and $\epsilon > 0$. There exists a constant $k = k(M, m, \epsilon)$ such that for every M-regular kernel $([0, 1[, \Lambda_1, W)$ and every polished box approximation $((P_i)_{i=1}^{m}, S, f)$ of W, there are*

- *a measure-preserving function $\rho : [0, 1[\to [0, 1[$ and*
- *a polished box approximation $((P'_i)_{i=1}^{k}, S', f')$ of W^ρ with k parts and error at most ϵ.*

Proof. Let $([0, 1[, \Lambda_1, W)$ be an M-regular kernel kernel and $((P_i)_{i=1}^{m}, S, f)$ a polished box approximation of $([0, 1[, \Lambda_1, W)$ as in the statement of the lemma. Since W is M-regular, there is a box approximation $((Q_i)_{i=1}^{n}, T, g)$ of W with $n = M(\frac{\epsilon}{2})$ parts and error $\frac{\epsilon}{2}$. Note that for $i \in \{1, \ldots, n\}$, the part Q_i may not be an interval.

We first construct a measure preserving function ρ such that for every $i \in \{1, \ldots, m\}$, $(\rho(P_i \cap Q_j))_{j=1}^{n}$ is a partition of P_i into intervals. Write $P_i = [a_i, a_{i+1}[$ where $a_i = \frac{i-1}{m}$ and set $b_{i,j} = a_i + \sum_{\ell=1}^{j-1} \Lambda_1(P_i \cap Q_\ell)$ for $i \in \{1, \ldots, m\}$ and $j \in \{1, \ldots, n\}$. For each $i \in \{1, \ldots, m\}$ and $j \in \{1, \ldots, n\}$, the function α is defined on $P_i \cap Q_j$ by

$$\alpha(x) = b_{i,j} + \Lambda_1([a_i, x] \cap Q_j).$$

Note that α maps $P_i \cap Q_j$ to a subset of $[b_{i,j}, b_{i,j+1}]$. The function α is measure preserving. Moreover, the restriction of α to $P_i \cap Q_j$ is increasing and measure preserving, therefore the function α is invertible almost everywhere on $P_i \cap Q_j$ for $i \in \{1, \ldots, m\}$ and $j \in \{1, \ldots, n\}$. Since the images of the form $\alpha(P_i \cap Q_j)$ have pairwise intersections of Lebesgue measure 0, the function α is invertible almost everywhere on $[0, 1[$, that is, there is $N \subseteq [0, 1[$ of 0 measure and a measure preserving function $\rho : [0, 1[\setminus N \to [0, 1[$ such that $\alpha \circ \rho(x) = x$ whenever $x \notin N$. Write $I_{i,j} = [b_{i,j}, b_{i,j+1}[$. We may assume that N contains $\{ b_{i,j} \mid 1 \leq i \leq m, 1 \leq j \leq n \}$, so that $\rho(I_{i,j} \setminus N) \subseteq P_i \cap Q_j$ for $i \in \{1, \ldots, m\}$ and $j \in \{1, \ldots, n\}$. We extend ρ to $[0, 1[$ by sending $\rho(x)$ arbitrarily to $P_i \cap Q_j$ whenever $x \in I_{i,j} \cap N$, so $\rho(I_{i,j}) \subseteq P_i \cap Q_j$ for $i \in \{1, \ldots, m\}$ and $j \in \{1, \ldots, n\}$.

Define $k = 6m \cdot \lceil \frac{n}{\epsilon} \rceil$. Let $(P'_i)_{i=1}^k$ be the partition defined by $P'_i = [\frac{i-1}{k}, \frac{i}{k}[$. Note that since m divides k, the family

$$\left\{ P'_j \;\middle|\; \frac{k}{m}(i-1) + 1 \leq j \leq \frac{k}{m} i \right\}$$

forms a partition of P_i for $i \in \{1, \ldots, m\}$. Let S' be the set of boxes $B' = P'_{i_1} \times P'_{i_2} \times P'_{i_3}$ such that W^ρ is constant on B'. For such a set B', define $f'(B')$ as the value of W^ρ on B'.

The triple $((P'_i)_{i=1}^k, S', f')$ is a polished box approximation of W^ρ. Moreover, $((P'_i)_{i=1}^k, S', f')$ is a refinement of $((P_i)_{i=1}^m, S, f)$. Indeed, if $B' = P'_{i_1} \times P'_{i_2} \times P'_{i_3}$ is included in $B \in S$, then W is constant and equal to $f(B)$ on B. Since $\rho^3(B') \subseteq \rho^3(B) \subseteq B$, the function W^ρ is constant on B', so $B' \in S'$ and $f'(B') = f(B)$.

It remains to show that $((P'_i)_{i=1}^k, S', f')$ has error at most ϵ. Let $(i_1, i_2, i_3) \in \{1, \ldots, k\}^3$ such that P_{i_q} is included in some I_{a_q, b_q} for $q \in \{1, 2, 3\}$. If $Q_{b_1} \times Q_{b_2} \times Q_{b_3} \in T$, then W is constant on $Q_{b_1} \times Q_{b_2} \times Q_{b_3} \supseteq \rho(B)$ where $B = P'_{i_1} \times P'_{i_2} \times P'_{i_3}$. It follows that W^ρ is constant on B, so $B \in S'$. Let U_1 be the union of the remaining boxes $P'_{i_1} \times P'_{i_2} \times P'_{i_3}$ of this type, that is such that P_{i_q} is included in an interval I_{a_q, b_q} for each $q \in \{1, 2, 3\}$ and $Q_{b_1} \times Q_{b_2} \times Q_{b_3} \notin T$. By definition $\rho(U_1)$ is contained in the union V of the boxes of the form $Q_{b_1} \times Q_{b_2} \times Q_{b_3}$ that are not in T. It follows that $\Lambda_1^3(U_1) = \Lambda_1^3(\rho(U_1)) \leq \Lambda_1^3(V) \leq \frac{\epsilon}{2}$.

The number of indices $i \in \{1, \ldots, k\}$ such that P'_i intersects more than one interval of $\{ I_{p,j} \mid 0 \leq p \leq m, 0 \leq j \leq n \}$ is at most mn. Indeed, such an interval P'_i must contain at least one bound $b_{i,j}$ for some $i \in \{1, \ldots, m\}$ and $j \in \{1, \ldots, n\}$. Therefore, it remains at most $3mnk^2$ boxes not included in U_1 and not in S'. The union U_2 of these boxes has Λ_1^3-measure at most $3\frac{mn}{k}$.

Consequently, the error of the box approximation $((P'_i)_{i=1}^k, S', f')$ is at most

$$\frac{\epsilon}{2} + 3\frac{mn}{k} \leq \frac{\epsilon}{2} + \frac{\epsilon}{2} = \epsilon.$$

$\square$

Theorem 1.28. *Let $M : \mathbf{R}_+^* \to \mathbf{N}$ be a decreasing function and $F \subseteq \mathbf{R}$ be a finite set. For every sequence $(J_n, \mu_n, W_n)_{n \in \mathbf{N}}$ of M-regular kernels with values in F, there are measure preserving functions $\rho_n : [0,1[\to J_n$, an extraction $\psi : \mathbf{N} \to \mathbf{N}$ and a kernel $([0,1[, \Lambda_1, W)$ such that*

$$\|W_{\psi(n)}^{\rho_n} - W\|_1 \underset{n \to \infty}{\longrightarrow} 0.$$

Proof. For each $m \geq 0$, we construct a subsequence $(J_{\phi_m(n)}, \mu_{\phi_m(n)}, W_{\phi_m(n)}^{\rho_m^n})_{n \in \mathbf{N}}$ where $\rho_m^n : [0,1[\to [0,1[$ is measure preserving and a polished box approximation $(\mathcal{P}_m, S_m, f_m)$ of the kernels of this subsequence with error $\epsilon_m = 1/m$, in such a way that $(\mathcal{P}_{m+1}, S_{m+1}, f_{m+1})$ is a refinement of $(\mathcal{P}_m, S_m, f_m)$.

Let $\mathcal{P}_0$ be the trivial partition $\{[0,1[\}$, S_0 be the empty set, $f : \varnothing \to \mathbf{R}$ be the empty function, ϕ_0 be the identity function and for every $n \in \mathbf{N}$, let ρ_0^n be the identity of $[0,1[$.

For $m \geq 0$, we construct $(\mathcal{P}_{m+1}, S_{m+1}, f_{m+1})$, ϕ_{m+1} and $\rho^n_{m+1} : [0,1[\to [0,1[$ from $(\mathcal{P}_m, S_m, f_m)$, ϕ_m and $\rho^n_m : [0,1[\to [0,1[$. For $n \in \mathbf{N}$, by Lemma 1.27 applied on $W_{\phi_m(n)}$ and $\mathcal{P}_m$ there is a measure preserving function ρ^n_{m+1} and a polished box approximation $(\mathcal{P}^n_{m+1}, S^n_{m+1}, f^n_{m+1})$ of the kernel $(J_m, \mu_m, W^{\rho^n_{m+1}}_{\phi_m(n)})$ with $k_m = k(M, m, \epsilon_m)$ parts and error at most ϵ_{m+1}.

Note that $\mathcal{P}^n_{m+1}$ is the partition $\mathcal{P}_{m+1} = \left\{ [\frac{i}{k_m}, \frac{i+1}{k_m}[\;\middle|\; 0 \leq i \leq k_m - 1 \right\}$ because $\mathcal{P}^n_{m+1}$ is a polished partition of P_m with size k_m. As a consequence, the set $\left\{ S^n_{m+1} \;\middle|\; n \in \mathbf{N} \right\}$ is finite and $\left\{ f^n_{m+1} \;\middle|\; n \in \mathbf{N} \right\}$ is finite because f^n_{m+1} has a finite domain and takes values in the finite set F. There exists an extraction σ such that the sequences $f^{\sigma(n)}_{m+1}$ and $S^{\sigma(n)}_{m+1}$ are constant. Let f_{m+1} and S_{m+1} be their respective values. Set $\phi_{m+1} := \phi_m \circ \sigma$ and $\rho_{m+1} := \rho^{\sigma(m+1)}_{m+1}$. Note that $(\mathcal{P}_{m+1}, S_{m+1}, f_{m+1})$ is a box approximation of $W^{\rho_{m+1}}_{\phi_{m+1}(m+1)}$ with error at most ϵ_{m+1}.

For $n \in \mathbf{N}$, define $\psi(n) = \phi_n(n)$ and consider the sequence $U_n = W^{\rho_n}_{\psi(n)}$. By construction, $(\mathcal{P}_n, S_n, f_n)$ is a polished box approximation of U_n with error at most ϵ_n.

We claim that if (a,b,c) belongs to a box $B \in S_{n_0}$ for some $n_0 \in \mathbf{N}$ then $(U_n(a,b,c))_{n \in \mathbf{N}}$ is eventually constant. Indeed, for every $n \geq n_0$, $(\mathcal{P}_n, S_n, f_n)$ is a refinement of $(\mathcal{P}_{n_0}, S_{n_0}, f_{n_0})$. It follows from the definition that $(a,b,c) \in \bigcup_{B \in S_{n_0}} B \subseteq \bigcup_{B' \in S_n} B'$. Therefore there exists $B' \in S_n$ with $(a,b,c) \in B'$ and $B' \subseteq B$. Consequently, $U_n(a,b,c) = f_n(B') = f_{n_0}(B) = U_{n_0}(a,b,c)$.

The measure $\Lambda^3_1([0,1[^3 \setminus \bigcup_{B \in S_n} B)$ is less than ϵ_n for every $n \in \mathbf{N}$ and therefore converges to 0 when n goes to infinity. This proves that U_n converges almost everywhere to a kernel $([0,1[, \Lambda_1, W)$. □

We deduce from Theorem 1.28 that a limit of order types can always be represented by a kernel.

Corollary 1.29. *Every limit ℓ of order types can be represented by a kernel $([0,1], \Lambda_1, W)$ in the sense that*

$$p(\omega, W) = \ell(\omega) \text{ for every } \omega \in \mathcal{O}.$$

Proof. Let $(\omega_n)_{n \in \mathbf{N}}$ be a sequence of order types that converges to ℓ. For each $n \in \mathbf{N}$, let W_{ω_n} be the kernel associated to ω_n. Recall that this kernel is $([0,1[, \Lambda_1, W_{\omega_n})$, where $W_{\omega_n}(x,y,z) = \omega_n(i,j,k)$ whenever $x \in [\frac{i-1}{|\omega_n|}, \frac{i}{|\omega_n|}[$, $y \in [\frac{j-1}{|\omega_n|}, \frac{j}{|\omega_n|}[$, and $z \in [\frac{k-1}{|\omega_n|}, \frac{k}{|\omega_n|}[$. It follows from Lemma 1.26 that these kernels are M-regular, for a common function M.

By Theorem 1.28, there is a sequence $(\rho_n)_{n \in \mathbf{N}}$ of measure-preserving function and an extraction $\phi : \mathbf{N} \to \mathbf{N}$ such that $W^{\rho_n}_{\omega_{\phi(n)}}$ tends to a kernel W for $\|\cdot\|_1$.

For every $\omega \in \mathcal{O}$, we know that $p(\omega, W_{\omega_n}) = p(\omega, \omega_n) + o(1)$. Consequently, $p(\omega, W^{\rho_n}_{\omega_{\phi(n)}}) = p(\omega, W_{\omega_{\phi(n)}}) = p(\omega, \omega_n) + o(1)$. Moreover, $|p(\omega, W^{\rho_n}_{\omega_{\phi(n)}}) - p(\omega, W)| \leq \|W^{\rho_n}_{\omega_{\phi(n)}} - W\|_1 = o(1)$. As a consequence, $p(\omega, \omega_n) \underset{n\to\infty}{\longrightarrow} \ell(\omega)$. □

1.10 Rigidity

In this section, we investigate the following question: What does the set of measures realizing a given limit look like?

For arbitrary probability measures, this set can be large. As mentioned in Example 1.5, the limit $\ell_\diamond$ of the sequence $(C_n)_{n\in\mathbf{N}}$ of order types in convex position is realized by every atom-free measure with support in convex position.

This allows arbitrarily disconnected support, as wells as support of any Hausdorff dimension between 0 and 1 (consider for instance a Cantor set on $[0,1]$ with that dimension and map the interval to the circle). The limit $\ell_\diamond$ is exceptionally simple, so one may wonder if this variety of realizations is also exceptional. The limit ℓ_E described in Section 1.10.1 gives a different example of limit that is realized by a wide family of measures that are not pairwise projectively equivalent.

However, it seems that for measure with *big enough* support, the situation is radically different. Theorem 1.39 answers this question for measures μ whose support has non-empty interior: there is, up to a spherical transform, only one measure that realizes ℓ_μ. Theorem 1.37 gives a similar conclusion on a different hypothesis. It states that two measures with Hausdorff dimension larger than 1 cannot generate the same measure unless they are projectively equivalent.

1.10.1 A limit with an interesting realization space

We present an example of limit with an interesting realization space, than looks less trivial than the limit $\ell_\diamond$.

Let us describe it via a family of Cantor measures that realize it in the plane (see to Figure 1.17). Fix some parameters a and b with $0 < b < a < \frac{1}{2}$, and define the rectangles $R = [0,1]^2$, $R_0 = [0,a] \times [0,b]$ and $R_1 = [1-a,1] \times [1-b,1]$. Let ϕ_0 (resp. ϕ_1) denote the homothetic transform fixing $(0,0)$ (resp. $(1,1)$) and mapping R to R_0 (resp. to R_1). To any word $w = i_1 i_2 \dots i_n \in \{0,1\}^*$ we associate a set

$$R_w = \phi_{i_n} \circ \phi_{i_{n-1}} \circ \cdots \circ \phi_{i_1}(R)$$

and let $m_{a,b}$ be the probability measure such that $m_{a,b}(R_w) = \frac{1}{2^{|w|}}$ for every $w \in \{0,1\}^*$. Remark that $R_w \subseteq R_v$ if and only if v is a prefix of w. Letting $A_n := \bigcup_{w\in\{0,1\}^n} R_w$ for $n \geq 0$, the support of $m_{a,b}$ writes $A = \bigcap_n A_n$. The measures $m_{a,b}$ with $b \leq (1-2a)(1-2b)a$ all realize the same limit of order types (Lemma 1.31), which we denote by ℓ_E. Moreover, the Hausdorff dimension of $\operatorname{supp} m_{a,b}$ is $\frac{\ln 2}{-\ln a}$ (Lemma 1.32). (by an application of [43, Theorem 2.12]). This construction shows that the Hausdorff dimension of the support of measures realizing ℓ_E is quite free.

Theorem 1.30. *For any $t \in \,]0,1[$, the limit ℓ_E can be realized by a measure whose support has Hausdorff dimension t.*

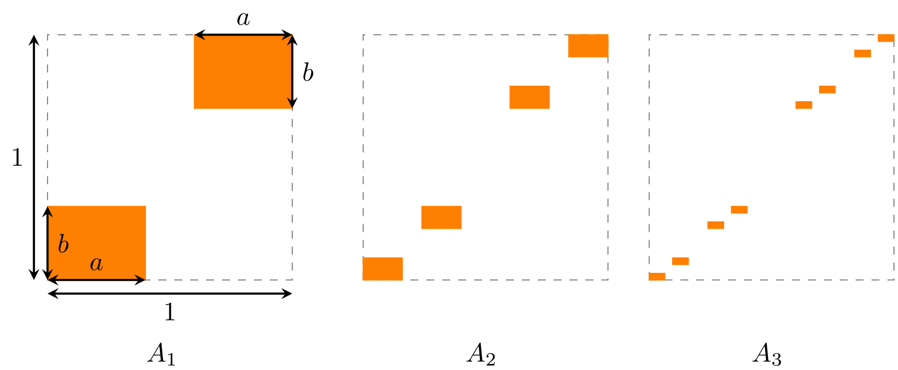


Figure 1.17 – Definition of ℓ_E.

Analysis of ℓ_E

We begin with a combinatorial description of the limit of order types ℓ_E as a kernel $(E, \mu_{\{0,1\}}^{\mathbf{N}}, \chi_E)$ on the measured space $E = \{0,1\}^{\mathbf{N}}$ with the coin-tossing distribution $\mu_{\{0,1\}}^{\mathbf{N}}$. For $u, v \in E$, let $u \wedge v$ denote the longest common prefix of u and v and let $\prec_{\text{lex}}$ be the lexicographic order on E. We first suppose that $u \prec_{\text{lex}} v \prec_{\text{lex}} w$, then define $\chi_E(u,v,w) := 0$ if u, v and w are not pairwise distinct, $\chi_E(u,v,w) := 1$ if $|u \wedge v| < |v \wedge w|$ and $\chi_E(u,v,w) := -1$ otherwise. The other cases are defined so that χ_E is antisymmetric, i.e. such that $\chi_E(u,v,w) = \operatorname{sgn}(\sigma)\chi_E(\sigma(u), \sigma(v), \sigma(w))$, where σ is the permutation of $\{u,v,w\}$ such that $\sigma(u) \prec_{\text{lex}} \sigma(v) \prec_{\text{lex}} \sigma(w)$ and $\operatorname{sgn}(\sigma) \in \{-1, 1\}$ is the signature of σ.

Let ℓ_E be the limit represented by the kernel $(E, \mu_{\{0,1\}}^{\mathbf{N}}, \chi_E)$. Recall that for any order type ω of size k, the number $\ell_E(\omega)$ is therefore the probability that the restriction of χ to k random elements of E chosen independently from μ_E has order type ω.

Lemma 1.31. *If $0 < b \leq a < \frac{1}{2}$ and $b \leq (1-2a)(1-2b)a$, then $\ell_{m_{a,b}} = \ell_E$.*

Proof. The measure $m_{a,b}$ is the image of the probability measure $\mu_{\{0,1\}}^{\mathbf{N}}$ on $\{0,1\}^{\mathbf{N}}$ by the function Ψ that assigns to $w \in \{0,1\}^{\mathbf{N}}$ the unique point of $\bigcap_{w_v} R_{w_v}$ where the intersection is taken over all prefixes w_v of w.

We claim that every point of $A \cap R_1$ lies above any line spanned by two points of $A \cap R_0$ provided that

$$b \leq (1-2a)(1-2b)a.$$

Since A is stable by the symmetry of center $(\frac{1}{2}, \frac{1}{2})$, it then follows that every point of $A \cap R_0$ lies below any line spanned by two points of $A \cap R_1$

In this case, $\ell_{m_{a,b}}$ is then fully determined. Let $\Psi(u)$, $\Psi(v)$ and $\Psi(w)$ be three pairwise distinct points in A with $u, v, w \in \{0,1\}^{\mathbf{N}}$ and assume that $u \prec_{\text{lex}} v \prec_{\text{lex}} w$. If $|u \wedge v| < |v \wedge w|$, set $p = u \wedge v$ and $pq = v \wedge w$. Since $u \prec_{\text{lex}} v$, the word $p.0$ is a prefix of u and $p.1$ is a prefix of v, and therefore of w. It follows that $\Psi(u) \in R_{p.0}$ and $\Psi(v), \Psi(w) \in R_{p.1}$, moreover, $\Psi(v)$ has smallest abscissa than $\Psi(w)$ because $v \prec_{\text{lex}} w$. Consequently, $\chi(\Psi(u), \Psi(v), \Psi(w)) = 1 = \chi_E(u, v, w)$. The proof of $\chi(\Psi(u), \Psi(v), \Psi(w)) = \chi_E(u, v, w)$ when $|u \wedge v| > |v \wedge w|$ is similar so we omit it.

It remains to prove the claimed property. For two distinct points $x, y \in A$, let $\alpha(x, y)$ be the angle between the line $h(x, y)$ and the abscissa axis (so $\alpha(x, y)$ is defined modulo π). If $x \in R_0$ and $y \in R_1$, then

$$1 - 2b \leq \tan \alpha(x, y) \leq \frac{1}{1 - 2a}.$$

as the minimum is obtained by $x = (0, b)$ and $y = (1, 1 - b)$ and the minimum is obtained by $x = (a, 0)$ and $y = (1 - a, 1)$ (See Figure 1.18). The application of a function ϕ_i for $i \in \{1, 2\}$ acts as follows: $\tan \alpha(\phi_i(x), \phi_i(y)) = \frac{b}{a} \tan \alpha(x, y)$ Since $\frac{b}{a} \leq 1$, it further holds that $\tan \alpha(\phi_i(x), \phi_i(y)) \leq \tan \alpha(x, y)$. By iterating this property, it follows that $\tan \alpha(x, y) \leq \frac{1}{1-2a}$ for every $x, y \in A$ such that $x \neq y$ and x has smaller abscissa than y.

Let $x, y \in A \cap R_0$ and $z \in A \cap R_1$ such that x has smaller abscissa than y. Note that z lies above the line $h(x, y)$ if and only if $\alpha(x, y) \leq \alpha(x, z)$, where the values of the angles are taken in $]-\frac{\pi}{2}, \frac{\pi}{2}[$ (See Figure 1.18), which is equivalent to $\tan \alpha(x, y) \leq \tan \alpha(x, z)$. Since $\tan \alpha(x, y) \leq \frac{b}{a} \frac{1}{1-2a}$ and $\tan \alpha(x, z) \geq 1 - 2b$, it suffices than $\frac{1}{1-2a} \frac{b}{a} \leq 1 - 2b$, i.e. $(1 - 2a)(1 - 2b)a$. This proves the claimed property. ☐

Lemma 1.32. *If* $0 < b < a < \frac{1}{2}$ *then* $\operatorname{supp} m_{a,b}$ *has Hausdorff dimension* $\frac{\ln 2}{-\ln a}$.

Proof. Let $\beta = \frac{\ln 2}{-\ln a}$. Set

$$I_\delta^\alpha := \inf \left\{ \sum_i (\operatorname{diam} B_i)^\alpha \;\middle|\; A \subseteq \bigcup_i B_i, \operatorname{diam}(B_i) \leq \delta \right\}.$$

By definition, $\mathcal{H}^\alpha(X) = \lim_{\delta \to 0} I_\delta^\alpha$. For any $\delta > 0$ there exists $n = n(\delta)$ such that $\operatorname{diam} R_w < \delta$ for every $w \in \{0, 1\}^n$. The family $\{R_w\}_{|w|=n}$ covers A, so $I_\delta^\beta \leq \sum_{w \in \{0,1\}^n} (\operatorname{diam} R_w)^\beta = 2^n (a^{2n} + b^{2n})^{\beta/2} \leq 2^n (2a^{\beta n}) \leq 2$. Thus, $\mathcal{H}^\beta(A)$ is finite and further $\dim_H(A) \leq \beta$. For the reverse inequality, consider the projection of A on the first coordinate. This is a Lipschitz map so it cannot increase Hausdorff dimension. The image is the Cantor set $\mathcal{C}_a$ of parameter a. See Example 1.11 for a construction of Cantor sets. This set $\mathcal{C}_a$ is known to have Hausdorff dimension $\frac{\ln 2}{-\ln a} = \beta$ (See [43, Theorem 2.12]). It follows that $\dim_H \operatorname{supp} m_{a,b} \geq \beta$. ☐

Proof of Theorem 1.30. The theorem follows from Lemmas 1.31 and 1.32. Indeed, for every $a \in]0, \frac{1}{2}[$, there is $b \in]0, a[$ small enough to satisfy $b \leq (1 - 2a)(1 - 2b)a$. By Lemma 1.31, the measure $m_{a,b}$ realizes ℓ_E and by Lemma 1.32, $\dim_H \operatorname{supp} m_{a,b} = \frac{\ln 2}{-\ln a}$. Note that $\frac{\ln 2}{-\ln a}$ ranges in $]0, 1[$ when a ranges in $]0, \frac{1}{2}[$. The theorem follows. ☐

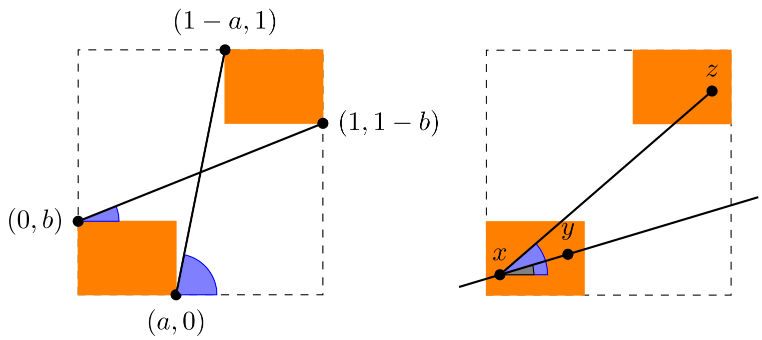


Figure 1.18 – Proof of Lemma 1.31. Left: Minimal and maximal value for $\alpha(x,y)$ when $x \in R_0$ and $y \in R_1$. Right: $z \in R_1$ is above $h(x,y)$ if $\alpha(x,y) < \alpha(x,z)$.

1.10.2 Cantor dust

Example 1.11 (Cantor dust)**.** Let p be a real number with $0 < p < \frac{1}{2}$. The *cantor set* $\mathcal{C}_p$ of parameter p is a subset of the interval $[0,1]$ defined inductively as follows.

Set $T_1^0 = [0,1]$. Now for $n \geq 0$ and $i \in \mathbf{N}$ with $1 \leq i \leq 2^n$, the sets T_{2i-1}^{n+1} and T_{2i}^{n+1} are defined from the interval $T_i^n = [a,b]$ by $T_{2i-1}^{n+1} = [a,(1-p)a+pb]$ and $T_{2i}^{n+1} = [pa+(1-p)b, b]$, so that $\Lambda_1(T_{2i-1}^{n+1}) = \Lambda_1(T_{2i}^{n+1}) = p\cdot\Lambda_1(T_i^n)$. The cantor set is $\mathcal{C}_p = \bigcap_{n\in\mathbf{N}} \bigcup_{i=1}^{2^n} T_i^n$. The cantor set is the support a probability measure ν_p on $\mathbf{R}$ uniquely defined by the property $\nu_p(T_i^n) = \frac{1}{2^n}$ for $i \in \{1,\dots,2^n\}$. The function

$$\Phi_p : \left\{ \begin{array}{ccc} \{0,1\}^{\mathbf{N}} & \to & \mathcal{C}_p \\ (a_i)_{i\in\mathbf{N}} & \mapsto & (1-p)\sum_{i\in\mathbf{N}} a_i \cdot p^i \end{array} \right.$$

is a bijection between the Cantor set C_p and the set $\{0,1\}^{\mathbf{N}}$ of sequences with values in $\{0,1\}$. The Hausdorff dimension of the Cantor set of parameter p is known to be $\frac{\log 2}{-\log p}$ (See Theorem 2.12 in [43]).

We now focus on the two-dimensional generalization of the Cantor set. The (two-dimensional) *cantor dust* $\mathcal{C}_p^2 \subseteq \mathbf{R}^2$ is the Cartesian product of the cantor set $\mathcal{C}_p$ with itself. See Figure 1.19 for an illustration of the two-dimensional Cantor dust.

The Cantor dust inherit several properties of the cantor set. In particular, $\mathcal{C}_p^2$ is the support on $\mathbf{R}^2$ of the probability measure $\mu_p = (\nu_p)^2$. The Cantor dust $\mathcal{C}_p^2$ has Hausdorff dimension $2\dim_H(C_p^2) = \frac{\log 4}{-\log p}$. The function $\Phi_p^2 : (\{0,1\}^{\mathbf{N}})^2 \to \mathcal{C}_p^2$ defined by $\Phi_p^2((a_n)_{n\in\mathbf{N}},(b_n)_{n\in\mathbf{N}}) = (\Phi_p((a_n)_{n\in\mathbf{N}}),\Phi_p((b_n)_{n\in\mathbf{N}}))$ is a bijection.

Note also that μ_p does not charge lines. For each $p \in \,]0,\frac{1}{2}[$, the probability measure μ_p gives a geometric limit of order types ℓ_{μ_p}.

As the different sets $\mathcal{C}_p^2$ share a common structure, one may wonder whether these limits are the same for different values of p. This question can be stated

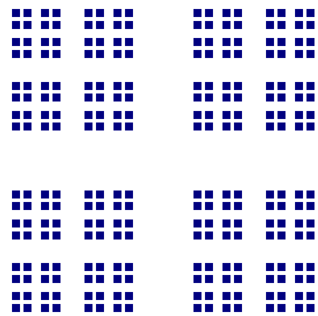


Figure 1.19 – Cantor dust.

as follows.

Question 1.1. For which pairs of parameters p_1 and p_2 with $0 < p_1 < p_2 < \frac{1}{2}$ do we have $\ell_{\mu_{p_1}} = \ell_{\mu_{p_2}}$?

For the Cantor dust sets with large enough Hausdorff dimension, the answer will be given by Theorem 1.37. Indeed, this result implies that $\ell_{\mu_{p_1}} \neq \ell_{\mu_{p_2}}$ as soon as the Hausdorff dimension of the measure μ_{p_1}[2] is strictly larger than one, which happens when $\dim_H \mathcal{C}^2_{p_1} > 1$, i.e. when $p_1 > \frac{1}{4}$.

What happens for Cantor sets of small dimensions is unknown to me. Corollary 1.23 shows that the equality $\ell_{\mu_{p_1}} = \ell_{\mu_{p_2}}$ is equivalent to the existence of an (almost everywhere) chirotope-preserving and measure-preserving bijection between $\operatorname{supp} \mu_{p_1} = \mathcal{C}^2_{p_1}$ and $\operatorname{supp} \mu_{p_2} = \mathcal{C}^2_{p_2}$. A good candidate for such a bijection is the function $\Phi^2_{p_2} \circ (\Phi^2_{p_1})^{-1}$, that is the function that sends each squares of $\mathcal{C}^2_{p_1}$ to the corresponding square of $\mathcal{C}^2_{p_1}$ (for instance the upper left one to the upper left one). This function $\Phi^2_{p_2} \circ (\Phi^2_{p_1})^{-1}$ is a measure preserving bijection from $(\mathcal{C}^2_{p_1}, \mu_{p_1})$ to $(\mathcal{C}^2_{p_2}, \mu_{p_2})$. The question of whether it also preserves chirotopes when p_1 and p_2 are small enough boils down to the following question.

Question 1.2. Is there a positive constant C such that for every power series $P_1, \ldots, P_6$ with coefficients in $\{0, 1\}$, the power series

$$Q = (P_1 - P_5)(P_2 - P_6) - (P_3 - P_5)(P_4 - P_6)$$

has no positive root smaller than C, unless Q is uniformly zero?

I do not even know how to disprove the generalization where Q is instead constructed as $Q = R_1 R_2 - R_3 R_4$ where $R_1, \ldots, R_4$ are any power series with coefficients in $\{-1, 0, 1\}$, nor how to prove it in the particular case where $P_1, \ldots, P_6$ are polynomials.

Proposition 1.33 formalizes the link between questions 1.1 and 1.2.

[2]The Hausdorff dimension *of a measure* is defined page 68.

Proposition 1.33. *For $C \in]0, \frac{1}{2}[$, the function $\Phi^2_{p_2} \circ (\Phi^2_{p_1})^{-1}$ from $\mathcal{C}^2_{p_1}$ to $\mathcal{C}^2_{p_2}$ is chirotope preserving for every $p_1, p_2 \in]0, c[$ if and only if c satisfies the property of Question 1.2.*

Proof. To see this, let $a = (a^1_n, a^2_n)_{n\in\mathbf{N}}$, $b = (a^1_n, a^2_n)_{n\in\mathbf{N}}$ and $c = (c^1_n, c^2_n)_{n\in\mathbf{N}}$ be three elements of $(\{0,1\}^{\mathbf{N}})^2$ (written as sequences of couples instead of couple of sequences for the sake of readability). For $p \in [0, \frac{1}{2}]$, the orientation $\chi(\Phi^2_p(a), \Phi^2_p(b), \Phi^2_p(c))$ of the triangle $(\Phi^2_p(a), \Phi^2_p(b), \Phi^2_p(c))$ is given by the sign of the determinant of the matrix

$$\begin{pmatrix} 1 & \Phi_p(a^1) & \Phi_p(a^2) \\ 1 & \Phi_p(b^1) & \Phi_p(b^2) \\ 1 & \Phi_p(c^1) & \Phi_p(c^2) \end{pmatrix},$$

which can be conveniently rewritten as

$$(\Phi_p(b^1) - \Phi_p(a^1))(\Phi_p(c^2) - \Phi_p(a^2)) - (\Phi_p(c^1) - \Phi_p(a^1))(\Phi_p(b^2) - \Phi_p(a^2)).$$

Taking $P_1 = \Phi_p(b^1)$, $P_2 = \Phi_p(c^2)$, $P_3 = \Phi_p(c^1)$, $P_4 = \Phi_p(b^2)$, $P_5 = \Phi_p(a^1)$ and $P_6 = \Phi_p(a^2)$, this becomes $Q = (P_1 - P_5)(P_2 - P_6) - (P_3 - P_5)(P_4 - P_6)$. Note that for $x \in \{a, b, c\}$ and $i \in \{1, 2\}$, the expression $\Phi_p(x^i) = \sum_{n\in\mathbf{N}} x^i_n p^n$ is a power series with coefficients in $\{0, 1\}$. The constant C satisfies the hypothesis of Proposition 1.33 if and only if for every choice of a, b and c, the sign of Q does not change when p varies in $]0, C[$. Recall that the sign of zero is zero. If $\operatorname{sign} Q = 0$ (i.e. $Q = 0$) whenever $p \in]0, C[$, then the power series Q is uniformly zero. Otherwise, the sign of Q is constant when p varies in $]0, C[$ if and only if Q has no zero on $]0, C[$ because Q is a continuous function of p. This concludes the proof of Proposition 1.33. □

1.11 Rigidity and Hausdorff dimension

1.11.1 Rigidity theorem

If x is a point in $\mathbf{R}^2$, the *radial projection to* x is the function $\pi_x : \mathbf{R}^2 \setminus \{x\} \to S_1$ that maps the point $y \in \mathbf{R}^2 \setminus \{x\}$ to $\pi_x(y) = \frac{y-x}{\|y-x\|}$. For $A \subseteq \mathbf{R}^2$, the set $\pi_x(A \setminus \{x\})$ is the *view* of A from x, that is the set of directions in which x can move to a point in A. Let $\rho : D_1 \to D_2$ be a chirotope-preserving function from $D_1 \subseteq \mathbf{R}^2$ to $D_2 \subseteq \mathbf{R}^2$. If x is a point of D_1, the chirotope-preserving property of ρ implies that the order of the directions of the points of D_1 around x is preserved by ρ. We formalize this idea as follows. First note that for $a, b \in D_1$, if $\pi_x(a) = \pi_x(b)$ then $\pi_{\rho(x)}(\rho(a)) = \pi_{\rho(x)}(\rho(b))$. Indeed, this means that a, b and x are aligned and that a and b lie on the same side of x on the line $ax = bx$, and these properties (alignment and order in a line) are preserved by ρ. This permits to consistently define a *deformation function* function θ_x from $\pi_x(D_1 \setminus \{x\})$ to $\pi_{\rho(x)}(D_2 \setminus \{\rho(x)\})$ satisfying the equality $\theta_x(\pi_x(a)) = \pi_{\rho(x)}(\rho(a))$ for every $a \in D_1 \setminus \{x\}$.

We need the following technical property.

Proposition 1.34. *Let s be a real parameter and let A be a Borel subset of $\mathbf{R}^n$ with $\mathcal{H}^s(A) > 0$. There is a measurable part $B \subseteq A$ with $0 < \mathcal{H}^s(B) < \infty$.*

Proof. Note that it suffices to set $B := A$ if $\mathcal{H}^s(A)$ is finite. Otherwise, we apply a theorem of Besicovitch and Davies that states that every analytic set (this includes Borel sets) of $\mathcal{H}^s$-positive measure has a closed subset of finite an positive $\mathcal{H}^s$-positive measure [13]. $\square$

Note that in Proposition 1.34, the set B has Hausdorff dimension s.

Lemma 1.35. *Let $D_1, D_2 \subseteq \mathbf{R}^2$ be two domains and let $\rho : D_1 \to D_2$ be a chirotope-preserving map. There is a subset $\mathcal{N} \subseteq D_1$ of Hausdorff dimension at most 1 such that ρ is differentiable on $D_1 \setminus \mathcal{N}$. That is, for every $a \in D_1 \setminus \mathcal{N}$, there is a linear function $\mathrm{d}\,\rho_a : \mathbf{R}^2 \to \mathbf{R}^2$ such that*

$$\rho(a+h) = a + \mathrm{d}\,\rho_a(h) + o(\|h\|)$$

whenever $a + h \in D_1$.

Proof. Let $\mathcal{N} \subseteq D_1$ be the set of points on which ρ is not differentiable. For the sake of contradiction, assume that $\dim_H \mathcal{N} > 1$.

Claim 1. For every $x \in D_1$, the function θ_x is differentiable almost everywhere (with respect to the one-dimensional Lebesgue measure Λ_1).

This is a consequence of the following classic Theorem of Lebesgue. A monotone function f from an interval I to $\mathbf{R}$ is differentiable almost-everywhere [31, p. 128]. Since a monotone function f from any set $I \subseteq \mathbf{R}$ can be extended to a monotone function on an interval, this theorem holds without hypothesis on I.

To properly apply this theorem on θ_x, we need to define an equivalent function θ'_x defined on a set of real numbers. Fix a direction $u_0 = \pi_x(y_0) \in \pi_x(D_1 \setminus \{x\})$, which will serve as a base direction. Let I is the set of angles $\{\ (u_0, u) \mid u \in \pi_x(D_1) \setminus \{x\}\ \}$, where the angles are taken in the range $[0, 2\pi[$. Define a function $\theta'_x : I \subseteq [0, 2\pi[\to [0, 2\pi[$ that for every vector $u \in S_1$, maps the angle (u_0, u) to the angle $(\rho(u_0), \rho(u))$, taken in $[0, 2\pi[$.

The function θ'_x is strictly increasing. Indeed, assume $(u_0, u_1) < (u_0, u_2)$, for some $u_1 = \rho(y_1) \in \pi_x(D_1 \setminus \{x\})$ and $u_2 = \rho(y_2) \in \pi_x(D_1 \setminus \{x\})$. The points y_0, y_1 and y_2 appears in this order around x. Since ρ preserves the chirotope, it follows that $\rho(y_0)$, $\rho(y_1)$, and $\rho(y_2)$ appears in this same order around $\rho(x)$. As a consequence, $\theta'_x((u_0, u_1)) = (\rho(u_0), \rho(u_1)) < (\rho(u_0), \rho(u_2)) = \theta'_x((u_0, u_2))$. This shows the monotonicity of θ'_x.

By Lebesgue's theorem, θ'_x is differentiable almost everywhere. Let $f : S_1[0, 2\pi[$ be the function that maps u to the angle (u_0, u). The function θ_x can be written as $\theta_x = f^{-1} \circ \theta'_x \circ f$. Moreover both f and f^{-1} are C^∞ and map a set of 0 measure to a set of 0 measure (for the 1-dimensional Lebesgue measure). As a consequence, θ_x is differentiable almost-everywhere. This finishes the proof of Claim 1.

Claim 2. If A is a subset of D_1 with Hausdorff dimension strictly larger than one, then there is a point $x \in A$ and a subset $B \subseteq A \setminus \{x\}$ of Hausdorff dimension strictly larger than 1 such that θ_x is differentiable on $\pi_x(B)$.

First, note that it is enough to prove this claim in the case where $0 < \mathcal{H}^s(A) < s$ with $s = \dim_H A$. Indeed, assume otherwise and take a real number $s \in]1, \dim_H(A)[$. It holds that $\mathcal{H}^s(A) = +\infty$. By Proposition 1.34, there is a subset A' of A with $0 < \mathcal{H}^s(A') < +\infty$. Now, if A' satisfies Claim 2 with a subset $B \subseteq A'$, then A satisfies Claim 2 for the same set $B \subseteq A$.

Marstrand proved [35] that if $s > 1$, and $A \subseteq \mathbf{R}^2$ is a measurable set with $0 < \mathcal{H}^s(A) < \infty$, then $A \cap (x + \mathbf{R}\vec{u})$ has Hausdorff dimension $s - 1$ for $\mathcal{H}^s \times \mathcal{L}^\infty$-almost all $(x, \vec{u}) \in A \times S_1$. (This has been generalized in higher dimensions by Matilla [36].) By Mastrand's Theorem, there exists $x \in A$ such that $A \cap (x + \mathbf{R}\vec{u})$ has Hausdorff dimension $s - 1$ for almost every $\vec{u} \in S_1$, so in particular $A \cap (x + \mathbf{R}\vec{u})$ is non-empty, so such a vector $\vec{u}$ belongs to $\pi_x(A)$. By Claim 1, the set $S' \subseteq S_1 \setminus \{x\}$ of directions $\vec{u} \in \pi_x(A)$ such that

(i) $A \cap (x + \mathbf{R}\vec{u})$ has Hausdorff dimension $s - 1$; and

(ii) θ_x is differentiable in $\vec{u}$

satisfies $\Lambda_1(S_1 \setminus S') = 0$. We define B to be the set of points of A that x sees in the directions of S', that is $B = \pi_x^{-1}(S') \cap A$. It follows from this last definition and Property (ii), that θ_x is differentiable on $\pi_x(B)$. To finish the proof of the claim, it remains to prove that B has Hausdorff dimension s: we use Property (i). To do so, we use the following fact [16, Lemma 3.1]. If $B \subseteq \mathbf{R}^2$ is a (Borel) set and there is $t \in]0, 1]$ such that $\dim_H(B \cap \{x_0\} \times \mathbf{R}) \geq t$ for x_0 in a subset of $\mathbf{R}$ of positive measure, then $\dim_H B \geq t + 1$. To reduce our statement to this theorem, define the function $\Phi : \mathbf{R}^2 \setminus \{x\} \to [0, 2\pi[\times \mathbf{R}^*$ that maps a point of $\mathbf{R}^2 \setminus \{x\}$ to its polar coordinates centered on x. Since Φ is an isometry on every line through x, $\Phi(B) \cap \{\alpha\} \times \mathbf{R}$ has Hausdorff dimension $s - 1$ for each angle α that corresponds to a direction of S', so any such α belongs to a subset of positive measure. So the theorem applies with $t = s - 1$ and shows that $\Phi(B)$ has Hausdorff dimension at least $t + 1 = s$. Since the function Φ is locally Lipschitz, it does not increase the Hausdorff dimension, so $\dim_H B \geq \dim_H(\Phi(B)) \geq s$, and further $\dim_H B = s$ because $B \subseteq A$. This finishes the proof of the claim.

Claim 3. There are $x, y \in D_1$ and a subset $B \subseteq \mathcal{N} \setminus h(x, y)$ of Hausdorff dimension strictly higher than 1 such that θ_x and θ_y are differentiable on $\pi_x(B)$ and $\pi_y(B)$, respectively.

We apply Claim 2 twice. A first application to N gives $x \in \mathcal{N}$ and a set $B_1 \subseteq \mathcal{N}$ with Hausdorff dimension $s_1 > 1$ such that θ_x is differentiable on $\pi_x(B_1)$ and a second application of Claim 2 to B_1 gives $y \in B_1$ (so $y \neq x$) and $B_2 \subseteq B_1$ with Hausdorff dimension $s_2 > 1$ and θ_y is differentiable on $\pi_y(B_2)$. To finish, it suffices to define $A = B_2 \setminus h(x, y)$, which gives a set satisfying the aforementioned properties with $\dim_H A = \dim_H B_2 > 1$ because $h(x, y)$ has Hausdorff dimension 1, this proves Claim 3.

Claim 4. ρ is differentiable on every $a \in B$.

Indeed, let $b \in D_1$ in a small neighborhood of a, so that $b \notin h(x, y)$. Note that by the definitions $\rho(b)$ is the intersection of the lines $x + \mathbf{R}\theta_x(\pi_x(b))$ and $y + \mathbf{R}\theta_y(\pi_x(b))$. By Claim 3, the function $b \mapsto (\theta_x(\pi_x(b)), \theta_y(\pi_y(b))$ is differentiable in b. The function that to a pair of vectors (θ_x, θ_y) assigns the intersection of the lines $x + \mathbf{R}\theta_x$ and $y + \mathbf{R}\theta_y$ is differentiable on $\mathbf{R}^2 \setminus h(x, y)$. This yields a contradiction and concludes the proof of Claim 4.

Note that the set B of Claim 4 is non-empty because it has positive Hausdorff dimension. This yields a contradiction with the definition of $\mathcal{N}$ and concludes the proof of the theorem. $\square$

Given three points a, b and c of $\mathbf{R}^2$, the *half-wedge* is the set $v(a, b, c) = h^+(c, a) \cap h^+(b, c)$. See Figure 1.20.

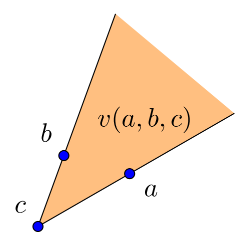


Figure 1.20 – The half-wedge $v(a, b, c)$.

A point a of a domain $D \subseteq \mathbf{R}^2$ is *omnidirectional* if for every $\epsilon > 0$ and every $x, y \in \mathbf{R}^2 \setminus \{a\}$ such that $v(x, y, a) \neq \varnothing$, the set $v(x, y, a) \cap B(a, \epsilon)$ is non-empty.

Equivalently, $a \in D$ is omnidirectional if and only if $\pi_a(B(a, \epsilon) \setminus \{a\})$ is dense in S_1 for every $\epsilon > 0$. In particular, if $u \in S_1$ is a unitary vector then there is a sequence $(h_n)_{n \in \mathbf{N}}$ of vectors such that $a + h_n \in D$ for every $n \in \mathbf{N}$, the sequence $\left(\frac{h_n}{\|h_n\|_2}\right)_{n \in \mathbf{N}}$ tends to u and $\|h_n\|_2$ tends to 0 when n goes to infinity.

Lemma 1.36. *Let $D \subseteq \mathbf{R}^2$. The set of points $x \in D$ that are not omnidirectional has Hausdorff dimension at most 1.*

Proof. By the definition, a point $a \in D$ is not omnidirectional if there exists there exists a positive real ϵ_a and two points $a + \vec{u}_a$ and $a + \vec{v}_a$ with $b, c \in \mathbf{R}^2$ such that the half-wedge $v(a + \vec{v}_a, b + \vec{v}_a, a)$ is non-empty and satisfies $v(a + \vec{v}_a, b + \vec{v}_a, a) \cap B(a, \epsilon_a) \cap D = \varnothing$. Moreover, it is possible to choose ϵ_a, $\vec{u}_a$ and $\vec{v}_a$ such that ϵ_a is rational and the vectors $\vec{u}_a$ and $\vec{v}_a$ are elements of $\mathbf{Q}^2$. To see this, note that $v(a + \vec{u}'_a, b + \vec{v}'_a, a) \cap B(a, \epsilon'_a)$ does not intersect D when $\epsilon'_a < \epsilon_a$ and $v(a + \vec{u}'_a, b + \vec{v}'_a, a)$ is a subset of $v(a + \vec{u}_a, a + \vec{v}_a, a)$.

Consequently, the set of $x \in D$ that are not omnidirectional is the countable union $\bigcup A_{\vec{u}, \vec{v}, \epsilon}$ over all $(\vec{u}, \vec{v}, \epsilon) \in \mathbf{Q}^2 \times \mathbf{Q}^2 \times \mathbf{Q}^+$ where $A_{\vec{u}, \vec{v}, \epsilon}$ stands for the set of non-omnidirectional points $a \in D$ such that $\vec{u}_a = \vec{u}$, $\vec{v}_a = \vec{v}$ and $\epsilon_a = \epsilon$. To prove the theorem, it suffices to show that for each triple $(\vec{u}, \vec{v}, \epsilon)$ of non-colinear rational vectors, the Hausdorff dimension of $A_{\vec{u},\vec{v},\epsilon}$ is at most one, since then the countable union of these set will also have Hausdorff dimension at most 1.

From now on, we fix $(\vec{u}, \vec{v}, \epsilon)$ in $\mathbf{Q}^2 \times \mathbf{Q}^2 \times \mathbf{Q}$. Up to applying an affine transformation, we may assume that $\vec{u} = (1, -1)$ and $\vec{v} = (1, 1)$. Again, since $\mathbf{R}^2$ can be covered by a countable union of balls of radius $\epsilon/2$ it is enough to prove that $A_{\vec{u},\vec{v},\epsilon} \cap B$ has Hausdorff dimension at most 1 for any ball B of radius $\epsilon/2$ to complete the proof. Let B be such a ball. The hypothesis implies that if a point $a = (x_a, y_a)$ belongs to $A_{\vec{u},\vec{v}} \cap B$, then any other point $b = (x_b, y_b) \in \operatorname{supp}\mu \cap B$ satisfies $|x_a - x_b| \leq |y_a - y_b|$. Indeed, in this case $\|a - b\|_2 \leq \epsilon$, so $b \notin w(a + \vec{u}, a + \vec{v}, a)$ and $a \notin w(b + \vec{u}, b + \vec{v}, b)$, which boils down to the condition above.

In particular, a is the only point of the support of abscissa x_a. It follows that there is a 1-Lipschitz function ϕ that maps the subset $S = \{ x \mid (x, y) \in A_{\vec{u},\vec{v},\epsilon} \cap B \}$ of $\mathbf{R}$ to $A_{\vec{u},\vec{v},\epsilon} \cap B$. Since Lipschitz functions do not increase Hausdorff dimensions, it follows that the Hausdorff dimension of $A_{\vec{u},\vec{v},\epsilon} \cap B$ is at most 1. This finishes the proof of the lemma. □

The (lower) *Hausdorff dimension* $\dim_H \mu$ *of a measure* μ is the smallest Hausdorff dimension of the sets charged by μ, that is

$$\dim_H \mu := \inf \{ \dim_H A \mid \mu(A) > 0 \} .$$

As an example, the Hausdorff dimension of the Hausdorff measure of dimension s is $\dim_H \mathcal{H}_s = s$.

Since the support $\operatorname{supp}\mu$ of a measure μ has positive μ-measure unless it is empty, the following holds:

$$\dim_H \mu \leq \dim_H \operatorname{supp}\mu.$$

This inequality can be strict, even if μ is a probability measure. Indeed, if $\mu = \mu_1 + \mu_2$, then $\dim_H \mu = \min(\dim_H \mu_1, \dim_H \mu_2)$ while $\dim_H \operatorname{supp}\mu = \max(\dim_H \operatorname{supp}\mu_1, \dim_H \operatorname{supp}\mu_2)$. Relying on this, an example can be obtained by a mixture of Lebesgue measures of different dimensions.

Theorem 1.37. *If μ_1 and μ_2 are measures with Hausdorff dimensions strictly larger than 1 with $\ell_{\mu_1} = \ell_{\mu_2}$, then μ_1 and μ_2 are projectively equivalent.*

Proof. Let μ_1 and μ_2 be such probability measures. First note that for $i \in \{1, 2\}$, the hypothesis on the Hausdorff dimension of μ_i implies that μ_i charges no line. Indeed, since $\dim_H \mu_i > 1$ and any line L has Hausdorff dimension one, it follows from the definitions that $\mu_i(L) = 0$.

In the first part of the proof, we apply the preceding results to show the existence of two domains $D_1 \subseteq \operatorname{supp}\mu_1$ and $D_2 \subseteq \operatorname{supp}\mu_2$ and a measure preserving homeomorphism $\rho : D_1 \to D_2$ that preserves the chirotopes such that

(i) $\mu_1(D_1) = 1 = \mu_2(D_2)$;

(ii) $\dim_H D_1 > 1$ and $\dim_H D_2 > 1$;

(iii) ρ is differentiable on every point of D_1;

(iv) ρ^{-1} is differentiable on every point of D_2; and

(v) every $x \in D_1$ is omnidirectional.

In a second part, we show that such a function ρ is the restriction of a projective transformation on D_1, which concludes the proof.

By Corollary 1.23, there are subsets $\mathcal{N}_1^1 \subseteq \operatorname{supp} \mu_1$ and $\mathcal{N}_2^1 \subseteq \operatorname{supp} \mu_2$ and a measure-preserving and chirotope-preserving homeomorphism $\rho_0 : \operatorname{supp} \mu_1 \setminus \mathcal{N}_1^1 \to \operatorname{supp} \mu_2 \setminus \mathcal{N}_2^1$, where $\mathcal{N}_1^1$ and $\mathcal{N}_2^1$ are the union of a countable set and at most one line. In particular, $\mathcal{N}_1^1$ and $\mathcal{N}_2^1$ have Hausdorff dimension at most 1 and $\mu_1(\mathcal{N}_1^1) = 0 = \mu_2(\mathcal{N}_2^1)$.

By Lemma 1.35 applied to ρ_0, there is a subset $\mathcal{N}_1^2 \subseteq \operatorname{supp} \mu_1$ of Hausdorff dimension at most 1 such that ρ_0 is differentiable on $\operatorname{supp} \mu_1 \setminus (\mathcal{N}_1^1 \cup \mathcal{N}_1^2)$. Similarly, by Lemma 1.35 applied to ρ_0^{-1}, there is a set $\mathcal{N}_2^2 \subseteq \operatorname{supp} \mu_2$ with $\dim_H \mathcal{N}_2^2 \leq 1$ such that ρ_0^{-1} is differentiable on every point of $\operatorname{supp} \mu_2 \setminus (\mathcal{N}_2^1 \cup \mathcal{N}_2^2)$.

We know from Lemma 1.36 that the set $\mathcal{N}_1^3$ of points in $\operatorname{supp} \mu_1 \setminus (\mathcal{N}_1^1 \cup \mathcal{N}_1^2)$ that are not omnidirectional has Hausdorff dimension at most one.

Set $\mathcal{N}_1 = \mathcal{N}_1^1 \cup \mathcal{N}_1^2 \cup \mathcal{N}_1^3$ and $\mathcal{N}_2 = \mathcal{N}_2^1 \cup \mathcal{N}_2^2$. For $i \in \{1, 2\}$, it holds that $\dim_H(\mathcal{N}_i) \leq 1$ and further $\mu_i(\mathcal{N}_i) = 0$ because $\dim_H \mu_i > 1$. Moreover, $\mu_2(\rho_0(\mathcal{N}_1^2 \cup \mathcal{N}_1^3)) = \mu_1(\mathcal{N}_1^2 \cup \mathcal{N}_1^3) = 0$ and $\mu_1(\rho_0^{-1}(\mathcal{N}_2^2)) = \mu_2(\mathcal{N}_2^2) = 0$ because ρ_0 and ρ_0^{-1} are measure preserving. It remains to take $D_1 = \operatorname{supp} \mu_1 \setminus (\mathcal{N}_1 \cup \rho_0^{-1}(\mathcal{N}_2^2))$ and $D_2 = \operatorname{supp} \mu_2 \setminus (\mathcal{N}_2 \cup \rho_0(\mathcal{N}_1^2 \cup \mathcal{N}_1^3))$. For $i \in \{1, 2\}$, it follows from the properties above that $\mu_i(D_i) = 1$ and since $\dim_H \mu_i > 1$, it holds that $\dim_H D_i > 1$. Also note that $\rho_0(D_1) = D_2$.

Letting ρ be the restriction of ρ_0 to D_1 finishes the first part of the proof. It remains to prove that ρ is a projective transformation.

For every $a \in D_1$, we show that the linear function $\mathrm{d}\,\rho_a : \mathbf{R}^2 \to \mathbf{R}^2$ is invertible. This essentially comes from the fact that ρ^{-1} is differentiable, but we have to be careful since D_1 is not an open set. For $h \in \mathbf{R}^2$ such that $a + h \in D_1$, the differentiability of ρ in a yields

$$\rho(a + h) = \rho(a) + \mathrm{d}\,\rho_a(h) + o(\|h\|_2).$$

Since ρ^{-1} is differentiable in $\rho(a)$ and $\rho(a + h) \in D_2$,

$$a + h = \rho^{-1}(\rho(a + h)) = a + \mathrm{d}\,\rho^{-1}_{\rho(a)} \circ \mathrm{d}\,\rho_a(h) + o(\|h\|_2),$$

and further $\mathrm{d}\,\rho^{-1}_{\rho(a)} \circ \mathrm{d}\,\rho_a(h) = h + o(\|h\|_2)$. Fix a unitary vector $u \in S_1$, since a is omnidirectional, there is a sequence h_n with $a + h_n \in D_1$ for every $n \in \mathbf{N}$ and such that

$$\|h_n\|_2 \underset{n\to\infty}{\longrightarrow} 0 \quad \text{and} \quad \frac{h_n}{\|h_n\|_2} \underset{n\to\infty}{\longrightarrow} u.$$

Applying the equality to h_n, we have $\mathrm{d}\,\rho_a \circ \mathrm{d}\,\rho^{-1}_{\rho(a)}(h_n) = h_n + o(\|h_n\|_2)$. Further,

$$\mathrm{d}\,\rho^{-1}_{\rho(a)} \circ \mathrm{d}\,\rho_a\left(\frac{h_n}{\|h_n\|_2}\right) = \frac{h_n}{\|h_n\|_2} + o(1).$$

Letting n go to infinity, we obtain $\mathrm{d}\,\rho^{-1}_{\rho(a)} \circ \mathrm{d}\,\rho_a(u) = u$ as $\mathrm{d}\,\rho_a$ and $\mathrm{d}\,\rho^{-1}_{\rho(a)}$ are both continuous on D'_1. Since this is true for every unitary vector $u \in S_2$, this proves that $\mathrm{d}\,\rho^{-1}_{\rho(a)} \circ \mathrm{d}\,\rho_a = \mathrm{id}$. In particular, $\mathrm{d}\,\rho_a$ and $\mathrm{d}\,\rho^{-1}_{\rho(a)}$ are invertible.

We claim that for every $a \in D_1$ and $u \in S_2$, the line $L = a + \mathbf{R}u$ through a satisfies

$$\rho(D_1 \cap L) \subseteq \rho(a) + \mathbf{R}\,\mathrm{d}\,\rho_a(u).$$

Since a is omnidirectional, there are two sequences $(h_n^+)_{n\in\mathbf{N}}$ and $(h_n^-)_{n\in\mathbf{N}}$ with $a + h_n^+ \in D_1$ and $a + h_n^- \in D_1$ for $n \in \mathbf{N}$ such that $\|h_n^+\|_2$ and $\|h_n^-\|_2$ both tend to 0 and the intersection of wedges $\bigcap_{n\in\mathbf{N}} w(a, a+h_n^-, a+h_n^+)$ equals $L \setminus \{a\}$.

Since ρ preserves the chirotope, we have

$$\rho(w(a, a+h_n^-, a+h_n^+) \cap D_1) \subseteq w(\rho(a), \rho(a+h_n^-), \rho(a+h_n^+))$$

for every $n \in \mathbf{N}$. Furthermore, $\rho(a+h_n^s) = \rho(a) + \mathrm{d}\,\rho_a(h_n^s) + o(\|h_n^s\|_2)$, for $s \in \{+,-\}$. Letting $u_n^s := \frac{1}{\|h_n\|_2}(\rho(a+h_n^s) - \rho(a))$, we have $u_n^s = \mathrm{d}\,\rho_a(\frac{h_n^s}{\|h_n^s\|_2}) + o(1) = \mathrm{d}\,\rho_a(u) + o(1)$. It follows that

$$\begin{aligned} w(\rho(a), \rho(a+h_n^-), \rho(a+h_n^+)) &= \rho(a) + w(0, \|h_n^-\|_2 \cdot u_n^-, \|h_n^+\|_2 \cdot u_n^+) \\ &= \rho(a) + w(0, u_n^-, u_n^+). \end{aligned}$$

As a consequence, $\rho(w(a, a+h_n^-, a+h_n^+)$ is included into $\rho(a) + w(0, u_n^-, u_n^+)$ for each $n \in \mathbf{N}$, which proves that $\rho(L \setminus \{a\}) \subseteq \rho(a) + \mathbf{R}\,\mathrm{d}\,\rho_a(u)$ since $(u_n^-)_{n\in\mathbf{N}}$ and $(u_n^+)_{n\in\mathbf{N}}$ tends to $\mathrm{d}\,\rho_a(u)$.

Let a_1, a_2, a_3 and a_4 be four points of D_1 in general position, that is without aligned triple. Such a quadruplet exists because D_1 is not contained into a finite union of lines, as a set of Hausdorff dimension greater than one. Let f be the unique projective transformation that maps a_i to $\rho(a_i)$ for every $i \in \{1,2,3,4\}$. We aim to prove that ρ equals f on D_1. The function f is differentiable, therefore for each $a \in \mathbf{R}^2$, there is a linear function $\mathrm{d}\,f_a$ such that $f(a+h) = f(a) + \mathrm{d}\,f_a(h) + o(\|h\|_2)$ and further for each line $L = a + \mathbf{R}u$, we have $f(L) = f(a) + \mathbf{R}\,\mathrm{d}\,f_a(u)$ because f preserves alignment.

We claim that for $i \in \{1,\dots,4\}$, we have $\mathrm{d}\,\rho_{a_i} = \lambda \cdot \mathrm{d}\,f_{a_i}$ for some $\lambda \in \mathbf{R}$. Let us prove this for $i = 1$, the other cases will follow by symmetry.

For $j \in \{2,3,4\}$, we know from the aforementioned property that $\rho(h(a_1, a_j))$ is included into a line containing $\rho(a_1)$. As moreover $\rho(a_j) \in \rho(h(a_1, a_j))$, it holds that

$$\rho(h(a_1, a_j)) \subseteq h(\rho(a_1), \rho(a_j)) = h(f(a_1), f(a_j)) = f(h(a_1, a_j)).$$

Consequently, the linear function $\mathrm{d}\,\rho_{a_1} \circ (\mathrm{d}\,f_{a_1})^{-1} : \mathbf{R}^2 \to \mathbf{R}^2$ has three non-colinear eigenvectors, namely $a_2 - a_1$, $a_3 - a_1$ and $a_4 - a_1$, and therefore $\mathrm{d}\,\rho_{a_1} \circ (\mathrm{d}\,f_{a_1})^{-1}$ is an homothetic transform, which proves the claimed property.

We now show that for every line L that contains a_i for some $i \in \{1, 2, 3, 4\}$, then

$$\rho(L \cap D_1) \subseteq f(L).$$

Indeed, write $L = a_i + \mathbf{R}u$, for some $u \in S_2$, we have

$$\rho(L \cap D_1) \subseteq \rho(a_i) + \mathbf{R}\,\mathrm{d}\,\rho_{a_i}(u) = f(a_i) + \mathbf{R}\,\mathrm{d}\,f_{a_i}(u) = f(L).$$

We are now ready to prove that $\rho(x) = f(x)$ whenever $x \in D_1$. Since a_1, a_2, a_3 and a_4 are not in convex position, for every $x \in D_1$, there exist $i, j \in \{1, \ldots, 4\}$ such that a_i, a_j and x are not aligned. Applying the preceding property with $L = a_i x$, we know that $\rho(x)$ holds on the line $h(f(a_i), f(x))$ and similarly, $\rho(x)$ is on the line $h(f(a_j), f(x))$, so $\rho(x) = f(x)$. □

Theorem 1.37 seems unsatisfying for me because its hypothesis is about the dimension of the measures involved while Corollary 1.23 suggests that rigidity should only depend on the geometry of the supports of these measures, independently from the distribution on this support. I expect the following statement to be correct.

Conjecture 1.38. *If μ_1 and μ_2 are two measures with $\ell_{\mu_1} = \ell_{\mu_2}$ such that μ_1 charges no line and $\dim_H \operatorname{supp} \mu_1 > 1$, then μ_1 and μ_2 are projectively equivalent.*

Note that the two differences with Theorem 1.37: first the hypothesis concerns the Hausdorff dimension of the support of μ_1 instead of the Hausdorff dimension of the measure μ_1 itself; second, there is no hypothesis on μ_2.

1.12 Rigidity when the support has nonempty interior

The following theorem states that under some assumption, two measures that give the same chirotope distribution are essentially the same.

Theorem 1.39. *Let ν_1 and ν_2 be two measures of S^2 that charge no line. If $\operatorname{supp} \nu_1$ contains an open set and $\ell_{\nu_1} = \ell_{\nu_2}$, then there exists a projective transformation f such that $\nu_2 \circ f = \nu_1$.*

Theorem 1.39 is a consequence of Theorem 1.22 and the following geometric property, that a function preserving alignments on a small ball acts as a projective (or spherical) transform on some dense grid.

Lemma 1.40. *There is a bounded and countable subset A of $\mathbf{R}^2$ whose adherence contains an open ball and with the following property.*

(∗) *For every set $T \subseteq \mathbf{R}^2$ every map $\rho : A \to T$ such that $\rho(a)$, $\rho(b)$ and $\rho(c)$ are aligned if and only if a, b and c are aligned for every $(a, b, c) \in A^3$, there is a projective transformation p such that $p_{|A} = \rho$.*

Moreover, this property is stable by homotetic transformations, that is for every $a_0 \in \mathbf{R}^2$ *and* $\lambda \in \mathbf{R}^*$, *the set* $a_0 + \lambda A$ *satisfies* $(*)$.

Proof. Fix a convex quadrilateral $Q_0 = a_0b_0c_0d_0$ such that the lines $h(a_0, b_0)$ and $h(c_0, d_0)$ in one hand, and $h(b_0, c_0)$ and $h(d_0, a_0)$ in the other hand, are not parallel. We now describe a construction that can be applied to any convex quadrilateral $Q = abcd$ contained in Q_0 and such that $h(a, b) \cap h(c, d) = \{e\}$ and $h(a, d) \cap h(b, c) = \{f\}$. Figure 1.21 illustrates the construction. *Processing* Q means constructing the following five points:

- the intersection point x of $h(a, c)$ and $h(b, d)$;
- the intersection point of $h(e, x)$ and $h(a, d)$;
- the intersection point of $h(e, x)$ and $h(b, c)$;
- the intersection point of $h(f, x)$ and $h(a, b)$; and
- the intersection point of $h(f, x)$ and $h(c, d)$.

This yields four new convex quadrilaterals contained in Q and all containing x, each of which can be processed. Starting from $Q_0 \cup \{e, f\}$, we recursively process Q_0 and all subsequent quadrilaterals emanating from the process. We define A to be the set of vertices of all obtained quadrilaterals. We observe that A is dense in the quadrilateral Q_0. To see this, we consider the projective transformation that maps a to $(0, 0)$, b to $(1, 0)$, c to $(1, 1)$ and d to $(0, 1)$. (The images of e and f are on the line at infinity.) The image of A by this transformation is the grid $\left\{ \left(\frac{i}{2^k}, \frac{j}{2^k}\right) \,\middle|\, k \in \mathbf{N} \text{ and } 0 \le i, j \le 2^k \right\}$.

We note that $h(\rho(a), \rho(b))$ and $h(\rho(c), \rho(d))$ intersect in a unique point, which is $\rho(e) \in T$. Similarly, $h(\rho(a), \rho(d))$ and $h(\rho(b), \rho(c))$ intersect in a unique point, which is $\rho(f) \in T$.

Let p be the unique projective transformation that maps a to $\rho(a)$, b to $\rho(b)$, c to $\rho(c)$ and d to $\rho(d)$. In particular, p preserves the colinearity of points in $\mathbf{R}^2$. This implies that if $x \in A$, then $\rho(x) = p(x)$ as all points in A is defined from Q_0 and from $(\rho(a), \rho(b), \rho(c), \rho(d)) = (p(a), p(b), p(c), p(d))$ using only colinearity properties.

The second part of the theorem directly follows from the fact that the affine transformation $f : x \mapsto a_0 + \lambda x$ preserves colinearities: if a function ρ on $A' = a_0\lambda A$ preserves colinarities then by $(*)$, $\rho \circ f = p_{|A}$ for some projective transformation f, and further $\rho = (p \circ f^{-1})_{|A'}$.

□

Proof of Theorem 1.39. Theorem 1.22 applies to ν_1 and ν_2. Indeed, these measures charge no line and $\ell_{\nu_1} = \ell_{\nu_2}$. Thus, there are two countable sets $\mathcal{N}_1 \subseteq \operatorname{supp} \nu_1$ and $\mathcal{N}_2 \subseteq \operatorname{supp} \nu_2$ and a measure-preserving map $\rho : \operatorname{supp} \nu_1 \setminus \mathcal{N}_1 \to \operatorname{supp} \nu_2 \setminus \mathcal{N}_2$ that preserves the chirotope (on *every* triple).

Now, let A be the set provided by Lemma 1.40. Let O be an open ball of $\operatorname{supp} \nu_1$ that is included into a hemisphere, so that we can consider O as an

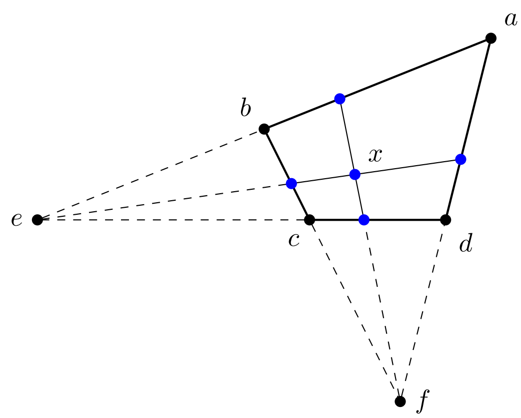


Figure 1.21 – The construction of four new quadrilaterals (meeting in x) from the quadrilateral $Q = abcd$.

open set of the plane. Such a ball exists by hypothesis. Up to shifting and rescaling A, we may assume that $A \subseteq O \setminus \mathcal{N}_1$. Indeed, since O is an open ball and A is bounded, there is a point a_0 and real numbers $t > 0$ and $\lambda > 0$ such that the set $A_s = a_0 + (s, 0) + \lambda A$ is included in O for every $s \in [0, t]$. Now, note that if $a \in A$ and $n \in \mathcal{N}_1$, then $a_0 + (s, 0) + \lambda a$ is equal to n for at most one value of s. Since A and $\mathcal{N}_1$ are countable, it follows that A_s intersects $\mathcal{N}_1$ for at most countably many s, which suffices to conclude that there is s such that $A_s \subseteq O \setminus \mathcal{N}_1$.

We know from Lemma 1.40 that there is a projective transform p, such that $\rho_{|A} = p_{|A}$. By continuity, it follows that Hence $p(x) = \rho(x)$ for every $x \in \overline{A} \setminus \mathcal{N}_1$.

Since p also preserves the orientation of the triangles in A, there is a spherical transform s such that $s(x) = p(x) = \rho(x)$ for every $x \in A$.

It remains to show that $s(x) = \rho(x)$ for every $x \in \operatorname{supp} \nu_1 \setminus \mathcal{N}_1$. Let $x \in S$ and consider two distinct lines L_1 and L_2 that both intersect the interior of $\overline{A}$ and such that $L_1 \cap L_2 = \{x\}$. For $i \in \{1, 2\}$, let a_i and b_i be two distinct (and not antipodal) points of $L_i \cap \overline{A}$ that are not in $\mathcal{N}_1$. These points exist because $\mathcal{N}_1$ is countable. Thus $s(a_i) = \rho(a_i)$ and $s(b_i) = \rho(b_i)$. Hence both $\rho(x)$ and $s(x)$ belong to $h(\rho(a_i), \rho(b_i))$ since both ρ and s preserve colinearity. As $h(\rho(a_1), \rho(b_1))$ and $h(\rho(a_2), \rho(b_2))$ intersect in two antipodal points, it follows that $\rho(x) \in \{s(x), -s(x)\}$.

To finish the proof, note that the line $h(a_1, a_2)$ does not contain x (since otherwise $L_1 = L_2$). Since both ρ and s are chirotope-preserving, the triangles $\rho(a_1)\rho(a_2)\rho(x)$ and $s(a_1)s(a_2)s(x) = \rho(a_1)\rho(a_2)s(x)$ have the same orientation, which is that of $a_1 a_2 x$. This suffices to conclude that $\rho(x) = s(x)$ and finish the proof.

Since, $\nu_1(\mathcal{N}_1) = 0$, it follows that $\nu_2 \circ f = \nu_1$. $\square$

Chapter 2

Property testing

2.1 Introduction

How to get information from big networks, such as the graph of the internet, when it is so big that even a mere exploration of this structure is impossible? In this case, classical graph algorithms are in general not usable.

The idea of property testing is to use algorithms that examine only a random sampling of the graph, whose size is very small in comparison to the size of the graph. This approach was initiated by Oded Goldreich, Shafi Goldwasser and Dana Ron [22]. It leads to very efficient algorithms, and in particular algorithms whose complexity does not depend on the size of their input, but instead only on a precision parameter.

Since two almost identical structures are likely to generate the same samplings, the drawback for such an algorithm is that it can only decide a property up to a certain distance, which corresponds to what is distinguishable for such an algorithm.

2.1.1 Tester

Assume $G = (V, E)$ is a that a (usually very large) graph whose vertex set is known and consider an algorithm $\mathcal{A}$ that accesses G via *queries* of the type

Does uv belongs to E?

for some vertices $u, v \in V$. The *query complexity* of $\mathcal{A}$ is the number of queries this algorithm asks. Note that at each step of the algorithm $\mathcal{A}$, the algorithm cannot distinguish between two vertices involved in no previous query, so we may assume that every new vertex used by $\mathcal{A}$ is chosen uniformly at random. Equivalently, if $X \subseteq V$ is the set of vertices involved in some query of a run of $\mathcal{A}$, we may assume that X is a random subset of V of some size s chosen uniformly among the subsets of V of size s.

The query complexity c of $\mathcal{A}$ for this run is at least $\frac{|X|}{2}$, which happens if each query involves completely new vertices, and at most $\binom{|X|}{2}$, which happens if $\mathcal{A}$ asked for the entire knowledge of $G[X]$, the subgraph of G induced by X. Up to changing the complexity of $\mathcal{A}$ from c to $2c^2$, we may assume that we are always in the second case, i.e. $\mathcal{A}$ asks for the whole subgraph $G[X]$ (up to not using all the available information), and we consider only such algorithms in the following. This simplification in particular does not change which problems can be solved with query complexity independent from the size of G, or which problems admit an algorithm with polynomial query complexity with respect to some parameter.

As a consequence, this approach boils down to study the following algorithmic scheme.

Algorithm 1 Tester

Input: A graph $G = (V, E)$
Output: Does G have some property $\mathcal{P}$?
1: Sample $X \subseteq V$ of size s uniformly at random for some s.
2: Construct $G[X]$. $\triangleright$ $\binom{|X|}{2}$ queries.
3: Accept or reject G depending on $G[X]$.

Algorithm 1 is a randomized algorithm, and it may give a wrong answer with some probability (and our concern is to find under which condition we can have some guarantee under this probability). A randomized decision algorithm that may reject a graph with the tested property and may accept a graph without the tested property is *two-sided*. It is *one-sided* if a rejection is always exact.

Since the number of queries made by Algorithm 1 depends only on the size of X, we make a small abuse of notation and we call *query complexity* of Algorithm 1 the size of the set X instead of the number of edges actually asked for, which is $\binom{|X|}{2}$.

Let $\mathcal{P}$ be a graph property, i.e. a set of graphs, such as the set of bipartite graphs or that of triangle-free graphs. The property $\mathcal{P}$ is called *hereditary* if it is closed under induced subgraph, i.e. if for every $G \in \mathcal{P}$ and $X \subseteq V(G)$, the induced subgraph $G[X]$ belongs to $\mathcal{P}$.

We want a one-sided algorithm of the type of Algorithm 1 that tests a property $\mathcal{P}$ that is supposed to be hereditary. To be one-sided, this algorithm cannot reject if $G[X] \in \mathcal{P}$. On the other side, a subgraph $G[X]$ that is not in $\mathcal{P}$ is a certificate that $G \notin \mathcal{P}$, since otherwise it would contradict the heredity of $\mathcal{P}$. This leads to the following algorithm.

Algorithm 2 One-sided tester for the property $\mathcal{P}$

Input: A graph $G = (V, E)$
Output: Does G belong to $\mathcal{P}$?

```
Sample X ⊆ V of size s uniformly at random for some s.
Construct G[X].
if G[X] ∈ P then
    Accept
else
    Reject
end if
```

This whole chapter is about the following question: which value should I take for s and what kind of guarantees do I have on the outcome of this algorithm?

Note on the query complexity

In Algorithm 2, we are not interested in the complexity of determining whether $G[X]$ is in $\mathcal{P}$. In particular, this recognition can be NP-complete, in which case the best algorithm known to perform the recognition has time complexity exponential in $|X|$. The first reason for this is that if $|X|$ is constant, then recognizing if $G[X] \in \mathcal{P}$ takes constant (although perhaps long) time. The second reason is that it isolates the testability problem from the classical algorithm complexity problem. If $\mathcal{P}$ is, for instance, the class of 3-colorable graphs, recognizing graphs of $\mathcal{P}$ is NP-complete and therefore knowing if there is a polynomial algorithm for this class is equivalent to arguably one of the most difficult problem in computer science. One of the main difficulties to the question P vs NP is that it is difficult to prove that no algorithm does something. When caring only for the query complexity of Algorithm 2, no purely algorithmic question remains as the algorithm is known. Instead, the question boils down to study the statistics of the random induced subgraph $G[X]$.

Formally speaking, this approach is also justified by a result attributed to Noga Alon and exposed by Goldreich and Trevisan [23, Proposition E.2] that shows that a heridetary graph property can be tested (as properly defined in the next section) by some algorithm of query complexity s, then it can be tested with query complexity polynomial in s by Algorithm 2.

2.1.2 A distance on graphs

Detecting edges

We start with the very simple example where $\mathcal{P}$ is the class of empty graphs, that is the set of graphs with no edge. Let $G = (V, E)$ be a graph with n vertices and m edges. For the class $\mathcal{P}$, Algorithm 2 rejects if and only if $G[X]$ has at least one edge, i.e.

$$\mathbf{P}(G \text{ is rejected}) = \mathbf{P}(G[X] \text{ has an edge})$$

Let us estimate this last probability. Recall first that X is a random set of size s. The probability for a random pair $\{u, v\}$ in V (chosen uniformly at random) that uv is an edge is $m/\binom{n}{2}$. As a first consequence, the expected number of edges in $G[X]$ is $\mathbf{E}(|E(G[X])|) = \binom{s}{2} \cdot m/\binom{n}{2}$. Using $\mathbf{E}(|E(G[X])|)$ as an upper bound on the probability that $|E(G[X])| \geq 1$, it holds that

$$\mathbf{P}(G[X] \text{ has an edge}) \leq \mathbf{E}(|E(G[X])|) = \binom{s}{2} \cdot \frac{m}{\binom{n}{2}} \leq s^2 \cdot \frac{m}{n^2}.$$

Moreover, since X contains $s/2$ independent pairs of vertices, the probability that none of them is an edge is at most

$$\begin{aligned}\mathbf{P}(G[X] \text{ has no edge}) &\leq (1 - m/\binom{n}{2})^{s/2} \\ &\leq (e^{-2\frac{m}{n^2}})^{\frac{s}{2}} \\ &= e^{-s \cdot \frac{m}{n^2}}\end{aligned}$$

Now set $\lambda = \frac{m}{n^2}$, and rewrite the inequalities above as

$$1 - e^{-\lambda s} \leq \mathbf{P}(G[X] \text{ has an edge}) \leq \lambda s^2. \tag{2.1}$$

Note that the bounds in (2.1) depend only on s and the ratio $\lambda = \frac{m}{n^2}$. It follows from the upper bound that for a fixed s, Algorithm 2 can fail with probability arbitrarily small provided that the input has small enough $\frac{m}{n^2}$. For instance, a big graph with only one edge is very unlikely to be detected as an non-empty graph by Algorithm 2. In particular, there is no hope that a constant s works well for *every* graph. This also shows that such an algorithm cannot distinguish between graphs that differ in a subquadratic number of edges. If, however, we consider only graphs with λ larger than some fixed value $\epsilon > 0$, the lower bound shows that taking $s = \frac{1}{\epsilon}$, then the algorithm gives the correct answer (i.e. rejects) with probability at least $1 - e^{-\frac{\lambda}{\epsilon}} \geq 1 - 1/e \geq \frac{1}{2}$.

Let us formalize this phenomenon.

Definition 2.1. The distance $d(G, H)$ between two graphs G and H on the same vertex set of size n is the smallest number λ such that one can get a graph isomorphic to H from G by adding or removing $\lambda \cdot n^2$ edges to G.

If $\mathcal{P}$ is a graph property and G is a graph, the *distance from G to $\mathcal{P}$* is

$$d(G, \mathcal{P}) := \min_{H \in \mathcal{P}, |H| = |G|} d(G, H).$$

For $\epsilon > 0$, the graph G is *ϵ-far* from $\mathcal{P}$ if $d(G, \mathcal{P}) \geq \epsilon$. The graph G is *ϵ-close* to $\mathcal{P}$ if $d(G, \mathcal{P}) \leq \epsilon$. In other words, G of size n is ϵ-far from $\mathcal{P}$ if one need to add and/or remove at least ϵn^2 edges to transform G into a graph of $\mathcal{P}$.

Definition 2.2. A hereditary graph class $\mathcal{P}$ is *testable* if for every $\epsilon > 0$ there is an integer $s(\epsilon)$ such that for every graph $G = (V, E)$ that is ϵ-far from $\mathcal{P}$,

$$\mathbf{P}(G[X] \notin \mathcal{P}) \geq \frac{1}{2}.$$

where X is chosen uniformly at random among the subsets of V of size $s(\epsilon)$.

For instance, Equation (2.1) proves that the class of empty graphs is testable with query complexity $s(\epsilon) = \frac{1}{\epsilon}$.

From an algorithmic point of view, a hereditary property $\mathcal{P}$ is testable if and only if for every $\epsilon > 0$ Algorithm 2 with $s = s(\epsilon)$ works on G with probability at least $\frac{1}{2}$ unless $d(G, \mathcal{P}) < \epsilon$. More precisely, Algorithm 2 has in this case the following outcome.

- $\mathbf{P}(G \text{ is accepted}) = 1$ if $G \in \mathcal{P}$.
- $\mathbf{P}(G \text{ is rejected}) \geq \frac{1}{2}$ if $d(G, \mathcal{P}) \geq \epsilon$.
- There is no guarantee if $d(G, \mathcal{P}) < \epsilon$ and $G \notin \mathcal{P}$.

This is summarized by Figure 2.1.

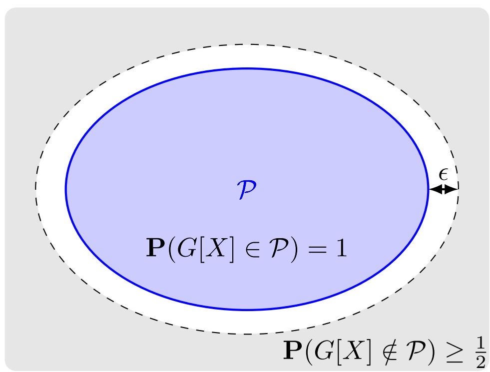


Figure 2.1 – Definition of testable.

2.1.3 Tricks on the algorithm

In Algorithm 2, the set of vertices X is taken uniformly at random among the subsets of some size s, which corresponds to a random sampling without replacement.

For the analysis, it is often convenient to consider that X is instead constructed as the union of a given number k random sets $X_1, \ldots, X_k$ with respective size $s_1, \ldots, s_k$ with $\sum_{i=1}^{k} s_i = s$, where each X_i is chosen uniformly among the sets of vertices of size s_i, independently from each other. We write $\widetilde{X} = \bigcup_{i=1}^{k} X_i$ this variant of X for the sake of readability. In other words, $\widetilde{X}$ is a random sampling with replacement between each X_i.

We argue that in this case, if $\mathbf{P}(G[\widetilde{X}] \notin \mathcal{P}) \geq \frac{1}{2}$ then $\mathbf{P}(G[X] \notin \mathcal{P}) \geq \frac{1}{2}$. Let us prove it if $k = 2$. For each non-negative integer $r \leq \min(s_1, s_2)$, the probability that $|X_1 \cap X_2| = r$ is some value $p_r = \binom{s_1}{r}\binom{n-s_1}{s_2-r}/\binom{n}{s_2}$ whose exact expression actually does not matter. Now, $\widetilde{X}$ can be simulated as a subset of X as follows: take a random value r, where $r \in \{0, \dots, \min(s_1, s_2)\}$ is chosen with probability p_r, this fixes the size of $|X_1 \cap X_2|$. Then, pick three disjoint subsets $Y_1, Y_2, Z \subseteq X$, uniformly at random so that $|Y_1| = s_1 - r$, $|Y_2| = s_2 - r$ and $|Z| = r$. Set $X_1 = Y_1 \cup Z$ and $X_2 = Y_2 \cup Z$. As a consequence to this construction, we can correlate X and $\widetilde{X}$ so that $\widetilde{X} \subseteq X$, and further $\mathbf{P}(G[X] \notin \mathcal{P}) \leq \mathbf{P}(G[\widetilde{X}] \notin \mathcal{P})$ since $G[X] \notin \mathcal{P}$ whenever $G[\widetilde{X}] \notin \mathcal{P}$. The case $k > 2$ can be treated similarly.

As another remark, note that the constant probability threshold of $\frac{1}{2}$ is arbitrary. It is easy to see that this constant can be arbitrarily improved by taking a larger sampling. More precisely, if $\mathbf{P}_{|X|=s}(G[X] \in \mathcal{P}) < \frac{1}{2}$, where X is taken with size s, then $\mathbf{P}_{|X|=ks}(G[X] \in \mathcal{P}) < 2^{-k}$. Indeed, by the remark above, taking $|X|$ of size ks gives at least the information of taking k independent samplings of size s in V.

2.1.4 Testability of Hereditary properties

The fundamental result on one-sided testability of graph properties is the following general theorem, proved by Noga Alon and Asaf Shapira in 2005.

Theorem 2.1 (Alon and Shapira [7])**.** *Every hereditary property is testable.*

This theorem is remarkable: it provides a one-sided algorithm of query complexity (and further computational complexity) that is *independent of the size of the graph* for *every* hereditary class.

Alon and Shapira actually proved that Theorem 2.1 is essentially an equivalence. A graph property $\mathcal{P}$ is *semi-hereditary* if $\mathcal{P}$ is contained in a hereditary graph property $\mathcal{H}$ such that for every $\epsilon > 0$ there is $M(\epsilon)$ such that every graph of size at least $M(\epsilon)$ which is ϵ-far from $\mathcal{P}$ has an induced subgraph not in $\mathcal{H}$. A graph property is testable (with one-sided tester) if and only if it is semi-hereditary [7, Theorem 2]. As a simple example, a hereditary property from which we remove a finite number of graphs is semi-hereditary, like for instance the class of bipartite graphs of size at least 42.

Theorem 2.1 uses a strengthening of Szemerédi regularity lemma and gives a query complexity $s(\epsilon)$ that is a tower of towers of exponentials of size polynomial in $1/\epsilon$. This has been improved by David Conlon and Jacob Fox [12] to a single tower of exponentials. This bound in the regularity lemma cannot be significantly improved since Timothy Gowers constructed a family of graphs whose regularity partition all have size at least a (single) tower of exponentials of height $\left(\frac{1}{\epsilon}\right)^c$ for some constant c [24].

This huge constant makes this theorem unusable in practice.

2.1.5 Easily testable classes

We know from Theorem 2.1 that every hereditary class is testable. The next step is to refine the analysis to determine which of these classes are testable with reasonable query complexity. A natural definition for *reasonable query* is to ask that the complexity is a polynomial in $\frac{1}{\epsilon}$.

Definition 2.3. A hereditary property $\mathcal{P}$ is *easily testable* if it is testable with a constant $s(\epsilon)$ that is polynomial in $\frac{1}{\epsilon}$.

The rest of this chapter is mainly devoted to the question of which classes are easily testable.

2.1.6 H-free classes

For any graph H, let H-FREE be the class of H-free graphs, that is graphs with no induced subgraph isomorphic to H. As will be seen below, the question of finding for which H the class H-FREE is easily testable is "nearly" solved.

If $G = (V, E)$ is a graph, then $\overline{G}$ denotes the complementary graph of G, i.e. the graph on V where uv is an edge if and only if $uv \notin E$, for every pair of different vertices u and v. Because of the symmetry of the role played by edges and non-edges in Algorithm 2, the complexity for testing a property $\mathcal{P}$ is the same as the complexity for testing that the class $\left\{ \overline{G} \mid G \in \mathcal{P} \right\}$. In particular, the classes H-FREE and $\overline{H}$-FREE are testable with same query complexity.

If k is a positive integer, let P_k be the path with k vertices. The class P_2-FREE is the set of empty graphs. It is an easy exercise to see that P_2-FREE is testable with query complexity $O(\frac{1}{\epsilon})$. We established this result in Section 2.1.2. The class P_3-FREE, that consists of disjoint unions of cliques, has been proved to be easily testable by Alon and Shapira [6]. On the other hand, they showed that the class H-FREE is not easily testable when H is different from P_2, P_3, P_4, C_4 and different from the complementary of one of these graphs, so in particular H-FREE is not easily testable when H has size at least 5. The case of the class P_4-FREE has been solved by Alon and Fox [4], who showed that P_4-FREE is easily testable. The graphs in P_4-FREE are also called co-graphs, they are the graphs that can be recursively obtained from single vertices by disjoint unions and (complete) joins.

As a consequence, all graphs H for which the class H-FREE is easily testable are known, except when H is C_4 or its complementary $\overline{C_4} = 2K_2$. The following question is still open.

Question 2.1. Is the class C_4-FREE easily testable?

Last year, Lior Gishboliner and Asaf Shapira [19] showed that every graph that is ϵ-far from the class of C_4-free graphs contains at least $n^4/2^{(1/\epsilon)^c}$ induced copies of C_4 for some constant c. This result implies that the class of C_4-free graphs can be tested with query complexity $2^{(1/\epsilon)^c}$ (recall that the general bound on the query complexity of hereditary classes is a tower of exponentials of polynomial size in $1/\epsilon$). We discuss Question 2.1 in the next section (Section 2.2).

Alon and Fox constructed for every small enough ϵ a graphs G_ϵ that is ϵ-far from being C_5-free, from which a sampled induced subgraph of size s is a comparability graph with probability at least $1 - s^3\epsilon^{\Omega(\log(1/\epsilon))}$, which tends to 1 when s is a polynomial in $1/\epsilon$ and ϵ tends to 0 [4]. This further implies that every class $\mathcal{P}$ such that

$$\text{COMPARABILITY} \subseteq \mathcal{P} \subseteq C_5\text{-FREE}$$

is not easily testable. In particular, the class of perfect graphs is not easily testable.

2.2 The case of C_4-free graphs

A possible first step to attack Question 2.1 is to look at subclasses of C_4-FREE that are more structured. This includes the following classes.

- **Split graphs.** A graph is a split graph if its vertices can be partitioned into a stable set and a clique. Equivalently, a graph is split if it is C_4, $\overline{C_4} = 2K_2$ and C_5-free. We prove in Section 2.4 that split graphs are testable with query complexity $O(\frac{1}{\epsilon} \cdot \ln \frac{1}{\epsilon})$.

- **$\{C_4, \overline{C_4}\}$-free graphs, or pseudo-split graphs.** It is the set of graphs without C_4 or its complementary $2K_2$ as induced subgraphs. Such a graph is also called a pseudo-split graph because of the following characterization. A graph $G = (V, E)$ is $\{P_4, 2K_2\}$-free if and only if G is split or V can be partitioned into a clique C, an independent set I and a set S that induces a C_5. In particular, a pseudo-split graph is split if it has no induced C_5 and has exactly one such C_5 otherwise.

  It follows that the class of pseudo-split graphs is testable with same asymptotic complexity as split graphs, i.e. $O(\frac{1}{\epsilon} \ln \frac{1}{\epsilon})$. To see this, consider a graph G that is ϵ-far from being pseudo-split, then G is in particular ϵ-far from being split. Let c be a constant such that the class of split graphs is testable with query complexity $c \cdot \frac{1}{\epsilon} \ln \frac{1}{\epsilon}$. Now sample a set $S = S_1 \cup S_2 \cup S_3$ of size $3c \cdot \frac{1}{\epsilon} \ln \frac{1}{\epsilon}$, partitioned into three parts of size $|S_i| = c \cdot \frac{1}{\epsilon} \ln \frac{1}{\epsilon}$. We know that with probability at least $\frac{1}{2}$, the subgraph $G[S_i]$ is not split. As a consequence, with probability at least $\frac{1}{2}$, more than half of these three sets, thus at least two of them – say S_1 and S_2 –, induce a graph that is not split. If $G[S_1]$ and $G[S_2]$ are pseudo-split graphs, then by the characterization above both of them contain an induced C_5 and therefore $G[S]$ is not pseudo-split. This proves that the class of pseudo-split graph is testable with query complexity $3c \cdot \frac{1}{\epsilon} \ln \frac{1}{\epsilon}$.

  Note also that a pseudo-split graph G on n vertices is $\frac{5}{n}$-close to a split graph since it suffices to transform the possible C_5 in G into an independent set disconnected from the rest of the graph by removing at most $5 + 5(n-5) \leq 5n$ edges. This distance tends to 0 when n goes to infinity, so these two classes generate the same limits of graphs.

- **$\{C_4, P_4\}$-free graphs, or trivially perfect graphs.** These graphs are also intersection graphs of nested interval families, that are families of intervals $(I_v)_{v \in V}$ such that $I_v \subseteq I_u$ or $I_u \subseteq I_v$ whenever I_u and I_v intersects. This class is easily testable by Theorem 2.11. It is also a consequence of Corollary 2.39 that we show at the end of this chapter or Theorem 6 in [20]. This class has a nice recursive structure similar to the structure of co-graphs (from which it is a subclass). In Section 2.5.2, we sketch a proof that the class of trivially perfect graphs are easily testable using this structure.

- **Interval graphs.** Interval graphs are intersection graphs of family of intervals. We prove in Section 2.6 that the class of interval graphs is easily testable.

  In the mean time , Gishboliner and Shapira [20] independently proved a far more general result: every semi-algebraic classes – a very large family of classes containing classes of intersection graphs of geometric object (such as interval graphs) – are easily testable. See Section 1.9.1 for a definition of semi-algebraic relation. The proof given by Gishboliner and Shapira is much shorter than the one presented in this thesis. Nevertheless, the techniques used are very different and this latest proof might give ideas for cases that are not solved by the general result of Gishboliner and Shapira.

- **Chordal graphs.** Chordal graphs are the graphs without induced cycle of size at least 4. The class of chordal graphs is a superclass of the class of interval graphs that is not semi-algebraic. It is an open question to know if this class is easily testable. As for C_4-free graphs, a recent result [19, Theorem 1.2] shows that chordal graph are testable with query complexity $2^{(1/\epsilon)^c}$.

2.3 Graph partition problem

The class of k-colorable graphs is an example of NP-complete problem that is easily testable. This result – k-COLORABILITY is testable – was first proved by Goldreich, Goldwasser and Ron [22] for the query complexity $\frac{k^2 \ln k}{\epsilon^2}$. The query complexity was improved to $\frac{36k \ln k}{\epsilon^2}$ by Noga Alon and Michael Krivelevich [5]. The proof of Alon and Krivelevich can be directly adapted to show a similar result on a larger family of graph classes.

Given a graph $G = (V, E)$ and two sets $V_1, V_2 \subseteq V$, let $E(V_1, V_2)$ be the set of vertices $uv \in E$ such that $u \in V_1$ and $v \in V_2$. Recall that if k is an integer, then $[k]$ is the set $\{1, \ldots, k\}$.

Definition 2.4. A *graph partition problem* is parameterized by $k + \binom{k}{2}$ pairs of numbers $(\ell_{ij}, u_{ij})_{1 \leq i \leq j \leq k}$ with $k \in \mathbf{N}$. It consists of the set of graph $G = (V, E)$ with a partition $V = \bigcup_{i=1}^{k} V_i$ such that:

1. For every i, j with $1 \leq i < j \leq k$,
$$\ell_{ij} \leq \frac{|E(V_i, V_j)|}{|V_i| \cdot |V_j|} \leq u_{ij}.$$

2. For every $i \in [k]$,
$$\ell_{ii} \leq \frac{|E(V_i, V_i)|}{\binom{|V_i|}{2}} \leq u_{ii}.$$

In this definition, k is the *number of parts* of this graph partition problem. The class of k-colorable graphs is a particular case of graph partition problem. To see this, set $\ell_{ij} = 0$ and $u_{ij} = 1$ whenever $i < j$, so that condition 1 in Definition 2.4 is always satisfied, and $\ell_{ii} = u_{ii} = 0$ for $i \in [k]$, so that condition 2 ensures that V_i is a stable set.

The graph partition problems in the literature [22, 23] uses constraints of the type

$$\ell'_{ij} \leq \frac{|E(V_i, V_j)|}{n^2} \leq u'_{ij} \tag{2.2}$$

instead of the constraints in Definition 2.4. These graph partition problems are known to be testable with two-sided tester and query complexity polynomial in $\frac{1}{\epsilon}$. Goldreich and Trevisan [23] characterized the graph partition problems defined in (2.2) that are one-sided.

The bound of Alon and Krivelevich in [5] extends directly to the following problems. These problems contain the one-sided testable classes characterized in [23].

Theorem 2.2. *Let $\mathcal{P}$ be a graph partition problem with k parts such that all parameters belongs to $\{0,1\}$. Then $\mathcal{P}$ is easily testable with query complexity is $\frac{36k\ln(k)}{\epsilon^2}$.*

First note that taking all parameters in $\{0,1\}$ means that every pairs (ℓ_{ij}, u_{ij}) is equal to one of $(0,0)$, $(0,1)$ and $(1,1)$. Note also that if $(\ell_{ii}, u_{ii}) = (0,1)$ for some $i \in [k]$ then the class $\mathcal{P}$ contains every graph, so in this case the property is trivial. Otherwise, the constraint can be restated as follows: each part is forced to be either a clique or an independent set and some pairs of parts may be forced to be complete or anti-complete.

Theorem 2.2 can be shown following the proof of Alon and Krivelevich in [5, Theorem 3]. As the adaptation is quite direct, we sketch the proof of Alon and Krivelevich for k-colorable graphs (see [5] for the complete proof), and we explain what needs to be modified to show Theorem 2.2.

Outline of the proof. [5] Let $\mathcal{P}$ be the class of k-colorable graphs. The goal of the proof is to show that if $G = (V, E)$ is a graph ϵ-far from the class $\mathcal{P}$ and R is a random subset of V of size $r = 36k \ln k/\epsilon^2$, then $G[X] \notin \mathcal{P}$ with probability at least $\frac{1}{2}$.

For the sake of the analysis, Alon and Krivelevich consider that R is generated in r rounds, each time choosing a new vertex v_j in V.

At each step of the algorithm, they consider a subset $S \subseteq R$ and a k-partition $\phi : S \to [k]$ for S, that is consistent with the constraints of the k-coloring, i.e. each part is an independent set. Given such an S and such a ϕ, they consider for every vertex $v \in V$ the list $L_\phi(v)$ of colors $i \in [k]$ such that extending ϕ to $S \cup \{v\}$ by $\phi(v) = i$ satisfy the constraints. In the case of k-coloring, $L_\phi(v)$ is the set of colors in $[k]$ such that v has no neighbor in S with this color. The value

$$E_\phi := \sum_{v \in V} |L_\phi(v)|$$

is a measure of the freedom we have when trying to extend the partition ϕ to V. A vertex v with $L_\phi(v) = \varnothing$ is called *colorless*. If the algorithm draws a colorless vertex v, then we know that ϕ does not extends properly to $S \cup \{v\}$. Let U be the set of colorless vertices.

A greedy coloring of the whole graph G extending ϕ consists of assigning to a vertex $v \in V \setminus (S \cup U)$ the color $c = \alpha_\phi(v)$ that minimizes the number $\delta^c_\phi(v)$ of neighbors of v with a color $c' \in L_\phi(v)$ such that the constraints force that u does not have color c and v color c' simultaneously. In the particular case of k-coloring, $\delta^c_\phi(v)$ is the number of $v \in N(u)$ with $c \in L_\phi(v)$. Let $\delta_\phi(v) := \min_{c \in L_\phi(v)} \delta^c_\phi(v)$ be this minimum. If u is colorless, then $\alpha_\phi(u)$ is chosen arbitrarily. Now, the number

$$\Delta_\phi := \sum_{u \in V} \delta_\phi(v) + n|U| \tag{2.3}$$

is an upper bound on the number of edges one has to change to make α_ϕ a valid partition. By the hypothesis, Δ_ϕ is at least ϵn^2.

It follows that with probability at least ϵ, either a colorless vertex is drawn, which proves that ϕ cannot be extended, or a vertex v with $\delta_\phi(v)$ at least ϵn, in which case E_ϕ decreases by at least $\delta_\phi(v) \geq \epsilon n$ independently of the color assigned to v. This last case cannot happen too often (less than $\frac{k}{\epsilon}$ times) since $0 \leq E_\phi \leq kn$, so the algorithm is likely to end up finding a colorless vertex, proving that the particular ϕ considered does not extend to the sampling.

A further analysis of the process then shows that with some good probability, *every* choice of ϕ ends with a contradiction. $\square$

Let us explain how to use this proof to show Theorem 2.2 (that is the same result extended to graph partition problems), doing only the following a few changes.

Fix a graph partition problem. Throughout the proof ϕ represents a k-partition $V = \cup_{i=1}^k V_i$ such that every part $V_i := \phi^{-1}(i)$ satisfy the condition in Definition 2.4. The set $L_\phi(v)$ is then defined as

$$L_\phi(v) := [k] \setminus \{1 \leq i \leq k| \quad \exists u \in S \setminus \{v\}, (v \in N(v) \text{ and } u_{i\phi(u)} = 0) \text{ or } \\ (v \notin N(v) \text{ and } u_{i\phi(u)} = 1)\}.$$

Then for every $c \in L_\phi(v)$, define

$$\begin{aligned} \delta^i_\phi(v) \ &:= |\{\, u \in N(v) \mid \exists j \in L_\phi(u), u_{ij} = 0 \,\}| \\ &+ |\{\, u \notin N(v) \mid \exists j \in L_\phi(u), \ell_{ij} = 1 \,\}|. \end{aligned}$$

and $\delta_\phi(v)$ as $\delta_\phi(v) := \min_{i \in L_\phi(v)} \delta^i_\phi(v)$. This allows us to define the set U of *colorless* vertices as the set of vertices $v \in V$ with $L_\phi(v) = \varnothing$. It is then easy to see that the number Δ_ϕ defined by (2.3) (with the new definitions) is an upper bound on the number of edges to change in G to satisfy the constraints and that defining a color $\phi(v)$ for v decreases the value E_ϕ (defined with the same expression) by $\delta_\phi(v)$. With these definitions, the rest of the proof follows similarly and provides the same bound on the query complexity.

2.4 SPLIT is easily testable

In this section we prove that the class of split graphs is testable with polynomial query complexity. Let us first recall the definition.

Definition 2.5. A graph G is a *split graph* if the set of vertices can be partitioned into two sets K and I such that K is a clique and I an independent set of G. In this case, (K, I) is a *split partition* of G.

The class SPLIT of split graphs corresponds to the graph partition problem with two parts and parameters $(\ell_{12}, u_{12}) = (0, 1)$, $(\ell_{11}, u_{11}) = (1, 1)$ and $(\ell_{22}, u_{22}) = (0, 0)$. Consequently, Theorem 2.2 applies to split graphs and shows that SPLIT is testable with query complexity $\frac{c}{\epsilon^2}$, where $c = 36 \cdot 2 \ln 2$.

We prove that SPLIT is testable with query complexity $O(\frac{1}{\epsilon} \ln \frac{1}{\epsilon})$. We first show the following structural result about split graphs that shows that the split partitions of a split graph are almost equal.

Property 2.3. *Let $G = (V, E)$ be a split graph, there is a split partition (K_0, I_0) of G such that one of the following holds.*

- *There is $U \subseteq K_0$ such that the split partitions of G are exactly (K_0, I_0) and $(K_0 \setminus \{v\}, I_0 \cup \{v\})$ for each $v \in U$; or*
- *there is $U \subseteq I_0$ such that the split partitions of G are exactly (K_0, I_0) and $(K_0 \cup \{v\}, I_0 \setminus \{v\})$ for each $v \in U$.*

Proof. First assume that G has two split partitions (K_1, I_1) and (K_2, I_2) such that $K_1 \nsubseteq K_2$ and $K_2 \nsubseteq K_1$. In this case, $K_1 \cap I_2$ and $K_2 \cap I_1$ have size at most one since they are both cliques and independent sets. It follows that $K_1 \cap I_2 = \{v_1\}$ and $K_2 \cap I_1 = \{v_2\}$ for some vertices $v_1, v_2 \in V$.

We may assume without loss of generality that $v_1 v_2 \in E$. Indeed, the statement of Property 2.3 is symmetrical toward graph complementary, that is it is satisfied by G if and only it is satisfied by $\overline{G}$. As a consequence, if $v_1 v_2 \notin E(G)$, it suffices to prove the property for $G' = \overline{G}$ instead, with the split partitions $(K'_1, I'_1) = (I_1, K_1)$ and $(K'_2, I'_2) = (I_2, K_2)$ (i.e. cliques and stable sets have been swapped) and then it holds that $v_1 v_2 \in E(G')$.

In this case, the set $K_0 = K_1 \cup K_2$ is a clique. Define $I_0 = V \setminus K_0 = (I_1 \cap I_2)$ and $U = \{ v \in K_0 \mid v \text{ has no edge to } I_0 \}$. We prove that (K_0, I_0) and U satisfy the second item of Property 2.3.

First note that (K_0, I_0) is a split partition of G, as $I_0 \subseteq I_1$ is an independent set. Moreover, for each $v \in U$, it follows from the definition of U that $(K_0 \setminus \{v\}, I_0 \cup \{v\})$ is a split partition.

It remains to show that every split partition (K, I) of G is of the form above. Let us prove that $I_0 \subseteq I$. Since $v_1 v_2 \in E$, one of v_1 and v_2 is in K. Further, since v_1 and v_2 have no edge to I_0 (because $I_0 \cup \{v_1\} \subseteq I_2$ and $I_0 \cup \{v_2\} \subseteq I_1$ respectively), no vertex of I_0 is in K, so $I_0 \subseteq I$.

Assuming that $I \neq I_0$, there is exactly one vertex v in $I \setminus I_0 = I \cap K_0$ since this set is both a clique and an independent set, and further $(K, I) =$

$(K_0 \setminus \{v\}, I_0 \cup \{v\})$. Moreover, $v \in U$ because $I_0 \cup \{v\} = I$ is an independent set, which prove the sought property.

Now assume that for every pair of split partitions (K_1, I_1) and (K_2, I_2) of G, it holds that $K_1 \subseteq K_2$ or $K_2 \subseteq K_1$. Since moreover, $|K_1 \setminus K_2| = |K_1 \cap I_2| \leq 1$ and $|K_2 \setminus K_1| \leq 1$, it follows G has either exactly one split partition or exactly two split partitions (K_0, I_0) and $(K_0 \setminus \{v\}, I_0 \cup \{v\})$ or $(K_0 \cup \{v\}, I_0 \setminus \{v\})$. This finishes the proof. □

As consequence of Property 2.3, a split graph with n vertices has at most $n+1$ split partitions. We are now ready to show the following theorem.

Theorem 2.4. *The class* SPLIT *is testable with query complexity* $O(\frac{1}{\epsilon} \ln \frac{1}{\epsilon})$.

Proof. **Algorithm**

Fix a graph $G = (V, E)$ with n vertices. Let s be a positive integer to be fixed later. We draw a subset R of size $2s$, chosen uniformly at random among the subsets of V of size $2s$. For the needs of the proof, we partition R into two sets S and T with $|S| = |T| = s$. We do as if S and T are chosen successively and independently. The set S is used to guess the structure of G as a split graph (i.e. what a split partition of G can be) and T is used to check this structure.

We assume that G is ϵ-far from being a split graph and we aim to prove that with probability at least $\frac{1}{2}$, the induced subgraph $G[R]$ is not a split graph.

The parameter s is chosen so that

1. $\frac{2}{\epsilon} e^{-\epsilon s} \leq 1/4$ and
2. $s e^{-s\epsilon/4} \leq 1/8$.

This assumptions appear naturally later in the proof. Note that $s = \frac{c}{\epsilon} \ln(\frac{1}{\epsilon})$ meets these conditions when c is a large enough real number.

Guessing the structure

Given S, assume the event

$$E_1 : G[S] \text{ is a split graph.}$$

As a consequence of Property 2.3, $G[S]$ has at most $s+1$ split partitions. Fixing such a partition (K_0, I_0), we classify the other vertices of V as follows, depending on their relations with I_0 and K_0:

- $\widetilde{K}$ is the set of $v \in V \setminus S_0$ complete to K_0.
- $\widetilde{I}$ is the set of $v \in V \setminus S$ with no neighbor in I_0.

Then define

- $K := \widetilde{K} \setminus \widetilde{I}$
- $I := \widetilde{I} \setminus \widetilde{K}$

- $J := \widetilde{K} \cap \widetilde{I}$
- $U := V \setminus (S \cup \widetilde{I} \cup \widetilde{K})$

which gives a partition $V = K \cup I \cup J \cup U \cup S$.

If we ignore J and U, then $(K_0 \cup K, I_0 \cup I)$ is a good candidate for a split partition of $G[V \setminus (J \cup U)]$. Let $A := E(G[I]) \cup \overline{E}(G[K])$ be the set of pairs that we have to change to indeed make it a partition, that is the edges of $G[E]$ and non-edges of $G[K]$.

Recall that G is assumed to be ϵ-far from SPLIT. Since G can be made split by changing the edges of A and removing all edges incident to a vertex of J or U, it follows that

$$|A| + n(|J| + |U|) \geq \epsilon n^2. \tag{2.4}$$

We shall use this inequality later.

Checking the guess

Observe that if (K_1, I_1) is a valid split partition of $G[R]$ extending (K_0, I_0), i.e. a partition with $K_0 \subseteq K_1$ and $I_0 \subseteq I_1$, then every vertex of $K_1 \cap T$ belongs to $\widetilde{K}$ and every vertex of $I_1 \cap T$ belongs to $\widetilde{I}$. It follows that in this case T does not intersect U and that for two vertices u and v of $K \cap T$ (resp. of $I \cap T$), we have $u, v \in K_1$ and further $uv \in E$ (resp. $u, v \in I_1$ and $uv \notin E$).

As a consequence of this observation, the partition (K_0, I_0) may extend to a split partition of $G[R]$ only if both of following events hold

- $E_2 : U \cap T = \varnothing$.
- $E_3 : \binom{T}{2} \cap A = \varnothing$.

The probability that E_2 does not hold is bounded from above by the probability that $|T| = s$ independent random vertices miss the set U. Therefore,

$$\mathbf{P}(\overline{E_2}) \leq \left(1 - \frac{|U|}{n}\right)^s \leq e^{-s\frac{|U|}{n}}.$$

Similarly, the probability that E_3 does not hold is at most the probability that $|T|/2 = s/2$ independent random pairs miss A, so

$$\mathbf{P}(\overline{E_3}) \leq \left(1 - \frac{2|A|}{n^2}\right)^{s/2} \leq e^{-s\frac{|A|}{n^2}}.$$

J is small

Fix a vertex $v \in V$ and observe that v belongs to J only if $N(v) \cap S$ is a clique and $S \setminus N(v)$ is an independent set. Let J' be the set of vertices with this last property, so that $J \subseteq J'$. The use of this set is that J' does not depend on a partition (K_0, I_0) (while J may).

Set $K_v = N(v) \cup \{v\}$ and $I_v = V \setminus (N(v) \cup \{v\})$ and let A_v be the set of edges that prevent (K_v, I_v) from being a split partition of G, i.e. the non-edges between neighbors of v and the edges between non-neighbors of v. By the remark above, $v \in J'$ if only if $\binom{S}{2}$ does not intersect A_v.

Since G is ϵ-far from SPLIT, it moreover holds that

$$|A_v| \geq \epsilon n^2$$

because G can be made split with split partition (K_v, I_v) by changing every edge of A_v.

The probability that $\binom{S}{2}$ intersects A_v is at least the probability that $s/2$ pairs chosen independently at random intersects A_v. It follows that

$$\mathbf{P}(v \in J') \leq \left(1 - \frac{2|A_v|}{n^2}\right)^{\frac{s}{2}} \leq (1 - 2\epsilon)^{\frac{s}{2}} \leq e^{-\epsilon s}.$$

It follows that

$$\mathbf{E}(|J'|) \leq n e^{-\epsilon s}.$$

Define the event

$$E_4 : |J'| \leq \frac{\epsilon}{2} n.$$

By Markov inequality, it holds that $\mathbf{P}(\overline{E_4}) \leq 2\frac{\mathbf{E}(|J'|)}{\epsilon n} \leq \frac{2}{\epsilon} e^{-\epsilon s} \leq \frac{1}{4}$ by our assumption on s. Note that if E_4 holds then $|J| \leq |J'| \leq \frac{1}{4}$ for every choice of (K_0, J_0).

Conclusion

Assuming E_4 and fixing a split partition (K_0, J_0), it follows from (2.4) that $\frac{|A|}{n^2} + \frac{|U|}{n} \geq \epsilon/2$. We distinguish two cases, depending on the largest term in this sum:

Case 1: if $\frac{|U|}{n} \geq \epsilon/4$, then $\mathbf{P}(\overline{E_2}) \leq e^{-s\frac{\epsilon}{4}} \leq \frac{1}{8s}$

Case 2: if $\frac{|A|}{n^2} \geq \epsilon/4$, then similarly $\mathbf{P}(\overline{E_3}) \leq e^{-s\frac{\epsilon}{4}} \leq \frac{1}{8s}$.

Recall that the events E_2 and E_3 are related to only *one* split partition (K_0, I_0) of $G[S]$. In both cases, the probability that (K_0, I_0) extends to a split partition of $G[R]$ is at most $\frac{1}{8s}$. Since S has at most $s+1$ split partitions, the probability that at least one of them extends to R is at most $\frac{s+1}{8s} \leq \frac{1}{4}$. Consequently,

$$\mathbf{P}(G[R] \in \text{SPLIT}) \leq \mathbf{P}(E_4) - \frac{1}{4} \leq \frac{1}{2}.$$

□

2.5 Testability of graph classes with an inductive decomposition

If $G_1 = (V_1, E_1)$ and $G_2 = (V_2, E_2)$ are two graphs with disjoint vertex sets, the *disjoint union* of G_1 and G_2 is the graph on $V = V_1 \cup V_2$ with edges $E = E_1 \cup E_2$.

The *join* of G_1 and G_2 is the graph on vertex set $V = V_1 \cup V_2$ obtained from the disjoint union of G_1 and G_2 by adding all edges between V_1 and V_2. Formally the edge set of this graph is $E = E_1 \cup E_2 \cup V_1 \times V_2$.

The graphs without induced P_4 are also called *co-graph*. Cographs have the following characterization.

Proposition 2.5. *A graph is a* cograph *if and only if it can be obtained from graphs with only one vertex by joins and disjoint unions.*

Using this inductive structure, Alon et Fox [4] showed that the class of cographs is easily testable. Their method can be adapted to simpler classes that have a similar recursive decomposition scheme.

2.5.1 THRESHOLD is easily testable

In a graph $G = (V, E)$, a vertex $v \in V$ is *isolated* if $\deg_G(v) = 0$, and *universal* if $\deg_G(v) = |V| - 1$, i.e. $uv \in E$ for every $u \in V$. The vertex v is *homogeneous* if it is either isolated or universal.

Definition 2.6. A graph G is a *threshold graph* if one can associate every $v \in V(G)$ with a number $a_v \in [0, 1]$ in such way that

$$uv \in E \Leftrightarrow a_u + a_v \geq 1.$$

Equivalently, G is a threshold graph if it can be built inductively from the empty graph by successive additions of homogeneous vertices. Let THRESHOLD denote the set of threshold graphs.

Let us introduce a relaxed version of homogeneous. Given a parameter $\beta > 0$, a vertex v is *β-homogeneous* if

$$\frac{\deg_G(v)}{n-1} \notin [\beta, 1-\beta].$$

Note that an homogeneous vertex is in particular β-homogeneous for every $\beta > 0$. The following lemma shows that a sample of a graph without β-homogeneous vertex is unlikely to be a threshold graph.

Lemma 2.6. *Let G be a graph and β a positive real number, then at least one of the following propositions is satisfied.*

1. *G has a β-homogeneous vertex.*
2. *$\mathbf{P}(G[X] \in \text{THRESHOLD}) \leq 2re^{\beta(r-1)}$ where X is a random set chosen uniformly among the subsets of V of size r.*

Proof. Assume that G has no β-homogeneous vertex and let us show the second item. Fix $v \in X$, the probability for a random vertex u chosen uniformly in $V \setminus \{v\}$ that $uv \in E$ is $\frac{\deg_G(v)}{n-1}$. Since v is not β-homogeneous, it holds that

$\beta \leq \frac{\deg_G(v)}{n-1} \leq 1-\beta$. Consequently the probability that $v \in X$ has no neighbor in X satisfies

$$\mathbf{P}(v \text{ is isolated in } G[X]) \leq \left(1 - \frac{\deg_G(v)}{n-1}\right)^{r-1} \leq e^{\beta(r-1)}.$$

Similarly,

$$\mathbf{P}(v \text{ is universal in } G[X]) \leq \left(\frac{\deg_G(v)}{n-1}\right)^{r-1} \leq e^{\beta(r-1)}.$$

Consequently, the probability that v is homogeneous in $G[X]$ is at most $2e^{\beta(r-1)}$. By the union bound, the probability that one of the r vertices in X is homogeneous in $G[X]$ is at most $2re^{\beta(r-1)}$. □

We are ready to show that THRESHOLD is easily testable.

Theorem 2.7. THRESHOLD *is testable with query complexity* $\frac{8}{\epsilon^{-2}}$.

Proof. Let $G = (V, E)$ be a graph ϵ-far from THRESHOLD. We peel the vertices of G as follows: as long as G has a ϵ-homogeneous vertex v, we remove this vertex from G. We obtain a decomposition of G constituted of a set of vertices $U \subseteq V(G)$ and an enumeration $v_1, \ldots, v_\ell$ of $V \setminus U$ such that v_i is ϵ-homogeneous to $U \cup \{v_{i+1}, \ldots, v_\ell\}$ for every $i \in [\ell]$ and such that $G[U]$ has no ϵ-homogeneous vertices.

The graph G can be transformed into a threshold graph G_0 by adding and removing edges as follows. First add up to $\binom{|U|}{2}$ edges to make U a clique. Then for each $i \in [\ell]$ make v_i homogeneous to $U \cup \{v_{i+1}, \ldots, v_\ell\}$. Since v_i is ϵ-homogeneous to $U \cup \{v_{i+1}, \ldots, v_\ell\}$, it can be done by adding or removing at most $\epsilon(n-i)$ edges. Let G_0 be the obtained graph. This graph is threshold graph because a clique is a threshold graph and the class of threshold graphs is stable by adding universal or isolated vertices. Moreover, we have

$$d(G, G_0) \leq \binom{|U|}{2} \cdot \frac{1}{n^2} + \sum_{i=1}^{\ell} \epsilon \frac{n-i}{n^2} \leq \frac{|U|^2}{2n^2} + \frac{\epsilon}{2}$$

and further

$$\frac{|U|^2}{2n^2} + \epsilon/2 \geq d(G, \text{THRESHOLD}) > \epsilon.$$

It follows that $|U|^2/n^2 \geq \epsilon$, so

$$|U| \geq \epsilon^{\frac{1}{2}} \cdot n.$$

The idea of the rest of the proof is the following: by the previous inequality, a random subset of V is likely to have a large intersection with U, and by Lemma 2.6, this intersection is unlikely to be a threshold graph.

Sample a subset Y of y vertices of V, where y will be fixed later. The probability for a vertex of Y to be in U is $\frac{|U|}{n}$. Moreover, the probability law of $|Y \cap U|$ is more concentrated that the binomial law with y samplings and parameter $p = \frac{|U|}{n}$, whose average is yp and variance is $yp(1-p)$. By Hoeffding inequality, it follows that

$$\mathbf{P}\left(\frac{|Y \cap U|}{y} - \frac{|U|}{n} < -a\right) \leq \exp(-2ya^2)$$

for every $a > 0$. Taking $a = \frac{|U|}{2n}$,

$$\mathbf{P}\left(\frac{|Y \cap U|}{y} < \frac{|U|}{2n}\right) \leq \exp\left(-\frac{y|U|^2}{2n^2}\right) \leq \exp\left(-\frac{y\epsilon}{2}\right).$$

Set $r = \frac{1}{2}y\epsilon^{\frac{1}{2}} \leq y\frac{|U|}{2n}$, so that

$$\mathbf{P}\left(|Y \cap U| \geq r\right) \geq 1 - \exp\left(-\frac{y\epsilon}{2}\right).$$

Let E_1 be the event that $|Y \cap U| \geq r$. Assuming E_1, we subsample a subset $X \subseteq Y$ uniformly at random in Y among subsets of size r. This set X is distributed as a random subset of U of size r chosen uniformly. By Lemma 2.6, the event E_2 that $G[X]$ is not a threshold graphs happens with probability at least $1 - 2r\epsilon^{r-1}$ assuming E_1. It follows that

$$\mathbf{P}(G[Y] \in \text{THRESHOLD}) \leq \mathbf{P}(\overline{E_1}) + \mathbf{P}(\overline{E_2}|E_1) \leq \exp\left(-\frac{y\epsilon}{2}\right) + 2re^{\epsilon(r-1)}.$$

Choosing $y = 8\epsilon^{-2}$, we have $r = 4\epsilon^{-3/2}$ and

- $\exp(-\frac{y\epsilon}{2}) = e^{-4\epsilon^{-1}} \leq e^{-4}$.
- $2re^{\epsilon(r-1)} \leq 8\epsilon^{-3/2}e^{3\epsilon^{-1/2}} = 8\left((\epsilon^{-1/2})e^{-(\epsilon^{-1/2})}\right)^3 \leq \frac{8}{e^3}$

Since the sum of these two numbers is less than $1/2$, it follows that $G[Y]$ is not a threshold graph with probability at least $1/2$. $\square$

2.5.2 Trivially perfect graphs

Let us use a similar approach to show that trivially perfect graphs are easily testable.

Definition 2.7. A graph G is *trivially perfect* if it is the intersection graph of a family of nested intervals $(I_v)_{v \in V}$, that is such that $I_u \subseteq I_v$ or $I_v \subseteq I_u$ whenever $I_u \cap I_v \neq \varnothing$.

Equivalently, a graph is trivially perfect if it can be inductively constructed from graphs on one vertex by disjoint union or addition of universal vertices. A *cut* of a graph $G = (V, E)$ is a partition of V into two non-empty sets A and B such that there are no edge between a vertex of A and a vertex of B. By the characterization above, a trivially perfect graph G with at least one vertex has either a cut or a universal vertex. Let us write TP the class of trivially perfect graphs.

The class TP can be shown to be is easily testable by a proof similar as for THRESHOLD and COGRAPH. We first relax the notions of disjoint union and universal vertex. In a graph $G = (V, E)$ with n vertices, a vertex $v \in V$ is β-universal if

$$\frac{\deg(v)}{n-1} > 1 - \beta$$

Similarly, a β-cut is a partition of V into two non-empty sets A and B such that

$$\frac{E(A, B)}{|A| \cdot |B|} < \beta.$$

The β-cut of G and the probability to have a cut in a random induced subgraph $G[X]$ are linked as follows.

Lemma 2.8 (Alon, Fox)**.** *There is a constant c such that for every $\beta > 0$ and $r \geq \frac{c}{\beta^3}$, if a graph $G = (V, E)$ has no β-cut, then $G[X]$ has no cut with probability $1 - re^{-c\beta^2 r}$, where X is sampled uniformly a random among the subsets of V of size r.*

Although not explicitly stated, this lemma can be directly derived from the argument of the proof of Theorem 2.1. in [4].

For universal vertices, a variant of Lemma 2.6 shows the following

Lemma 2.9. *For $\beta > 0$, if G is a graph without β-universal vertex, then $G[X]$ has no universal vertex with probability $re^{-\beta(r-1)}$, where X is a random set chosen uniformly among the subsets of V of size r.*

The inductive structure of trivially perfect graphs gives the following decomposition lemma.

Lemma 2.10. *Let G be a graph of size n that is ϵ-far from* TP*, then G has an induced subgraph of size at least ϵn with no ϵ-cut and no ϵ-universal vertex.*

Proof. Assume otherwise that G is a graph whose induced subgraphs of size at least ϵn all have an ϵ-cut or ϵ-universal vertex and let us prove that G is at distance ϵ from a trivially perfect graph.

We actually prove inductively the following stronger property: If H is an induced subgraph of G of size k, then H can be transformed into a trivially perfect graph by adding or removing at most $\frac{1}{2}(kn + k^2)\epsilon$ edges.

If $k \leq \epsilon n$ then H can be changed into a complete graph by adding at most $\binom{k}{2} < \frac{1}{2}k\epsilon n$ vertices. We now assume that $k > \epsilon n$, so H has an ϵ-cut or an ϵ-universal vertex.

If H has an ϵ-universal vertex v, we transform H as follows. We add or remove edges so that $H \setminus \{v\}$ is trivially perfect, by the induction hypothesis applied on $H \setminus \{v\}$, this is possible by adding or removing at most $\frac{1}{2}(n(k-1) + (k-1)^2)\epsilon$, then we add at most ϵk edges so that v is a universal vertex.

The obtained graph H' is therefore trivially perfect. The number of added or removed edges is at most

$$\frac{1}{2}(n(k-1)+(k-1)^2)\epsilon+\epsilon k=\frac{1}{2}(n(k-1)+k^2-1)\epsilon\leq\frac{1}{2}(nk+k^2)\epsilon.$$

If H has an ϵ-cut $V(H)=A_1\cup A_2$, set $k_1=|A_1|$ and $k_2=|A_2|$. For each $i\in\{1,2\}$, we transform $H[A_i]$ into a trivially-perfect graph changing at most $\frac{1}{2}(nk_i+k_i^2)\epsilon$ edges, then we remove every of the at most ϵk_1k_2 edge from A_1 to A_2. The graph we obtain is trivially perfect and the number of edges added or removed is at most

$$\frac{1}{2}(nk_1+k_1^2)\epsilon+\frac{1}{2}(nk_2+k_2^2)\epsilon+\epsilon k_1k_2=\frac{1}{2}(nk+k^2)\epsilon.$$

This concludes the proof of the Lemma. □

We are now ready to deduce that TP is easily testable.

Theorem 2.11. TP *is testable with query complexity* $O(\frac{1}{\epsilon^4})$.

Proof. Let $G=(V,E)$ be a graph ϵ-far from being a trivially perfect graph. By Lemma 2.10 there is a set $U\subseteq V$ of size at least ϵn that induces a subgraph $G[U]$ without ϵ-cut nor ϵ-universal vertex. By the Azuma-Hoeffding inequality, a random set X chosen uniformly among the subsets of V of size $m:=\frac{2}{\epsilon^4}$ intersects U in a set of size at least $r:=\epsilon\frac{m}{2}=\epsilon^{-3}$ with probability at least $1-p_1$ with $p_1:=\exp(-\frac{1}{2}\epsilon^2m)=1-\exp(-\epsilon^{-2})$.

If this happens, it follows from Lemmas 2.8 and 2.9 that the probability that $G[U\cap X]$ contains a universal vertex is at most $p_2:=re^{-\epsilon(r-1)}=\epsilon^{-3}\exp(-\epsilon^{-2}+\epsilon)$ and the probability that it has a cut is at most $p_3:=re^{-c\epsilon^2r}=\epsilon^{-3}e^{-c\epsilon^{-1}}$ when ϵ is large enough. If this last two events do not happen, then $G[U\cap X]$ – and thus $G[X]$ – is not a trivially perfect graph.

It follows that $\mathbf{P}(G[X]\in\mathrm{TP})\leq p_1+p_2+p_3$. To finish the proof, it suffices to notice that if ϵ is small enough, so that p_1, p_2 and p_3 are arbitrarily small, so in particular we can ensure that $1-p_1-p_1-p_3\geq\frac{1}{2}$. □

2.5.3 Classes with an inductive decomposition

One can prove similarly that classes defined inductively as the closure of the set containing only the graph with one vertex by a subset of the following operations: adding an universal vertex, adding an isolated vertex, disjoint union and join, are easily testable are easily testable. Unfortunately, there is essentially no other such classes than cograph, trivially perfects graphs, threshold graph and cliques.

The following table summarizes this fact. Note that adding a universal or isolated vertex are particular cases of respectively join and disjoint union as they correspond to the case where one class has size one.

	None	Add universal vertex	Join
None	$\{K_1\}$	$\{K_n\}$	$\{K_n\}$
Add universal vertex	$\{E_n\}$	THRESHOLD	co-TP
Disjoint union	$\{E_n\}$	TP	COGRAPH

The class co-TP is the set of complementary graphs of trivially perfect graphs. For $n \in \mathbf{N}$, K_n is the complete graph and E_n is the empty graph of size n.

2.6 Interval graphs are easily testable

A graph G is an *interval graph* is there is a family of $(I_v)_{v\in V}$ of intervals of $\mathbf{R}$ such that for every pair $u, v \in V$ of distinct vertices, $uv \in E$ if and only if I_u and I_v have a non-empty intersection. The family $(I_v)_{v\in V}$ is called an *interval representation* of G. We denote by $\mathcal{R}(G)$ the set of the interval representations of G.

This section aims to prove the following result.

Theorem 2.12. *The class of interval graphs is easily testable.*

The query complexity that appears in the proof is $s_{2.12}(\epsilon) = 2^{140}\left(\frac{1}{\epsilon}\right)^{20}$. We did not try to optimize the constant nor the exponent.

In the following, we use the letter ϵ to denote a (small) positive real. We always implicitly assume that $\epsilon < 1/2$.

2.6.1 Extending an interval representation and C-interval Graphs

The purpose of this subsection is to introduce a tool to deal with the following problem: given a (typically big) graph G far from being an interval graph and an interval representation $\mathcal{I}$ of interval induced subgraph of G, how likely am I to sample a certificate that this particular representation $\mathcal{I}$ does not extend to the whole graph G? As it happens to be, this question is strongly related to the testability of a stronger type of interval graphs.

Definition 2.8 (C-interval graphs)**.** Let $C = (V_1, V_2, (a_v)_{v\in V_1})$ where $V_1 \cup V_2 = V$ is a partition and a_v is a real number for every $v \in V_1$. Such a C is called a *constraint* on V. A graph $G = (V, E)$ is a *C-interval graph*, if G is the intersection graph of a family of intervals $(I_v)_{v\in V}$ such that for every $v \in V$,

- $a_v \in I_v$ if $v \in V_1$ and
- I_v is a singleton if $v \in V_2$.

The family $(I_v)_{v\in V}$ is called a *representation* of G as a C-interval graph. A C-interval graph is in particular an interval graph.

Remark 2.1 (Induced sub-constraint)**.** Let $G = (V, E)$ be a C-interval graph with $C = (V_1, V_2, (a_v)_{v\in V_1})$ and let $X = X_1 \cup X_2$ be a subset of V where $X_1 \subseteq V_1$ and $X_2 \subseteq V_2$. Define $C[X] := (X_1, X_2, (a_v)_{v\in X_1})$ and observe that the subgraph $G[X]$ of G induced by X is a $C[X]$-interval graph.

Remark 2.2. Every C-interval graph has a representation (I_v) where the bounds of the intervals I_v for all $v \in V_1$ are pairwise distinct. In what follows we consider only such representations.

Our proof of Theorem 2.12 is based on the following weaker property. It says that one can separate C-interval graphs from graphs ϵ-far from being an interval graph with polynomial query complexity.

Theorem 2.13. *Let $G = (V, E)$ be a graph and C be a constraint on V. Suppose that G is ϵ-far from interval graphs. The probability on an uniformly chosen subset $X \subseteq V$ of size $s_{2.13}(\epsilon) = 2^{64}\left(\frac{1}{\epsilon}\right)^{10}$ that $G[X]$ is a $C[X]$-interval graph is less than $\frac{1}{2}$.*

The proof of Theorem 2.13 is postponed to Section 2.6.3. Before this, we explain how to deduce Theorem 2.12 from Theorem 2.13.

Extending interval representations to induced subgraphs.

The following lemma shows that testing if an interval representation of an induced subgraph of a graph G extends to G nearly boils down to testing if G is a C-interval graph for some constraint C.

Lemma 2.14. *Let $G = (V, E)$ be a graph, $S \subseteq V$ a set of vertices and $\mathcal{I} = (I_v)_{v\in S} \in \mathcal{R}(G[S])$ an interval representation of $G[S]$. There is a constraint $C = C(G, S, \mathcal{I})$ such that for every $T \subseteq V$ if*

- *for every $v \in T$, if $N(v) \cap S$ is a clique then $N(v) \cap T$ is a clique; and*
- *$\mathcal{I}$ extends to $G[S \cup T]$*

then $G[T]$ is a $C[T]$-interval graph.

Proof. Let V_2 be the set of vertices $v \in V$ whose neighbourhood in S is a clique and set $V_1 = V \setminus V_2$. Define $C = (V_1, V_2, (a_v)_{v\in V_1})$, where for every $v \in V_1$, the parameter a_v is defined as follows: we choose u_v and w_v in $S \cap N(v)$ such that $u_v w_v \notin E$ (such vertices exist by definition of V_1). As the intervals $I_{u_v} = [\ell_{u_v}, r_{u_v}]$ and $I_{w_v} = [\ell_{w_v}, r_{w_v}]$ are disjoint, we may assume without loss of generality that $r_{u_v} < \ell_{w_v}$. Then we set $a_v := r_{u_v}$.

Now, take $T \subseteq V$ with the hypotheses of the lemma and an extension $\mathcal{I}' = (I_v)_{v\in S\cup T}$ of $\mathcal{I}$ to $G[S \cup T]$. For every $v \in T$, the interval I_v intersects I_{u_v} and I_{w_v} with $r_{u_v} < \ell_{w_v}$, therefore $I_v \supseteq [r_{u_v}, \ell_{w_v}] \ni r_{u_v} = a_v$.

By what precedes, the collection $(I_v)_{v\in T}$ is an interval representation of $G[T]$ that satisfies $a_v \in I_v$ for every $v \in V_1$. To prove that $G[T]$ is a $C[T]$-interval graph and finish the proof, it suffices to prove that for every $v \in V_2 \cap S$, we can transform I_v into a singleton without changing the intersections between the intervals of $(I_v)_{v\in T}$.

Let $v \in V_2 \cap S$, the definition of V_2 and the hypothesis of the theorem implies that the neighbourhood $N_{G[T]}(v)$ of v in $G[T]$ is a clique. Consequently, $\bigcap_{u\in N_{G[T]}(v)} I_u$ is non-empty because intervals have the Helly property. Let $x \in \bigcap_{u\in N_{G[T]}(v)} I_u$, substituting $I_v = [\ell_v, r_v]$ by $I_v = \{x\}$ preserves the fact that $(I_v)_{v\in T}$ is an interval representation of $G[T]$. Indeed, x belongs to I_u if $uv \in E$ and $x \in I_v \subseteq \mathbb{R} \setminus I_u$ if $uv \notin E$ for every $u \in T$.

Iterating this last process for each vertex of V_2 gives a representation of $G[T]$ as $C[T]$-interval graph. □

For a graph G and $S, T \subseteq V(G)$, let $E_{\text{clique}}(S, T)$ denote the event that

$$\forall v \in T,\ N(v) \cap S \text{ is a clique } \Rightarrow N(v) \cap T \text{ is a clique.}$$

Combining Lemma 2.14 and Theorem 2.13 gives the following statement.

Lemma 2.15. *Let $G = (V, E)$ be a graph ϵ-far from being an interval graph. Let $S \subseteq V$ a be set of vertices and $\mathcal{I}$ be an interval representation of $G[S]$ (so $G[S]$ is an interval graph). Consider a random set T chosen uniformly at random among the subsets of V of size $s_{2.13}(\epsilon)$. With probability at least $\frac{1}{2}$, if $E_{clique}(S, T)$ occurs then $\mathcal{I}$ does not extend to $G[S \cup T]$.*

Proof. Let $C = C(G, S, \mathcal{I})$ be the constraint given by Lemma 2.14, so that a subgraph $G[T]$ induced by $T \subseteq V$ is a $C[T]$-interval graph whenever $\mathcal{I}$ extends to $G[S \cup T]$ and $E_{\text{clique}}(S, T)$ holds.

Let T be random subset of V of size $s_{2.13}(\epsilon)$. Since G is ϵ-far from being an interval graph, we know from Theorem 2.13 that $G[T]$ is not a $C[T]$-interval graph with probability at least $\frac{1}{2}$.

Moreover, if this occurs (i.e. $G[T]$ is not a $C[T]$-interval graph) and $E_{\text{clique}}(S, T)$ holds, then it follows from Lemma 2.14 that $\mathcal{I}$ does not extend to $G[S \cup T]$. □

In order to use Lemma 2.15, we now estimate the probability of $E_{c\ell ique}(S, T)$.

Lemma 2.16. *Assume $s = k(t-1)$ with $k, s, t \in \mathbf{N}$, then*

$$\mathbf{P}_{S,T}(E_{clique}(S, T)) \geq 1 - \frac{t^2}{s}.$$

where S and T are independent random sets chosen uniformly among the subsets of V of size respectively s and t

Proof. Let $v \in V$ and let p denote the probability for a random set T_0 taken uniformly among the subsets of V of size $t-1$ that $N(v) \cap T_0$ is not a clique. Note that $\mathbf{P}(N(v) \cap T \text{ is a clique} | v \in T) = 1 - p$. Since S contains $k = \frac{s}{t-1}$ independent subsets of $t-1$ vertices, it follows that $\mathbf{P}(N(v) \cap S \text{ is not a clique}) \leq$

p^k. As S and T are independent, assuming $v \in T$, the probability that $N(v) \cap S$ is a clique but $N(v) \cap T$ is not a clique is at most $p^k(1-p) \leq \frac{1}{k}$. This last inequality can be obtained by computing the maximum of the function $x \mapsto x^k(x-1)$ on $[0,1]$. This maximum is $\frac{k^k}{(k+1)^{k+1}}$, which is smaller than $\frac{1}{k}$.

By the union bound, we deduce that

$$\mathbf{P}(E_{\text{clique}}(S,T)) \geq 1 - \frac{t}{k} = 1 - \frac{t(t-1)}{s} \geq 1 - \frac{t^2}{s}.$$

□

Linear specialization of interval representations

Recall that $\mathcal{R}(G)$ denotes the set of interval representations of a the graph $G = (V,E)$, that is the set of interval sequences $(I_v)_{v \in V}$ such that $uv \in E \Leftrightarrow I_u \cap I_v \neq \varnothing$. Note that G is an interval graph if and only if $\mathcal{R}(G) \neq \varnothing$. If G is an induced subgraph of some graph H and $\mathcal{I} = (I_v)_{v \in V} \in \mathcal{R}(G)$ is an interval representation of G, we say that $\mathcal{I}$ *extends to* H if there is a representation $(I'_v)_{v \in V(H)} \in \mathcal{R}(H)$ that coincides with (I_v) on $V(G)$, i.e. such that $I_v = I'_v$ for every $v \in V(G)$. Let $\mathcal{R}_G(H) = \left\{ (I_v)_{v \in V(G)} \,\middle|\, (I_v)_{v \in V(H)} \in \mathcal{R}(H) \right\}$ be the set of representations of G that extend to H.

An important ingredient of our proof of Theorem 2.12 is the following property.

Lemma 2.17. *Let G be a graph on n vertices. Then every sequence $H_1, \ldots, H_\ell$ of induced subgraphs of G such that*

$$\mathcal{R}_G(H_1) \supsetneq \cdots \supsetneq \mathcal{R}_G(H_\ell)$$

has length at most $m_{2.17}(n) = 16n$.

Note that by a cardinality argument, it is trivial to prove this lemma for $m_{2.17}(n) = (2n)!$ since $(2n)!$ is an upper bound of the number of interval representations of G with pairwise distinct bounds (up to homeomorphisms, see Section 2.7). In order to prove that interval graphs are easily testable, we however need a polynomial bound.

A proof of this lemma is exposed in Section 2.7. We can now proceed to the proof of Theorem 2.12.

2.6.2 Proof of Theorem 2.12

Proof of Theorem 2.12. Let $G = (V,E)$ be a graph that is ϵ-far from being an interval graph. Set $t = s_{2.13}(\epsilon)$, $s = 16t^2$ and $p = 8m_{2.17}(s)$.

We sample uniformly at random in V a set S of size s and p sets $T_1, \ldots, T_p$, each of size t. The total number of vertices sampled is therefore $s + pt$. Let $E_1(S)$ be the event that $\mathbf{P}_T(E_{\text{clique}}(S,T)) \geq \frac{3}{4}$, where T is chosen uniformly at random among the subsets of V of size t. By Lemma 2.16,

$$1 - \frac{t^2}{s} \leq \mathbf{P}_{S,T}(E_{\text{clique}}(S,T)) \leq \mathbf{P}_S(E_1(S)) + \frac{3}{4}(1 - \mathbf{P}_S(E_1(S)))$$

provided that $t^2/s \in \mathbf{N}$. It follows that $1 - 4\frac{t^2}{s} \leq \mathbf{P}_S(E_1(S))$, which gives

$$\frac{3}{4} \leq \mathbf{P}_S(E_1(S))$$

since $s = 16t^2$.

For each i from 1 to p, choose $(I_v^i) \in \mathcal{R}_{G[S]}(G[S \cup T_1 \cup \cdots \cup T_{i-1}])$ an interval representation of $G[S]$ that extends to $G[S \cup T_1 \cup \cdots \cup T_{i-1}]$, which is possible whenever this last graph is an interval graph.

Let E_2^i denote the event that $(I_v^i)_{v \in S}$ does not extend to $G[S \cup T_i]$. Let us bound from below the probability of E_2^i. By Lemma 2.15 applied to S and $T = T_i$,

$$\mathbf{P}(E_2^i \cup \overline{E(S,T_i)} | E_1) \geq \frac{1}{2}.$$

Thus $\mathbf{P}(E_2^i | E_1) \geq \frac{3}{4} - \mathbf{P}(\overline{E(S,T_i)} | E_1) \geq \frac{1}{4}$.

Let J denote the set of indices i such that E_2^i occurs. We claim that $|J| < m_{2.17}(s)$. Indeed, let $i_1 < \cdots < i_{|J|}$ be an enumeration of J and set $H_j := G[S \cup T_{i_1} \cup \cdots \cup T_{i_j}]$ for $j \in \{1, \ldots, |J|\}$. We have $\mathcal{R}_{G[S]}(H_{j+1}) \subsetneq \mathcal{R}_{G[S]}(H_j)$, where the inclusion holds because H_j is an induced subgraph of H_{j+1} and the difference is given by $(I_v^{i_j}) \in \mathcal{R}_{G[S]}(H_j)$, which does not extend to $G[S \cup T_{i_j}]$ thus in particular not to $H_j = G[S \cup T_1 \cup \cdots \cup T_{i_j}]$. Hence Lemma 2.17 applies and gives $|J| < m_{2.17}(s)$.

Fix S such that E_1 holds. We proved that in this case $\mathbf{P}_{T_i}(E_2^i) \geq \frac{1}{4}$ unless $G[S \cup T_1 \cup \cdots \cup T_{i-1}]$ is not an interval graph. As long as the sampled vertices induce an interval graph, the size of J is therefore larger than a random variable $\mathbf{B}(\frac{1}{4}, p)$ with a binomial distribution of parameter $\frac{1}{4}$ and p experiments. Note moreover that by the Azuma-Hoeffding inequality,

$$\begin{aligned}\mathbf{P}(\mathbf{B}(1/4,p) \leq m_{2.17}(s)) &\leq \exp\left(-\frac{2(p/4 - m_{2.17}(s))^2}{p}\right) \\ &= \exp\left(-\frac{1}{4} m_{2.17}(s)\right) < \frac{1}{3}\end{aligned}$$

Since $|J| \leq m_{2.17}(s)$, it follows that with probability (on $T_1, \ldots, T_q$) that $G[S \cup T_1 \cup \cdots \cup T_p]$ is not an interval graph is at least $\frac{2}{3}$.

As a consequence, the probability on $S, T_1, \ldots, T_p$ that $G[S \cup T_1 \cup \cdots \cup T_p]$ is not an interval graph is at least $\mathbf{P}(E_1(S)) \cdot \frac{2}{3} \geq \frac{3}{4} \cdot \frac{2}{3} = \frac{1}{2}$.

The query complexity proved for the class of interval graphs is $s_{2.12}(\epsilon) = s + pt = 16t^2 + 8t \cdot m_{2.17}(16t^2) \leq 2^{12} t^2 = 2^{12}(s_{2.13}(\epsilon))^2 = 2^{140} \cdot \left(\frac{1}{\epsilon}\right)^{20}$.

$\square$

2.6.3 Proof of Theorem 2.13

The structure of V_1

In Section 2.6.3, we consider a graph G and a constraint $C = (V_1, V_2, (a_v)_{v \in V_1})$ on G.

For each vertex v of V_1, we define two formal variables r_v and ℓ_v that represent respectively the right and left ends of the interval associated to v in a representation. If $\mathcal{I} = (I_v)_{v \in V}$ is an interval representation of G, we denote by $\ell_v(\mathcal{I})$ and $r_v(\mathcal{I})$ the actual values of these bounds in $\mathcal{I}$, that is the real numbers such that $I_v = [\ell_v(\mathcal{I}), r_v(\mathcal{I})]$.

Let $L(V_1) = \{ \ell_v \mid v \in V_1 \}$ and $R(V_1) = \{ r_v \mid v \in V_1 \}$ be respectively the sets of left and right bounds of V_1 and set $B(V_1) = L(V_1) \cup R(V_1)$. Conversely, if $A \subseteq B(V_1)$, we denote by $V(A)$ the set of vertices v such that $\ell_v \in A$ or $r_v \in A$.

Notice that if $a_u < a_v$ for some $u, v \in V_1$ and $\mathcal{I} = (I_v)_{v \in V}$ is an interval representation of G satisfying $a_x \in I_x$ for every $x \in V_1$, then

$$\ell_u(\mathcal{I}) \leq a_u < a_v \leq r_v(\mathcal{I})$$

and

$$uv \in E \text{ if and only if } r_u(\mathcal{I}) \geq \ell_v(\mathcal{I}).$$

From this remark, we deduce a relation $\prec_G$ on every pair $\{r, \ell\}$ where $r = r_v \in R(V_1)$ and $\ell = \ell_u \in L(V_1)$ intended so that $b_1(\mathcal{I}) \leq b_2(\mathcal{I})$ whenever $b_1 \prec_G b_2$ for $b_1, b_2 \in B(V_1)$.

Definition 2.9. The relation $\prec_G$ is defined by the following rules:

1. $\ell_u \prec_G r_v$ if $a_u \leq a_v$.
2. $\ell_u \prec_G r_v$ if $a_u > a_v$ and $uv \in E(H)$.
3. $\ell_u \succ_G r_v$ if $a_u > a_v$ and $uv \notin E(H)$.

The cases of Definition 2.9 are illustrated in Figure 2.2. Note that every $\ell \in L(V_1)$ is "comparable" to every $r \in R(V_1)$, that is either $\ell \prec_G r$ or $r \prec_G \ell$ holds. As explained above, the fundamental property of $\prec_G$ is the following.

Observation 2.18. *For every $x, y \in B(V)$ and every representation $\mathcal{I}$ of G as a C-interval graph, if $x \prec_G y$ then $x(\mathcal{I}) \leq y(\mathcal{I})$.*

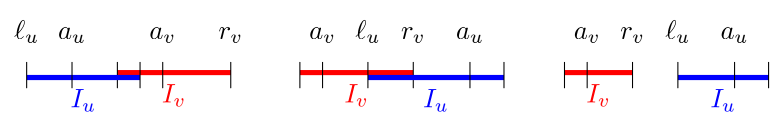


Figure 2.2 – Cases in the definition of $\prec_G$.

Another important property of this relation is that determining whether $r_v \prec_G \ell_u$ or $\ell_u \prec_G r_v$ only depends on a_u, a_v and whether $uv \in E(G)$. Consequently, the relation $\prec_H$ of the subgraph $H = G[X]$ induced by a set $X \subseteq V$ associated to the constraint $C[X]$ is the restriction of $\prec_G$ to X.

Remark 2.3. The relation $\prec_G$ does not only depend on G but also on C. As the constraint it refers to will always be obvious, we omit C in the notation.

Definition 2.10. Let $C = (V_1, V_2, (a_v)_{v \in V_1})$ be a constraint and G a graph on the vertex set $V = V_1 \cup V_2$. Define $D(\ell, \ell') = \{ r \in R(V_1) \mid \ell \prec_G r \prec_G \ell' \}$ for $\ell, \ell' \in L(A)$ and symmetrically $D(r, r') = \{ \ell \in L(V_1) \mid r \prec_G \ell \prec_G r' \}$ for $r, r' \in R(V_1)$. For $(b, b') \in L(V_1)^2 \cup R(V_1)^2$ we define $\mathrm{Between}(b, b') = D(b, b') \cup D(b', b)$.

The following Lemma allows us to control the structure of the part V_1 of the graph. It shows that in a graph G where a random sampling is likely to induce a C-interval graph, (most of) the bounds of V_1 can be gathered into blocks $M_1, \dots, M_p$ of bounds of same type that behave similarly.

Lemma 2.19. *Fix two parameters $0 < \epsilon < \frac{1}{2}$ and $k \geq 3$. Let $C = (V_1, V_2, (a_v)_{v \in V_1})$ be a constraint and $G = (V, E)$ be a graph on $V = V_1 \cup V_2$.*

Let S be a subset chosen uniformly at random among the subsets of V of size $s = s_{2.19}(\epsilon, k) = 3 \cdot 2^{11} \frac{(k+4)^3}{\epsilon^8}$. We write $S_1 = S \cap V_1$. With probability at least $1 - 2^{-k}$ the following occurs.

If $G[S]$ is a $C[S]$-interval graph then there exists a subset $B_0 = \{m_1, \dots, m_p\}$ of $B(S_1)$, a set $\mathcal{N} \subseteq V_1$ such that

- *$|\mathcal{N}| \leq \epsilon n$; and*
- *$m_i \in R(S_1)$ if i is even and $m_i \in L(S_1)$ if i is odd for $i \in [p]$;*
- *$m_i \prec_G m_{i+1}$ for every $1 \leq i \leq p - 1$;*

and further defining $M_i := D(m_{i-1}, m_{i+1})$ for $i \in [p]$,

- *$(M_i)_{i=1}^p$ form a partition of $B(V_1 \setminus \mathcal{N})$ and*
- *for every $i \in [p]$ and every $x \in M_i$, $|\mathrm{Between}(m_i, x)| \leq \epsilon n$.*

Proof. Let $0 < \epsilon_0 < \frac{1}{2}$ be a real number whose value will be fixed later. We partition S into two parts T and U of respective size t and u, to be defined later as well. Set $T_1 = T \cap V_1$ and $U_1 = U \cap V_1$.

For $(x, y) \in L(V_1)^2 \cup R(V_1)^2$, let $P(x, y)$ be the property

$$D(x, y) \text{ intersects } B(T_1) \text{ or } |D(x, y)| \leq \epsilon_0 n$$

and set

$$\mathcal{N}_1 := \left\{ x \in B(V_1) \;\middle|\; \exists y \in B(T_1),\ \overline{P(x, y)} \right\}.$$

Similarly, let $Q(x, y)$ be the property

$$D(x, y) \cap B(U_1) \neq \varnothing \text{ or } |D(x, y)| \leq \frac{\epsilon_0}{2t^2} \cdot n.$$

Let E_1 be the event that $|\mathcal{N}_1| \leq \epsilon_0 n$ and E_2 the event that $Q(x, y)$ occurs for all $(x, y) \in L(T_1)^2 \cup R(T_1)^2$. We show that both of these events have high probability.

Let $\mathcal{N}_2$ denote the union of $D(x, y)$ over all $(x, y) \in L^2(T_1) \cup R^2(T_1)$ such that $D(x, y)$ does not intersect $B(U_1)$. Note that if E_2 occurs then $|\mathcal{N}_2| \leq 2t^2 \cdot \frac{\epsilon_0}{2t^2} n \leq \epsilon_0 n$.

For $(x, y) \in L^2(T_1) \cup R^2(T_1)$, let us bound the probability $\mathbf{P}(\overline{P(x, y)})$ that $P(x, y)$ does not happen. If $|D(x, y)| \leq \epsilon_0 n$, then $P(x, y)$ is always true. Otherwise, the probability for an element s of S_1 that $s \notin V(D(x, y))$ is at most $\frac{1}{n}(n - |V(D(x, y))|) \leq 1 - \epsilon_0$. As a consequence,

$$\mathbf{P}(\overline{P(x, y)}) \leq (1 - \epsilon_0)^t \leq e^{-\epsilon_0 t}.$$

By the union bound, it follows that for every $x \in B(V_1)$,

$$\mathbf{P}(x \in \mathcal{N}_1) \leq t \cdot e^{-\epsilon_0 t}.$$

Consequently, $\mathbf{E}(|\mathcal{N}_1|) \leq 2nt \cdot e^{-\epsilon_0 t}$. Set $t = \frac{2(k+4)}{\epsilon_0^2}$, then Markov Inequality gives

$$\begin{aligned}\mathbf{P}(\overline{E_1}) &\leq \frac{2t}{\epsilon_0} e^{-\epsilon_0 t} \leq \frac{4(k+4)}{\epsilon_0^3} \cdot e^{-\frac{2k+8}{\epsilon_0}} \leq \frac{4(k+4)}{\epsilon_0^3} \cdot \epsilon_0^{2k+8} \\ &\leq (k+4)2^{-2k-3} \leq 2^{-k-1}.\end{aligned}$$

Similarly, for $(x, y) \in L(T)^2 \cup R(T)^2$, the probability $\mathbf{P}(\overline{Q(u, v)})$ is at most $(1 - \frac{\epsilon_0}{2t^2})^u \leq e^{-\frac{\epsilon_0}{2t^2} u}$. Set $u = 2(k+4)\frac{t^3}{\epsilon_0}$, then

$$\mathbf{P}(\overline{E_2}) \leq 2t^2 e^{-\frac{\epsilon_0}{2t^2} u} \leq 2t^2 e^{-(k+4)t} \leq 2t^{-k-2} \leq 2^{-k-1}.$$

This shows that $E_1 \cup E_2$ occur with probability at least $1 - 2^{-k}$. Note also that with these values of t and u, the size of S is

$$s = t + u \leq 2u = 4(k+4)\frac{t^2}{\epsilon_0} = 16\frac{(k+4)^3}{\epsilon_0^5}.$$

In the rest of the proof, we assume E_1, E_2 and that $G[S]$ is a C-interval graph and we construct $\mathcal{N}$ and B_0 as in the statement of the lemma.

Let $\mathcal{I}$ be a representation of $G[S]$ as a $C[S]$-interval graph. We order the elements of $B(T_1)$ into a sequence $b_1, \ldots, b_{2|S_1|}$ according to the order given by $\mathcal{I}$, i.e. such that $b_1(\mathcal{I}) < \cdots < b_{2|T_1|}(\mathcal{I})$. Note that by Observation 2.18, $i < j$ whenever $b_i \prec_G b_j$.

Claim 1. If $i < j$ then $D(b_j, b_i) \subseteq \mathcal{N}_2$ whenever $(b_i, b_j) \in L(V_1)^2 \cup R(V_1)^2$.

Proof. If there is an element $x \in B(U_1) \cap D(b_j, b_i)$ then by Observation 2.18, $b_j(\mathcal{I}) < x(\mathcal{I}) < b_i(\mathcal{I})$, which is impossible since $b_i(\mathcal{I}) < b_j(\mathcal{I})$. □

From the sequence $(b_i)_{1 \leq i \leq 2|T_1|}$, we extract a subsequence $m_1, \ldots, m_p$ by deleting every term $b_i \in L(T_1)$ (resp. $b_i \in R(T_1)$) such that $b_{i-1} \in L(T_1)$ (resp. $b_{i-1} \in R(T_1)$). In other words, we keep only the first element of each (maximal) block of consecutive elements of $L(T_1)$ (resp. $R(T_1)$). As a consequence, the sequence $(m_i)_{1 \leq i \leq p}$ alternates the bounds of $L(T_1)$ with the bounds of $R(T_1)$. It follows that this sequence satisfies $m_i \prec_G m_{i+1}$ for all $i \in \{1, \ldots, p-1\}$. Since the leftmost bound of $\mathcal{I}$ is a left bound, $m_1 \in L(T_1)$. Moreover, the following property holds.

Claim 2. For every $i \in [p]$ and $x \in B(V_1) \setminus \mathcal{N}_2$, if $D(m_i, x)$ intersects $B(S_1)$ then $D(m_i, x)$ intersects $B_0 = \{m_1, \dots, m_p\}$.

Proof. We take $b_j \in B(S_1)$ such that $m_i \prec_G b_j \prec_G x$ and we prove that $m_i \prec_G m_{i+1} \prec_G x$. Let j_1 and j_2 be the indices such that $m_i = b_{j_1}$ and $m_{i+1} = b_{j_2}$. With this notation, the hypothesis gives $b_{j_1} \prec_G b_j$, so $j_1 < j$. It follows from the construction of $(b_k)_{k=1}^p$ that every b_k with $j_1 \leq k < j_2$ (i.e. the bounds of $B(S_1)$ between m_i and m_{i+1} in $\mathcal{I}$) belongs to $L(T_1)$ (resp. $R(T_1)$) if $m_i \in L(T_1)$ (resp. $m_i \in R(T_1)$). Consequently, $j_2 \leq j$, and by the preceding claim $D(b_j, b_{j_2}) \subseteq \mathcal{N}_2$. It follows that $x \notin D(b_j, b_{j_2})$. Then $m_{i+1} = b_{j_2} \prec_G x$ (because $b_j \prec_G x$), which finishes the proof of the claim. □

Define M_i as the set of bounds $a \in B(V_1)$ such that for every j such that $i - j$ is odd, $m_j \prec_G a$ if $j < i$ and $a \prec_G m_j$ if $i < j$.

Claim 3. If $x \in B(V_1) \setminus \mathcal{N}_2$ then $x \in M_i$ for some $i \in [p]$.

Proof. If the claim does not hold, there is $x \in B(V_1)$ with $m_i \prec_G x \prec_G m_j$ for some $i, j \in [p]$ and $i > j$. In this case, $D(m_i, m_j)$ does not intersect $B(S)$ so by E_2, $x \in D(m_i, m_j) \subseteq \mathcal{N}_2$. □

For $x \in B(V_1) \setminus \mathcal{N}_2$, let $i(x)$ denote the index for which $x \in M_{i(x)}$.

Claim 4. If $y \in \text{Between}(m_{i(x)}, x)$, then either $y \in D(m_{i(x)}, x)$ or $x \in D(m_{i(y)}, y)$.

Proof. If $y \in \text{Between}(m_{i(x)}, x) \setminus D(m_{i(x)}, x) = D(x, m_{i(x)})$ then $x \prec_G y \prec_G m_{i(x)}$. We deduce that $i(y) < i(x)$ (since $y \in M_{i(y)}$ and $y \prec_G m_{i(x)}$) and consequently $m_{i(y)} \prec_G x \prec_G y$ (since $x \in M_{i(x)}$), which proves the claim. □

We now estimate the sum $\sum_{x \in B(V_1) \setminus \mathcal{N}_2} |\text{Between}(m_{i(x)}, x)|$. It follows from Claim 4 that

$$\sum_{x \in B(V_1) \setminus \mathcal{N}_2} |\text{Between}(m_{i(x)}, x) \setminus \mathcal{N}_2| \leq 2 \sum_{x \in B(V_1) \setminus \mathcal{N}_2} |D(m_{i(x)}, x)|.$$

As a consequence of Claim 1 and the definition of $M_{i(x)}$, the set $D(m_{i(x)}, x)$ does not intersect B_0 and further by Claim 2, $D(m_i, x)$ does not intersect $B(T_1)$ neither. It follows that $|D(m_i, x)| < \epsilon_0 n$ whenever $x \notin \mathcal{N}_1$. Thus by E_1 and E_2,

$$\sum_{x \in B(V_1) \setminus \mathcal{N}_2} |\text{Between}(m_{i(x)}, x)| \leq |\mathcal{N}_1| n + (n - |\mathcal{N}_1|)\epsilon_0 n \leq 2\epsilon_0 n^2.$$

By Markov inequality, the set $\mathcal{N}_3$ of elements x such that $|D(x, m_{i(x)})|$ is at least $2\epsilon_0^{\frac{1}{2}} n$ has size at most $\epsilon_0^{\frac{1}{2}} n$. To conclude, defining $\mathcal{N} := V(\mathcal{N}_1 \cup \mathcal{N}_2 \cup \mathcal{N}_3)$ and $\epsilon_0 := \frac{\epsilon^2}{4}$ (so that $2\epsilon_0^{\frac{1}{2}} = \epsilon$) suffices to yield the lemma. Indeed, $|\mathcal{N}| \leq |\mathcal{N}_1| + |\mathcal{N}_2| + |\mathcal{N}_3| \leq 2\epsilon_0 n + \epsilon_0^{\frac{1}{2}} n \leq (\frac{\epsilon^2}{2} + \frac{\epsilon}{2}) n \leq \epsilon n$. The size s depends on ϵ as $s = 24 \frac{(k+4)^3}{\epsilon_0^4} = 3 \cdot 2^{11} \frac{(k+4)^3}{\epsilon^8}$. □

Structure with V_2

A *pseudometric* d on a set E is symmetric function $d : E^2 \to \mathbb{R}$ such that $d(x,x) = 0$ for every $x \in E$ and that satisfies the triangular inequality. For $x \in E$ and $\rho > 0$, the *open ball* centered on x is the set $B_d(x,\rho) := \{ y \in E \mid d(x,y) < \rho \}$. If S is a subset of E, define $B_d(S,\rho) = \bigcup_{x\in S} B_d(x,\rho)$.

The following lemma shows that if a pseudometric subspace is not nearly covered by m balls, a small sample of this space is unlikely to be covered by m balls.

Lemma 2.20. *Fix parameters $0 < \epsilon$, $0 < \rho$, k and m. Let E be a finite set, F be a subset of E and d be a pseudo-metric on E. We take $S \subseteq E$ uniformly at random among all subsets of size $s \geq s_{2.20}(\epsilon,k,m) = \frac{km}{\epsilon}$. With probability at least $1 - me^{-k}$, at least one of the two following events occurs.*

1. *There is $A \subseteq S$ of size m such that for every nonequal $x, y \in A$, it holds $d(x,y) > \rho$.*
2. *$F \setminus B_d(S,\rho)$ has size at most $\epsilon|E|$.*

Proof. We make a proof by induction on m.

The case $m = 1$ is straightforward since Item 1 always holds whenever S is nonempty.

Suppose $m \geq 2$. Partition S into $S = S_0 \cup S_1$, where $|S_0| = s(\epsilon,k,m-1) = (m-1)\frac{k}{\epsilon}$ and $|S_1| = \frac{k}{\epsilon}$. By induction hypothesis, we know that with probability at least $1-(m-1)e^{-k}$, either $\epsilon n \geq |F \setminus B_d(S_0,\rho)| \geq |F \setminus B_d(S,\rho)|$, in which case Item 2 holds, or there is $A_0 \subseteq S_0$ of size $m-1$ satisfying $\forall x, y \in A, d(x,y) > \rho$.

Assume this last case holds. The negation of 2 implies that $|F \setminus B_d(S_0,\rho)| \geq \epsilon$. Hence the probability that S_1 does not intersect $F \setminus B_d(S_0,\rho)$ is at most

$$\left(1 - \frac{|F \setminus B_d(S_0,\rho)|}{n}\right)^{|S_1|} \leq (1-\epsilon)^{|S_1|} \leq e^{-\epsilon|S_1|} = e^{-k}.$$

Assuming these sets intersect, take $a \in (F \cap S_1) \setminus B_d(S_0,\rho)$. The set $A := A_0 \cup \{a\}$ is a witness of item 1. By the union bound, this happens with probability at least $1 - (m-1)e^{-k} - e^{-k} = 1 - me^{-k}$ ☐

As a 1-dimensional object for the neighbourhood distance, the set V_2 of a C-interval graphs is covered by balls of radius ρ that are in linear number in $\frac{1}{\rho}$.

Proposition 2.21. *Let H be a C-interval graph with $C = (V_1, V_2, (a_v)_{v\in V_1})$. Let d be the pseudometric defined by $d(u,v) := |N(u) \triangle N(v)|/|V|$ then for every $\rho > 0$ there is a subset $A \subseteq V_2$ of size at most $\left\lceil \frac{2}{\rho} \right\rceil$ such that $B_d(A,\rho)$ covers H.*

Proof. As an interval graph, we know that H has an interval representation $(I_v)_{v\in V}$ where each interval end is distinct (note that in such a representation, the vertices of V_2 are not necessarily represented by points. Further, we may also assume that these interval ends are exactly the elements of $\{1,\ldots,2n\}$.

Since H is a C-interval graph, the neighbourhood of every vertex $v \in V_2$ is a clique so it corresponds to the set of intervals intersecting for instance the lower bound x_v of I_v.

We claim that for $u, v \in V_2$, it holds $d(u,v) \leq |x_u - x_v|/n$. Indeed, if w belongs to $N(u) \triangle N(v)$ then I_w has an end between x_u and x_v, which is possible only for at most $|x_u - x_v| - 1$ vertices.

Define the clusters $C_i := \{ u \in V_2 \mid x_u \in]y_i, y_{i+1}] \}$ where $y_i := i\rho n$ for each integer $1 \leq i \leq \left\lceil \frac{2}{\rho} \right\rceil$. By the above property, the diameter (for d) of each C_i is at most ρ. To finish the proof, it suffices to construct A that contains one element of each nonempty C_i. □

Proof of Theorem 2.13. We partition X into three sets $X = S \cup T \cup U$. Instead of sampling X directly, we sample three successive sets S, T and U, chosen independently from each other.

Let E_1 be the event that $G[X] = G[S \cup T \cup U]$ is a $C[X]$-interval graph. If $\mathbf{P}(E_1) < \frac{1}{2}$ then the theorem holds. In the rest of the proof, we assume that $\mathbf{P}(E_1) \geq \frac{1}{2}$ and we prove that G is ϵ-close to be an interval graph.

Let $\epsilon_0 < \frac{1}{2}$ be a small positive number that will be chosen later in the proof.

The subsets S and T are chosen uniformly at random among all subsets of V of size $s = t = s_{2.19}(\epsilon_0, 4) = \frac{3 \cdot 2^{20}}{\epsilon_0^8}$. The set $|U|$ is chosen uniformly at random among the subsets of V of size $\frac{4}{\epsilon_0} t \ln t$. In particular $|U| \geq t = s$.

Let E_2 be the event that there are $\mathcal{N}_1 \subseteq V_1$, $(m_i)_{i=1}^{\ell}$ and $(M_i)_{i=1}^{\ell}$ as in Lemma 2.19, with $|P_1| \leq \epsilon_0 n$ and $\ell \leq s$. Lemma 2.19 ensures that $\mathbf{P}(E_1 \Rightarrow E_2) \geq 1 - 2^{-4} > 1 - \frac{1}{12}$, so

$$\mathbf{P}(E_1 \cap E_2) > \mathbf{P}(E_1) - \frac{1}{12} \geq \frac{1}{2} - \frac{1}{10}. \tag{2.5}$$

For $i \in [\ell]$, define

$$N(i) = \{ v \in V_1 \setminus \mathcal{N}_1 \mid \ell_v \in M_{j_1} \text{ and } r_v \in M_{j_2} \text{ with } j_1 < i < j_2 \}.$$

The set $N(i)$ corresponds to the intervals a point in the block M_i intersects. Given a node $V \in T_2$ and a subset $A \subseteq V$, we define

$$\alpha_i^A(v) = |(N(i) \triangle (N(v) \setminus (V(M_i) \cup \mathcal{N}_1))) \cap A|.$$

This is a measure of how suitable is the block M_i to place v. To control the relations between vertices inside a block define $\delta_i(x,y) := (N(x) \setminus N(y)) \cap V(M_i)$.

Note that assuming E_1, every connected component of $G[T_2]$ is a clique because in this case $G[T_2]$ is an intersection graph of points. Let Bad be the set of vertices $v \in V_2$ such that $N(x) \cap T_2$ is not a connected component (and a clique) of $G[T_2]$ and BadEdge be the set of pairs $uv \in V_2^2$ such that either $uv \in E$ and $N(u) \cap T \neq N(v) \cap T$ or $uv \notin E$ and $N(u) \cap N(v) \cap T_2 \neq \varnothing$. Note that if U intersect Bad or contains a pair of BadEdge then $G[T \cup U]$ is not a $C[T \cup U]$-interval graph.

Let d be the distance on V defined by $d(u,v) = |N(u) \triangle N(v)|$. We define the following events that depend on T and U.

E_3 : $|V_2 \setminus B_d(T_2, \epsilon_0)| \leq \epsilon_0 n$

E_4 : $\left|\frac{1}{|U|}\alpha_i^U(v) - \frac{1}{n}\alpha_i^V(v)\right| \leq \epsilon_0$ for every $v \in T_2$ and $0 \leq i \leq \ell$.

E_5 : For every $x, y \in T_2$ and $b \in B(S_1 \cup T_1 \cup U_1)$ with $b \in M_i$, the set $\delta_i(x, y) \setminus V(\text{Between}(m_i, b))$ intersects U unless $|\delta_i(x, y) \setminus V(\text{Between}(m_i, b))| \leq \eta n$ with $\eta = \frac{\epsilon_0}{t}$.

E_6 : $|\,\text{Bad}\,| \leq \epsilon_0 n$ and $|\,\text{BadEdge}\,| \leq \epsilon_0 n^2$.

E_7 : For every $b \in B(S_1 \cup T_1 \cup U_1)$ with $b \in M_i$, $|V(\text{Between}(m_i, b)) \cap U| \leq 2\epsilon_0|U|$.

Note that E_4 and E_5 are well defined only when E_2 occurs. The purpose of events $E_1, \ldots, E_6$ is to ensure the following property.

Claim 1. If E_i occurs for every $1 \leq i \leq 7$ then G is $17\epsilon_0$-close to be an interval graph.

To prove the theorem, it suffice to prove Claim 1 and that the event E_i occurs with good enough probability for every $2 \leq i \leq 6$. Before showing Claim 1, we need the following claim, that gives a suitable place to each vertex of V_2.

Claim 2. Assume E_1, E_2, E_4 and E_7 occur, then there is a function $\phi : T_2 \to [p]$ and for each $i \in [p]$, there are two bounds $a_i, b_i \in M_i \cap B(S_1 \cup T_1 \cup U_1)$ such that

- for every $v \in T_2$, $\alpha^V_{\phi(v)}(v) \leq 5\epsilon_0 n$;
- for every $i \in [p]$, $\phi^{-1}(i)$ can be ordered into $x_1^i, \ldots, x^i_{|\phi^{-1}(i)|}$ such that $\delta_i(x^i_j, x^i_{j+1}) \cap U_1 \subseteq \mathcal{N}_2^i$ for every $1 \leq j < |\phi^{-1}(i)|$; and
- every maximal clique of $G[T_2]$ is a subsequence $x_a^i, \ldots, x_b^i$ for some $i \in [p]$.

where $\mathcal{N}_2^i = V(\text{Between}(a_i, b_i))$.

Proof. Let $\mathcal{I} = (I_v)_{v \in S \cup T \cup U}$ be a representation of $G[S \cup T \cup U]$ as $C[S \cup T \cup U]$-interval graph. Such a representation exists because of E_1.

For $i \in [p]$, let a_i be the bound of $M_i \cap B(S_1 \cup T_1 \cup U_1)$ with minimal value $a_i(\mathcal{I})$. In other words, a_i is the leftmost bound of M_i in the representation $\mathcal{I}$. Similarly, let b_i be the the bound of $M_i \cap B(S_1 \cup T_1 \cup U_1)$ with maximal value $b_i(\mathcal{I})$. We set $\mathcal{N}_2^i := V(\text{Between}(a_i, b_i))$. Note that a_i and b_i indeed exists because $m_i \in B(S_1) \cap M_i$. Moreover, $a_i(\mathcal{I}) \leq m_i(\mathcal{I}) \leq b_i(\mathcal{I})$. Since $m_{i-1} \prec a_i$ and $b_i \prec m_{i+1}$, we have

$$m_{i-1}(\mathcal{I}) < a_i(\mathcal{I}) \leq m_i(\mathcal{I}) \leq b_i(\mathcal{I}) < m_{i+1}(\mathcal{I}).$$

In particular, $a_i(\mathcal{I}) \leq m_i(\mathcal{I}) < a_{i+1}(\mathcal{I})$. Given $v \in T_2$, we define $\phi(v) = i$ if $v(\mathcal{I}) \in [a_i(\mathcal{I}), a_{i+1}(\mathcal{I})[$.

Let $v \in T_2$ and take $i \neq \phi(v)$ and $x \in M_i \cap B(U_1)$. We show that if $i < \phi(v)$ (resp. $\phi(v) < i$), then $x(\mathcal{I}) \leq v(\mathcal{I})$ (resp. $v(\mathcal{I}) \leq x(\mathcal{I})$) unless $x \in \text{Between}(a_i, b_i)$. To see this, we consider two cases.

- If $\phi(v) - i$ is even then x and m_{i-1} (resp. m_{i+1}) are of different types. In this case, $x \prec m_{i-1}$ (resp. $m_{i+1} \prec x$) and further $x(\mathcal{I}) < m_{i-1}(\mathcal{I}) < a_i(\mathcal{I}) \leq v(\mathcal{I})$ (resp. $v(\mathcal{I}) \leq a_{i+1}(\mathcal{I}) \leq m_{i+1}(\mathcal{I}) < x(\mathcal{I})$).
- If $\phi(v) - i$ is odd, then x and m_i are of different types. Consequently, $x \prec m_i$ (resp. $m_i \prec x$), and further $x(\mathcal{I}) < m_i(\mathcal{I}) \leq b_i(\mathcal{I})$ (resp. $a_i(\mathcal{I}) \leq m_i(\mathcal{I}) < x(\mathcal{I})$). Assuming for the sake of contradiction that $v(\mathcal{I}) < x(\mathcal{I})$ (resp. $x(\mathcal{I}) < v(\mathcal{I})$), we deduce that $a_i(\mathcal{I}) < x(\mathcal{I})$ (resp. $x(\mathcal{I}) < b_i(\mathcal{I})$, so $x(\mathcal{I}) \in [a_i(\mathcal{I}), b_i(\mathcal{I})]$. In this case, $a_i \prec x \prec b_i$, so $x \in \text{Between}(a_i, b_i) \subseteq B(\mathcal{N}_2^i)$.

Let us bound the size of $\mathcal{N}_2^i$ for $i \in [\ell]$. It holds that

$$\text{Between}(a_i, b_i) \subseteq \text{Between}(a_i, m_i) \cup \text{Between}(b_i, m_i).$$

Indeed, if for instance, $a_i \prec c \prec b_i$ for some $c \in B(V_1)$ then either $c \prec m_i$, so $c \in \text{Between}(a_i, m_i)$, or $m_i \prec c$ so $c \in \text{Between}(b_i, m_i)$. The case $b_i \prec c \prec a_i$ is symmetrical. It follows that

$$|\mathcal{N}_2^i| = |\,\text{Between}(a_i, b_i)| \leq |\,\text{Between}(a_i, m_i)| + |\,\text{Between}(b_i, m_i)| \leq 2\epsilon_0 n.$$

We now deduce the first item of the claim. Let $v \in T_2$ and set $i = \phi(v)$. Take $w \in U_1$ that contributes to $\alpha_i^U(v)$, i.e. such that $w \notin V(M_i) \cup \mathcal{N}_1$ and $w \in N(i) \triangle N(v)$. Let j_1 and j_2 be the integers of $[k] \setminus \{i\}$ such that $\ell_w \in M_{j_1}$ and $r_w \in M_{j_2}$. Assume moreover that $w \notin \mathcal{N}_2^i$, it then follows from the remark above that $\ell_w(\mathcal{I}) < v(\mathcal{I})$ if and only if $j_1 < i$ and, similarly, $r_w(\mathcal{I}) < v(\mathcal{I})$ if and only if $j_2 < i$. As a consequence, if $w \in N(i)$, then $j_1 < i < j_2$ and further $\ell_w(\mathcal{I}) < v(\mathcal{I}) < r_w(\mathcal{I})$, so $v \in N(x)$; and if $w \notin N(i)$, then $j_1, j_2 < i$ or $i < j_1, j_2$ so $\ell_w(\mathcal{I}), r_w(\mathcal{I}) < v(\mathcal{I})$ or $v(\mathcal{I}) < \ell_w(\mathcal{I}), r_w(\mathcal{I})$ and further $w \notin N(x)$. It follows from E_7 that

$$\alpha_i^U(v) \leq |\mathcal{N}_2^i \cap T| \leq |\,\text{Between}(a_i, m_i) \cap T| + |\,\text{Between}(b_i, m_i) \cap T| \leq 4\epsilon_0 t.$$

It then follows from E_4 that $\alpha_i^V(v) \leq \frac{n}{|U|}\alpha_i^U(v) + n \cdot |\frac{1}{n}\alpha_i^V(v) - \frac{1}{|U|}\alpha_i^U(v)| \leq 5\epsilon_0 n$.

It remains to prove the second and the third item. Let $i \in [\ell]$ and order $\phi^{-1}(i)$ into a sequence $x_1^i, \ldots, x_{|\phi^{-1}(i)|}^i$ such that the real number sequence $(x_j^i(\mathcal{I}))_j$ is increasing if $M_i \subseteq L(V_1)$ and decreasing otherwise, i.e. if $M_i \subseteq R(V_1)$.

Note that the third item of the claim directly follows from the construction. Indeed, a maximal clique C of $G[T_2]$ is exactly a set of the form $C = \{ v \in T_2 \mid v(\mathcal{I}) = r \}$ for some $r \in \mathbf{R}$. As a consequence, all elements of C are mapped to a same index $i \in [\ell]$ by ϕ and by construction they appear consecutively in the sequence $(x_j^i)_{j=1}^{|\phi^{-1}(i)|}$.

It remains to show the third item. Fix $1 \leq j < |\phi^{-1}(i)| - 1$, and take y in $(N(x_j) \setminus N(x_{j+1})) \cap U_1$ that is not in $V(\mathcal{N}_2^i)$.

Let us assume that $M_i \subseteq L(V_1)$, so $x_j(\mathcal{I}) \leq x_{j+1}(\mathcal{I})$. The other case can be proved symmetrically. In this case, we have $x_j(\mathcal{I}) \leq r_y(\mathcal{I}) \leq x_{j+1}(\mathcal{I})$. Indeed, $x_j \in N(y)$ so the interval $[\ell_y(\mathcal{I}), r_y(\mathcal{I})]$ contains $x_j(\mathcal{I})$ and ends before $x_{j+1}(\mathcal{I})$ since $x_{j+1} \notin N(y)$. So by the property stated above, $y \in N_2^i$. $\square$

We deduce the following property.

Claim 3. Assume events E_1, E_2, E_5 and E_6. For every $i \in [p]$, there is $\mathcal{N}_3^i$ such that $|\mathcal{N}_3^i| \leq 2\epsilon_0 n$ and there is a partition $X_0^i, \dots, X^i_{|\phi^{-1}(i)|}$ of $M_i \setminus B(\mathcal{N}_2^i \cup \mathcal{N}_3^i)$ satisfying

$$(N(x_j^i) \cap V(M_i)) \setminus (\mathcal{N}_2^i \cup \mathcal{N}_3^i) = \bigcup_{r=0}^{j-1} V(X_r^i)$$

for every $j \in [p]$. Moreover, $X_j^i = \varnothing$ whenever $x_j^i x_{j+1}^i$ is an edge of G.

Proof. First note that both $\delta_i(x_j, x_{j+1}) \setminus V(\text{Between}(a_i, m_i))$ and $\delta_i(x_j, x_{j+1}) \setminus V(\text{Between}(b_i, m_i))$ does not intersect U_1. If moreover $x_j^i x_{j+1}^i$ is an edge of G, then $v \in \text{Bad}$ whenever $v \in N(x_{j+1}^i) \setminus N(x_j^i)$ because in this case the neighborhood of v is not a connected component of $G[T_2]$. It follows from E_5 that $|\delta_i(x_j, x_{j+1}) \setminus \mathcal{N}_2^i| \leq 2\eta n$ and from E_6 that $|\,\text{Bad}\,| \leq \epsilon_0 n$.

Set $\mathcal{N}_3^i = \text{Bad} \cup \left(\bigcup_{j=1}^{|\phi^{-1}(i)|} \delta_i(x_j, x_{j+1}) \setminus \mathcal{N}_2^i\right)$. Since $|\phi^{-1}(i)| \leq t$, it holds that $|\mathcal{N}_3^i| \leq \epsilon_0 n + 2t\eta n = 2\epsilon_0 n$. Moreover,

$$(N(x_j^i) \cap V(M_i)) \setminus (\mathcal{N}_2^i \cup \mathcal{N}_3^i) \subseteq (N(x_{j+1}^i) \cap V(M_i)) \setminus (\mathcal{N}_2^i \cup \mathcal{N}_3^i)$$

for every index j with $1 \leq j \leq |\phi^{-1}(i)| - 1$, with equality whenever $x_j^i x_{j+1}^i \in E(G)$.

Consequently, it suffices to set $X_0^i = (N(x_j^i) \cap V(M_i)) \setminus (\mathcal{N}_2^i \cup \mathcal{N}_3^i)$, $X_j^i = \delta_i(x_{j+1}, x_j) \setminus (\mathcal{N}_2^i \cup \mathcal{N}_3^i)$ for $j \in [|\phi^{-1}(i)|]$ and $X_{|\phi^{-1}(i)|^i} = V(M_i)) \setminus (\mathcal{N}_2^i \cup \mathcal{N}_3^i \cup \bigcup_{r=0}^{|\phi^{-1}(i)|} X_r^i)$. $\square$

Given a graph H, a vertex $v \in V(H)$, we write $N_H(v)$ the set of neighbors of v in H. The following claim shows that assuming the events $E_1, \dots, E_7$, we can construct a graph close to $G[V_1 \cup T_2]$ whose error is small on *every* vertex of T_2.

Claim 4. Assume E_i for every $i \in \{1, \dots, 7\}$. Then there is an interval representation $\mathcal{I} = (I_v)_{v \in V_1 \cup T_2}$ of a graph H on $V_1 \cup T_2$ such that

- $|E(H[V_1] \triangle G[V_1])| \leq 3\epsilon_0 n^2$;
- for every $t \in T_2$, $|N_{G[V_1]}(t) \triangle N_{H[V_1]}(t)| \leq 10\epsilon_0 n$; and
- there is no error inside T_2, i.e. $G[T_2] = H[T_2]$ as a labeled graph.

Proof. To describe $\mathcal{I}$, it suffices to choose one value for each bounds in $B(V_1)$ and each vertex of T_2. We place these elements with the following rules.

- For each left or right bound $b \in M_i$ its value $b(\mathcal{I})$ is in the interval $]i, i+1[$;
- For every $v \in T_2$, the value $v(\mathcal{I})$ is in $]\phi(v), \phi(v)+1[$;
- For a fixed i, the bounds in $]i, i+1[$ are placed such in the order (resp. reversed order of) $X_0^i, x_1^i, X_1^i, \dots, x^i_{|\phi^{-1}(i)|}, X^i_{|\phi^{-1}(i)|}$ given by Claims 2 and 3 if $M_i \subseteq R(V_1)$ (resp. $M_i \subseteq L(V_1)$). We add the condition that $x_j(\mathcal{I}) = x_{j+1}(\mathcal{I})$ if $x_j x_{j+1} \in E(G)$ and $x_j(\mathcal{I}) \neq x_{j+1}(\mathcal{I})$ otherwise, recall that this is possible since $X_j^i = \varnothing$ whenever $x_j x_{j+1} \in E(G)$; the bounds of $\mathcal{N}_2^i \cup \mathcal{N}_3^i$ are placed anywhere in $]i, i+1[$; and
- the bounds of $\mathcal{N}_1$ are placed anywhere.

Let H be the interval graph on $V_1 \cup T_2$ represented by $\mathcal{I}$. Let us prove that H satisfies the claimed properties.

We first show that $G[T_2]$ and $H[T_2]$ are equal as labeled graphs. This boils down to showing that $u(\mathcal{I}) = v(\mathcal{I})$ if and only if $uv \in E(G)$. As previously stated, it follows from E_1 that $G[T_2]$ is a disjoint union of clique. Let C be connected component (and a clique) of $G[T_2]$. By Claim 2, $C = \{x_a^i, \dots, x_b^i\}$ for some a, b and i. By the construction of $\mathcal{I}$ it holds that $x_a^i(\mathcal{I}) = x_{a+1}^i(\mathcal{I}) = \dots = x_b^i(\mathcal{I})$, so C is a clique of H. In the other direction, if $uv \in E(H[T_2])$ then by the construction there is a path from u to v in G, so u and v are in a same clique of G. This proves that $G[T_2] = H[T_2]$.

Let $x \in T_2$ and let us bound the size of $N_{G[V_1]}(x) \triangle N_{H[V_1]}(x)$ from above. Set $i = \phi(x)$. Observe that $N_{H[V_1]}(x) \setminus (\mathcal{N}_1 \cup V(M_i)) = N(i)$. By the construction and Claim 3, the sets $N_G(t)$ and $N_H(t)$ have same intersection with $V(M_i) \setminus (\mathcal{N}_2^i \cup \mathcal{N}_3^i)$, (these intersections are equal to $\bigcup_{r=1}^{j-1} X_r^i$ if $t = x_j^i$). It follows that

$$N_{H[V_1]}(x) \triangle N_{G[V_1]}(x) \subseteq \mathcal{N}_1 \cup \mathcal{N}_2^i \cup \mathcal{N}_3^i \cup (N_{G[V_1]} \triangle N(i)) \setminus (V(M_i) \cup \mathcal{N}_1).$$

Consequently,

$$\begin{aligned} |N_{H[V_1]}(x) \triangle N_{G[V_1]}(x)| &\leq |\mathcal{N}_1| + |\mathcal{N}_2^i| + |\mathcal{N}_3^i| + \alpha_i^V(x) \\ &\leq \epsilon_0 n + 2\epsilon_0 n + 2\epsilon_0 n + 5\epsilon_0 n \\ &\leq 10\epsilon_0 n. \end{aligned}$$

It remains to estimate the error on V_1. For $u, v \in V_1$, we know that if $b_1(\mathcal{I}) < b_2(\mathcal{I})$ whenever $b_1 \prec b_2$ and $b_1, b_2 \in \{r_u, \ell_u, r_v, \ell_v\}$ then $uv \in E(G) \Leftrightarrow I_u \cap I_v \neq \varnothing$. It follows that the size of $E(G[V_1] \triangle H[V_1])$ is at most the number of pairs $(b_1, b_2) \in B(V_1)^2$ such that $b_1 \prec b_2$ and $b_1(\mathcal{I}) > b_2(\mathcal{I})$. Let b_1 and b_2 be such bounds and assume that b_1 and b_2 are not in $B(\mathcal{N}_1)$. In this case, $b_1 \in M_{i_1}$ and $b_2 \in M_{i_2}$ for some indices $i_1, i_2 \in [p]$. We know that $i_1 \neq i_2$ because $b_1 \prec b_2$ (so b_1 and b_2 are of different types) and it follows from the construction that $i_1 > i_2$ since $b_1(\mathcal{I}) > b_2(\mathcal{I})$. As a consequence, $b_2 \prec m_{i_1}$ so $b_2 \in \text{Between}(m_{i_1}, b_1)$. It follows that

$$\begin{aligned}
|E(G[V_1]\triangle H[V_1])| &\le |\mathcal{N}_1|\cdot n + \sum_{\substack{b\in B(V_1\setminus\mathcal{N}_1)\\ b\in M_i}} \mathrm{Between}(m_i,b)\\
&\le \epsilon_0 n^2 + 2n\cdot\epsilon_0 n\\
&\le 3\epsilon_0 n^2.
\end{aligned}$$

This concludes the proof of Claim 4. □

We can now proceed to the proof of Claim 1.

Proof of Claim 1. Let $(I_v)_{v\in V_1\cup T_2}$ be the interval representation given by Claim 4. We now extend $(I_v)_{v\in V_1\cup T_2}$ to an interval representation $\mathcal{I}$ on V to construct an interval graph close to G. Let $v\in V_2\setminus T_2$.

Let $V_2' = V_2\cap B_d(T_2,\epsilon_0)\setminus \mathrm{Bad}$. Since E_3 and E_6 occur, we have $|V_2\setminus V_2'|\le 2\epsilon_0 n$. Now, for each $u\in V_2'$, we choose a vertex $w_u\in T_2$ such that $d(u,w_u)\le\epsilon_0$. If $uw_u\in E(G)$, set $u(\mathcal{I}) := w_u(\mathcal{I})$. Otherwise we set $u(\mathcal{I}) := w_u(\mathcal{I})+\epsilon_u$ where ϵ_u is real number that is chosen small enough so that non bound of $B(V_1)$ has a value between $w_u(\mathcal{I})$ and $u(\mathcal{I})$ and such that $u(\mathcal{I})\ne v(\mathcal{I})$ for every $v\in V_2$. If $u\in V_2\setminus V_2'$, we give an arbitrary value to $u(\mathcal{I})$.

Let H be the interval graph on V represented by $\mathcal{I}$. It remains to show that G and H are at distance at most $17\epsilon_0$.

Let A_1 be the set of edges uv where $u\in V_2$ and $v\in N_{G[T_1]}(w_u)\triangle N_{H[T_1]}(w_u)$. Let A_2 be the set of edges uv with $u\in V_2$ and $v\in V$ such that $v\in N_G(w_u)\triangle N_G(u)$.

Note that these sets are small. Indeed, $|A_1|\le 10\epsilon_0 n^2$ because the size of $N_{G[T_1]}(w)\triangle N_{H[T_1]}(w)$ is at most $10\epsilon_0$ for every $w\in T_2$ and $|A_2|\le\epsilon_0 n^2$ because $|N_G(w_u)\triangle N_G(u)| = d(u,w_u)\cdot n\le\epsilon_0 n$.

It therefore suffices to show that

$$E(G\triangle H)\subseteq E(H[V_1]\triangle G[V_1])\cup A_1\cup A_2\cup \mathrm{BadEdge}\cup\{\, uv \mid u\notin V_2\setminus V_2', v\in V\,\} \tag{2.6}$$

to deduce the claim. Indeed, we know that $|V_2\setminus V_2'|\le 2\epsilon_0 n$ and $|E(H[V_1]\triangle G[V_1])|\le \epsilon_0 n^2$, so assuming (2.6),

$$|E(G\triangle H)|\le (3\epsilon_0 + 10\epsilon_0 + \epsilon_0 + \epsilon_0 + 2\epsilon_0)n^2 = 17\epsilon_0 n^2.$$

Let us prove (2.6). Let $uv\in E(H\triangle G$ be an edge that is in none of the sets A_1, A_2, $E(H[V_1]\triangle G[V_1])$, BadEdge and $\{\, uv \mid u\in V_2\setminus V_2', v\in V\,\}$. In particular, we may assume that $u\in V_2'$ and $v\in V_1\cup V_2'$. Since $uv\notin A_2$ we know that $uv\in E(G)\Leftrightarrow w_u v\in E(G)$.

If $v\in V_1$, then since $uv\notin A_1$ it holds that $w_u v\in E(G)\Leftrightarrow w_u v\in E(H)$. Recall that we also know from the construction that u ans w_u have same neighbors in V_1 in H (Recall that $u(\mathcal{I}) = w_u(\mathcal{I})$ or $u(\mathcal{I}) = w_u(\mathcal{I})+\epsilon_u$). It follows that $uv\notin E(G\triangle H)$.

Assume now that $v\in V_2'$, then since $uv\notin A_2$, we have $uv\in E(G)\Leftrightarrow uw_v\in E(G)$. If $uv\in E(G)$, then $uw_v\in E(G)$ and $w_u v\in E(G)$. Since $uv\notin$

BadEdge, the vertices u and v have same neighborhood in T_2 in G, so $uw_u \in E(G)$ and $vw_v \in E(G)$. Moreover, $u, v \notin$ Bad, so the neighborhood of u in T_2 is a clique of G. It follows that $w_u w_v \in E(G)$ and further $w_u w_v \in E(H)$ because $G[T_2] = H[T_2]$, which finishes the proof in the case $uv \in E(G)$. Now assume that $uv \notin E(G)$, then $uw_v \notin E(G)$ and $w_u v \notin E(G)$. If $uw_u \notin E(G)$, then it follows from the construction of $u(\mathcal{I}) = w_u(\mathcal{I}) + \epsilon_u$ that u has no neighbor in V_2 in H, so in particular $uv \notin E(G)$. We now assume that $uw_u \in E(G)$. Since $u \notin$ Bad, the neighborhood of u in T_1 in G is a connected component, so $w_u w_v \notin E(G)$. It follows that $w_u(\mathcal{I}) \neq w_v(\mathcal{I})$ and further $u(\mathcal{I}) \neq v(\mathcal{I})$. This proves that $uv \notin E(H)$ and finishes the proof. □

Claim 5. $Pr(E_3|E_1) \geq 1 - \frac{1}{12}$.

Proof. Lemma 2.20 applied to the random set T for $E = V$, $F = V_2$ with $\rho = \epsilon_0$, $m = \frac{3}{\epsilon_0} \geq \left\lceil \frac{2}{\epsilon_0} \right\rceil + 1$ and $k = 8 \ln 1/\epsilon_0$ ensures that provided that $t \geq s_{2.20}(\epsilon_0, k, m)$, with probability at least $1 - me^{-k}$, either E_3 holds or there is a set $A \subseteq T$ of size m with $d(x, y) > \epsilon$ for every $x, y \in S$. In this last case, Proposition 2.21 implies that $G[S]$ is not a $C[S]$-interval graph so E_1 does not hold.

This proves that $\mathbf{P}(E_3 \cup \overline{E_1}) \geq 1 - me^{-k}$ i.e. $\mathbf{P}(\overline{E_3} \cap E_1) \leq me^{-k}$. So $\mathbf{P}(\overline{E_3}|E_1) = \mathbf{P}(\overline{E_3} \cap E_1)/\mathbf{P}(E_1) < 2me^{-k} = \frac{6}{\epsilon_0}e^{-k} = 6\epsilon_0^7 \leq \frac{6}{2^7} \leq \frac{1}{12}$. It can be checked that $s_{2.20}(\epsilon_0, k, m) = \frac{km}{\epsilon_0} = 24\frac{1}{\epsilon_0^2} \ln \frac{1}{\epsilon_0} \leq t$. □

Claim 6. $\mathbf{P}(E_4|E_2) \geq 1 - \frac{1}{12}$.

Proof. We use the following property. Given a subset $A \subseteq V$, we have $|\frac{|A|}{n} - \frac{|A \cap U|}{|U|}| \leq \epsilon_0$ with probability at least $1 - 2e^{-\epsilon_0^2 |U|}$ by Azuma-Hoeffding inequality. For $i \in [\ell]$ and $v \in T_2$, it suffices to take $A = (N(i) \triangle (N(v) \setminus (V(M_i) \cup \mathcal{N}_1)))$ to deduces that $|\frac{1}{|U|}\alpha_i^U(v) - \frac{1}{n}\alpha_i^V(v)| \leq \epsilon_0 n$ with probability at least $1 - 2e^{-\epsilon_0^2 |U|}$.

By the union bound, it holds

$$\mathcal{P}(\overline{E_2}|E_1) \leq |U|\ell \cdot 2e^{-\epsilon_0^2|U|} \leq tse^{-\epsilon_0^2|U|} \leq t^2 e^{-\epsilon_0^2|U|} \leq \frac{1}{t} \leq \frac{1}{12}$$

since $|U| \geq 3\frac{1}{\epsilon_0^2} \ln t$ and $t \geq 12$. □

Claim 7. $\mathbf{P}(E_5|E_2) \geq 1 - \frac{1}{12}$.

Proof. For $x, y \in T_2$ and $b \in M_i$ for some $i \in [\ell]$, assume that the set $\delta_i(x, y) \setminus V(\text{Between}(m_i, b))$ has size at least ηn. In this case, the probability that U does not intersect $\delta_i(x, y) \setminus V(\text{Between}(m_i, b))$ is at most $(1 - \eta)^{|U|} \leq e^{-\eta|U|}$. Hence by the union bound,

$$\mathbf{P}(\overline{E_5}|E_2) \leq \ell t^2 e^{-\eta|U|} \leq t^3 e^{-\frac{\epsilon_0}{t}|U|} \leq \frac{1}{t} \leq \frac{1}{12}$$

since $\ell \leq s = t$ and $|U| = \frac{4}{\epsilon_0} t \ln t$. □

Claim 8. $\mathbf{P}(E_6) \geq 1 - \frac{1}{12}$.

Proof. If T intersects Bad then $G[X]$ is not a $C[X]$-interval graph. It follows that if $|\,\text{Bad}\,| > \epsilon_0 n$ then the probability that $G[X]$ is a $C[X]$-interval graph is at most

$$\mathbf{P}_T(T \cap \text{Bad} = \varnothing) \leq (1 - \epsilon_0)^{|T|} \leq e^{-\epsilon_0 |T|}.$$

Moreover, $G[X]$ is not a $C[X]$-interval graph if there is a pair $uv \in$ BadEdge with $u, v \in T$. So if $|\,\text{BadEdge}\,| \geq \epsilon_0 n^2$ then the probability that $G[X]$ is a $C[X]$-interval graph is at most

$$\left(1 - \frac{|\,\text{BadEdge}\,|}{\binom{n}{2}}\right)^{t/2} \leq (1 - 2\epsilon_0)^{t/2} \leq e^{-\epsilon_0 t}.$$

This shows that $\mathbf{P}(E_1|\overline{E_6}) \leq e^{-\epsilon_0 t}$. Recall that we assumed $\mathbf{P}(E_1) \geq \frac{1}{2}$. It follows that

$$\mathbf{P}(\overline{E_6}|E_1) = \frac{\mathbf{P}(\overline{E_6})}{\mathbf{P}(E_1)} \cdot \mathbf{P}(E_1|\overline{E_6}) \leq 2e^{-\epsilon_0 t} \leq \frac{1}{12}$$

since $t > \frac{\ln 24}{\epsilon_0}$. □

Claim 9. $\mathbf{P}(E_7|E_2) \geq 1 - \frac{1}{12}$.

Proof. Assume E_2 occurs. We use the property described in the proof of Claim 6. Let $b \in S_1 \cup T_1 \cup U_1$ with $b \in M_i$. By the Azuma-Hoeffding inequality, $|\frac{|V(\text{Between}(m_i,b))|}{n} - \frac{|V(\text{Between}(m_i,b)) \cap U|}{|U|}| \leq \epsilon_0$ with probability at least $1 - 2e^{-\epsilon_0^2 |U|}$. Note that in this case $|V(\text{Between}(m_i, b)) \cap U| \leq 2\epsilon_0 |U|$ since $|V(\text{Between}(m_i, b))| = |\,\text{Between}(m_i, b)| \leq \epsilon_0 n$. Recall that $|S_1 \cup T_1 \cup U_1| \leq s + t + |U| \leq 3|U|$. By the union bound,

$$\mathbf{P}(\overline{E_7}|E_2) \leq 3|U| \cdot 2e^{-\epsilon_0^2 |U|} \leq 6\epsilon_0^{-8} < \frac{1}{12}.$$

since $|U| \geq 10\frac{\ln(1/\epsilon_0)}{\epsilon_0^2}$ □

Recall that $\mathbf{P}(E_1 \cup E_2) > \frac{1}{2} - \frac{1}{12}$. It follows from Claims 5 to 9 that

$$\mathbf{P}(\bigcup_{i=1}^{7} E_i) > \frac{1}{2} - \frac{1}{12} - \frac{1}{12} - \frac{1}{12} - \frac{1}{12} - \frac{1}{12} - \frac{1}{12} = 0.$$

By Claim 1, the graph G is therefore $17\epsilon_0$-close to an interval graph.

Now set $\epsilon := 17\epsilon_0$, so that G is ϵ-close to being an interval graph. The size of X is at most

$$3|U| = \frac{12}{\epsilon_0} t \ln t = 9 \cdot \frac{2^{22}}{\epsilon_0^9} (\ln 3 + 20 \ln 2 + 8 \ln \frac{1}{\epsilon_0}) \leq 9 \cdot \frac{2^{22}}{\epsilon_0^9} \cdot 30 \ln \frac{1}{\epsilon_0} \leq \frac{2^{31}}{\epsilon_0^{10}}.$$

With respect to ϵ, it gives

$$|X| \leq \frac{2^{31} \cdot 17^8}{\epsilon^{10}} \leq \frac{2^{64}}{\epsilon^{10}},$$

which concludes the proof of Theorem 2.13. $\square$

2.7 Linear specialization of interval graphs

Recall that if G is an induced subgraph of H, the set $\mathcal{R}_G(H)$ is the set of representations of G that extend to H. We recall Lemma 2.17.

Lemma 2.17. *Let G be a graph on n vertices. Then every sequence $H_1, \ldots, H_\ell$ of induced subgraphs of G such that*

$$\mathcal{R}_G(H_1) \supsetneq \cdots \supsetneq \mathcal{R}_G(H_\ell)$$

has length at most $m_{2.17}(n) = 16n$.

2.8 PQ-trees and proof of Lemma 2.17

Let $G = (V, E)$ be an interval graph and $(I_v)_{v \in V}$ be an interval representation of G, that is a set of intervals such that

$$uv \in E \Leftrightarrow I_u \cap I_v \neq \varnothing$$

for every two distinct vertices $u, v \in V$. To each point x of the real line $\mathbf{R}$, the set Q_x of vertices $v \in V$ such that I_v contains x is a clique because all these intervals intersect in x. Note that $Q_x = Q_y$ if $]x, y[$ contains no bound of an interval in $(I_v)_{v \in V}$. As a consequence, it is possible to list the cliques given by the function $x \mapsto Q_x$ as a finite sequence $Q_{x_1}, \ldots, Q_{x_p}$ such that $(x_i)_{i=1}^p$ is an increasing sequence with $Q_{x_i} \neq Q_{x_{i+1}}$, $Q_{x_1} = \varnothing = Q_{x_p}$ and $Q_y \in \{Q_{x_i}, Q_{x_{i+1}}\}$ whenever $x_i < y < x_{i+1}$. The remark above even shows that in this case $p \leq 2|V| + 1$.

The sequence $(Q_{x_i})_{i=1}^p$ characterizes the representation $(I_v)_{v \in V}$ up to an homeomorphism of $\mathbf{R}$. That is, another collection of intervals $(J_v)_{v \in V}$ representing G does not yield the same sequence $(Q_{x_i})_{i=1}^p$ unless there is a strictly increasing function $f : \mathbf{R} \to \mathbf{R}$ such that $f(J_v) = I_v$ for every $v \in V$. It is convenient to use these sequences of cliques of G instead of the usual interval representation.

In the following, we consider that $\mathcal{R}_G(H)$ is a set of clique sequences instead of a set of families of intervals.

2.9 Clique orderings

A *clique ordering* is a sequence $S = Q_1, \ldots, Q_p$ of sets, that we call *cliques*, such that for every vertex v, the cliques in the sequence that contain v appear consecutively. Note that a clique can be empty. Let $C^{\max}(S)$ be the set of maximal elements of S with respect to set inclusion. The sequence S is a *clique ordering of the graph G* if the elements in S are actual cliques of G and if S contains all the maximal cliques of G. In this case, $C^{\max}(S) = C^{\max}(G)$ is the set of maximal cliques of the graph G.

A *subclique* is a non-maximal clique. The insertion of a subclique Q in S between Q_i and Q_{i+1} is a *subclique insertion* if Q satisfies the property that $Q_i \cap Q_{i+1} \subseteq Q$ and either $Q \subseteq Q_i$ or $Q \subseteq Q_{i+1}$. It is easy to check that the sequence obtained by a subclique insertion is a clique ordering. A clique ordering S' is an *extension* of S if it can be obtained from S by subclique insertions.

A clique ordering $S = Q_1, \ldots, Q_p$ is *closed* if $Q_i \subseteq Q_{i+1}$ or $Q_i \supseteq Q_{i+1}$ for every $i \in \{1, \ldots, p-1\}$. This definition comes from the following observation.

Observation. If G is an interval graph represented by a family $(I_v)_{v \in V}$ of closed intervals, and $(Q_{x_i})_{i=1}^{p}$ is the sequence of cliques described in the previous section, then $(Q_{x_i})_{i=1}^{p}$ is a closed clique ordering.

If Q_m is a maximal clique belonging to S, the *right tail* of Q_m in S is the longest decreasing sequence of the form $Q_m, \ldots, Q_{m+r}$ (with respect to set inclusion). Similarly, the *left tail* of Q_m is the longest increasing sequence of the form $Q_{m-r}, Q_{m-r+1}, \ldots, Q_m$.

If $Q_i = Q_{i+1}$ for some $i \in [p-1]$, the cliques ordering $S' = Q_1 \ldots Q_i Q_{i+2} \ldots Q_p$ (where Q_{i+1} has been removed) and S correspond to the same interval representations. We identify S and S', that is we consider them as equal. Since Q_i and Q_{i+1} belong to the same (left or right) tails and the deletion of Q_{i+1} does not change the possibilities of clique insertion, this identification is compatible with the notions defined above. In particular, we may assume that for every clique ordering $S = Q_1 \ldots Q_p$ we consider, $Q_i \neq Q_{i+1}$whenever $i \in [p-1]$.

Let us first prove some basic properties regarding the structure of a clique ordering.

Property 2.22. *The following properties are satisfied by every clique ordering $S = Q_1, \ldots, Q_p$ with maximal cliques $Q_{m_1}, \ldots, Q_{m_k}$.*

1. *If $1 \leq i \leq \ell \leq j \leq p$ then $Q_i \cap Q_j \subseteq Q_\ell$.*
2. *If $m_i \leq j \leq m_{i+1}$, then Q_j is a subset of Q_{m_i} or $Q_{m_{i+1}}$.*
3. *If $m_i < j < m_{i+1}$, then Q_j is in the right (resp. left) tail of Q_{m_i} (resp. $Q_{m_{i+1}}$) if and only if Q_j is a subset of Q_{m_i} (resp. $Q_{m_{i+1}}$).*
4. *For $i \in [k-1]$, the sequence $Q_{m_i}, Q_{m_i+1}, \ldots, Q_{m_{i+1}}$ is covered by the union of the right tail of Q_{m_i} and the left tail of $Q_{m_{i+1}}$.*
5. *For $i \in [k-1]$, the right tail of Q_{m_i} and the left tail of $Q_{m_{i+1}}$ intersects only in the clique $Q_{m_i} \cap Q_{m_{i+1}}$, if it is present in S.*
6. *The clique ordering S is closed if and only if for every $i \in [k-1]$ there an index r with $m_i < r < m_{i+1}$ such that $Q_r = Q_{m_i} \cap Q_{m_{i+1}}$.*

Proof. 1. Let x be an element of $Q_i \cap Q_j$. By the definition of a clique ordering, the cliques of S that contain x are consecutive. Since Q_ℓ is between Q_i and Q_j in S, it follows that $x \in Q_\ell$.

2. This is clear if $j = m_i$ or $j = m_{i+1}$. Assume that $m_i < j < m_{i+1}$. In this case Q_j is a subclique, so Q_j is the subset of a maximal clique Q_{m_q} for some

$q \in [k]$. If $q \leq i$ then $m_q \leq m_i \leq j$ so by Item 2.22.1, $Q_{m_i} \supseteq Q_{m_q} \cap Q_j = Q_j$. If $i+1 \leq q$, it follows similarly from Item 2.22.1 that $Q_{m_{i+1}} \supseteq Q_j \cap Q_{m_q} = Q_j$.

3. By symmetry of the notions of left and right tails, it is enough to prove that Q_j is in the right tail of Q_{m_i} if and only if $Q_j \subseteq Q_{m_i}$. If Q_j is in the right tail of Q_{m_j}, it is straightforward from the definitions that $Q_j \subseteq Q_{m_i}$. Let us prove the other direction. Assume that $Q_j \subseteq Q_{m_i}$, it suffices to prove that $Q_{q-1} \supseteq Q_q$ for every $q \in \{m_i+1, \ldots, j\}$ to deduce that Q_j is in the right tail of Q_{m_i}. We first claim that $Q_{m_i} \supseteq Q_q$. To see this, assume otherwise $Q_{m_i} \not\supseteq Q_q$. In particular $q < j$, so by Item 2.22.2 we have $Q_q \subseteq Q_{m_{i+1}}$. By Item 2.22.1, it follows that $Q_j \supseteq Q_q \cap Q_{m_{i+1}} = Q_q$ and further $Q_j \not\subseteq Q_{m_i}$, which yields a contradiction. This proves that $Q_{m_i} \supseteq Q_q$. We conclude using Item 2.22.1 that $Q_{q-1} \supseteq Q_q \cap Q_{m_i} = Q_q$.

4. This is a straightforward consequence of items 2.22.2 and 2.22.3.

5. Let A be a clique that is both in the right tail of Q_{m_i} and the left tail of $Q_{m_{i+1}}$. It follows from the definition that A is a subset of Q_{m_i} and $Q_{m_{i+1}}$. By Item 2.22.1, we also have $Q_{m_i} \cap Q_{m_{i+1}} \subseteq A$. Consequently, $A = Q_{m_i} \cap Q_{m_{i+1}}$.

6. Assume first that S is closed and take $i \in [k-1]$. Let Q_r be the last clique of the right tail of Q_{m_i}, so $Q_r \not\supseteq Q_{r+1}$. Since S is closed, $Q_r \subseteq Q_{r+1}$. Note that $r < m_{i+1}$ as otherwise $Q_{m_{i+1}} \subseteq Q_{m_i}$, so by Item 2.22.4 $Q_{m_i+1} \supseteq Q_{m_i} \cap Q_{m_{i+1}} = Q_{m_{i+1}}$, and further $Q_{m_i} = Q_{m_{i+1}} = Q_{m_i+1}$ because $Q_{m_{i+1}}$ and Q_{m_i} are maximal cliques. By Item 2.22.4 we moreover know that Q_{r+1} is in the left tail of $Q_{m_{i+1}}$. Consequently, Q_r is also in the left tail of $Q_{m_{i+1}}$ and further by Item 2.22.5 it holds that $Q_r = Q_{m_i} \cap Q_{m_{i+1}}$. For the other direction, assume that for $i \in [k-1]$, the clique $Q_{m_i} \cap Q_{m_{i+1}}$ is present in S between Q_{m_i} and $Q_{m_{i+1}}$. First note that if Q_i and Q_{i+1} are contained in the same left or right tail of a maximal clique, then it is true that $Q_i \subseteq Q_{i+1}$ or $Q_i \supseteq Q_{i+1}$. It follows from Item 2.22.2 and Item 2.22.3 that S is covered by the tails of the maximal cliques. Further, the hypothesis shows that each tail intersects the following tail in at least one clique. More precisely, the left and right tails of C_{m_i} intersects in C_{m_i} for $i \in [k]$ and the right tail of C_{m_i} intersects the left tail of $C_{m_{i+1}}$ in $C_{m_i} \cap C_{m_{i+1}}$ for $i \in [k-1]$. Consequently, every pair (Q_i, Q_{i+1}) is contained in a common (left or right) tail of a maximal clique. This suffices to show that S is closed. □

As a consequence of Property 2.22.6, every clique ordering has a closed extension, that can be constructed by insertions of the intersections of the pairs of consecutive maximal cliques as described above.

2.10 PQ-trees

The main tool of the proof of Lemma 2.17 is the notion of PQ-tree. PQ-trees have been introduced by Booth and Lueker [9] and used by these authors to design a linear-time recognition algorithm for the class of interval graphs and planar graphs.

A *PQ-tree* is a tree whose leaves are labeled with elements of an alphabet Σ and with two types of inner nodes, *Q-nodes* and *P-nodes*. Each inner node is endowed with an ordering of its children. This order can be changed under some rules that depend on the type of the node. The order $c_1 \dots c_k$ of a Q-node can be reversed into $c_k \dots c_1$. The order $c_1 \dots c_k$ of a P-node can be freely reordered by any permutation σ into $c_{\sigma(1)} \dots c_{\sigma(k)}$. A *leaf order* of a PQ-tree $\mathcal{T}$ is an ordering of the leaves of $\mathcal{T}$ corresponding to a depth-first exploration of $\mathcal{T}$ following the order given by the inner nodes, after a possible reordering of $\mathcal{T}$ according to the rules stated above.

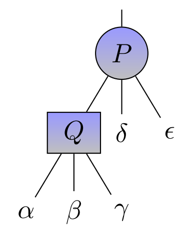


Figure 2.3 – A PQ-tree.

As an example, the leaf orders of the PQ-tree in Figure 2.3 are $\alpha\beta\gamma\delta\epsilon$, $\alpha\beta\gamma\epsilon\delta$, $\delta\alpha\beta\gamma\epsilon$, $\epsilon\alpha\beta\gamma\delta$, $\delta\epsilon\alpha\beta\gamma$, $\epsilon\delta\alpha\beta\gamma$, $\gamma\beta\alpha\delta\epsilon$, $\gamma\beta\alpha\epsilon\delta$, $\delta\gamma\beta\alpha\epsilon$, $\epsilon\gamma\beta\alpha\delta$, $\delta\epsilon\gamma\beta\alpha$ and $\epsilon\delta\gamma\beta\alpha$. A P-node is usually represented by a circle and a Q-node by a rectangle. We follow this graphic rule throughout this chapter.

Note that if n is a node of a PQ-tree $\mathcal{T}$ with two children, the same permutations of the children of n (swapping them or not) are allowed regardless of whether n is a P-node or a Q-node, so we choose to consider that a node with exactly two children is always a Q-node. In particular, we assume that every P-node has at least three children.

The main property of PQ-trees is that they can represent *consecutive ones properties*, defined as follows. Let $S_1, \dots, S_k$ be subsets of an alphabet Σ. The associated *consecutive ones property* is the set of orderings of Σ in which for every $i \in [k]$ the elements of S_i appear in a consecutive order. More precisely, with the notation above, there is a PQ-tree $\mathcal{T} = \mathcal{T}(\Sigma, S_1, \dots, S_k)$ such that a permutation σ of Σ satisfies the consecutive ones property if and only if σ is a leaf order of $\mathcal{T}$. This PQ-tree is constructed by an algorithm that starts from the PQ-tree formed by a unique P-node whose children are $|\Sigma|$ leaves labeled by the elements of Σ (this PQ-tree generates every permutation of Σ) and iteratively modifies this PQ-tree to integrate the constraints one by one. A consecutive ones problem can therefore be solved by trying to construct the corresponding PQ-tree.

In their paper [9], Booth and Lueker use the fact that a graph G is an interval graph if and only if there is an ordering of the maximal cliques of G such that for each vertex v, the (maximal) cliques that contain v appear consecutively in this ordering. In other words, deciding whether G is an interval graph boils down to solve the consecutive one problem where Σ is the set of maximal cliques of G and with a constraint $S_v = \{ C \text{ maximal clique} \mid v \in C \}$ for each $v \in V(G)$. This problem defines a PQ-tree on the maximal cliques of G that represents the set of interval representations of G. This tree is *the* PQ-tree of G.

Example. Figure 2.4 illustrates a graph G and the PQ-tree of G. The maximal cliques of G are the sets $\{1,2,4\}$, $\{1,3,4\}$, $\{5\}$, $\{6,7\}$, $\{6,8\}$ and $\{6,9\}$. See [9] for more details on the construction of such a PQ-tree.

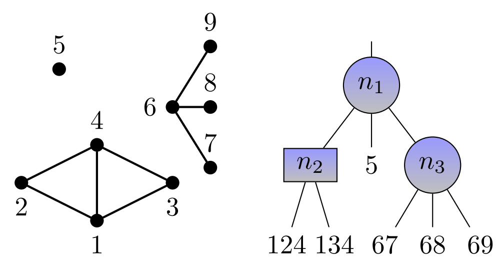


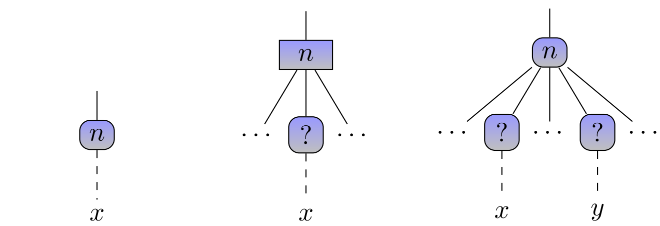
Figure 2.4 – Left: a graph G. Right: the PQ-tree of G, that encodes the interval representations of G.

Two graphs H_1 and H_2 sharing some vertices $U = V(H_1) \cap V(H_2)$ are *simultaneous interval graphs* if there exist interval representations $(I_v)_{v \in V(H_1)}$ and $(I'_v)_{v \in V(H_2)}$ of H_1 and H_2 respectively such that $I_v = I'_v$ whenever $v \in U$. Equivalently, H_1 and H_2 are simultaneous interval graph if there is an interval representation of the graph $G := H_1[U] = H_2[U]$ that extends to an interval representation of H_1 and an interval representation of H_2. Jampani and Lubiw [28] gave a $O(n^2 \log n)$ algorithm that decides whether two given graphs are simultaneous interval graphs. To that purpose, they introduced a *U-reduced PQ-tree* of a graph H, which is a PQ-tree whose leaves are cliques of $H[U]$ and that generates the set of clique ordering corresponding to the set of interval representations of $G = H[U]$ that extend to H. This PQ-tree is labeled by (possibly non-maximal) cliques of G.

If $\mathcal{T}$ is a PQ-tree whose leaves are labeled with cliques, a clique ordering is *generated* by $\mathcal{T}$ if it is a closed extension of a leaf order of $\mathcal{T}$. The set of clique orderings generated by $\mathcal{T}$ is denoted by $\pi(\mathcal{T})$. Note that in particular, every element of $\pi(\mathcal{T})$ is closed. Recall that by Property 2.22.6, a leaf order $S = Q_1, \dots, Q_p$ with maximal cliques $Q_{m_1} \dots Q_{m_k}$ is closed if and only it contains the clique $Q_{m_i} \cap Q_{m_{i+1}}$ between Q_{m_i} and $Q_{m_{i+1}}$ for every $i \in [p-1]$, so the clique ordering S' obtained from S by inserting these cliques (the clique $Q_{m_i} \cap Q_{m_{i+1}}$ between Q_{m_i} and $Q_{m_{i+1}}$) in S is closed. The closed extensions of S are exactly the extensions of S'. Recall that $\mathcal{R}_G(H)$ was defined as the set of interval representations of G that extend to H. The result of Jampani and Lubiw can then be formulated as follows.

Lemma 2.23 ([28])**.** *Let G be an induced subgraph of H. There is a PQ-tree $\mathcal{T}_G(H)$ such that $\mathcal{R}_G(H)$ is the set of interval representations $(I_v)_{v \in V(G)}$ (of G) whose clique ordering belongs to $\pi(\mathcal{T}_G(H))$.*

This tree $\mathcal{T}_G(H)$ is the PQ-tree whose leaves are cliques of G obtained from the PQ-tree of H by relabeling every leaf labeled by a maximal clique A into $A \cap V(G)$. The leaves of this PQ-tree are cliques of $G = H[V(G)]$ that may not be maximal. The strategy implemented in the algorithm of [28] for solving the simultaneous interval problem on H_1 and H_2 is to try to unify step-by-step the PQ-trees $\mathcal{T}_G(H_1)$ and $\mathcal{T}_G(H_2)$ until they are equal or a contradiction appears.



Essential node. Simple Q-node. Skeleton node.

In the following, every PQ-tree will have leaves labeled by (possibly non maximal) cliques of a graph.

2.10.1 Notations and properties

In the rest of this section we fix an interval graph G. Every PQ-tree considered will be a PQ-tree whose leaves are labeled with (possibly non-maximal) cliques of G.

Let $\mathcal{T}$ be a PQ-tree representing a subset of interval representations of G. A node n is *essential* if n is a maximal clique leaf or has a maximal clique descendant (later, Lemma 2.24 shows that we may assume that every inner node is essential). A Q-node n is a *simple Q-node* if n has exactly one essential child. A node n is a *skeleton node* if n is a maximal clique leaf or if n has at least two essential children. The *skeleton parent* of a node n is the nearest ancestor of n that is a skeleton node. Similarly, a *skeleton child* of n is a nearest descendant of n that is a skeleton node. If n is a non-skeleton essential inner node, then n has a unique skeleton child.

Example. Consider the graphs G and H in Figure 2.5. The graph G is an induced subgraph of H. The PQ-tree $\mathcal{T}$ on the same figure represents the set $\mathcal{R}_G(H)$ of interval representations of G that extend to a representation of H. The maximal cliques of G are $\{1,2,4\}$, $\{1,3,4\}$, $\{5\}$, $\{6,7\}$, $\{6,8\}$ and $\{6,9\}$. On the tree $\mathcal{T}$ of Figure 2.5, n_3 and n_5 are simple Q-nodes, n_1, n_2, n_4 and n_6 and the six maximal clique leaves are skeleton nodes. The node n_4 is the skeleton child of n_2 and n_3; and n_1 is the skeleton parent of n_2, n_5 and n_6.

2.10.2 First simplification

A same set of clique orderings may be generated by different PQ-trees. Two PQ-trees $\mathcal{T}_1$ and $\mathcal{T}_2$ are *equivalent* if the generate the same clique ordering, that is $\pi(\mathcal{T}_1) = \pi(\mathcal{T}_2)$.

Lemma 2.24 ([28], Corollary 1)**.** *Let $\mathcal{T}$ be a PQ-tree, there is an equivalent PQ-tree where every inner node is an essential node and every P-node has at least 3 children, and all these children are essential nodes.*

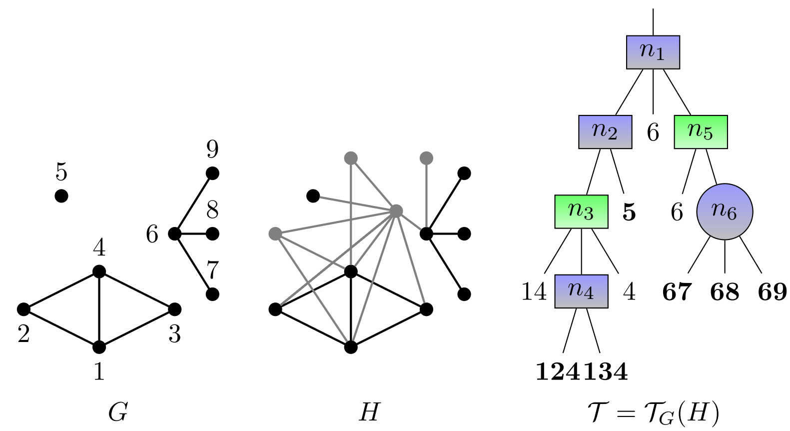


Figure 2.5 – A PQ-tree representing $\mathcal{R}_G(H)$.

To obtain the reduced PQ-tree of Lemma 2.24, proceed in two steps. First, replace every node n whose descendant leaves all represent the same clique A by a leaf A. After this operation, every inner node of the tree is an essential node. Then, delete the non-essential children of each P-node with at least two essential children, and remove every non-essential children except one (it can be shown that they all represent the same clique) of P-nodes with exactly one essential node. Since this last node has only two children left, it is then considered as a Q-node (recall that nodes with exactly two children are always considered as Q-nodes). The correctness of these operations are proved in Lemma 6 and Lemma 7 in [28].

2.10.3 Universal sets

If n is a node, let $C(n)$ be the set of (possibly non-maximal) cliques that are labels of the leaf descendants of n. If n is an essential leaf labeled with a clique A, we adopt the convention that $C(n) = \{A\}$. Similarly, let $C^{\max}(n)$ be the set of *maximal* cliques that are labels of the leaf descendants of n. Set $C^{\max}(n) = \{x\}$ if n is an essential leaf labeled by a (necessarily maximal) clique x. Define

$$\Omega(n) = \bigcap_{A \in C(n)} A$$

and similarly, if n is an essential node,

$$\Omega^{\max}(n) = \bigcap_{A \in C^{\max}(n)} A.$$

In other words, $\Omega(n)$ is the set of the elements that appear in every leaf of the subtree of $\mathcal{T}$ rooted at n; it is called the *universal set* of n. Note that $C^{\max}(n) \subseteq C(n)$, and further $\Omega(n) \subseteq \Omega^{\max}(n)$.

If n is an essential node, with an essential child c, set

$$\mho_c(n) = \bigcup_{\substack{A \in C(n) \\ A \notin C(c)}} A \cap \Omega(c).$$

and, similarly, if n is a skeleton node

$$\mho_c^{\max}(n) = \bigcup_{\substack{A \in C^{\max}(n) \\ A \notin C^{\max}(c)}} A \cap \Omega(c).$$

The set $\mho_c(n)$ contains the elements that are forced to be in every clique descendant of c by the structure of the rest of the tree. This property is proved by Proposition 2.28.2. Lemma 2.26 shows simplified expressions for $\mho_c(n)$.

Lemma 2.25 (Lemma 6 in [28])**.** *In a PQ-tree $\mathcal{T}$,*

- *If n is a P-node, then for any two children c_1 and c_2 of n (which are essential by Lemma 2.24), we have $\Omega(n) = \Omega(c_1) \cap \Omega(c_2)$.*
- *If n is a Q-node and c_1 and c_2 are the first and last children of n then $\Omega(n) = \Omega(c_1) \cap \Omega(c_2)$.*

We prove similar characterizations for the set $\mho_c(n)$.

Lemma 2.26. *In a PQ-tree $\mathcal{T}$, let n be an essential node with an essential child c.*

- *For every $x \in C(c)$,*

$$\mho_c(n) = \bigcup_{\substack{A \in C(n) \\ A \notin C(c)}} A \cap x \quad \text{and} \quad \mho_c^{\max}(n) = \bigcup_{\substack{A \in C^{\max}(n) \\ A \notin C^{\max}(c)}} A \cap x.$$

- *If n is a P-node, $\mho_c(n) = \mho_c(n)^{\max} = \Omega^{\max}(n) = \Omega(n)$.*
- *If n is a simple Q-node with children $A_1 \dots A_k c A_{k+1} \dots A_p$, then $\mho_c(n) = A_k \cup A_{k+1} = \bigcup_{i=1}^{p} A_i$.*

Proof. We first prove that for $A \in C(n) \setminus C(c)$ and $x, y \in C(c)$, we have $x \cap A = y \cap A$. We claim that there is a leaf order $\mathcal{O}$ of $\mathcal{T}$ such that x appears between A and y in $\mathcal{O}$. To see this, reorder the node n so that the child of n that is an ancestor of the leaf A (or that is A itself) appears on the left of c. This is possible up to reversing the order of n, which is allowed regardless of the type of n. Then, reorder similarly the nearest common ancestor of x and y (which is c or a descendant of c) such that x appears on the left of y in the corresponding leaf order $\mathcal{O}$. Note that the clique x is between A and y in $\mathcal{O}$, which proves the claimed property. By Property 2.22.1, it follows that $y \cap A \subseteq x$. Further, $y \cap A \subseteq x \cap A$. By symmetry of the roles played by x and y, it also holds that $x \cap A \subseteq y \cap A$. Consequently, $x \cap A = y \cap A$.

Let us now prove the first item of the lemma. Take $x \in C(c)$. For every $A \in \mathcal{C}(n) \setminus C(c)$, we have $A \cap \Omega(c) = \bigcap_{y \in C(c)} A \cap y = A \cap x$ by the property above. The stated expressions of $\mho_c(n)$ and $\mho_c^{\max}(n)$ follow directly.

Let us prove the second item. Let n be a P-node. It follows from the property above and Lemma 2.25 that for any two different children c_1 and c_2 of n, and every $x \in C(c_1)$ and $y \in C(c_2)$, it holds that $\Omega(n) = x \cap y$. It first follows that $\Omega^{\max}(n) = \left(\bigcap_{x \in C^{\max}(c)} x\right) \cap \left(\bigcap_{y \in C^{\max}(n) \setminus C^{\max}(c)} y\right) = \Omega(n)$. Take $x \in C(c)$, we know from the first item and Lemma 2.25 that $\mho_c(n) = \bigcup_{A \in C(n) \setminus C(c)} A \cap x = \bigcup_{A \in C(n) \setminus C(c)} \Omega(n) = \Omega(n)$, and similarly $\mho_c^{\max}(n) = \bigcup_{A \in C^{\max}(n) \setminus C^{\max}(c)} A \cap x = \Omega(n)$

It remains to prove the third item. For $i \in [p]$, we claim that A_i is a subset of a maximal clique $x_i \in C^{\max}(c)$.

It would directly follows for $x \in \mathcal{C}(c)$ that that $\mho_c(n) = \bigcup_{i=1}^{p} A_i \cap x = \bigcup_{i=1}^{p} A_i$ since then $A_i \cap x = A_i \cap x_i = A_i$ by the property above. Indeed, the subclique A_i is a subset of a maximal clique y. The claimed property is satisfied by $x_i = y$ if $y \in C^{\max}(c)$, so assume $y \notin C^{\max}(c)$ and take $x \in C^{\max}(c)$. Up to reversing n, there is a leaf order of $\mathcal{T}$ where x is between A_i and y, so by Property 2.22.1 applied to this leaf order, it holds that $x \supseteq A_i \cap y = A_i$.

It remains to show that $\bigcup_{i=1}^{p} A_i = A_k \cup A_{k+1}$. For $i \in \{1, \ldots, k-1\}$, we know there exists $x \in C(c)$ with $A_i \subseteq x$. Note that in every leaf order of $\mathcal{T}$, the subclique A_k appears between A_i and x, so by Property 2.22.1, $A_k \supseteq A_i \cap x = A_i$. One can prove similarly that $A_{k+1} \supseteq A_j$ whenever $j \in \{k+2, \ldots, p\}$. It follows that $\bigcup_{i=1}^{p} A_i = A_k \cup A_{k+1}$, which concludes the proof of the third item.

This finishes the proof of the lemma.

$\square$

If n is a simple Q-node, n has exactly one essential child c so we can write $\mho(n)$ and $\mho^{\max}(n)$ instead of respectively $\mho_c(n_1)$ and $\mho_c^{\max}(n)$ without ambiguity.

Observation 2.27. *If n is a Q-node of $\mathcal{T}$ with children $c_1, \ldots, c_p$, then the sequence*

$$\mathcal{O} = \Omega(c_1), \ldots, \Omega(c_p)$$

is a clique ordering.

Indeed, if there is $x \in \Omega(c_i) \cap \Omega(c_j)$ for some integers $i, j \in [p]$ with $i < j$, then considering a leaf order of $\mathcal{T}$ proves that x belongs to every leaf descendant of c_k whenever $i < k < j$, and thus $x \in \Omega(c_k)$.

As a consequence of this observation, $\mathcal{O}$ satisfies Property 2.22. Recall that $\Omega(c) = c$ if c is a leaf. This allows us to define the right and left tails of an essential child of a node as follows (recall that so far, left and right tails have only been defined for maximal cliques in a given clique ordering).

If n is a Q-node with children $c_1, \ldots, c_p$, and $c = c_i$ is an essential child of n for some $i \in [p]$, the *left tail* (resp. *right tail*) of c in n is defined to be the left (resp. right) tail of $\Omega(c_i)$ in the clique ordering $\mathcal{O} = \Omega(c_1), \ldots, \Omega(c_p)$. Equivalently, it is the sequence of subclique children of n that appear to the immediate left (resp. right) of c, that are subsets of $\Omega(c)$. The left (resp. right) tail of a node is an increasing (resp. decreasing) sequence for the inclusion. Note that the notion of left and right is relative to the order chosen for the Q-node n.

If n is an essential node and $\mathcal{O}$ a clique ordering of $\pi(\mathcal{T})$, no maximal clique x that is not in $C^{\max}(n)$ appears in $\mathcal{O}$ between two maximal cliques in $C^{\max}(n)$. The *left tail* (resp. *right tail*) of $C^{\max}(n)$ in $\mathcal{O}$ is the left tail (resp. right tail) of the leftmost (resp. rightmost) clique of $C^{\max}(n)$ in $\mathcal{O}$. Note that this notion depends on a particular clique ordering, on the contrary to the previous one that depends only on the order of a Q-node.

If n is a skeleton Q-node, a subclique child A of n is on the *boundary* of n if $A \subseteq \Omega^{\max}(n)$.

A subclique leaf A that is a child of a node n_1 is *dominated* by the skeleton node n_2 if the leaf A participates in the left or right tail of $C^{\max}(n_2)$ in every leaf order of $\mathcal{T}$. This happens if $n_1 = n_2$ and A is on the boundary of n_1, if n_1 is a simple Q-node and n_2 is the skeleton child of n_1, or if n_1 is skeleton Q-node and A in on a tail of the child of n_1 that is an ancestor (or equal to) n_2.

For a skeleton node n, a clique A is *forced* for n if A is on the left tail and in the right tail of $C^{\max}(n)$ in any clique ordering of $\pi(\mathcal{T})$.

The *descending path* to a node n of $\mathcal{T}$ is the sequence of nodes $n_0, \ldots, n_k$ where n_0 is the root of $\mathcal{T}$, $n_k = n$ and n_{i+1} is a child of n_i for $i \in \{0, \ldots, k-1\}$. A subclique leaf ℓ is *dominated* by this descending path if ℓ is a child of n_i in the left tail or right tail of n_{i+1} for some $i \in \{0, \ldots, k-1\}$. See Figure 2.6 for an example.

We first show some general properties for the sets defines above.

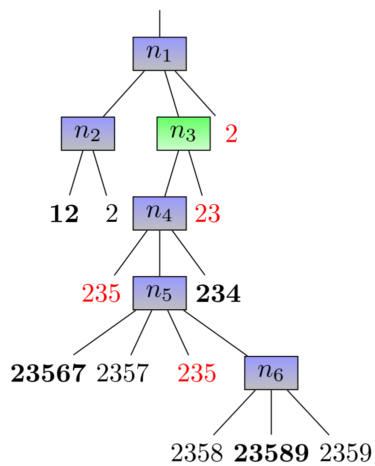


Figure 2.6 – A PQ-tree with maximal cliques $\{1, 2\}$, $\{2, 3, 5, 6, 7\}$, $\{2, 3, 5, 8, 9\}$ and $\{2, 3, 4\}$. In red: subcliques dominated by the descending path to n_6. The node n_4 dominates 23 and the child 2 of n_1. The child 235 of n_6 is dominated both by 23567 and by 23589.

Proposition 2.28. *Let n be a Q-node of a PQ-tree $\mathcal{T}$ and c be an essential child of n. The following properties hold.*

1. *If A is a subclique child of n in the left or right tail of c, then $\Omega(n) \subseteq A \subseteq \mho_c(n)$.*

2. *$\Omega(n) \subseteq \mho_c(n) \subseteq \Omega(c)$.*

3. *If n is a skeleton node, then $\Omega(n) \subseteq \Omega^{\max}(n) \subseteq \mho_c^{\max}(n) \subseteq \mho_c(n)$.*

4. *If n is a non-root skeleton node with skeleton parent p and A is forced for n then $\mho_n^{\max}(p) \subseteq A \subseteq \Omega^{\max}(n)$.*

Proof. 1. Since $A \in C(n)$, it directly follows from the definition that $\Omega(n) \subseteq A$. Since A is in a tail of $C(c)$, it moreover holds that $A \subseteq \Omega(c)$ and $A \in C(n) \setminus C(c)$, so $A = A \cap \Omega(c) \subseteq \mho_c(n)$.

2. This directly follows from the definitions. Indeed, $\Omega(n)$ is a subset of every clique involved in the definition of $\mho_c(n)$, so $\Omega(n) \subseteq \mho_c(n)$, and $\mho_c(n)$ is defined as the intersection of $\Omega(c)$ with another set, so $\mho_c(n) \subseteq \Omega(c)$.

3. The properties that $\Omega(n) \subseteq \Omega^{\max}(n)$ and $\mho_c^{\max}(n) \subseteq \mho_c(n)$ are easily deduced from the definitions using $C^{\max}(n) \subseteq C(n)$.

Let us prove that $\Omega^{\max}(n) \subseteq \mho_c^{\max}(n)$. To see this, write $\mho_c^{\max}(n) = E \cap \Omega(c)$, where E is the (non-empty) union of the elements of $C^{\max}(n) \setminus C^{\max}(c)$. It is clear from the definition that $\Omega^{\max}(n) \subseteq E$. It remains to prove that $\Omega^{\max}(n) \subseteq \Omega(c)$. Fix a clique descendant A of c, let x be a maximal clique descendant of c (such a leaf exists because c is an essential node) and a maximal clique descendant y of n which is not a descendant of c (such a leaf exists because n is a skeleton node). Let us prove that $x \cap y \subseteq A$, which would finish the proof since $\Omega^{\max}(n) \subseteq x \cap y$.

The property above is clear if $x = A$. Otherwise, let n' be the nearest common ancestor of A and x. Reorder n such that c lies on the left of y, then reorder n' such that x is on the left of A. This gives a leaf order of $\mathcal{T}$ where A appears between x and y, which implies that $x \cap y \subseteq A$ by Property 2.22.1.

4. Let us first prove $A \subseteq \Omega^{\max}(n)$. Take a clique ordering $\mathcal{O} \in \pi(\mathcal{T})$. Since A is forced for n, the subclique A appears in $\mathcal{O}$ at least once on the left and once on the right of the cliques in $C^{\max}(n)$. By Property 2.22.1, it follows that $A \subsetneq x$ for every $x \in C^{\max}(n)$ and further $A \subseteq \Omega^{\max}(n)$.

It remains to show $\mho_n^{\max}(p) \subseteq A$. Recall that $\mho_n^{\max}(p) = (\bigcup_B B) \cap (\bigcap_D D)$, where the union is taken over all cliques $B \in C^{\max}(p) \setminus C^{\max}(n)$ and the intersection over the cliques $D \in C(c)$. Fix $x \in C^{\max}(n)$, we show that for every such $B \in C^{\max}(p) \setminus C^{\max}(n)$, we have $x \cap B \supseteq A$. Indeed, consider a clique ordering $\mathcal{O} \in \pi(\mathcal{T})$. Without loss of generality we may assume that the maximal clique A is on the left of x in $\mathcal{O}$. Since A is forced in $\mathcal{T}$, there is A in the left tail of $C^{\max}(n)$ in $\mathcal{O}$, so in particular A is on the left of x and on the right of B. By Property 2.22.1, it follows that $A \supseteq B \cap x$. $\square$

The plan of the proof of Lemma 2.17 is the following. We define a positive integer weight function $w(.)$ on PQ-trees designed such that $w(\mathcal{T}_1) < w(\mathcal{T}_2)$ whenever $\mathcal{T}_2$ is a restriction of $\mathcal{T}_1$, i.e. $\pi(\mathcal{T}_1) \subsetneq \pi(\mathcal{T}_2)$. The main task of the proof is to show this last inequality. To do so, we define a set of basic operations that specialize PQ-trees, i.e. from a PQ-tree $\mathcal{T}$, such an operation produces a modified version $\mathcal{T}'$ of $\mathcal{T}$ such that $\pi(\mathcal{T}) \supsetneq \pi(\mathcal{T}')$ and we show that if $\pi(\mathcal{T}_1) \supsetneq \pi(\mathcal{T}_2)$ then $\mathcal{T}_2$ can be obtained from $\mathcal{T}_1$ by a sequence of these basic operations. It then suffices to show that these operations increase $w(.)$ to deduce the sought property.

2.10.4 Labeled PQ-trees

To simplify the analysis, it is convenient to remove the forced cliques from the tree. As a downside, we need to keep track of this information by adding some decoration on the PQ-tree.

A *labeled PQ-tree* $\mathcal{T}$ is a PQ-tree $\mathcal{T}_0$ where each non-root skeleton node n is labeled by a set of cliques $\mathrm{Forced}(n)$. We say that a clique ordering $\mathcal{O}$ is *generated* by $\mathcal{T}$ if $\mathcal{O}$ is generated by $\mathcal{T}_0$ and A is in the left and right tails of $C^{\max}(n)$ in $\mathcal{O}$ for every skeleton node n and every $A \in \mathrm{Forced}(n)$. As for standard PQ-trees, we denote by $\pi(\mathcal{T})$ the set of clique orderings that are generated by $\mathcal{T}$. Note that, consistently, if n is a node of $\mathcal{T}$ with $A \in \mathrm{Forced}(n)$ then A is forced for n in $\mathcal{T}$.

A labeled PQ-tree has the following reduction rules. The purpose of the first rule is to record a clique that is forced and the purpose of the second rule is to perform the subsequent simplifications.

(a) If A is forced in n and $A \notin \mathrm{Forced}(n)$, add A in $\mathrm{Forced}(n)$.

(b) If A is a leaf dominated by the skeleton node n with $A \in \mathrm{Forced}(n)$, then delete the leaf A.

Applying iteratively reduction rules (a) and (b) allows us to transform a non-labeled PQ-tree $\mathcal{T}_1$ (considered as a labeled PQ-tree whose labels are all empty) into a labeled PQ-tree $\mathcal{T}_2$ that generates the same clique orderings.

In the reverse direction, a labeled PQ-tree $\mathcal{T}_2$ can be transformed into a non-labeled PQ-tree $\mathcal{T}_1$ that generates the same clique orderings. To do so, for each skeleton node n of $\mathcal{T}_2$ and each subclique $A \in \mathrm{Forced}(n)$, insert A in the tree as follows. Let p denote the skeleton parent of n and let $p = n_0, \ldots, n_k = n$ be the path from p to n in $\mathcal{T}$, with $k \geq 1$. We know from Proposition 2.28 that

$$\mho_n(p) \subseteq \Omega(n_1) \subseteq \mho(n_1) \subseteq \cdots \subseteq \Omega(n_{k-1}) \subseteq \mho(n_{k-1}) \subseteq \Omega(n).$$

and that

$$\mho_n^{\max}(p) \subseteq A \subseteq \Omega^{\max}(n).$$

If there are two consecutive nodes n_i and n_{i+1} with $\mho_n(n_i) \subseteq A \subseteq \Omega(n_{i+1})$, then insert a simple Q-node n'_{i+1} with three children A, n_{i+1}, A, ordered in this way. In the new tree, n'_{i+1} takes the place of n_{i+1} among the children of n_i. If there is a node n_i with $1 \leq i \leq k-1$ and $\Omega_n(n_i) \subsetneq A \subsetneq \mho(n_i)$, or $i = 0$ and $A \subsetneq \mho(n_0)$, or $i = p$ and $\Omega_n(n_p) \subsetneq A$, then insert A in the left and right tails of n in n_i.

Proposition 2.29 (Consistency)**.**

1. *If $\pi(\mathcal{T}_2) \neq \varnothing$, the construction above is always possible, that is either $\mho_n(n_i) \subseteq A \subseteq \Omega(n_{i+1})$ or $\Omega_n(n_i) \subsetneq A \subsetneq \mho(n_i)$ for some index i, and in this last case the insertion is possible.*

2. *If $\mathcal{T}_1$ is reduced to $\mathcal{T}_2$ by rules (a) and (b), then $\pi(\mathcal{T}_1) = \pi(\mathcal{T}_2)$.*

3. If $\mathcal{T}_2$ is transformed into $\mathcal{T}_1$ by the operation above, then $\pi(\mathcal{T}_1) = \pi(\mathcal{T}_2)$.

Proof. 1. We know that $|\mho_n^{\max}(p)| \leq |A| \leq |\Omega^{\max}(n)|$ and

$$|\mho_n(p)| \leq |\Omega(n_1)| \leq |\mho(n_1)| \leq \cdots \leq |\Omega(n_{k-1})| \leq |\mho(n_{k-1})| \leq |\Omega(n)|.$$

Since A is forced in n, the clique A is in the left and right tails of $C^{\max}(n)$ in every clique ordering $\mathcal{O} \in \pi(\mathcal{T})$. It follows that every clique B in the left or right tail of such an ordering $\mathcal{O}$ satisfies $A \subseteq B$ or $B \subseteq A$. In particular, if B is a leaf of n_i in the left or right tail of n in the PQ-tree $\mathcal{T}$ for some $i \in [k]$, then $A \subseteq B$ or $B \subseteq A$.

For $i \in [k]$, let $L_1^i, \ldots, L_{\ell_i}^i$ and $R_1^i, \ldots, R_{r_i}^i$ be respectively the left and right tails of n in n_i. Recall that if $1 < i < k$, then $\Omega_n(n_i) = \bigcap_{j=1}^{\ell_i} L_j^i \cap \bigcap_{j=1}^{r_i} R_j^i$. and $\mho_n(n_i) = \bigcup_{j=1}^{\ell_i} L_j^i \cup \bigcup_{j=1}^{r_i} R_j^i$. Take $\mathcal{O} \in \pi(\mathcal{T}_1)$. In the clique ordering $\mathcal{O}$, the left tail of $C^{\max}(n)$ is a subsequence of the increasing sequence

$$\mathcal{L} = L_1^0, \ldots, L_{\ell_0}^0, L_1^1, \ldots, L_{\ell_1}^1, \ldots, L_1^p, \ldots, L_{\ell_p}^p$$

where, in particular, A is inserted. A similar situation happens for the right tail of $C^{\max}(n)$ in $\mathcal{O}$.

As a consequence, if $|A| \leq |\Omega(n_i)|$ for some i, then A is inserted before each L_j^i in $\mathcal{L}$, so $A \subseteq L_j^i$, and similarly, $A \subseteq R_j^i$ if $j \in [r_j]$. It follows that $A \subseteq \Omega(n_i)$. One can prove similarly that if $|\mho_n(n_i)| \leq |A|$, then $\mho_n(n_i) \subseteq A$.

We now consider two cases.

If there is $i \in \{0, \ldots, p-1\}$ such that $|\mho_n(n_i)| \leq |A| \leq |\Omega_n(n_{i+1})|$, then we have $\mho_n(n_i) \subseteq A \subseteq \Omega(n_{i+1})$ by the remark above.

Otherwise, there is $i \in \{0, \ldots, p\}$ such that $|\Omega(n_i)| < |A| < |\mho_n(n_i)|$. In this case, we know that in $\mathcal{O}$, the subclique A is inserted in the increasing sequence $L_1^i, \ldots, L_{\ell_i}^i$ and the decreasing sequence $R_1^i, \ldots, R_{r_i}^i$, and it follows from the bounds on $|A|$ that $\Omega(n_i) \subsetneq A \subsetneq \mho_n(n_i)$.

2. It suffices to prove this property for a single application of one of the rules (a) and (b).

Let $\mathcal{T}_2$ be the tree obtained from a tree $\mathcal{T}_1$ by a single application of Rule (a) on the subclique A forced in n. Since A is forced in n (in $\mathcal{T}_1$ and $\mathcal{T}_2$), the set of clique orderings generated by $\mathcal{T}_1$ with A in the left and right tail of $C^{\max}(n)$ is exactly the set of clique orderings generated by $\mathcal{T}_1$, i.e. $\pi(\mathcal{T}_2) = \pi(\mathcal{T}_1)$.

Let now $\mathcal{T}_2$ be the tree obtained from a tree $\mathcal{T}_1$ by a single application of Rule (b). It is clear that $\pi(\mathcal{T}_1) \subseteq \pi(\mathcal{T}_2)$ since the missing subclique A can still be added to a leaf order of $\mathcal{T}_2$ by clique insertion. In the other direction, every clique ordering $\mathcal{O} \in \pi(\mathcal{T}_2)$ contains A in the right and left tail of $C^{\max}(n)$ (because $A \in \mathrm{Forced}(n)$ in $\mathcal{T}_2$), so $\mathcal{O}$ can be generated by $\mathcal{T}_1$ using the same orientations of the nodes as when generating $\mathcal{O}$ with $\mathcal{T}_2$. This proves that $\pi(\mathcal{T}_1) = \pi(\mathcal{T}_2)$.

3. The construction exactly ensures that A is present on both tails of $C^{\max}(n)$ in a Q-node dominated by the node n, and further A is forced to appear on the left and right tails of $C^{\max}(n)$ in every leaf order of the resulting tree $\mathcal{T}_1$.

$\square$

2.10.5 Simplified PQ-tree

To be able to define a good measure of the PQ-tree, we need to eliminate redundancy in the tree. For that purpose, we need further reduction rules that we apply repeatedly until no such rule can be applied.

(c) If there is a Q-node n with two consecutive child leaves n_1 and n_2 representing the same clique, delete n_2.

(d) Let n be a Q-node and let c_1 and c_2 be different essential children of n such that c_2 is the first essential child of n on the right of c_1. If there is a leaf $A = \Omega(c_1) \cap \Omega(c_2)$ between c_1 and c_2, then remove this leaf A.

Let n is a simple Q-node with only one subclique leaf A (therefore two children). If the rest of the tree forces the existence of a single leaf A in this part of the tree, the node n brings no further information. The following reduction rules permit to remove such a node in every of these cases.

(e) Let n be a simple Q-node with two children, whose unique child subclique represents A and let p be the parent of n. If p is a Q-node where A appears adjacently to n, then remove A from n.

(f) With the same definition, let c be the unique essential child of n. If p is a Q-node where A appears adjacently at the leftmost or rightmost position, then remove A from n.

(g) With the same definition, if the parent p of n is a P-node with $\mho(p) = A$, then remove A from n..

(h) With the same definition, if the parent p of n is Q-node with one essential sibling n' of n with no other essential node between n and n', and if further $A = \Omega(n) \cap \Omega(n')$ (or equivalently $A \subseteq \Omega(n')$), then remove A from n..

After a reduction, if a node n has only one child, replace n by the child of n. This in particular happens after every application of Rule (e), (f), (g) or (h). During the algorithm, a node with exactly two children is always considered as a Q-node.

Property 2.30 (Consistency of the simplification rules)**.** *If $\mathcal{T}'$ is obtained from $\mathcal{T}$ by an application of a rule (c) to (h), then $\pi(\mathcal{T}') = \pi(\mathcal{T})$.*

Proof. Regardless of which rule is applied, the tree $\mathcal{T}'$ is obtained from $\mathcal{T}$ by deleting a leaf of $\mathcal{T}$. As a consequence, every leaf order O' of $\mathcal{T}'$ is a subsequence of a leaf order O' of $\mathcal{T}'$, so O can be obtained from O' by a clique insertion. Consequently, $\pi(\mathcal{T}') \subseteq \pi(\mathcal{T})$.

It remains to prove that $\pi(\mathcal{T}) \subseteq \pi(\mathcal{T}')$. Let $\mathcal{O}_1 = Q_1 \dots Q_n \in \pi(T)$ be a closed clique ordering generated from a leaf order $\mathcal{O}_0 = Q_{i_1} \dots Q_{i_m}$ of $\mathcal{T}$. We show that there is a leaf order $\mathcal{O}'_0$ of $\mathcal{T}'$ that generates $\mathcal{O}_1$.

We consider different cases corresponding to the rule used to construct $\mathcal{T}'$.

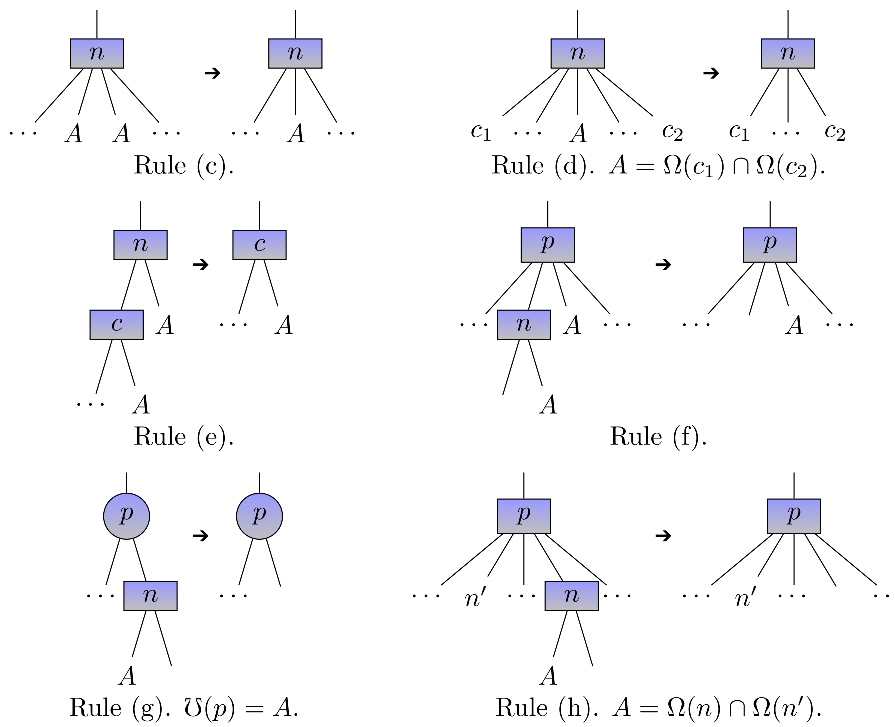


Figure 2.7 – Reduction rules (c) to (h).

- Rule (c). Let Q_{i_j} and $Q_{i_{j+1}}$ be the clique in $\mathcal{O}_0$ corresponding to the labels of n_1 and n_2 respectively and let A their (common) label. We have $\mathcal{O}_0 = Q_{i_1} \dots Q_{i_{j-1}} A A Q_{i_{j+1}} \dots Q_{i_m}$. Note that the trees $\mathcal{T}$ and $\mathcal{T}'$ have the same inner nodes. Keeping the same ordering for each node, the tree $\mathcal{T}'$ yields the leaf order $\mathcal{O}'_0 = Q_{i_1} \dots Q_{i_{j-1}} A Q_{i_{j+1}} \dots Q_{i_m}$, which is equivalent to $\mathcal{O}_0$, so $\mathcal{O}'_0$ generates $\mathcal{O}_1$.

- Rule (d). Write $A = \Omega(c_1) \cap \Omega(c_2)$. Keeping the same ordering of the nodes as when generating $\mathcal{O}_0$ with $\mathcal{T}$, $\mathcal{T}'$ gives the leaf order $\mathcal{O}'_0 = Q_{i_1} \dots Q_{i_j} A Q_{i_{j+1}} \dots Q_{i_m}$ for some $j \in [m]$. Let Q_{i_a} and Q_{i_b} be respectively the rightmost clique of $C^{\max}(c_1)$ and the leftmost clique of $C^{\max}(c_2)$. By the structure of $\mathcal{T}'$, there is no maximal clique between Q_{i_a} and Q_{i_b} in $\mathcal{O}_0$, and thus in $\mathcal{O}_1$. Since $\mathcal{O}_1$ is closed and by Property 2.22.5, the clique $Q_{i_a} \cap Q_{i_b}$ is present between Q_{i_a} and Q_{i_b} in $\mathcal{O}_1$. Moreover, $Q_{i_a} \cap Q_{i_b} = \Omega(c_1) \cap \Omega(c_2)$. Indeed, it is clear from the definition that $Q_{i_a} \cap Q_{i_b} \supseteq \Omega(c_1) \cap \Omega(c_2)$ and by Property 2.22.1, $Q_{i_a} \cap Q_{i_b} \subseteq \Omega(c_1) \cap \Omega(c_2)$. Recall that $\mathcal{O}_1$ is a subsequence of $\mathcal{O}_0$. Consequently, $\mathcal{O}'_0$ is a subsequence of $\mathcal{O}_1$ so $\mathcal{O}'_0$ generates $\mathcal{O}_1$.

- Rules (e) and (f). In both cases, assuming that Q_{i_j} is the clique A coming from a subclique child of p (for (e)) or c (for (f)), we have $\mathcal{O}_0 = Q_{i_1} \dots Q_{i_{j-1}} A Q_{i_{j+1}} \dots Q_{i_m}$. Ordering n such that its subclique A is adjacent in the leaf order and keeping the orders of the other nodes, $\mathcal{T}'$ gives a clique order $\mathcal{O}'_0 = Q_{i_1} \dots Q_{i_{j-1}} A A Q_{i_{j+1}} \dots Q_{i_m}$, which is equivalent to $\mathcal{O}_0$.

- Rule (g). It follows from the structure of $\mathcal{T}$ that there is a sibling c' of c such that the cliques of $C^{\max}(c) \cup C^{\max}(c')$ are consecutive as maximal cliques in $\mathcal{O}_0$. Let Q_{i_a} be a clique of $C^{\max}(c)$ that is adjacent as a maximal clique to a clique $Q_{i_b} \in \mathcal{C}^{\max}(c')$. We know that $A = Q_{i_a} \cap Q_{i_b}$. It is possible to order the tree $\mathcal{T}'$ to generate the leaf order $\mathcal{O}'_0 = Q_{i_1} \dots Q_{i_j} A Q_{i_{j+1}} \dots Q_{i_m}$ where j is between i_a and i_b. Since $\mathcal{O}_1$ is closed, it contains A between Q_{i_a} and Q_{i_b}, so $\mathcal{O}'_0$ generates $\mathcal{O}_1$.

- Rule (h). There is a clique $Q_{i_a} \in C^{\max}(n)$ and a clique $Q_{i_b} \in C^{\max}(n')$ with no maximal clique between Q_{i_a} and Q_{i_b} in $\mathcal{O}_0$. Up to reversing $\mathcal{O}_0$, the PQ-tree $\mathcal{T}'$ generates the clique order

  $$\mathcal{O}'_0 = Q_{i_1} \dots Q_{i_a} A \dots Q_{i_b} Q_{i_m}.$$

  Since $A = Q_{i_a} \cap Q_{i_b}$ and $\mathcal{O}_1$ is closed, $\mathcal{O}'_0$ generates $\mathcal{O}_1$.

□

2.10.6 Weight of a simplified PQ-tree

If $\mathcal{T}$ is a (labeled) PQ-tree, the *weight* $w(n)$ of a node n is defined as follows.

- A non-maximal clique leaf has weight 0.

- Let n be a maximal clique leaf, then $w(n) = 10|\operatorname{Forced}(n)|$
- Let n be simple Q-node with $d+1$ children, such that d of them are non-maximal leaves. Set $w(n) = 1$ if $d = 1$ and $w(n) = 4d - 5$ if $d \geq 2$.
- Let n be a P-node with d children (these children are essential nodes). Set $w(n) = 1 + 10|\operatorname{Forced}(n)|$.
- Let n be a Q-node with $d \geq 2$ essential children, m_1 non-essential children on the boundary of n and m_2 other non-essential children. Set $w(n) = 6d - 11 + 4m_1 + 5m_2 + 10|\operatorname{Forced}(n)|$.

The *weight* of $\mathcal{T}$ is then $w(\mathcal{T}) = \sum_n w(n)$, where the sum is taken over all the nodes of $\mathcal{T}$.

The next statement will use the following hypothesis on a PQ-tree $\mathcal{T}$.

($*_1$) For every node n of $\mathcal{T}$ and every subclique A, there are at most two cliques dominated by the descending path to n that are labeled with A.

($*_2$) For every skeleton Q-node n and every essential child c of n, the leaves of the left and right tails of c in n are pairwise labeled with distinct cliques. In other words, if $Q_1, \ldots, Q_k$ and $Q_{k+1}, \ldots, Q_p$ are respectively the left and right tails of c in n then the cliques of $(Q_i)_{i=1}^p$ are pairwise distinct.

Lemma 2.31. *Let $\mathcal{T}_1$ be a PQ-tree reduced by the rules (b) (c) (d) (e) (f) (g) and (h) satisfying ($*_1$). Let $\mathcal{T}_2$ be the PQ-tree obtained by reducing $\mathcal{T}_1$ by all the rules ((a) to (h)), it holds that $w(\mathcal{T}_1) \leq w(\mathcal{T}_2)$. Moreover, if Rule (a) applies at least once on $\mathcal{T}_1$ and $\mathcal{T}_1$ satisfies ($*_2$) then $w(\mathcal{T}_1) < w(\mathcal{T}_2)$.*

Proof of Lemma 2.31. We prove this result inductively on the number of applications of Rule (a). If Rule (a) does not apply to $\mathcal{T}_1$, then $\mathcal{T}_1$ is reduced and $\mathcal{T}_1 = \mathcal{T}_2$, in which case the lemma holds. Assume otherwise that Rule (a) applies on a node n and a subclique A, and let $\mathcal{T}_{1.1}$ be the PQ-tree produced by this single application of (a). Let $\mathcal{T}_{1.2}$ be the PQ-tree obtained by further reducing $\mathcal{T}_{1.1}$ by the rules (b) to (h) (i.e. all but (a)). Note that $\mathcal{T}_2$ is the reduction of $\mathcal{T}_{1.2}$ by all the reduction rules.

By induction hypothesis, $w(\mathcal{T}_2) - w(\mathcal{T}_{1.2}) \geq 0$. Further, it follows from the definition of the weight that $w(\mathcal{T}_{1.1}) - w(\mathcal{T}_1) = 10$. It remains to compute $w(\mathcal{T}_{1.2}) - w(\mathcal{T}_{1.1})$.

By hypothesis, the tree $\mathcal{T}_1$ satisfies ($*_1$), so $\mathcal{T}_{1.1}$ also satisfies ($*_1$). Since a leaf A dominated by n is dominated by the descending path to n, it follows that there are at most 2 leaves of $\mathcal{T}_{1.1}$ on which Rule (b) applies. For such a leaf ℓ (so ℓ is labeled with A), let m be the parent of ℓ. We distinguish three cases.

- If m is a simple Q-node with 2 children, then m is deleted, the weight decreases by $w(m) = 1$. After this, no other rule (from (b) to (h)) is triggered.

- If m is a simple Q-node with $d \geq 2$ subclique children, we have $w(m) = 4d - 5$ in $\mathcal{T}_1$. After the deletion of ℓ, one of the rules (e), (f), (g) and (h) may then apply on m when $d = 2$ (so m is a simple Q-node with exactly one subclique leaf child after the deletion of ℓ). In this case, m is deleted, so the weight decreases by $w(m) = 4 \cdot 2 - 5 = 3$. Otherwise the weight decreases by $(4d - 5) - (4(d-1) - 5) = 4$.

- If m is a skeleton Q-node (that is $m \in \{n, p\}$), then the deletion of ℓ makes the weight decrease by 4 if ℓ is not on the boundary of n and by 5 otherwise. Further, this decrease can be 5 only if $m = p$. Indeed, if $m = n$ then A is on the boundary of m since $\mho(n) \subseteq A$.

It follows that $w(\mathcal{T}_{1.2}) - w(\mathcal{T}_{1.1}) \geq -5 \cdot 2 = -10$. As a consequence,

$$\begin{aligned} w(\mathcal{T}_2) - w(\mathcal{T}_1) &= (w(\mathcal{T}_2) - w(\mathcal{T}_{1.2})) + (w(\mathcal{T}_{1.2}) - w(\mathcal{T}_{1.1})) + (w(\mathcal{T}_{1.1}) - w(\mathcal{T}_1)) \\ &\geq 10 - 10 + 0 = 0. \end{aligned}$$

This proves the first part of the Lemma.

In the discussion above, the equality $w(\mathcal{T}_{1.2}) - w(\mathcal{T}_{1.1}) = -10$ may happen only if the applications of Rule (b) remove two leaves ℓ_1 and ℓ_2 labeled with A that are both children of the node p. Since ℓ_1 and ℓ_2 are both dominated by n, they are both leaves of the left or right tail of the child of p that is an ancestor of n (or equal to n). Since ℓ_1 and ℓ_2 are labeled with the same clique A, this cannot happen if $\mathcal{T}$ satisfies $(*_2)$. Therefore, $w(\mathcal{T}_{1.2}) - w(\mathcal{T}_{1.1}) > -10$, so $w(\mathcal{T}_2) > w(\mathcal{T}_1)$. This proves the second case of the lemma. $\square$

Lemma 2.32. *If $\mathcal{T}$ is a reduced PQ-tree, then $\mathcal{T}$ satisfies $(*_1)$ and $(*_2)$.*

Proof. Let $\mathcal{T}$ be a reduced PQ-tree. We first prove that $\mathcal{T}$ satisfies $(*_2)$.

Assume otherwise that there is a node n with an essential child c and two subclique children ℓ_1 and ℓ_2 of n labeled with A that are in the left or right tail of c.

Assume that ℓ_1 and ℓ_2 are on the same tail of c, say the left tail. Write $Q_1, \ldots, Q_k$ the left tail of c, we have $\ell_1 = Q_{i_1}$ and $\ell_2 = Q_{i_2}$ for some $i_1, i_2 \in [k]$. We may also assume that $i_1 < i_2$. Since this sequence is a left tail, we have

$$A = Q_{i_1} \subseteq \cdots \subseteq Q_{i_2} = A.$$

As a consequence, $Q_{i_1+1} = A$ so Rule (c) applies on ℓ_1 and Q_{i_1+1} (which may be ℓ_2).

If one of ℓ_1 and ℓ_2 is in the left tail of c and the other is in the right tail of c, then A is forced for the node n_0 that dominates ℓ_1 and ℓ_2. As a consequence, Rule (a) applies if $A \notin \mathrm{Forced}(n_0)$ and Rule (a) applies on ℓ_1 (and ℓ_2) otherwise.

In both cases, it contradicts the hypothesis that $\mathcal{T}$ is reduced.

It remains to show that $\mathcal{T}$ satisfies $(*_1)$. Assume for the sake of contradiction that there are three leaves ℓ_1, ℓ_2 and ℓ_3 dominated by the descending path to n. For $i \in \{1, 2, 3\}$, let m_i be the parent of ℓ_i. Since $\mathcal{T}$ satisfies $(*_2)$, we know that m_1, m_2 and m_3 are pairwise distinct. Since these three Q-nodes are on the path to n, we may assume that m_i is the ancestor of m_{i+1} for $i \in \{1, 2\}$. By Item 2 of Proposition 2.28,

$$\mho_n(m_1) \subseteq \Omega(m_2) \subseteq \mho_n(m_2) \subseteq \Omega(m_3).$$

Moreover by Item 1 of Property 2.28, $A \subseteq \mho_n(m_1)$ and $\Omega(m_3) \subseteq A$. Consequently, $\Omega(m_2) = \mho(m_2) = A$, so the only leaf of m_2 is A and by hypothesis A is present only once. As a consequence, Rule (e) applies to $\mathcal{T}$, which is a contradiction. $\square$

2.10.7 Operations on PQ-trees

We define operations that specialize PQ-trees. From a PQ-tree $\mathcal{T}_1$, such an operation produces a PQ-tree $\mathcal{T}_1'$ satisfying $\pi(\mathcal{T}_1) \subseteq \pi(\mathcal{T}_1')$. After each of theses operations, the reduction rules are applied to $\mathcal{T}_1'$ as long as it is possible, which produces a PQ-tree $\mathcal{T}_2$. Property 2.30 then ensures $\pi(\mathcal{T}_1) \subseteq \pi(\mathcal{T}_2)$. For each of these rules, we show that $w(\mathcal{T}_1) < w(\mathcal{T}_2)$. We keep these notations in the description of the rules.

First operation: Force. Let n be a skeleton node of $\mathcal{T}_1$ with skeleton parent p and A be a subclique satisfying $\mho_n^{\max}(p) \subseteq A \subseteq \Omega^{\max}(n)$ and $A \notin \text{Forced}(n)$. The tree $\mathcal{T}_1'$ produced by this operation is obtained from $\mathcal{T}_1$ by adding A to the label set $\text{Forced}(n)$.

By Lemma 2.32, the PQ-tree $\mathcal{T}_1$ satisfies $(*_1)$ and $(*_2)$ because $\mathcal{T}_1$ is reduced. Thus, Lemma 2.31 applies, so the PQ-tree $\mathcal{T}_2$ reduced from $\mathcal{T}_1'$ satisfies $w(\mathcal{T}_1) < w(\mathcal{T}_2)$.

Second operation: Add a leaf. Let n be a Q-node of $\mathcal{T}_1$ dominated by n_0 and $A \notin \text{Forced}(n_0)$ a subclique. The tree $\mathcal{T}_1'$ produced from $\mathcal{T}_1$ by this operation is obtained by inserting a leaf ℓ labeled with A at some position in n, in the left or right tail of a children c of n. We assume that $\pi(\mathcal{T}_1') \subsetneq \pi(\mathcal{T}_1)$.

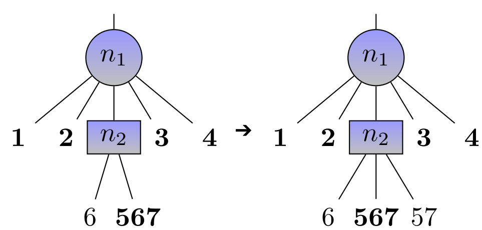


Figure 2.8 – Example of leaf addition.

If A is already present in the other tail of c in n, then this operation boils down to a Force operation. Indeed, note that A is forced for n in $\mathcal{T}_1'$. Let $\mathcal{T}_{1.5}$ be the PQ-tree obtained from $\mathcal{T}_1$ by a force operation on A and n (without further reduction). Note that applying Rule (a) (to A) followed by Rule (b) (to remove the new leaf) reduces $\mathcal{T}_1'$ to the PQ-tree $\mathcal{T}_{1.5}$. Let $\mathcal{T}_2$ be the (common) reduction of $\mathcal{T}_1'$ and $\mathcal{T}_{1.5}$, it holds that $w(\mathcal{T}_1) < (\mathcal{T}_2)$, as proved in the previous paragraph. Otherwise, rules (a) and (b) do not apply.

In the other cases, (a) does not apply on $\mathcal{T}_1'$. Moreover, the hypothesis $\pi(\mathcal{T}_1') \subsetneq \pi(\mathcal{T}_1)$ ensures that rules (c) and (d) do not apply to $\mathcal{T}_1'$. Indeed, these rules may apply only on ℓ because $\mathcal{T}_1$ is reduced, which would give $\mathcal{T}_1 = \mathcal{T}_2$, and further $\pi(\mathcal{T}_1') = \pi(\mathcal{T}_1)$. Also note that n has at least three children in $\mathcal{T}_1'$.

As a consequence, the reduction rules that may apply on $\mathcal{T}_1'$ are rules (f) and (e). These two last rules may only apply for the subclique A, to remove respectively the node c and on the parent p of n.

Let us compute $w(\mathcal{T}_2) - w(\mathcal{T}_1)$. The weight increase $w(\mathcal{T}_1') - w(\mathcal{T}_1)$ is 2 (if n is a simple Q-node with a single subclique leaf), 3 (if n is a simple Q-node with at least 2 subclique leaves) or 4 (if n is a skeleton Q-node). We claim that the rules (e) and (f) cannot both apply on $\mathcal{T}_1'$. This would show that $w(\mathcal{T}_2) - w(\mathcal{T}_1') \geq -1$, and further $w(\mathcal{T}_2) - w(\mathcal{T}_1) \geq 1$.

Let us prove the claimed property. Let p be the parent of n, and assume for a contradiction that rules (e) and (f) apply on p and c respectively. It follows from Proposition 2.28 that $A = \Omega(p) \subseteq \Omega(n) \subseteq \mho_c(n) \subseteq \Omega(c) = A$, so $\Omega(n) = \mho_c(n) = A$. As a consequence, the left and right tails of c in n (for the tree $\mathcal{T}_1$) are empty. Indeed, every subclique B in the right or left tail of c in n satisfies $\Omega(n) \subseteq B \subseteq \mho_c(n)$ so $B = A$. This shows that $\mathcal{T}_1$ does not satisfy $(*_1)$ (the leaf A attached to the descending path to c), and contradicts Lemma 2.32.

Third operation: Add a node. Let p and c be nodes of $\mathcal{T}_1$ such that p is the parent of c. The tree $\mathcal{T}_1'$ produced by this operation is obtained by replacing c in the children of p by a simple Q-node n whose essential child is c. We assume that $\mho_c(p) \subseteq \Omega(n)$, $\mho(n) \subseteq \Omega(c)$ and that if c' is the skeleton child of n (so the subclique leaves of n are dominated by c'), then no subclique A among the children of n belongs to Forced(c'). We also assume that $\pi(\mathcal{T}_1) \supsetneq \pi(\mathcal{T}_1')$, so that n cannot be deleted by Rule (f) or (e).

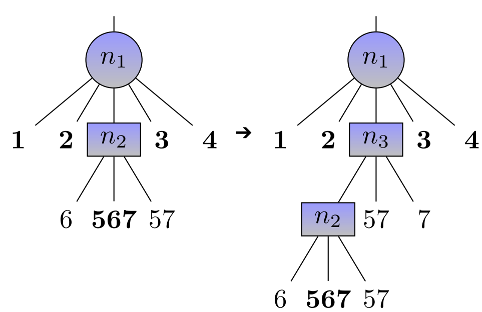


Figure 2.9 – Example of node addition.

Let $\mathcal{T}_2$ be the tree reduced from $\mathcal{T}_1'$. If n has a single subclique child, then no rule applies and $\mathcal{T}_2 = \mathcal{T}_1'$ and the weight increases by $w(\mathcal{T}_2) - w(\mathcal{T}_1) = 1$. If n has

$d \geq 2$ subclique children, this increases the weight by $w(\mathcal{T}_1) - w(\mathcal{T}_1') = 4d - 5 \geq 3$. After this, the only rules that may apply are rules (f) (on the node p) and (e) (on the node c). Consequently, $w(\mathcal{T}_1) - w(\mathcal{T}_2) \geq w(\mathcal{T}_1) - w(\mathcal{T}_1') - 2 \geq 1$.

Fourth operation: Merge two nodes. Let n and p be two Q-nodes such that p is the parent of n in the PQ-tree $\mathcal{T}_1$. Write

$$Q_1, \ldots, Q_{k-1}, n, Q_{k+1}, \ldots, Q_q$$

the children of p and $R_1, \ldots, R_r$ the children of n. The PQ-tree $\mathcal{T}_1'$ given by this operation is obtained from $\mathcal{T}_1$ replacing p by a node p' with children

$$Q_1, \ldots, Q_{k-1}, R_1, \ldots, R_r, Q_{k+1}, \ldots, Q_q,$$

possibly after reversing n.

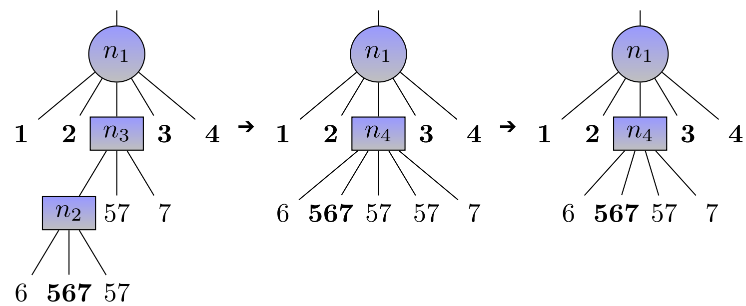


Figure 2.10 – Example of node merging. Here, Rule (c) applies.

Let $\mathcal{T}_{1.5}$ be the PQ-tree obtained by reducing by the rules (b) to (h) (i.e. all but (a)). We claim for now that $w(\mathcal{T}_1) < w(\mathcal{T}_{1.5})$. By Lemma 2.32, the PQ-tree $\mathcal{T}_1$ satisfies $(*_1)$. As no node is added to a descending path by the merging nor by the reduction rules, $\mathcal{T}_{1.5}$ also satisfies $(*_1)$. By Lemma 2.31, the PQ-tree $\mathcal{T}_2$ obtained by reducing $\mathcal{T}_{1.5}$ (with all rules) satisfy $w(\mathcal{T}_{1.5}) \leq w(\mathcal{T}_2)$, so $w(\mathcal{T}_1) < w(\mathcal{T}_2)$.

It remains to discuss the claim $w(\mathcal{T}_1) < w(\mathcal{T}_{1.5})$. First note that the rules (f), (e), (g) and (h) do not apply on $\mathcal{T}_1'$, as otherwise such a rule would apply on $\mathcal{T}_1$. Moreover, the rules (c) and (d) may apply in $\mathcal{T}_1'$ on only on the clique $A = \Omega(n)$. Indeed, Rule (c) is applied on A in $\mathcal{T}_1'$ if $R_1 = Q_{k-1} = A$ or $R_r = Q_{k+1} = A$. By Items 1 and 2 of Property 2.28,

$$A \subseteq \mho_n(p) \subseteq \Omega(n) \subseteq A,$$

so $A = \Omega(n)$. Now, if Rule (d) applies to a leaf $A = R_i$ in $\mathcal{T}_1'$, then $A = \Omega(Q_{m_1}) \cap \Omega(R_{m_2})$, where Q_{m_1} and R_{m_2} are essential nodes (in $\mathcal{T}_1'$). It follows that $A = R_i \supseteq \Omega(c)$ and by Lemma 2.26 and Proposition 2.22 $A = \Omega(Q_{m_1}) \cap \Omega(R_{m_2}) \subseteq \mho_c(n) \subseteq \Omega(c)$, so again $A = \Omega(c)$.

Since $\mathcal{T}_1$ is reduced by (a) and (b), this subclique $\Omega(c)$ is present at most once in p and n, so only one of rules (c) and (d) may apply on $\mathcal{T}_1'$.

- If n is a simple Q-node with 2 children in $\mathcal{T}_1$, let A be the subclique represented by the unique subclique leaf of n_1. Since $\mathcal{T}_1$ is reduced by Rule (e), the subclique A does not appear in p in the left or right tail of n. As a consequence, Rule (c) does not apply to $\mathcal{T}_1'$. Similarly, if n' is the next essential child of p on the left or right of n in p, then $A \neq \Omega(n) \cap \Omega(n')$ in $\mathcal{T}_1$ because $\mathcal{T}_1$ is reduced by Rule (h). As a consequence, Rule (d) does not apply to $\mathcal{T}_1'$. This shows the $\mathcal{T}_{1.5} = \mathcal{T}_1'$.

  The weight increase $w(\mathcal{T}_1') - w(\mathcal{T}_1)$ is $3 - 1 - 1 = 1$ if p is a simple node with 2 children and at least $3 - 1 = 2$ otherwise.

- If p is a simple Q-node with 2 children, it can be shown similarly that the weight increases by at least 2 using Rule (f) instead of Rule (e).

- If n and p are simple Q-nodes with respectively $d_n \geq 3$ and $d_p \geq 3$ children, then $w(\mathcal{T}_1') - w(\mathcal{T}_1) = 4(d_n + d_p) - 5 - (4d_n - 5) - (4d_p - 5) = 5$. There is at most one reduction by Rule (c) when reducing $\mathcal{T}_1'$ to $\mathcal{T}_{1.5}$, in which case the weight is decreased by 4. It follows that $w(\mathcal{T}_{1.5}) - w(\mathcal{T}_1) \geq 5 - 4 = 1$.

- If n is a simple Q-node with $d \geq 3$ children and p is a skeleton node, the only subclique of $\mathcal{T}_1'$ that may be remove by Rule (c) or (d) is $\Omega(n)$. It follows that $w(\mathcal{T}_{1.5}) - w(\mathcal{T}_1) \geq 5 - 4 = 1$.

- One proves similarly that if n is a skeleton node and p a simple Q-node with $d \geq 3$ children, then $w(\mathcal{T}_{1.5}) - w(\mathcal{T}_1) \geq 1$.

- If n and p are both skeleton nodes, then for each $A \in \text{Forced}(p)$, we insert a leaf A to p on both the left and right tail of n in p. After this, we replace n in p by its children. For the sake of clarity, let us call q the node of $\mathcal{T}_1'$ that takes the role of p. For a node $x \in \{n, p, q\}$, let m_1^x, m_2^x and d^x be respectively the number of subclique leaves in the boundary of x, the number of other subclique leaves in x and the number of essential children of x, as in the definition of the weight (taken in $\mathcal{T}_1$ if $x \in \{n, p\}$ or $\mathcal{T}_1'$ if $x = q$). Every node in the boundary of q (in $\mathcal{T}_1'$) is exactly in the boundary of p (in $\mathcal{T}_1$), so $m_1^q = m_1^p$. Further, it follows from the construction that $d_q = d_n + d_p$ and $m_2^q = m_2^p + m_1^n + m_2^n + |\text{Forced}(n)|$. Applying the formula of the weight, $w(\mathcal{T}_1') - w(\mathcal{T}_1) = w(q) - w(n) - w(p) = 11$.

  After this, the reduction rules that may apply are rules (c) and (d). The only clique that can be removed this way is $\Omega(n)$, which happens at most twice (once in each tail) and each reduces the weight by 5. Consequently, $w(\mathcal{T}_2) - w(\mathcal{T}_1') \geq -10$, so $w(\mathcal{T}_2) - w(\mathcal{T}_1) \geq 1$

Fifth operation: Split. Let n be a P-node and let $P_1 \cup P_2$ be a partition of the children of n with $|P_1| \geq 1$ and $|P_2| \geq 2$. This operation replace n by a P-node n' whose children are the element of P_1 and a new node c' whose children are the elements of P_2. Set $\text{Forced}(n') := \text{Forced}(n)$ and $\text{Forced}(c') := \varnothing$.

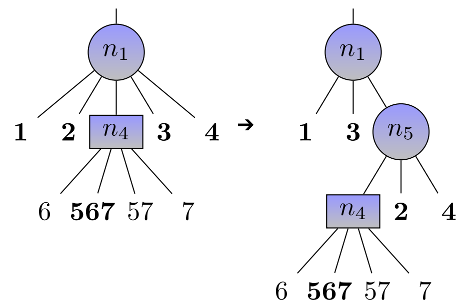


Figure 2.11 – Example of splitting.

This operation adds a P-node, so $w(\mathcal{T}_1') - w(\mathcal{T}_1) \geq 1$. Note that the particular case where a node n_i has only two (essential) children, then it is considered as a Q-node, which does not change the weight of the tree.

We know from Lemma 2.32 that $\mathcal{T}_1$ satisfies $(*_1)$, so $\mathcal{T}_1'$ also satisfies $(*_1)$. By Proposition 2.31, we conclude

$$w(\mathcal{T}_1) < w(\mathcal{T}_1') \leq w(\mathcal{T}_2).$$

Other operation: Transform a P-node into a Q-node Let n be a P-node with children $n_1, \ldots, n_m$ ordered in some arbitrary way. Note that the operation that transforms the node n into a Q-node endowed with the order $n_1, \ldots, n_m$ can be obtained by a composition of split and merge operations. To do so, split $m - 1$ times until $n_1, \ldots, n_m$ are the leaves of a binary tree, then apply $m - 1$ Merge operations so that $n_1, \ldots, n_m$ are children of a common P-node.

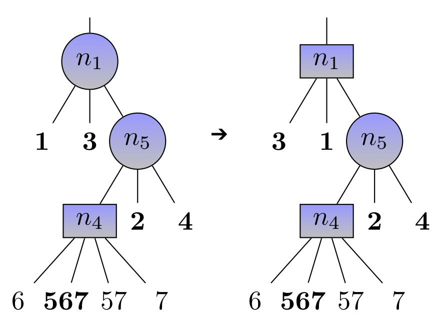


Figure 2.12 – Example: P-node into Q-node.

2.11 Specializing the PQ-tree

The operations defined in the previous section transform a PQ-tree in a way that restricts the set of clique orderings this tree generates. The following lemma shows that this set of operation is complete in the sense that restricting the clique orders generated by a PQ-tree is always a composition of these operations.

Lemma 2.33. *If $\mathcal{T}_1$ and $\mathcal{T}_2$ are reduced PQ-tree satisfying $\pi(\mathcal{T}_1) \supseteq \pi(\mathcal{T}_2)$, then there is a series of operations described above that transforms $\mathcal{T}_1$ into $\mathcal{T}_2$.*

The proof of Lemma 2.33 is postponed to section 2.13. Since each operation increases the value of $w(\mathcal{T})$, Lemma 2.33 has the following direct consequence.

Corollary 2.34. *If $\mathcal{T}_1$ and $\mathcal{T}_2$ are reduced PQ-tree satisfying $\pi(\mathcal{T}_1) \supsetneq \pi(\mathcal{T}_2)$, then $w(\mathcal{T}_1) < w(\mathcal{T}_2)$.*

Lemma 2.17 will be easily deduced from Corollary 2.34 once we have the following bounds.

Lemma 2.35. *Let $\mathcal{T}$ be a reduced PQ-tree of an interval graph G with n vertices, then $1 \leq w(\mathcal{T}) \leq 16n$.*

Assuming Lemma 2.35, we are ready to prove Lemma 2.17.

Proof of Lemma 2.17. Let G be a graph and $H_1, \ldots, H_\ell$ be a sequence of graphs as in the statement of Lemma 2.17. For $0 \leq i \leq \ell$, let $\mathcal{T}_G(H_i)$ be the reduced PQ-tree representing the interval representations of G that extend to H_i. The hypothesis of the lemma translates into the following strict inclusions

$$\pi(\mathcal{T}_G(H_1)) \supsetneq \cdots \supsetneq \pi(\mathcal{T}_G(H_\ell)).$$

Applying Corollary 2.34, it follows

$$w(\mathcal{T}_G(H_1)) < \cdots < w(\mathcal{T}_G(H_\ell)).$$

Since $(w(\mathcal{T}_G(H_1)))_{i=1}^{\ell}$ is sequence of integers,

$$\ell \leq w(\mathcal{T}_G(H_\ell)) - w(\mathcal{T}_G(H_1)) + 1.$$

By Lemma 2.35, we conclude

$$\ell \leq 16n - 1 + 1 = 16n.$$

□

2.12 Proof of Lemma 2.35

Proof of Lemma 2.35. It is clear from the definition that $w(\mathcal{T}) \geq 1$ because the weight of an essential node is at least 1. Note that this value is reached when $\mathcal{T}$ has only a P-node with the maximal cliques of G as leaves.

Let us prove the upper bound. Let $A_1, \dots, A_k$ be a clique order generated by $\mathcal{T}$. Let $\mathcal{T}^0_{\max}$ be the PQ-tree with only one Q-node, whose children are $A_1, \dots, A_k$, ordered this way. Let $\mathcal{T}_{max}$ be the PQ-tree obtained by reducing $\mathcal{T}^0_{max}$. In this reduction, only Rules (a), (b) and (d) may apply. Note that these rules keep the tree in the form with only one Q-node and leaves. Moreover, $\pi(\mathcal{T}) \supseteq \pi(\mathcal{T}^0_{\max}) = \pi(\mathcal{T}_{\max})$. In particular, Corollary 2.34 implies that $w(\mathcal{T}_{\max}) \geq w(\mathcal{T})$.

Let d be the number of maximal cliques and m the number of subcliques in the sequence $A_1, \dots, A_k$, and write $A_{m_1}, \dots, A_{m_d}$ these maximal cliques. In the reduction from $\mathcal{T}^0_{\max}$ to $\mathcal{T}_{\max}$, we choose to apply Rule (d), and then rules (a) and (b). Let $r_{(d)}$, $r_{(a)}$ and $r_{(b)}$ be the number of times the rules (d), (a) and (b) are respectively applied. Rule (d) is applied once on each $A_{i_j} \cap A_{i_{j+1}}$, so $r_{(d)} = d - 1$. Each time Rule (a) is applied on a subclique B and maximal clique leaf A_{m_j}, Rule (b) is applied on B once on the left tail and once on the right tail of A_{m_j}, except when $B = A_{i_{j-1}} \cap A_{i_j}$ (for the left tail) or $B = A_{i_j} \cap A_{i_{j+1}}$ (for the right tail), which can happen at most twice for each maximal clique. It follows that

$$r_{(b)} \geq 2r_{(a)} - 2(d-1).$$

Moreover, the number of subclique leaves of $\mathcal{T}_{\max}$ is

$$m - r_{(b)} - r_{(d)} \leq m - 2r_{(a)} + d - 1.$$

Therefore,

$$\begin{aligned} w(\mathcal{T}_{\max}) &\leq 6d - 11 + 5(m - 2r_{(a)} + d - 1) + 10r_{(a)} \\ &= 11d + 5m - 16. \end{aligned}$$

Since G is an interval graph, its number of maximal clique is at most n, i.e. $d \leq n$. Moreover, the length $d+m$ of the initial clique ordering is at most $2n+1$. we have $w(\mathcal{T}_{\max}) \leq 16n$. As a consequence,

$$11d + 5m - 16 = 6d + 5(d+m) - 16 \leq 6n + 10n + 5 - 16 \leq 16n$$

This proves $w(\mathcal{T}) \leq 16n$. □

2.13 Proof of Lemma 2.33

Lemma 2.33 can be inductively derived from the following lemma.

Lemma 2.36. *If $\mathcal{T}_1$ and $\mathcal{T}_2$ are reduced PQ-trees such that $\pi(\mathcal{T}_1) \supsetneq \pi(\mathcal{T}_2)$, then there is a sequence of at least one operation that transforms $\mathcal{T}_1$ into a (reduced) PQ-tree $\mathcal{T}_1'$ with $\pi(\mathcal{T}_1') \supseteq \pi(\mathcal{T}_2)$.*

Proof.

Claim 1. For every node n_1 of $\mathcal{T}_1$ and every node n_2 of $\mathcal{T}_2$, if $C^{\max}(n_1)$ and $C^{\max}(n_2)$ intersect then

$$C^{\max}(n_1) \subseteq C^{\max}(n_2) \quad \text{or} \quad C^{\max}(n_2) \subseteq C^{\max}(n_1).$$

Proof. Assume otherwise that there are two cliques $x_1 \in C^{\max}(n_1) \setminus C^{\max}(n_2)$ and $x_2 \in C^{\max}(n_2) \setminus C^{\max}(n_1)$. We know from the hypothesis that there also exists $y \in C^{\max}(n_1) \cap C^{\max}(n_2)$. Note that in $\mathcal{T}_2$, the leaf x_1 lies outside the subtree of $\mathcal{T}_2$ rooted at n_2, which contains x_2 and y. (See Figure 2.13). It follows that $\mathcal{T}_2$ generates a clique ordering $S \in \pi(\mathcal{T}_2)$ where x_2 is between x_1 and y. To do so, reverse every node of $\mathcal{T}_2$ if x_1 is not on the left of the descendants of n_2, then reverse every node of the subtree of $\mathcal{T}_2$ rooted at n_2 if y is not on the right of n_1. Since in the tree $\mathcal{T}_1$ the leaf x_2 is outside of the subtree rooted at n_1, which contains x_1 and y, it hold that $S \notin \pi(\mathcal{T}_1)$. □

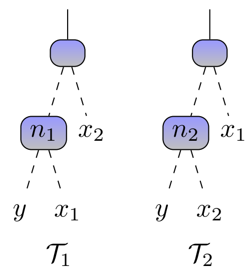


Figure 2.13 – Proof of Claim 1.

If n_1 is a skeleton node of $\mathcal{T}_1$, let $\alpha(n_1)$ be the nearest common ancestor of the leaves of $\mathcal{T}_2$ labeled by the elements of $C^{\max}(n_1)$.

Claim 2. If n_1 is a skeleton node of $\mathcal{T}_1$, then $\alpha(n_1)$ is a skeleton node.

Proof. This follows directly from the definition of $\alpha(n_1)$. If $\alpha(n_1)$ has a unique essential child c, then $C^{\max}(c) = C^{\max}(\alpha(n_1)) \supseteq C^{\max}(n_1)$, so c is a common ancestor of the cliques of $C^{\max}(n_1)$ with larger depth than $\alpha(n_1)$. □

Claim 3. If n_1 is a skeleton Q-node of $\mathcal{T}_1$, then $\alpha(n_1)$ is a skeleton Q-node.

Proof. Set $n_2 = \alpha(n_1)$. Let $c_1, \ldots, c_k$ be the essential children of n_1. For $i \in [k]$, choose a maximal clique $x_i \in C^{\max}(c_i)$. Now assume for the sake of contradiction that n_2 is a P-node. It follows that $k \geq 3$ and every child of n_2 is an essential node. Giving n_2 the order $c_2c_1c_3c_4 \ldots c_k$, we construct a clique ordering $S \in \pi(\mathcal{T}_2)$ in which the maximal clique x_1 appears after x_2 and before x_3, which is impossible for a clique ordering of $\mathcal{T}_1$. It follows that $S \notin \pi(\mathcal{T}_1)$, which yields a contradiction with our hypothesis $\pi(\mathcal{T}_2) \subseteq \pi(\mathcal{T}_1)$. This proves that n_2 is a skeleton Q-node. □

Claim 4. Let n_1 be a Q-node of $\mathcal{T}_1$ and let $c_1, \ldots, c_k$ be the essential children of $\alpha(n_1)$ (in $\mathcal{T}_2$), in the order given by $\alpha(n_1)$. Then $C^{\max}(n_1)$ correspond to consecutive skeleton children of $\alpha(n_1)$, i.e.

$$C^{\max}(n_1) = \bigcup_{i=p}^{q} C^{\max}(c_i)$$

for some $p, q \in [k]$ with $p < q$.

Proof. Let p and q be respectively the smallest and largest index of $[k]$ for which both $C^{\max}(c_p)$ and $C^{\max}(c_q)$ intersect $C^{\max}(n_2)$. This definition ensures that $C^{\max}(n_1) \subseteq \bigcup_{i=p}^{q} C^{\max}(c_i)$ and that there exist $x \in C^{\max}(c_p) \cap C^{\max}(n_1)$ and $y \in C^{\max}(c_q) \cap C^{\max}(n_1)$. It remains to prove the other inclusion.

For every $i \in [k]$, we know from the definition of $\alpha(n_1)$ that $C^{\max}(n_1) \nsubseteq C^{\max}(c_i)$, as otherwise c_i is a common ancestor of the leaves $C^{\max}(n_1)$ in $\mathcal{T}_2$ that is deepest than $\alpha(n_1)$. For $i = p$ or $i = q$, we moreover know that $C^{\max}(n_1)$ intersects $C^{\max}(c_i)$. It follows from Claim 1 that $C^{\max}(c_i) \subseteq C^{\max}(n_1)$.

Let $i \in [k]$ with $p < i < q$ and assume for a contradiction that there is $z \in C^{\max}(c_i) \setminus C^{\max}(n_1)$. In every clique ordering $S \in \pi(\mathcal{T}_2)$, the maximal clique z appears after x and before y. It follows that $S \notin \pi(\mathcal{T}_1)$, which gives a contradiction. $\square$

The next series of claims form an algorithmic proof. We show that we may assume certain properties on $\mathcal{T}_1$ and $\mathcal{T}_2$ unless a PQ-tree $\mathcal{T}_1'$ can be constructed from $\mathcal{T}_1$ by an operation that preserves the property $\pi(\mathcal{T}_1') \supseteq \pi(\mathcal{T}_2)$.

Claim 5. If there are two skeleton nodes n_1 and n_2 of respectively $\mathcal{T}_1$ and $\mathcal{T}_2$ with $C^{\max}(n_1) = C^{\max}(n_2)$, then we can assume that $\text{Forced}(n_1) = \text{Forced}(n_2)$.

Proof. Every $A \in \text{Forced}(n_1)$ appears in the left and right tails of $C^{\max}(n_1) = C^{\max}(n_2)$ in any clique order of $\pi(\mathcal{T}_1)$. Further, since $\pi(\mathcal{T}_1) \supseteq \pi(\mathcal{T}_2)$, the subclique A is forced in n_1. Since $\mathcal{T}_1$ is reduced, it follows from Rule (a) that $A \in \text{Forced}(n_2)$. This proves that $\text{Forced}(n_1) \subseteq \text{Forced}(n_2)$. Further, if there is $A \in \text{Forced}(n_1) \setminus \text{Forced}(n_2)$. We construct $\mathcal{T}_1'$ by applying a Force operation on n_1 and A. Doing this, the orders of $\pi(\mathcal{T}_1')$ are exactly the orders of $\pi(\mathcal{T}_1)$ in which A is in both the left and right tails of $C^{max}(n_1)$, which is a property that every order of $\pi(\mathcal{T}_1)$ satisfies, so $\pi(\mathcal{T}_1') \supseteq \pi(\mathcal{T}_2)$. $\square$

Claim 6. Let n_1 be a P-node with (essential) children $c_1, \ldots, c_k$. For every node n_2 of $\mathcal{T}_2$ with $C^{\max}(n_1) \cap C^{\max}(n_2) \neq \varnothing$ and $C^{\max}(n_1) \nsubseteq C^{\max}(n_2)$, we may assume that

$$C^{\max}(n_2) \subseteq C^{\max}(c_i)$$

for some $i \in [k]$.

Proof. By Claim 1, the hypothesis implies $C^{\max}(n_2) \subsetneq C^{\max}(n_1)$. Now define

$$X = \{\ i \mid C^{\max}(c_i) \cap C^{\max}(n_2) \neq \varnothing \text{ and } i \in [k]\ \}.$$

To prove the claim, it suffices to show that $|X| = 1$. Assume that $|X| \geq 2$. For each $i \in X$, we know from Claim 1 applied to c_i and n_2 that $C^{\max}(c_i) \subseteq C^{\max}(n_2)$. As a consequence, $|X| \leq k - 1$ because $C^{\max}(n_1) \nsubseteq C^{\max}(n_2)$.

In this case, we can apply a split operation to the node n_1 to separate the children $(c_i)_{i \in X}$. Indeed, in every clique ordering S generated by $\mathcal{T}_2$, there is no other maximal clique between the elements of $C^{\max}(n_2)$, so the children $(c_i)_{i \in X}$ are grouped in every leaf order of $\mathcal{T}_1$ that generates S. $\square$

Claim 7. Let n_1 be a P-node with essential children $c_1, \dots, c_k$. We may assume that $\alpha(n_1)$ has exactly k children $c'_1, \dots, c'_k$ that can be ordered so that $C^{\max}(c_i) = C^{\max}(c'_i)$ whenever $i \in [k]$. In this case, $C^{\max}(n_1) = C^{\max}(\alpha(n_1))$.

Proof. Let $c'_1, \dots, c'_{k'}$ be the essential children of $\alpha(n_1)$ (in $\mathcal{T}_2$). For $i \in [k']$ such that $C^{\max}(c'_i)$ intersects $C^{\max}(n_1)$, we prove that $C^{\max}(c'_i) \subseteq C^{\max}(c_{j(i)})$ for some $j(i) \in [k]$. By Claim 6, this is true unless $C^{\max}(n_1) \subseteq C^{\max}(c'_i)$. This last case is impossible since it contradicts the definition of $\alpha(n_1)$.

If j is not injective, that is if the set $X = \{ i \mid j(i) = j_0 \}$ has at least two elements for some $j_0 \in [k]$, then the children $(c_i)_{i \in X}$ of n_1 are consecutive when generating any clique ordering of $\pi(\mathcal{T}_2)$. So we isolate the children $(c_i)_{i \in X}$ of n_1 by a split operation.

We now assume that j is an injection. It follows that $C^{\max}(c'_i) = C^{\max}(c_{j(i)})$ for every $i \in [k']$.

If $\alpha(n_1)$ is a Q-node, then the order of $\alpha(n_1)$ induces an order with $j(.)$. When generating any clique ordering of $\pi(\mathcal{T}_2)$, the node n_1 is ordered increasingly or decreasingly according to $j(.)$. This allows us to change n_1 into a Q-node.

Otherwise, $\alpha(n_1)$ is a P-node, then for every $i \in [k']$, the set $C^{\max}(c'_i)$ intersects $C^{\max}(n_1)$, so j is defined on $[k']$. If $k' \geq k$, then there is $i_0 \in [k']$ such that $C^{\max}(c'_{i_0}) \cap C^{\max}(n_1) = \varnothing$. Letting $x \in C^{\max}(c'_{i_0})$, we can construct a clique ordering of $\pi(\mathcal{T}_2)$ where x is between two cliques of $C^{\max}(n_1)$, which gives a contradiction. It follows that and $k = k'$ and j is bijective. $\square$

Claim 8. We may assume that α induces a bijection between the P-nodes of $\mathcal{T}_1$ and the P-nodes of $\mathcal{T}_2$.

Proof. It follows directly from Claim 7 that α maps injectively P-nodes of $\mathcal{T}_1$ to P-nodes of $\mathcal{T}_2$. Take a P-node n_2 of $\mathcal{T}_2$ and let n_1 be the nearest common ancestor of the leaves labeled with $C^{\max}(n_2)$ in $\mathcal{T}_1$. It suffices to show that $\alpha(n_1) = n_2$ to deduce the claim. We know from Claim 7 that $C^{\max}(\alpha(n_1)) = C^{\max}(n_1) \supseteq C^{\max}(n_2)$. Assume that $n_2 \neq \alpha(n_1)$, so in particular $C^{\max}(n_2) \subsetneq C^{\max}(\alpha(n_1)) = C^{\max}(n_1)$. By Claim 6 applied to n_1 and n_2, it follows that $C^{\max}(n_2) \subseteq C^{\max}(c)$ for some child c of n_2. This contradicts the definition of n_1. $\square$

We now want to match simple Q-nodes of $\mathcal{T}_1$ with nodes of $\mathcal{T}_2$. To state the next lemma, we need some definitions.

Let n_1 be a simple Q-node of $\mathcal{T}_1$ and let n_2 be a Q-node of $\mathcal{T}_2$ with $C^{\max}(n_1) \subseteq C^{\max}(n_2)$. Let $c_1, \dots, c_k$ be the essential children of n_2. We know that there are a and b with $1 \leq a \leq b \leq k$ such that $C^{\max}(n_1) \subseteq \bigcup_{i=a}^{b} C^{\max}(c_i)$.

Let $\mathcal{L}_0 = L_1, \dots, L_\ell$ be the left tail of c_a in n_2 (so it corresponds to the left tail of $C^{\max}(n_1)$ in the leaf order of $\mathcal{T}_2$). We complete $\mathcal{L}_0$ as follows.

- If $\mho(n_1) \in \text{Forced}(c_a)$ and $L_\ell \neq \mho(n_1)$, we add $\mho(n_1)$ at the end of $\mathcal{L}$.
- If either $a \geq 2$ and $\mho(n_1) = \Omega(c_{a-1}) \cap \Omega(c_a)$ or $a = 1$ and $\mho(n_1) \in \text{Forced}(n_2)$, we add we add $\mho(n_1)$ at the beginning of $\mathcal{L}$ unless $L_1 = \mho(n_1)$.

The resulting sequence $\mathcal{L}$ is the *completed left tail* of $C^{\max}(n_1)$ in n_2. The *completed right tail* $\mathcal{R}$ of $C^{\max}(n_1)$ in n_2 is defined symmetrically.

The following claim extends the definition of α to the simple Q-nodes of $\mathcal{T}_1$.

Claim 9. Let n_1 be a simple Q-node of $\mathcal{T}_1$ with skeleton child s_1 and skeleton parent p_1. There is a node $n_2 = \alpha(n_1)$ on the path from $\alpha(p_1)$ to $\alpha(s_1)$ in $\mathcal{T}_2$ with the following property. Up to reversing the order of n_1, the left tail of n_1 is a subsequence of the completed left tail of $C^{\max}(n_1)$ in n_2 and, similarly, the right tail of n_1 is a subsequence of the completed right tail of $C^{\max}(n_1)$ in n_2. Further, $\Omega(n_2) \subseteq \Omega(n_1)$ and $\mho(n_1) \subseteq \mho(n_2)$.

Proof. We proceed by contradiction. Assume otherwise that for every node n_2 of $\mathcal{T}_2$ on the path from $\alpha(p_1)$ to $\alpha(s_1)$ and for both orientations of n_2, the left and right tails of n_1 are not subsets of the completed left and right tails of $C^{\max}(n_1)$ respectively. We show that there is a clique ordering $S \in \pi(\mathcal{T}_2)$ where one of the left and right tail of n_1 is not a subset of the corresponding – left or right – tail of $C^{\max}(n_1)$ in S. It will follow that S is not generated by $\mathcal{T}_1$, which yields a contradiction.

Let c_1 be the unique essential child of n_1 and write $A_1, \ldots, A_k, c_1, A_{k+1}, \ldots, A_p$ the children of n_1, in the order given by n_1.

First assume that $p \geq 2$. Choose A_i and A_j with respectively minimum and maximum size. Note that $A_i \neq A_j$ because $\mathcal{T}_1$ is reduced by Rules (a), (b) and (c), so $|\Omega(n_1)| \leq |A_i \cap A_j| < |A_i \cup A_j| \leq |\mho(n_1)|$. In particular, there is at most one node n_2 of $\mathcal{T}_2$ with $C^{\max}(n_2) \supseteq C^{\max}(n_1)$ that may contains both A_i and A_j in its completed tails of $C^{\max}(n_1)$. If there is no such node n_2, then we get a contradiction as follows.

If A_i and A_j are both on the same (let us say left) tail of c in n_1, then order each node n_2 of $\mathcal{T}_2$ such that A_i is on the left completed tail of n_2 and A_j is always on the right completed tail of n_2 (if present).

Similarly, if A_i and A_j are not on the same tails of c in n_1, then we order all the nodes of $\mathcal{T}_2$ such that A_i and A_j are only on the same tail of $C^{\max}(n_1)$. In both cases, $\mathcal{T}_2$ generates a clique ordering that $\mathcal{T}_1$ cannot generate.

Now assume that there is such a node n_2. The only subcliques A_i that may appear outside of n_2 are A_i if $A_i = \Omega(n_1)$ and A_j if $A_j = \mho(n_1)$. If n_2 does not satisfy the claim, then, again, it is possible to order the tree $\mathcal{T}_2$ such that A_i and A_j does not complete the completed left and right tails of $C^{\max}(n_1)$ into sequences that contain respectively $A_1, \ldots, A_k$ and $A_{k+1}, \ldots, A_p$. This gives a contradiction.

If $p = 1$, let $A = A_1$. If no node n_2 of $\mathcal{T}_2$ satisfies the claim, then it is possible to order each node such that the completed leaf order $\mathcal{O} \in \pi(\mathcal{T}_1)$ obtained does not contains A in the left or right tail of $C^{\max}(n_1)$. $\square$

Claim 10. Let n_1 and n_1' be two skeleton nodes of $\mathcal{T}_1$ with $\alpha(n_1) = \alpha(n_1')$, then every essential node n of $\mathcal{T}_1$ on the path from n_1 to n_1' in $\mathcal{T}_1$ satisfies $\alpha(n) = \alpha(n_1)$.

Proof. Set $n_2 = \alpha(n_1) = \alpha(n_1')$. Let a_1 be the nearest common ancestor of n_1 and n_1' in $\mathcal{T}_1$. We first show that $\alpha(a_1) = n_2$. This is follows from the hypothesis if $a_1 = n_1$ or $a_1 = n_1'$, so we assume that $a_1 \notin \{n_1, n_1'\}$ and $\alpha(a_1) \neq n_2$.

Note that a_1 is an ancestor of n_2 because $C^{\max}(\alpha(a_1)) \supseteq C^{\max}(a_1) \supseteq C^{\max}(n_1)$. Let c_1 be the child of a_1 that is an ancestor or equal to n_1 and, similarly, let c_1' be the child of a_1 that is an ancestor or equal to n_1'. On the other side, let c_2 be the child of $\alpha(a_1)$ that is an ancestor or equal to n_2. By Lemma 7, we know that a_1 is not a P-node because $C^{\max}(c_1)$ and $C^{\max}(c_1')$ are both included into $C^{\max}(c_2) = C^{\max}(n_2)$. By the definition of $\alpha(a_1)$, we know that $\alpha(a_1)$ has a leaf descendant labeled by some maximal clique $x \in C^{\max}(a_1)$ that is not a descendant of c_2. Take $y \in C^{\max}(c_1)$ and $y' \in C^{\max}(c_1')$, the order of the cliques in the clique ordering in $\pi(\mathcal{T}_1)$ induce only two different permutations of $\{x, y, y'\}$ (depending only on the orientation of a_1) while $\mathcal{T}_2$ generates four of them (depending on the orientations of $\alpha(a_1)$ and n_2). This contradicts $\pi(\mathcal{T}_1) \supseteq \pi(\mathcal{T}_2)$.

Let n be a node of $\mathcal{T}_1$ on the path from n_1 to n_1' in $\mathcal{T}_1$. Without loss of generality, we assume that n is on the path from n_1 to a_1.

If n is an essential node, note that n_2 is a common ancestor of $C^{\max}(n)$ because $C^{\max}(n) \subseteq C^{\max}(a_1)$ and $\alpha(a_1) = n_2$ and the children of n_2 are not ancestor of every clique of $C^{\max}(n_1) \subseteq C^{\max}(n)$ because $\alpha(a_1) = n_2$, so $\alpha(n) = n_2$.

If n is a simple Q-node with skeleton parent p and skeleton child s, then since both of p and s are on the path from a_1 to n_1 in $\mathcal{T}_1$, we know that $\alpha(p) = n_2 = \alpha(s)$. Claim 9 then ensures that $\alpha(n)$ is on the path from $\alpha(p) = n_2$ to $\alpha(s) = n_2$ in $\mathcal{T}_2$, so $\alpha(n_1) = n_2$. □

Claim 11. We may assume that α (defined on every node of $\mathcal{T}_2$) is injective. Further, α is a bijection between skeleton nodes of $\mathcal{T}_1$ and skeleton nodes of $\mathcal{T}_2$.

Proof. For a skeleton Q-node or single P-node n_1, we may always assume that the orientation of n_1 is consistent with the orientation of $\alpha(n_1)$ when $\mathcal{T}_1$ and $\mathcal{T}_2$ are ordered to generate a same clique ordering. Indeed, if n_1 is a skeleton node, it has to be order constantly with $\alpha(n_1)$ and if n_1 is a single Q-node, this orientation is not forced but we can assume it is consistent with $\alpha(n_1)$.

As a consequence, if there is a node n_1 of $\mathcal{T}_1$ with parent p_1 such that $\alpha(n_1) = \alpha(p_1)$ (so n_1 and p_1 are Q-nodes), we can merge n_1 and p_1.

We now assume that for two nodes n_1 and p_1 such that p_1 is the parent of n_1, it holds that $\alpha(n_1) \neq \alpha(p_1)$.

If there are two skeleton nodes n_1 and p_1 in $\mathcal{T}_1$ such that $\alpha(n_1) = \alpha(p_1)$, then by Claim 10, we can choose n_1 and p_1 such that p_1 is the parent of n_1, which contradicts our hypothesis above. It follows that α induces a bijection between skeleton nodes.

Since every node n_1 on the path from a skeleton node s_1 to its skeleton parent p_1 in $\mathcal{T}_1$ is sent by α to a node on the path from $\alpha(s_1)$ to $\alpha(p_1)$, it follows that if $n_1 \neq n_1'$ and $\alpha(n_1) = \alpha(n_1')$ then n_1 and n_1' have same skeleton child and skeleton parent in $\mathcal{T}_1$.

To conclude the proof, it suffices to show that for every nodes n_1 and n_1' of $\mathcal{T}_1$ such that n_1 is an ancestor of n_1' in $\mathcal{T}_1$, then $\alpha(n_1)$ is an ancestor of $\alpha(n_1')$ in $\mathcal{T}_2$. To see this, it suffices to note that $\Omega(\alpha(n_1)) \subseteq \Omega(n_1)$ and $\Omega_s(n_1) \subseteq \Omega_s(\alpha(n_1))$. It follows that if $\alpha(n_1) = \alpha(n_1')$ then it is possible to choose n_1 and n_1' such that n_1 is the parent of n_1', which gives a contradiction. $\square$

We are now ready to finish the proof. Claims 11 and 5 show that $\mathcal{T}_2$ can be obtained from $\mathcal{T}_1$ as follows.

First add leaves to each node n_1 of $\mathcal{T}_1$ so that it is equal to $\alpha(n_1)$. Then, insert the nodes of $\mathcal{T}_2$ that are not image of α. This concludes the proof of Lemma 2.33. $\square$

2.14 Testing subclasses

It is in general not true that if $\mathcal{P}$ is easily testable then every subclass $\mathcal{Q}$ of $\mathcal{P}$ is also easily testable. Indeed, the class of every graph is trivially easily testable while some graph classes are not. This part was an attempt to develop ways to extend the testability of $\mathcal{P}$ to the testability of $\mathcal{Q}$ modulo some hypotheses on $\mathcal{P}$ and $\mathcal{Q}$, in particular in the case of semi-algebraic graphs. Most of these remarks are now covered by the result proven in the mean time by Gishboliner and Shapira [19] on the testability of semi-algebraic classes.

2.14.1 Property testing chain relation

In this section, we define a chain relation for property testing. To that purpose, we define a parametrized version of testability of a class *inside* another one.

Definition 2.11. Fix $\epsilon > 0$ and $q \in \mathbf{N}$. Let $\mathcal{P}$ and $\mathcal{Q}$ be hereditary graph properties. The property $\mathcal{P}$ is (ϵ, q)*-testable in* $\mathcal{Q}$ if $\mathcal{P} \subseteq \mathcal{Q}$ and for every graph $G \in \mathcal{Q}$ of size at least q,

$$d(G, \mathcal{P}) > \epsilon \Rightarrow \mathbf{P}(G[X] \in \mathcal{P}) < \frac{1}{2}$$

where the probability is taken uniformly over all subsets X of $V(G)$ of size q.

In this language, a property $\mathcal{P}$ is easily testable if and only if there is a polynomial Q such that $\mathcal{P}$ is $(\epsilon, Q(1/\epsilon))$-testable in the class of all graphs for every $\epsilon > 0$.

It may not be true that if $\mathcal{P}$ is (ϵ, p)-testable in $\mathcal{Q}$ and $\mathcal{Q}$ is (η, q)-testable in $\mathcal{R}$ then $\mathcal{P}$ is $(f(\epsilon, \eta), g(p, q))$ for some reasonable functions f and g. Instead, we need that $\mathcal{P}$ is testable in a small extension of $\mathcal{Q}$.

If η is a positive number and Q is a class of graphs, let Q_η be the set of graphs that are at distance at most η from Q. Then the following chain relation holds.

Property 2.37. *Let $\mathcal{P}$, $\mathcal{Q}$ and $\mathcal{R}$ be hereditary properties such that $\mathcal{P} \subseteq \mathcal{Q} \subseteq \mathcal{R}$ and assume that for some ϵ, η, p, and q,*

1. *$\mathcal{P}$ is (ϵ, p)-testable in $\mathcal{Q}_\eta$ and*

2. *$\mathcal{Q}$ is (η, q)-testable in $\mathcal{R}$.*

Then $\mathcal{P}$ is (ϵ, k)-testable in $\mathcal{R}$ where $k = \max(p, q)$.

Proof. First notice that if $\mathcal{S}$ is a hereditary property, G is a graph, and p_k denotes the probability that the subgraph $G[X_k]$ induced by a subset X_k chosen uniformly at random among the subsets of $V(G)$ of size k belongs to $\mathcal{S}$, then $p_k \leq p_\ell$ whenever $k \geq \ell$. Indeed, deleting $k - \ell$ elements (chosen uniformly at random) of X_k gives a random set X'_ℓ with same distribution as X_ℓ and that is correlated to X_k in such a way that $G[X'_\ell] \in \mathcal{S}$ whenever $G[X_k] \in \mathcal{S}$ since $X'_\ell \subseteq X_k$ and $\mathcal{S}$ is hereditary.

Fix a graph $G \in \mathbf{R}$ that is ϵ-far from $\mathcal{P}$ and let $X = X_k$ be a subset chosen uniformly at random among all subsets of $V(G)$ of size k. We consider two cases :

Case 1: If $G \in \mathcal{Q}_\eta$, then by the remark above applied to $\mathcal{S} = \mathcal{P}$ and Hypothesis 1, $\mathbf{P}(G[X] \in \mathcal{P}) \leq \frac{1}{2}$ since $k \geq p$.

Case 2: Otherwise, $G \notin \mathcal{Q}_\eta$. By Hypothesis 2 and the remark above applied to $\mathcal{Q}$, $\mathbf{P}(G[X] \in \mathcal{Q}) \leq \frac{1}{2}$ because $k \geq q$. As $\mathcal{P} \subseteq \mathcal{Q}$, it follows that $\mathbf{P}(G[X] \in \mathcal{P}) \leq \mathbf{P}(G[X] \in \mathcal{Q}) \leq \frac{1}{2}$.

□

2.14.2 Semi-algebraic subclasses

Let us specialize the definitions of Section 1.9.2 for graphs. If $G = (V, E)$ is a graph, a vertex partition $V = V_1 \cup \cdots \cup V_r$ is *equitable* if $||V_i| - |V_j|| \leq 1$ for every $i, j \in [r]$. Two parts V_i and V_j with $1 \leq i, j \leq r$ are *complete* if every pair uv of distinct vertices with $u \in V_i$ and $v \in V_j$ is an edge of G and *empty* if no pair uv with $u \in V_i$ and $v \in V_j$ is an edge of G. The parts V_i and V_j are *homogeneous* if they are either complete or empty.

For a decreasing function $m : \mathbf{R}_*^+ \to \mathbf{N}$, a graph property $\mathcal{P}$ has a *random-free regularity lemma* with m parts if for every $\epsilon > 0$, every graph $G = (V, E)$ of $\mathcal{P}$ has an equitable partition $V_1 \cup \cdots \cup V_r$ with $r \leq m(\epsilon)$ such that all but at most ϵr^2 pairs (V_i, V_j) are homogeneous.

A result of Fox, Pach and Suk [18] exposed in Section 1.9.2 (Theorem 1.24) shows that semi-algebraic classes of graphs, such as intersection graphs of geometrical objects have this property have a random-free regularity lemma with $\epsilon \mapsto P(1/\epsilon)$ parts, where P is a polynomial. (See Section 1.9.1 for more details on semi-algebraic structures).

Given a graph F on vertex set $[r]$ and a set V with a partition $V = V_1 \cup \cdots \cup V_r$, the *blow-up* B of F on V is the graph on V such that for $u \in V_i$ and $v \in V_j$, uv is an edge of B if and only if ij is an edge of F. An *extension* of B is a graph G on V that agrees with B on every $V_i \times V_j$ with $i \neq j$. A property $\mathcal{P}$ is *extendable* if for every graph $F = (V, E) \in \mathcal{P}$ and every vertex $v \in V$, at least one of the two graphs obtained from F by adding a vertex v' with the same neighborhood as v and adding the edge vv' or not has property $\mathcal{P}$. By iterating the definition, it is easy to see that if $\mathcal{P}$ is extendable then for every $F \in \mathcal{P}$, every blow-up of F has an extension in $\mathcal{P}$.

Property 2.38. *Let $\mathcal{Q}$ be a hereditary graph property with a strong regularity lemma with constant $m(\epsilon)$. Then every extendable and hereditary property $\mathcal{P} \subseteq \mathcal{Q}$ is $(\epsilon, q(\epsilon))$-testable in $\mathcal{Q}_{\eta(\epsilon)}$, where*

- $m'(\epsilon) = \max(m(\epsilon), \frac{8}{\epsilon})$;
- $q(\epsilon) = m'(\epsilon)\ln(3m'(\epsilon))$; *and*
- $\eta(\epsilon) = \frac{1}{4(m'(\epsilon))^2}$

for every $0 < \epsilon < 1/2$.

Proof. For the sake of readability, we write m, m', η and q instead of $m(\epsilon)$, $m'(\epsilon)$, $\eta(\epsilon)$ and $q(\epsilon)$ respectively.

It suffices to show that for every graph $G = (V, E)$ with $d(\mathcal{Q}, G) \leq \eta$ and $d(\mathcal{P}, G) \geq \epsilon$, the subgraph $G[X]$ induced by the set X taken uniformly at random among all subsets of V of size q has property $\mathcal{P}$ with probability less than $\frac{1}{2}$. Let G' be a graph of $\mathcal{Q}$ satisfying $d(G, G') \leq \eta$. The strong regularity lemma of $\mathcal{Q}$ applied to G' gives a balanced partition $V = V_1 \cup \cdots \cup V_r$ with $r \leq m = m(\epsilon)$ parts such that all but an ϵ-fraction of pairs (V_i, V_j) are homogeneous in G'.

We have to take care of the case where m is too small. Let $m' = \max(m, \frac{8}{\epsilon})$, up to splitting each part into two parts of equitable size as many times as necessary, we may assume that $\frac{1}{2}m' \leq r \leq m'$.

Let $\mathcal{F}$ be the set of graphs F with vertex set $[r]$ such that $ij \in E(F)$ if the pair (V_i, V_j) is complete and $ij \notin E(F)$ if (V_i, V_j) is empty in G'. There is no constraint for pairs ij with (V_i, V_j) not homogeneous (We only know that there is an ϵ-fraction of them) .

To prove the theorem, it suffices to show that $\mathcal{F} \cap \mathcal{P} = \varnothing$ and that with probability at least $\frac{1}{2}$, the graph $G[X]$ has an induced subgraph F that belongs to $\mathcal{F}$. Indeed, if this last event occurs, then $G[X] \notin \mathcal{P}$ since $\mathcal{P}$ is hereditary.

We first prove by contradiction that $\mathcal{F}$ and $\mathcal{P}$ do not intersect. Suppose on the contrary that there is a graph F in $\mathcal{F} \cap P$. Since $\mathcal{P}$ is extendable, the blow-up of F on $V_1 \cup \cdots \cup V_r$ has an extension H, i.e. a graph on V such that (V_i, V_j) is complete in H if $ij \in E(F)$ and (V_i, V_j) is empty otherwise for every $i \neq j$. The graphs G' and H only differ inside the parts V_i and in an ϵ-fraction of the $\binom{r}{2}$ pairs (V_i, V_j). Consequently,

$$d(H, G') \leq \frac{1}{n^2}\left(\epsilon\binom{r}{2} + r\right) \cdot \frac{n^2}{r^2} \leq \frac{\epsilon}{2} + \frac{1}{r}.$$

Using the triangular inequality and that $\frac{1}{r} \leq \frac{2}{m'} \leq \frac{\epsilon}{4}$ and $\eta = \frac{1}{4(m')^2} \leq \frac{\epsilon^2}{256}$, it follows that $d(\mathcal{P}, H) \leq d(H, G) \leq d(H, G') + d(G', G) \leq \frac{\epsilon}{2} + \frac{2}{m'} + \eta \leq$

$\frac{\epsilon}{2} + \frac{\epsilon}{4} + \frac{\epsilon^2}{256} < \epsilon$, which gives a contradiction with the hypothesis that G is ϵ-far from $\mathcal{P}$.

It remains to prove that $G[X]$ has an induced subgraph in $\mathcal{F}$ with probability at least $\frac{1}{2}$. We first estimate the probability that X contains a vertex in V_i for each i. For a fixed $i \in [r]$, the probability that X does not intersect V_i is at most $(1 - \frac{|V_i|}{|V|})^q \leq (1 - \frac{1}{m'})^q \leq e^{-q/m'}$. Taking the union bound over all i, the probability that there exists $i \in [r]$ such that X does not intersect every V_i is at most $re^{-q/m'} \leq m'e^{-q/m'} = \frac{1}{3}$ since $q = m' \ln(3m')$. Hence, X hits all V_i's with probability $\frac{2}{3}$.

Conditioning on this event, we take y_i uniformly at random in $X \cap V_i$ for each $1 \leq i \leq r$ and we estimate the probability that $G[Y] = G'[Y]$, where $Y = \{ y_i \mid 1 \leq i \leq r \}$. Note that $G[Y] = G'[Y]$ if and only if no edge of $E(G) \triangle E(G')$ with both end vertices in Y. The probability that there is such an edge is bounded by the expected number of them, which is $\eta r^2 \leq \eta (m')^2 \leq \frac{1}{4}$ because $\eta \leq \frac{1}{4(m')^2}$. It follows that $\mathbf{P}(G[Y] \neq G'[Y]) \geq 1 - \frac{1}{4} = \frac{3}{4}$. Consequently, $G[X]$ has an induced subgraph $F = G[Y]$ that belongs to $\mathcal{F}$ is at least $\frac{2}{3} \cdot \frac{3}{4} = \frac{1}{2}$. □

Properties 2.37 and 2.38 implies that if a semi-algebraic class of graphs $\mathcal{P}$ is easily testable, then every extendable subclass $\mathcal{Q} \subseteq \mathcal{P}$ is also easily testable.

In particular, we prove starting on the next section that the class of interval graphs is easily testable (Theorem 2.12). This leads to the following corollary.

Corollary 2.39. *Every extendable subclass of interval graphs is easily testable.*

Proof. Let $\mathcal{Q}$ be the class of interval graph and $\mathcal{P}$ be an extendable subclass of $\mathcal{Q}$. By Theorem 2.12, there is a polynomial P_1 such that $\mathcal{Q}$ is $(\epsilon, P_1(1/\epsilon))$-testable (in the set of all graphs). By Theorem 1.24, there is a polynomial P_2 such that $\mathcal{Q}$ has a random-free regularity lemma with $P_2(1/\epsilon)$ parts. Further, by Properties 2.38 and 2.37, $\mathcal{P}$ is easily testable. □

This corollary gives another proof that the class of trivially perfect graphs and the class of threshold graphs are easily testable.

Corollary 2.39 is covered by the earlier mentioned recent result of Gishboliner and Shapira [20]. Interestingly, the most general form of their result [20, Theorem 5] states that a class of graph is easily testable if it has bounded VC-dimension and if it has another property they call that *blowup quality*, which is a weaker –i.e more general– version of being extendable.

2.15 Open questions

Chordal graphs

Among the classes of graphs for which whether they are easily testable is not known, the class of chordal graph is of particular interest since it is both a natural extension of interval graphs with unbounded VC-dimension and a subclass of C_4-FREE.

Question 2.2. Is the class of chordal graphs easily testable?

Chordal graphs are intersection graphs of subtrees of a tree, hence we call a *representation* of a chordal graph G a family of trees that are subsets of a common tree and whose intersection graph is G. A possible strategy to solve this question is to apply the technique we use for interval graphs. Given a chordal graph H and an induced subgraphs G of H, let $\mathcal{R}_G(H)$ be the set of representations of G that can be completed in a representation of H. Then if

1. an equivalent of Lemma 2.17 holds for chordal graphs with this definition with a polynomial bound on the number of specializations of the set of representations; and

2. testing if a representation of a small subgraph extends to the whole graphs up to some error can be done with polynomial query complexity

then one could deduce that chordal graphs are easily testable. The elements of the set $\mathcal{R}_G(H)$ may need to be considered modulo some symmetries and restricted to a particular set of representations to make 1 work.

Testability of intersections

It is very to see that if $\mathcal{P}_1$ is testable with query complexity s_1 and $\mathcal{P}_2$ is testable with query complexity s_2 then their union $\mathcal{P}_1 \cup \mathcal{P}_2$ is testable with query complexity at most $2\max(s_1, s_2)$. This comes from the fact that if G is ϵ-far from $\mathcal{P}_1 \cup \mathcal{P}_2$ then G is ϵ-far from both of $\mathcal{P}_1$ and $\mathcal{P}_2$.

Interestingly, no such relation is known for the intersection of two classes. The problem that prevent a naive proof of such a relation is that a graph ϵ-close to both $\mathcal{P}_1$ and $\mathcal{P}_2$ could still be very far from $\mathcal{P}_1 \cap \mathcal{P}_2$. This leads to the following question.

Question 2.3. If $\mathcal{P}_1$ and $\mathcal{P}_2$ are easily testable graph classes, can we prove that the class $\mathcal{P}_1 \cap \mathcal{P}_2$ is easily testable?

Extension of semi-algebraic graphs

As previously mentioned, Gishboliner and Shapira showed that every semi-algebraic class of graphs is easily testable [19]. They proved actually a more general result whose essential (but not sufficient) hypothesis is that the class considered has bounded VC-dimension. The class of k-colorable graphs is a non-trivial example of class known to be easily testable on which this theorem does not apply. Indeed, this class has a lot of freedom (the graph between the color classes is arbitrary), and in particular has unbounded VC-dimension, but it still has a simple structure. Let us define a family of classes that covers both behaviors of semi-algebraic and k-colorable graphs.

Definition 2.12. A graph property $\mathcal{P}$ is *nearly semi-algebraic* if there is an integer n, a sequence of polynomials $f_1, \ldots, f_k$ on the $2n$ coordinates of $\mathbf{R}^n \times \mathbf{R}^n$ and a function $\psi : \{T, F\}^k \to \{T, F, ?\}$ such that the following holds. The graph

$G = (V, E)$ belongs to $\mathcal{P}$ if and only if there exists a vector $x_v \in \mathbf{R}^n$ for each $v \in V$ such that for every $u, v \in V$, if

$$\psi(f_1(x_u, x_v) \geq 0, \ldots, f_k(x_u, x_v) \geq 0) = \begin{cases} T & \text{then } uv \in E \\ F & \text{then } uv \notin E \end{cases}$$

There is no constraint if this value is ?, i.e. $uv \in E$ or $uv \notin E$.

This leads to the following question.

Question 2.4. Are nearly semi-algebraic graphs properties easily testable?

This family contains the semi-algebraic classes, the graph partition properties and, for instance, the class of graphs that can be partitioned into two interval graphs.

Chapter 3

Implementation of Flag algebras

3.1 Introduction

The theory of flag algebras is a framework introduced by Alexander Razborov [40] in 2007 to derive and manipulate inequalities on densities of some sub-structures. Among others, it formalizes the use of some type of double counting and averaging arguments, which are very often used in combinatorics. It is closely correlated with the notion of graphs limits as described by László Lovász [32]. This framework is very flexible and permits to manipulate densities in different kinds of structures like graphs, hypergraphs, permutations or order types. The most important feature of the flag algebra framework is the semi-definite method, where semi-definite programming is used to automatize the search of an extremal density under a set of constraints.

In this chapter, we first introduce flag algebras and the semi-definite method in Section 3.3. In Section 3.4, we detail how to apply this method to the Caccetta-Häggkvist conjecture, following a work of Jan Hladký, Daniel Král and Sergey Norin [27]. In Section 3.5, I present a program I wrote to manipulate flag algebras.

3.1.1 Flags

The main purpose the flag algebra framework is to manipulate densities. The *density* of a graph G with k vertices in another graph H with $n \geq k$ vertices is the probability $p(H, G)$ that a random set $X \subseteq V(H)$ chosen uniformly among the $\binom{n}{k}$ subsets of $V(H)$ of size k induces a graph isomorphic to G. For instance, the density of the single edge K_2 in a graph $G = (V, E)$ is the ratio $|E|/\binom{|V|}{2}$.

The notion of density has a meaning for a wide class of objects in combinatorics. The theory of flag algebras, is centered in this notion of density. In the seminal paper [40], the generic objects for which the density is defined (and thus flag algebras apply), that are called *flags*, are defined in term of model theory as follows. Choose a first order theory $\mathcal{T}$ over a language without function (so without constant) and without quantifiers, which means that an axioms of $\mathcal{T}$ consists of predicates over free variables (that are implicitly universally quantified) combined by some logical functions. A *flag* is then a finite model G of $\mathcal{T}$, that is a finite set V, that we call the *vertex set*, with a truth table of predicates satisfying the axioms.

As an example, a theory of triangle-free graphs would have a single binary predicate $e(.,.)$ for the edge relation and the three following axioms.

(A_1) $e(x,y) \Rightarrow e(y,x)$ Edges are symmetrical

(A_2) $\neg e(x,x)$ There is no loop

(A_3) $\neg(e(x,y) \wedge e(y,z) \wedge e(z,x))$ There is no triangle

A finite model G of this theory is a finite set V and a relation $e_G \subseteq V^2$ satisfying the above axioms, i.e. a triangle-free graph, as expected.

If G is a model and $\phi : V(G) \to V'$ is a bijection, then G^ϕ is the model on vertex set V' obtained from G by renaming every vertex $v \in V(G)$ by $\phi(v)$. The function ϕ is an *isomorphism* (similarly to graphs). Flags are considered modulo isomorphisms, that is that G and G^ϕ are considered as equal.

We can assume that V is a set of integers. The very property that is needed from these structures is some notion of induced substructure. For a subset $U \subseteq V$ of vertices, the *subflag of G induced by U* is the model $G[U]$ obtained by removing all vertices of $V \backslash U$ in G. All flags in $\mathcal{F}$ are considered up to isomorphism.

The definition of Razborov is concise and captures well the generality of flags. For readers without background in logics, we give an alternative definition of flags, in the form of an axiomatization, intended to emphasize the properties we need for flags, though it may looks more complicated. This definition will help us to understand how a generic program dealing with flag algebras, like the one I implemented, can work.

Definition 3.1. A *set of flags* is a set $\mathcal{F}$ with the following elements

1. every $F \in \mathcal{F}$ has a finite *vertex set* $V(F)$;
2. for every $X \subseteq V(F)$, there is an element $F[X] \in \mathcal{F}$ with vertex set X, this element is called an *induced sub-flag of F*; and
3. for every bijection $\phi : V(F) \to V'$, there is a flag F^ϕ with $V(F^\phi) = V'$.

that satisfy the following few consistency properties

1. $F[V(F)] = F$;
2. $F[X] = (F[Y])[X]$ whenever $X \subseteq Y \subseteq V(F)$;
3. $(\phi, F) \mapsto F^\phi$ is a group action, i.e. $(F^\phi)^\psi = F^{\psi\circ\phi}$ and $F^{\mathrm{id}_{V(F)}} = F$; and

4. $F^{\phi}[\phi(X)] = F[X]$ for every $F \in \mathcal{F}$, $X \subseteq V(F)$ and $\phi : V(F) \to V'$.

A *flag* is then an element of $F \in \mathcal{F}$. The *size* of a flag F is the cardinal of $V(F)$ and is written $|F|$.

If $n \in \mathbf{N}$, let $\mathcal{F}_n$ be the set of flags of size n. We always assume that $\mathcal{F}_n$ is finite for every $n \in \mathbf{N}$.

As examples, the set $\mathcal{F}$ can be one of the following combinatorial objects.

- the set of every (finite) graphs;
- more generally, any hereditary class of graphs, for instance triangle-free graphs;
- the set of oriented graphs;
- the set of k-regular hypergraphs;
- the set of permutations;
- the set of order types (see Chapter 1); or
- any superposition of the components above, for instance, graphs with both symmetric and directed edges where each vertex is given a color.

The notion of density is defined on flags similarly to the particular case of graphs.

Definition 3.2. The *density* of a flag $F \in \mathcal{F}$ of size k in a flag $G \in \mathcal{F}$ of size $n \geq k$ is the probability for a random subset X of $V(G)$ of size k that $G[X]$ is isomorphic to F.

Equivalently, $p(F, G)$ can be understood as the normalized number of copies of F in G:

$$p(F, G) = \frac{1}{\binom{n}{k}} \cdot |\{X \subseteq V(G) \text{ of size } k \text{ such that } G[X] = F\}| \,.$$

If F and G are flags with $|F| > |G|$, we adopt the convention that $p(F, G) = 0$.

3.2 Convergent sequences of flags

The convergence of a sequence of flags is defined as the convergence in terms of densities.

Definition 3.3. A sequence $(G_n)_{n \in \mathbf{N}}$ of flags whose size tends to infinity is *convergent* if for every flag F, the real number sequence $(p(F, G_n))_{n \in \mathbf{N}}$ is convergent.

The *limit* of a converging sequence $(G_n)_{n\in\mathbf{N}}$ is the function

$$\ell : \left\{ \begin{array}{ccc} \mathcal{F} & \to & [0,1] \\ F & \mapsto & \lim_{n\to\infty} p(F,G_n) \end{array} \right.$$

As a simple example, if $\mathcal{F}$ is the set of graphs, then the sequence of cliques $(K_n)_{n\in\mathbf{N}}$ is convergent since $p(F,K_n)$ is equal to 1 whenever $|F| \le n$ if F is a clique and is equal to 0 otherwise. A compactness argument shows that every sequence $(G_n)_{n\in\mathbf{N}}$ whose size tends to infinity has a converging subsequence, which in particular shows the existence of a wide variety of limits. Indeed, a limit $\ell : \mathcal{F} \to [0,1]$ can be seen as an element of $[0,1]^{\mathcal{F}}$, which is compact by Tychonoff's theorem.

3.3 Flag algebra

3.3.1 Linear combinations of flags

Let $\mathbf{R}\mathcal{F}$ be the set of all formal linear combinations of flags provided with its natural vector space structure. An element $\sum_{i=1}^{k} \lambda_i F_i$ of $\mathbf{R}\mathcal{F}$ represents a linear combination of densities of the involved flags $(F_i)_{i=1}^{k}$.

A limit of flags $\ell : \mathcal{F} \to [0,1]$ can be extended into a linear mapping ℓ from $\mathbf{R}\mathcal{F}$ to $\mathbf{R}$ defined by

$$\ell\left(\sum_{i=1}^{m} \lambda_i F_i\right) = \sum_{i=1}^{m} \lambda_i \ell(F_i)$$

for every $\sum_{i=1}^{m} \lambda_i F_i \in \mathbf{R}\mathcal{F}$.

Let ℓ be the limit of a sequence $(G_n)_{n\in\mathbf{N}}$ of flags. Let k and m be integers with $k \le m$ and let n be large enough so that $m \le |G_n|$. For every $F \in \mathcal{F}_k$, the density of F in G_n can be expressed in terms of the densities $p(H,G_n)$ of the flags H of size m using the following relation.

$$p(F,G_n) = \sum_{H\in\mathcal{F}_m} p(F,H)p(H,G_n). \tag{3.1}$$

Indeed, (3.1) follows from a simple conditioning argument. Note that first picking a random subset $Y \subseteq V(G_n)$ of size m uniformly at random and subsampling a random subset $X \subseteq Y$ of size k uniformly gives a random variable subset X distributed uniformly among subsets of $V(G_n)$ of size k. The left side of (3.1) is therefore $p(F,G_n) = \mathbf{P}(G_n[X] = F)$, and the right side can be rewritten as $\sum_{H\in\mathcal{F}_m} \mathbf{P}(G_n[X] = F | G_n[Y] = H)\mathbf{P}(G_n[Y] = H)$. Equation (3.1) thus follows from the conditioning formula.

Letting n go to infinity in (3.1), it holds that

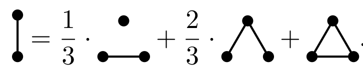
$$\ell(F) = \sum_{H\in\mathcal{F}_m} p(F,H)\ell(H). \tag{3.2}$$

Since Equation (3.2) holds for every flag limit ℓ we identify a flag and its projections in $\mathbf{R}\mathcal{F}$. Let $\mathcal{K}$ be the subspace of $\mathbf{R}\mathcal{F}$ generated by

$$\left\{ \sum_{H \in \mathcal{F}_m} p(F,H)H - F \;\middle|\; F \in \mathcal{F}_k \text{ and } k \leq m \right\}. \tag{3.3}$$

Consider the quotient space $\mathcal{A} = \mathbf{R}\mathcal{F}/\mathcal{K}$. This set $\mathcal{A}$ is called a *flag algebra*. We later define a product for this algebra. Equation (3.2) shows that a limit ℓ is still well defined on $\mathcal{A}$.

Example 3.1. If $\mathcal{F}$ is the set of graphs, it holds in $\mathcal{A}$ that



3.3.2 Product

We aim to define a product on the flag algebra $\mathcal{A}$. This product is meant to satisfy $\ell(f_1 \cdot f_2) = \ell(f_1)\ell(f_2)$ for every $f_1, f_2 \in \mathcal{A}$.

If $F_1 \in \mathcal{F}_{k_1}$, $F_2 \in \mathcal{F}_{k_2}$ and $G \in \mathcal{F}_{k_1+k_2}$, the *split probability* of F_1 and F_2 in G is the probability for a random partition of $V(G)$ into two parts X_1 and X_2 or respective size k_1 and k_2 that $G[X_1]$ is isomorphic to F_1 and $G[X_2]$ is isomorphic to F_2.

Note that the value $\alpha = p(F_1, G)p(F_2, G)$ is the probability for two random subsets $X_1, X_2 \subseteq V(G)$ of respective sizes k_1 and k_2 chosen uniformly and independently at random that $F_1 = G[X_1]$ and $F_2 = G[X_2]$. If now we condition X_1 and X_2 on the event that $X_1 \cap X_2 = \varnothing$, then the experience above is equivalent to first taking a random subset X of $V(G)$ with size $k_1 + k_2$ and then taking a random partition $X_1 \cup X_2$ of X with $X_i = k_i$ for $i \in \{1, 2\}$. The probability that $F_1 = G[X_1]$ and $F_2 = G[X_2]$ knowing $X_1 \cap X_2 = \varnothing$ is therefore

$$\beta = \sum_{H \in \mathcal{F}_{k_1+k_2}} p(F_1, F_2; H)p(H, G).$$

Note that $\mathbf{P}(X_1 \cap X_2 \neq \varnothing) \leq \mathbf{E}(|X_1 \cap X_2|) = \frac{k_1 \cdot k_2}{|G|}$, so

$$|\alpha - \beta| \leq \frac{k_1 \cdot k_2}{|G|}.$$

Let $(G_n)_{n \in \mathbf{N}}$ be a sequence of flags converging to ℓ, we know that $\frac{k_1 \cdot k_2}{|G_n|}$ tends to 0, so

$$\ell(F_1)\ell(F_2) = \sum_{H \in \mathcal{F}_{k_1+k_2}} p(F_1, F_2; H)\ell(H). \tag{3.4}$$

The product of the flag algebra $\mathcal{A}$ is then defined as

$$F_1 \cdot F_2 := \sum_{H \in \mathcal{F}_{k_1+k_2}} p(F_1, F_2; H)H,$$

and extended by bilinearity to $\mathcal{A}$ as follows

$$\left(\sum_{i=1}^{p} \lambda_i F_i\right) \cdot \left(\sum_{i=1}^{q} \mu_i G_i\right) = \sum_{i=1}^{p} \sum_{j=1}^{q} \lambda_i \mu_j F_i \cdot G_j$$

for every pair of elements of $\mathcal{A}$. By Equation (3.4), it holds that $\ell(F_1)\ell(F_2) = \ell(F_1 \cdot F_2)$ for every $F_1, F_2 \in \mathcal{F}$. Further, $\ell(f_1)\ell(f_2) = \ell(f_1 \cdot f_2)$ for every $f_1, f_2 \in \mathcal{A}$.

Flag limits are therefore algebra homomorphisms. This is actually an equivalence. Indeed, let $\mathrm{Hom}^+(\mathcal{A}, \mathbf{R})$ be the set of linear functions $\ell : \mathcal{A} \to \mathbf{R}$ such that $\ell(f \cdot g) = \ell(f)\ell(g)$ for every $f, g \in \mathcal{A}$ and $\ell(F) \geq 0$ for every $F \in \mathcal{F}$. The following holds.

Theorem 3.1. ([40, Theorem 3.3], [34]) A function $\ell : \mathcal{A} \to \mathbf{R}$ is an element of $\mathrm{Hom}^+(\mathcal{A}, \mathbf{R})$ if and only if there is a sequence of flags $(G_n)_{n \in \mathbf{N}}$ converging to ℓ.

3.3.3 Rooted flags

The effectiveness of flag algebra is improved by considering partially labeled flags. Rooted flags make it possible to use arguments of the form: fix some vertices in a flag, then count something related to these vertices and derive some inequality, and finally take the average on every suitable choice of the fixed vertices to obtain a general inequality on the structure. See 3.5.

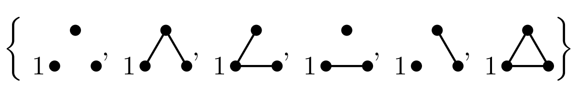
Let us extend the construction of the previous part to flags where some particular labeled subflag σ is fixed. For each such σ, this will define a new sigma-algebra $\mathcal{A}^\sigma$ with only linear combinations of flags rooted on σ.

A *rooted flag* $F = (F_0, \theta)$ is a flag F_0 together with an injective map $\theta : [k] \to V(F_0)$. In other words, it is a flag with k labeled vertices, where $\theta(i)$ is labeled by i, for every $i \in [k]$.

The *type* of such a flag is the rooted flag $(F_0[\mathrm{Im}\,\theta], \theta)$ obtained by keeping only the labeled vertices. An *isomorphism* of rooted flags between (F_0, θ_F) and (G_0, θ_G) is a flag isomorphism $\phi : V(F_0) \to V(G_0)$ that preserves the labels, i.e. such that $F_0^\phi = G_0$ and $\phi \circ \theta_F = \theta_G$. Rooted flags are considered modulo isomorphisms.

A rooted flag with type σ is called a *σ-flag*. Let $\mathcal{F}^\sigma$ be the set of all σ-flags and $\mathcal{F}^\sigma_n$ the set of all σ-flags of size n.

Example 3.2. In Graph Theory, the set of the rooted flags of size 3 with type the labeled point $\sigma = \bullet_1$ is

$$\mathcal{F}_3^\sigma =$$


If $\mathrm{Im}\,\theta \subseteq U \subseteq V(F_0)$, the *induced subflag* $F[U]$ is the rooted flag $(F_0[U], \theta)$. Note that U must contain $\mathrm{Im}\,\theta$ so that $F[U]$ has type σ. The density function p is defined in $\mathcal{F}^\sigma$ as in $\mathcal{F}$ except that the (common) type σ of every flag always has to be mapped to while looking at isomorphisms.

The *density* $p(F, G)$ of a σ-flag $F = (F_0, \theta_F)$ into another σ-flag $G = (G_0, \theta_G)$ of larger or equal size, is the probability for a random a subset $X \subseteq V(G)$ chosen uniformly at random among the $\binom{|G|-|\sigma|}{|F|-|\sigma|}$ subsets of $V(G)$ containing $\operatorname{Im}\theta$ to induce a rooted subflag $G[U]$ isomorphic to G.

The flag algebra $\mathcal{A}^\sigma$ of σ-flags is then the set $\mathbf{R}\mathcal{F}^\sigma$ of formal linear combinations of σ-flags quotiented by the subspace $\mathcal{K}^\sigma$ generated by

$$\left\{ \sum_{H \in \mathcal{F}_m^\sigma} p(F, H)H - F \;\middle|\; F \in \mathcal{F}_k^\sigma \text{ and } k \leq m \right\}.$$

The *split probability* of $F = (F_0, \theta_F) \in \mathcal{F}_k^\sigma$ and $F' = (F'_0, \theta'_F) \in \mathcal{F}_{k'}^\sigma$ in another σ-flag $G = (G_0, \theta_G) \in \mathcal{F}^\sigma$ of size $k + k' - |\sigma|$ is the probability for a random partition of $V(G) \setminus \operatorname{Im}\theta_G$ into two sets Y and Y' of respective size $k - |\sigma|$ and $k' - |\sigma|$ that the sets $X := Y \cup \operatorname{Im}\theta_G$ and $X' := Y' \cup \operatorname{Im}\theta_G$ induce σ-flags $G[X]$ and $G[X']$ respectively isomorphic to F and F'.

The algebra $\mathcal{A}^\sigma$ is endowed by its natural addition and the product defined on σ-flags by

$$F \cdot G \quad = \sum_{H \in \mathcal{F}_{m+n-|\sigma|}^\sigma} p(F, G; H)H$$

for every $F \in \mathcal{F}_n$ and $G \in \mathcal{F}_m$ and extended by linearity to $\mathcal{A}^\sigma$.

3.3.4 Rooted homomorphism

A sequence of $(G_n)_{n\in\mathbf{N}} \in (\mathcal{F}^\sigma)^{\mathbf{N}}$ of σ-flags whose size tends to infinity *converges* if the real number sequence $(p(F, G_n))_{n\in\mathbf{N}}$ converges for every σ-flag F. The *limit* of such a sequence is the function

$$\ell^\sigma : \begin{cases} \mathcal{F}^\sigma & \to \quad [0, 1] \\ \quad F & \mapsto \quad \lim_{n\to\infty} p(F, G_n) \end{cases}$$

This function ℓ^σ is an algebra homomorphism. Let $\operatorname{Hom}^+(\mathcal{A}^\sigma, \mathbf{R})$ be the set of limits of σ-flags. Equivalently, $\operatorname{Hom}^+(\mathcal{A}^\sigma, \mathbf{R})$ is the set of linear functions $\ell^\sigma : \mathcal{A}^\sigma \to \mathbf{R}$ satisfying $\ell^\sigma(F) \geq 0$ for every $F \in \mathcal{F}^\sigma$ and $\ell^\sigma(f \cdot g) = \ell^\sigma(f)\ell^\sigma(g)$ for every $f, g \in \mathcal{A}^\sigma$ [40, Theorem 3.3].

3.3.5 Averaging operation

A standard method to obtain inequalities in extremal combinatorics is to count something around a some particular structure and take the average over all possible choices of this structure. The unlabeling operation is a tool that formalizes this argument. It consists in unlabeling the type σ of a flag to have a non-rooted flag and multiplying by some normalization factor.

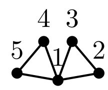
If σ is a type of size k, the *averaging operator*

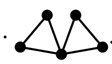
$$[\![\cdot]\!]_\sigma : \mathcal{A}^\sigma \to \mathcal{A}.$$

is defined as follows. For a σ-flag $F = (F_0, \theta) \in A^\sigma$, let $q_\sigma(F)$ be the probability that a function θ', chosen uniformly at random among the $\frac{|F|!}{(|F|-k)!}$ injections from $[k]$ to $V(F_0)$ induces a typed flag (F_0, θ') isomorphic to F (so in particular with type σ). The averaging operator is defined by

$$[\![F]\!]_\sigma = q_\sigma(F).F_0$$

and extended by linearity to $\mathcal{A}^\sigma$.

Example 3.3. If $\sigma =$ 

, then $[\![\sigma]\!]_\sigma = \frac{1}{15} \cdot$ .

Note that the rooted flag $\sigma = (\sigma_0, \theta)$, where θ is a bijection from $[|\sigma|]$ to $V(\sigma_0)$, is a σ-flag. We have $[\![\sigma]\!]_\sigma = q_\sigma(\sigma).\sigma_0$ and $q_\sigma(\sigma)$ is equal to the number of automorphisms of σ_0, normalized by $|\sigma_0|!$.

3.3.6 Random homomorphism

Let G be a non-labeled flag and σ a type of size $k \leq |G|$ such that $p([\![\sigma]\!]_\sigma, G) > 0$.

Consider a random map $\theta : [k] \to V(G)$ chosen uniformly among all injections such that the rooted flag (G, θ) has type σ. This associates to G a random σ-flag $G^\sigma = (G, \theta)$.

If $F = (F_0, \theta)$ is a σ-flag satisfying $|F_0| \leq |G|$, then the value $p([\![F]\!]_\sigma, G) = q_\sigma(F)p(F_0, G)$ is equal to the probability, for a random set of vertices $U \subseteq V(G)$ of size k and an injection $\theta' : [|\sigma|] \to U$ chosen uniformly at random, that $(G[U], \theta)$ is a σ-flag isomorphic to F. It is also the probability, if you first choose the injection $\theta : [|\sigma|] \to V(G)$ at random and then a random U of size k such that $\operatorname{Im}\theta \subseteq U \subseteq V(G)$, that $(G[U], \theta) = F$. Note that the probability $\mathbf{P}(G[U] = F | (G, \theta)$ has type $\sigma)$ on U and θ that $G[U]$ is isomorphic to F knowing that (G, θ) has type σ is equal to the average value $\mathbf{E}(p(F, G^\sigma))$ over the choice of θ of the density of F in $G^\sigma = (G, \theta)$. It follows that

$$p([\![F]\!]_\sigma, G) = p([\![\sigma]\!]_\sigma, G)\mathbf{E}(p(F, G^\sigma)). \tag{3.5}$$

If $(G_n)_{n\in\mathbf{N}}$ is a flag sequence converging to a limit ℓ such that $\ell(\sigma) > 0$, it turns out that the sequence of random rooted flags $(G_n^\sigma)_{n\in\mathbf{N}}$ also converges almost surely to a random homomorphism ℓ^σ satisfying Equation 3.5, taken at the limit.

Theorem 3.2 ([40])**.** *The random sequence (G_n^σ) converges to a random homomorphism $\ell^\sigma \in \operatorname{Hom}(\mathcal{A}^\sigma, \mathbf{R})$ such that*

$$\mathbf{E}(\ell^\sigma(F)) = \frac{\ell([\![F]\!]_\sigma)}{\ell([\![\sigma]\!]_\sigma)}.$$

In particular, assume that for every possible value of G_n^σ (i.e. every flag obtained by rooting G_n in an induced copy of σ), some inequality of the form

$$\sum_{i=1}^{k} \lambda_i p(F_i, G_n^\sigma) \geq o(1)$$

holds, where $F_i \in \mathcal{F}_i^\sigma$ for every $i \in [k]$. We can deduce that

$$\sum_{i=1}^{k} \lambda_i \ell^\sigma(F_i) \geq 0$$

with probability 1. By Theorem 3.2,

$$\sum_{i=1}^{k} \lambda_i \ell(\llbracket F_i \rrbracket_\sigma) \ell(\llbracket \sigma \rrbracket_\sigma) \geq 0.$$

Since $\ell(\llbracket \sigma \rrbracket_\sigma) \geq 0$, it follows that

$$\ell\left(\sum_{i=1}^{k} \lambda_i \llbracket F_i \rrbracket_\sigma\right) \geq 0.$$

This is formalized by the following result.

Corollary 3.3 ([40])**.** *Let σ be a type a $\ell \in \mathrm{Hom}^+(\mathcal{A}^\sigma, \mathbf{R})$ be a limit of flags and $f \in \mathcal{A}^\sigma$. Assume that $\ell^\sigma(f) \geq 0$ with probability 1. Then*

$$\ell(\llbracket f \rrbracket_\sigma) \geq 0.$$

3.3.7 The Cauchy-Schwarz inequality

A particularly important case of Corollary 3.3 is when f is a square. Indeed, for every linear combination of flags $g \in \mathcal{A}^\sigma$, the square $f = g^2$ (i.e. $f = g \cdot g$ for the product of $\mathcal{A}^\sigma$) satisfies

$$\ell^\sigma(g^2) = (\ell^\sigma(g))^2 \geq 0$$

for every algebra homomorphism $\ell^\sigma \in \mathrm{Hom}^+(\mathcal{A}^\sigma, \mathbf{R})$. It then follows from Corollary 3.3 that

$$\ell(\llbracket f^2 \rrbracket) \geq 0 \tag{3.6}$$

for every $\ell \in \mathrm{Hom}^+(\mathcal{A}, \mathbf{R})$.

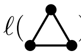
These inequalities can be used to generate a large number of non-trivial relations between densities, which makes them essential ingredients of the semi-definite method. This inequality is known in the literature as the *Cauchy-Schwarz* inequality since it is a consequence of the Cauchy-Schwarz inequality applied to the scalar product $\langle f, g \rangle = \mathbf{E}(\ell^\sigma(f \cdot g))$. Indeed,

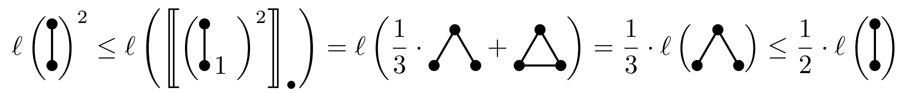
$$\ell(\llbracket f^2 \rrbracket_\sigma) = \mathbf{E}(\ell^\sigma(f^2)) = \|f\|^2 \geq \langle f, 1 \rangle^2 = \mathbf{E}(\ell^\sigma(f))^2 = \ell(\llbracket f \rrbracket_\sigma)^2 \geq 0. \tag{3.7}$$

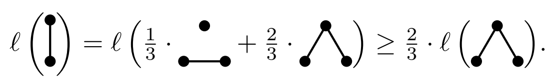
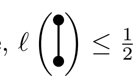
Note that the above inequality $\ell(\llbracket f^2 \rrbracket_\sigma) \geq \ell(\llbracket f \rrbracket_\sigma)^2$ is stronger than Relation (3.6) whenever $\ell(\llbracket f \rrbracket_\sigma)^2$ is positive.

However, if f is a linear combination of σ-flags of size up to n, then $\ell(\llbracket f \rrbracket_\sigma^2)$ is a linear combination of flags of size up to $2|n|$, while $\llbracket f^2 \rrbracket_\sigma$ can be expressed with flags of size (at most) $2|n| - |\sigma|$. As a consequence, only form (3.6) is used.

We give an example of application of the Cauchy-Schwarz inequality.

Example 3.4 (Asymptotic Mantel Theorem)**.** Let $\mathcal{F}$ be the set of graphs and let $\ell \in \mathrm{Hom}^+(\mathcal{A}, \mathbf{R})$ be a limit of graphs. Suppose that $\ell($$) = 0$ and let us bound the density of edges. Cauchy-Schwarz (3.7) gives

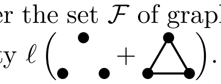
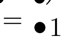
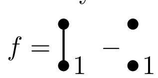


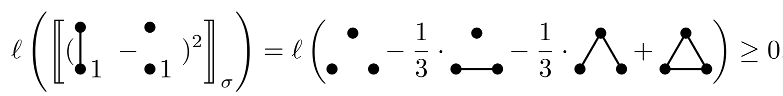
since . Therefore, .

3.3.8 The semi-definite method

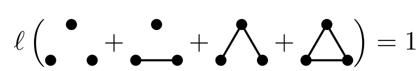
The purpose of this method is to solve a problem of maximization of a linear combination of densities under some constraints, that hopefully can be written in the flag algebra framework.

Let us first give an example in Graph Theory.

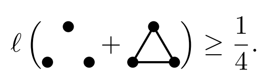
Example 3.5 (Asymptotic Goodman bound)**.** Consider the set $\mathcal{F}$ of graphs as the set of flags. Let us find a lower bound for the density . The Cauchy-Schwarz inequality (3.6) applied with the labeled point $\sigma =$  and  gives for all homomorphisms $\ell^\sigma \in \mathrm{Hom}^+(\mathcal{A}, \mathbf{R})$,

 (3.8)

The total sum probability gives

 (3.9)

By taking the combination $\frac{3}{4} \cdot (3.8) + \frac{1}{4} \cdot (3.9)$, we obtain

 (3.10)

In Example 3.5, other values may have been chosen for f (actually infinitely many), probably leading to a weaker conclusion. The purpose of the semi-definite method is to automatize the search for which inequalities can be combined to obtain the best bound possible.

Setting the problem

The semi-definite method solves optimization problems of the following form.

Problem 3.1. Given $f \in \mathcal{A}$, $k \in \mathbf{N}$ and $g_i \in \mathcal{A}$ and $c_i \in \mathbf{R}$ for every $i \in [k]$, find the largest possible $\alpha \in \mathbf{R}$ such that

$$\ell(f) \geq \alpha$$

for every $\ell \in \mathrm{Hom}^+(\mathcal{A}, \mathbf{R})$ satisfying

$$\forall i \in [k],\ \ell(g_i) \geq c_i.$$

The constraints of the type $\ell(g_i) \geq c_i$ typically expresses the hypothesis we have on the structure under consideration. In addition to these inequalities, we have at our disposal the following inequalities that are free, in the sense that we can use them regardless of the context.

- $\ell(F) \geq 0$ for every flag $F \in \mathcal{F}$, since the density of a flag is non-negative.
- $\sum_{F \in \mathcal{F}_n} \ell(F) = 1$ for all $n \in \mathbf{N}$, given by the total probability formula.
- $[\![f^2]\!]_\sigma \geq 0$ for every type σ and $f \in \mathcal{A}^\sigma$, given by the Cauchy-Schwarz inequality.

These inequalities constitute an infinite number of constraints expressed in the countable dimensional vector space $\mathcal{A}$. In order to handle the problem with a computer, we need to bound the space of work to the space generated by flags of size at most some integer $n \in \mathbf{N}$ and drop inequalities that cannot be expressed in this space. Let $\mathcal{A}_n$ be the subspace of $\mathcal{A}$ generated by the flags of size at most n. Since $\mathcal{A}$ is quotiented by the relations in $\mathcal{K}$ (defined by (3.3)), it follows that every flag of size $k < n$ can be expressed in $\mathcal{A}_n$ as a linear combination of flags of size exactly n. It follows that the (finite) family $\mathcal{F}_n = \{F_1, \ldots, F_m\}$ of flags with size n spans $\mathcal{A}_n$.

Note that if F is a σ-flag of size s, then the flag $[\![F^2]\!]_\sigma$ has size $2s - |\sigma|$. As a consequence, we use Cauchy-Schwarz inequalities only with the types σ satisfying $2s - |\sigma| = n$ for some positive integer s.

Generalized Cauchy-Schwarz

By the Spectral Theorem, an arbitrary linear combination of squares is equivalent to the application of a semi-definite positive (abbreviated s.d.p. in the remainder) matrix.

Let $f = (F_1, \ldots, F_k) \in \mathcal{F}^k$ and $g = (F_1, \ldots, F_k) \in \mathcal{F}^k$ be two vectors of flags, and $A = (a_{ij})_{1 \leq i,j \leq k}$ be a $k \times k$ matrix. The matrix A applies to f and g as follows

$$f^T A g = \sum_{i=1}^{k} \sum_{j=1}^{k} a_{ij} F_i \cdot G_j$$

where the product is the product of $\mathcal{A}$.

For a type σ and an integer $n \geq |\sigma|$, fix an enumeration $G_1, \ldots, G_m$ of the σ-flags of size n. Let g_n^σ be the vector $(G_1, \ldots, G_m) \in (\mathcal{F}^\sigma)^m$. Note that the result $g^T A g$ is an element of $\mathcal{A}^\sigma$.

Theorem 3.4 (Cauchy-Schwarz)**.** *Let σ be a type and n be an integer greater than or equal to $|\sigma|$, then for every s.d.p. matrix A and every $\ell \in \mathrm{Hom}^+(\mathcal{A}, \mathbf{R})$,*

$$\ell\left(\llbracket g^T A g \rrbracket_\sigma\right) \geq 0.$$

with $g = g_n^\sigma$

Proof. For every flag limit $\ell^\sigma \in \mathrm{Hom}^+(\mathcal{A}^\sigma, \mathbf{R})$,

$$\ell^\sigma(g^T A g) = \sum_{i,j} A_{i,j} \ell^\sigma(G_i) \ell^\sigma(G_j) = \ell^\sigma(g)^T A \ell^\sigma(g) \geq 0$$

where $\ell^\sigma(g) = (\ell^\sigma(G_1), \ldots, \ell^\sigma(G_k)) \in \mathbf{R}^k$, since ℓ^σ is an algebra homomorphism and A is semi-definite positive. The result then follows from Corollary 3.3. $\square$

Semi-definite problem

For two $m \times m$ matrices A and B, let $A * B = \sum_{1 \leq i,j \leq m} A_{ij} B_{ij}$. A *semi-definite problem* is a problem of the form:

Problem 3.2. Given an integer p, a sequence $b_1, \ldots, b_p$ of p real numbers and $p+1$ matrices $A, C_1, \ldots, C_p$ of size $m \times m$, maximize $A * X$ over all choices of positive $m \times m$ semi-definite matrices X satisfying $C_\ell * X = b_i$ whenever $1 \leq i \leq p$.

Reformulation

Let us transform the problem defined above into a semi-definite problem. The element $\llbracket g^T A g \rrbracket_\sigma$ can be expressed by:

$$\begin{aligned} \llbracket g^T A g \rrbracket_\sigma &= \left\llbracket \sum_{ij} A_{ij} G_i G_j \right\rrbracket_\sigma \\ &= \sum_{ij} A_{ij} \llbracket G_i G_j \rrbracket_\sigma . \end{aligned}$$

Let us decompose $\llbracket G_i G_j \rrbracket_\sigma \in \mathcal{A}_n$ in the basis $\mathcal{F}_n$ as $\llbracket G_i G_j \rrbracket = \sum_{\ell=1}^m u_{i,j,\ell} F_\ell$ for every indices i and j, where $u_{i,j,\ell}$ is a real number. Then

$$\begin{aligned} \llbracket g^T A g \rrbracket &= \sum_{ij} A_{ij} \sum_\ell u_{i,j,\ell} F_\ell \\ &= \sum_\ell (\sum_{ij} A_{ij} u_{i,j,\ell}) F_\ell \\ &= \sum_\ell (A * U_\ell^\sigma) F_\ell \end{aligned}$$

where U_ℓ^σ is the $k \times k$ matrix defined by $(U_\ell^\sigma)_{ij} = u_{i,j,\ell}$. So our problem can be rewritten as follows.

Problem 3.3. Given $(\gamma_i)_{i=1}^m \in \mathbf{R}^m$ and $(\mu_{i,j})_{\substack{1\le i\le m\\1\le j\le j}} \in \mathbf{R}^{m\times k}$, find the largest α such that

$$\sum_{i=1}^m \gamma_i \ell(F_i) \geq \alpha \tag{3.11}$$

using k inequalities

$$\sum_{i=1}^m \mu_{ij} \ell(F_i) \geq \alpha_j \tag{3.12}$$

for $j \in [k]$, and if $\frac{n}{2} \le m \le n$ then for every type σ with $|\sigma| = 2m - n$,

$$\sum_{i=1}^m (A^\sigma * U_i^\sigma)\ell(F_i) \geq 0 \tag{3.13}$$

for every s.d.p. matrix A^σ with the same size as U_ℓ^σ.

We now construct an inequality of type (3.11) in Problem 3.3 as a linear combination of inequalities of types (3.12) and (3.13). For every sequence of positive non-negative coefficient $(\lambda_j)_{j=1}^k \in (\mathbf{R}^+)^k$ and every sequence of s.d.p. matrices $(A^\sigma)_\sigma$. The inequalities of types (3.12) and (3.13) implies

$$\sum_{i=1}^m \left(\sum_{j=1}^k \lambda_j \mu_{i,j} + \sum_\sigma A^\sigma * U_i^\sigma \right) \ell(F_i) \geq \sum_{j=1}^k \lambda_j \alpha_j.$$

Problem 3.3 can therefore be stated as follows.

Problem 3.4 (s.d.p. form)**.** Maximize $\alpha = \sum_j \lambda_j \cdot \alpha_j$ under the constraint

$$\gamma_i = \sum_{j=1}^k \lambda_j \mu_{i,j} + \sum_\sigma A^\sigma * U_i^\sigma$$

over all choices of non-negative numbers $(\lambda_i)_{i=1}^k$ and semi-definite matrices $(A^\sigma)_\sigma$.

As a semi-definite problem, Problem 3.4 can then be solved by a solver ad-hoc.

3.4 The Caccetta-Häggkvist conjecture

This part is a joint work with Jean-Sébastien Sereni and Jan Volec.

3.4.1 The Caccetta-Häggkvist conjecture

A directed graph is a graph $G = (V, E)$ with oriented edges, with no edge from a vertex to itself, and without the edge uv whenever $vu \in E$. The *out-neighborhood* $N^+(v)$ of a vertex v is the set $\{u \in V | vu \in E\}$. The *out-degree*

$d^+(v)$ of v is the size of $N^+(v)$. The graph G is *out-regular* if all of the vertices of G have same out-degree. The *minimum out-degree* $\delta^+(G)$ of G is the minimum out-degree among all its vertices. A *directed cycle* of size n is a directed graph with n vertices $v_1, \dots, v_n$ and edges $v_i v_{i+1}$where indices are understood modulo n. A *triangle* is a directed cycle of size 3.

One of the most intriguing open problems regarding directed graphs is arguably the following conjecture, made by Louis Caccetta and Roland Häggkvist [10] about 40 years ago.

Conjecture 3.5 (Caccetta-Häggkvist, 1978)**.** *Let n and r be two integers such that $n \geq r \geq 2$. If G is a directed graph with n vertices and minimum out-degree at least r, then G contains a directed cycle of length at most $\lceil n/r \rceil$.*

In addition, the statement of Conjecture 3.5 would be best possible if true, because there exist n-vertex directed graphs with minimum out-degree r and no directed cycles of length $\lceil n/r \rceil$. The easiest way to obtain such graphs is maybe to set $n := kr+1$, take n vertices $v_0, \dots, v_{n-1}$ and add, for each $i \in \{0, \dots, n-1\}$, an arc from v_i to v_{i+j} whenever $1 \leq j \leq r = \lceil n-1/k \rceil$, where the second index is understood modulo n. The obtained graph contains no directed cycle of length at most $k = \lceil n/r \rceil$ and yet every vertex has out-degree (exactly) r.

In particular, the case where $r = \lceil n/3 \rceil$ has spawned a plethora of works. Despite a considerable amount of efforts, this case has so far been confirmed only for some special classes of graphs (for instance, a striking result is that of Yahya Ould Hamidoune [26], who established in 1981 the statement restricted to Cayley graphs using arguments from additive number theory). Written differently, Conjecture 3.5 for $r = 3$ reads as follows.

Conjecture 3.6 (Caccetta-Häggkvist, *the triangle case*)**.** *Fix an integer n at least 3. Let G be a directed graph with n vertices and no directed triangles. Then G contains a vertex with out-degree less than $n/3$.*

One way to study Conjecture 3.6 is to turn it into an optimization problem as follows.

Problem 3.5. Determine the infimum c_0 of the real numbers c such that every n-vertex simple directed graph with minimum out-degree at least $c \cdot n$ contains a directed triangle.

If true, Conjecture 3.6 would imply that the answer to Problem 3.5 is $\frac{1}{3}$. Caccetta and Häggkvist [10] proved in their seminal paper that the infimum value of c in Problem 3.5 is smaller than 0.382. An interesting and partially fruitful approach to Problem 3.5 has been initiated by Adrian Bondy [8], who proved the bound $c_0 \leq 0.379$. It relies on exploiting relations between densities of small subgraphs in a (minimal) simple directed graph with no directed triangles and with minimum out-degree at least $c \cdot n$, for some fixed number c. This bound was pushed to 0.3543 by Shen [41], next to 0.3532 by Hamburger, Haxell and Kostochka [25].

The best published bound so far is $c_0 \leq 0.3465$ and it has been proved by Hladký, Král' and Norin [27] using flags algebras and the s.d.p. method.

Their computations have been done on flags of size 4 to keep a proof of human size. Moreover, one of the inequalities is derived from a result of Chudnovsky, Seymour and Sullivan about the elimination of cycles in triangle-free directed graphs.

Theorem 3.7 (Chudnovsky, Seymour, Sullivan [11])**.** *Let G be a directed graph with no triangle. Let $\gamma(G)$ be the number of pairs of non-adjacent vertices and $\beta(G)$ the minimal number of edges to remove to obtain a graph without cycle. Then $\beta(G) \leq \gamma(G)$.*

Chudnovsky, Seymour and Sullivan conjectured that their inequality can be improved to $\beta(G) \leq \frac{1}{2}\gamma(G)$, which would be tight. This bound is for instance reached by an iterated blow-up $\Phi^n(\vec{C}_4)$ as defined later, in Definition 3.4.

An improvement of this bound can be directly reported to improve an inequality used by Hladký, Král' and Norin. The minimal value that can be taken for a has been conjectured to be $\frac{1}{2}$ by Chudnovsky, Seymour and Sullivan. Dunkum, Hamburger and Pór [14] proved the following improvement of Theorem 3.7.

Theorem 3.8 (Dunkum, Hamburger, Pór [14])**.**

$$\beta(G) \leq 0.88\gamma(G).$$

By using the arguments of Hladký, Král' and Norin [27] improved by the bound of Theorem 3.8, and using a computation on flags of size 6, we obtain a better bound for Problem 3.5.

Theorem 3.9. *Any directed graph on n vertices with minimum out-degree at least $0.339n$ has a directed triangle.*

3.4.2 Arguments for the Caccetta-Häggkvist conjecture

The purpose of this section is to prove inequalities we can assume for counter-examples to Problem 3.5. The extremal value c_0 for Problem 3.5 can be expressed as follows.

$$c_0 = \sup \left\{ \frac{\delta^+(G)}{n} \;\middle|\; G \text{ is a digraph of size } n \text{ without directed triangle.} \right\}$$

Recall that the Caccetta-Häggkvist conjecture would imply that $c_0 = \frac{1}{3}$.

Lemma 3.10. *If $c \leq c_0$, there is an infinite set $\mathbf{H}_c$ of out-regular triangle-free directed graphs such that every $G = (V, E) \in \mathbf{H}_c$ satisfies $d^+(v) = cn + o(n)$ for every $v \in V$, where n is the size of G.*

Lemma 3.10 relies on the following constructions, that is used to construct counter-examples of arbitrary sizes.

Definition 3.4. The blow-up $\Phi(G)$ of a directed graph $G = (V, E)$ is the graph on $|V|^2$ vertices obtained by replacing each vertex $v \in V$ with a copy G_v of G and each edge $uv \in E$ with a complete bipartite graph oriented from $V(G_u)$ to $V(G_v)$.

Note that if G is triangle-free then $\Phi(G)$ is triangle-free too. A node $w \in V(\Phi(G))$ corresponding to the vertex u in the copy G_v has out-degree $d^+_{\Phi(G)}(w) = d^+_G(u) + |V| \cdot d^+_G(v)$. Thus the minimum out-degree of $\Phi(G)$ is

$$\delta^+(\Phi(G)) = (1 + |V|)\delta^+(G).$$

Therefore,

$$\frac{\delta^+(\Phi(G))}{|V(\Phi(G))| - 1} = \frac{(1 + |V|)\delta^+(G)}{|V|^2 - 1} = \frac{\delta^+(G)}{|V| - 1}. \tag{3.14}$$

This construction gives a candidate for an optimal example. Let $\vec{C}_4$ be the directed cycle of size 4 and note that $\frac{\delta^+(\vec{C}_4)}{|V(\vec{C}_4)|-1} = \frac{1}{3}$. Then iterated blow-ups $G_n = \Phi^n(\vec{C}_4)$ form a growing sequence of examples such that $\frac{\delta^+(G_n)}{|V(G_n)|} = \frac{1}{3}$, so $\delta^+(G_n) = \frac{1}{3}|V(G_n)| + o(1)$.

Proof of Lemma 3.10. It follows from the definition of c_0 that for every positive ϵ, there is a directed graph G_ϵ without directed triangle satisfying

$$\delta^+(G_\epsilon) \geq (c_0 - \epsilon)|V(G_\epsilon)| \geq c - \epsilon.$$

For each $m \in \mathbf{N}$, set $\epsilon = \frac{1}{m}$ and consider the graph $H^0_m := \Phi^p(G_{\epsilon_m})$, where p is large enough to ensure that $|V(H^0_m)| \geq m$. Note that H^0_m has no directed triangle because G_{ϵ_m} has not. Moreover,

$$\frac{\delta^+(H^0_m)}{|V(H^0_m)| - 1} = \frac{\delta^+(G_{\epsilon_m})}{|V(G_{\epsilon_m})| - 1} \geq c - \epsilon_m.$$

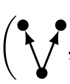
So

$$\delta^+(H^0_m) \geq (c - \epsilon_m)(|V(H_m)| - 1) = c \cdot |V(H^0_m)| + o_m(|V(H^0_m)|).$$

Now, construct H_m from H^0_m by removing edges until each vertex has out-degree exactly $\delta^+(H_m) = \delta^+(H^0_m)$, It remains to define $\mathbf{H}_c$ as the set of all such graphs H_m for $m \in \mathbf{N}$. □

The following inequality comes from an argument of Hamburger, Haxell and Kostochka [25] proved by Hladky, Kral and Norine [27] in the language of flag algebras.

Lemma 3.11. *In every $G \in \mathbf{H}_c$,*

$$p\left(\right. $$$$\left. , G\right) \geq 3(3c - 1) + o(1)$$

It is derived using the following relation

$$\beta(G) \leq a \cdot \gamma(G) \tag{3.15}$$

as provided by Theorem 3.7, that is with $a = 1$. As previously noted, Theorem 3.8 proves (3.15) for $a = 0.88$.

In the remaining, we fix a constant a such that (3.15) holds for every directed graph G without directed triangle.

Proposition 3.12 ([25]). *Let $G = (V, E)$ be a triangle-free digraph, then G has a vertex v such that*

$$d^+(v) < \sqrt{2a \cdot \gamma(G)}.$$

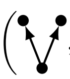
Proof. It follows from Relation (3.15) and the definition of β that there is a subset of edges $E_0 \subseteq E$ with $|E \backslash E_0| \leq a \cdot \gamma(G)$ such that the graph $G_0 = (V, E_0)$ is acyclic. The vertices of G_0 (that also are the vertices of G) can be arranged into a topological order $v_1, \ldots, v_n$, i.e. such that $i < j$ whenever $v_i v_j \in E_0$. It follows in particular that for every $i \in [n]$, the vertex v_i has out-degree at most $n - i$ in G_0. Let $\delta^+ = \delta^+(G)$ be the minimum out-degree in G. For every $j \in \{0, \ldots, \delta^+(G)\}$, the vertex v_{n-j} has out-degree at most j in G_0 and out-degree at least δ^+ in G. Consequently, $E \backslash E_0$ contains at least $\delta^+ - j$ edges starting from v_{n-j}. It follows that

$$|E| - |E_0| \geq \sum_{j=0}^{\delta^+} (\delta^+ - j) = \binom{\delta^+ + 1}{2}.$$

Further $a \cdot \gamma(G) \geq \binom{\delta^+ + 1}{2} > \frac{1}{2}(\delta^+)^2$, which proves that $\delta^+ < \sqrt{2a\gamma(G)}$. $\square$

Using Proposition 3.12, Hladký, Král' and Norin proved the following. We give a basic graph theory version of the proof of this theorem. See [27] for a version of this proof directly written in the langage of flag algebras.

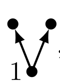
Theorem 3.13 ([27], Lemma 4.7). *In every graph $G \in \mathbf{H}_c$,*

$p\Big($  $, G\Big) \geq \dfrac{3(3c-1)^2}{a} + o(1).$

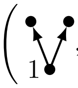
Proof. Let $G = (V, E) \in \mathbf{H}_c$, fix a vertex $v \in V$. Let $H = G[N^+(v)]$ be the subgraph of G induced by the out-neighborhood of v. Since G has no directed triangle, its subgraph H has also no directed triangle. Consequently, Proposition 3.12 applies to H and gives the existence of a vertex $u \in N^+(v)$ with

$$d_H^+(u) \leq \sqrt{2a\gamma(H)}.$$

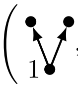
Let G_v be the flag G rooted on v, which is labeled by 1. With this notation,

$\gamma(H) = p\Big($  $, G_v\Big) \cdot \binom{n-1}{2}$

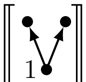
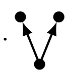
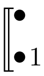
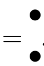
Note that a vertex $w \in N^+(u) \setminus N^+(v)$ has no edge to v, as otherwise vuw form a directed triangle. Since there are $d^+(u) - d^+_H(u)$ such edges, it follows that

$p\Big($  $, G_v\Big)\, n \geq d^+(u) - d^+_H(u) \geq cn - \sqrt{2a \cdot p\Big(}$  $\overline{, G_v\Big) \cdot \binom{n-1}{2}} + o(n).$

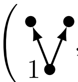
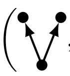
Normalizing by n,

$p\Big($  $, G_v\Big) \geq c - \sqrt{a \cdot p\Big(}$  $\overline{, G_v\Big)} + o(1). \qquad (3.16)$

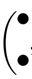
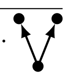
We now take the average over all choices of v. Let $\lambda := 1\bullet$ be the rooted flag with only one vertex that is rooted. Note that $p\left([\![\lambda]\!]_\lambda, G\right) = 1$. It holds that

$\Big[\!\!\Big[$  $\Big]\!\!\Big]_\lambda = \frac{1}{3} \cdot$  and $\Big[\!\!\Big[$  $\Big]\!\!\Big]_\lambda =$  .

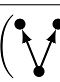
By (3.5), it follows that

$\mathbf{E}_v\Big(p\Big($  $, G_v\Big)\Big) = \frac{1}{3} p\Big($  $, G\Big)$ and $\mathbf{E}_v\Big(p\Big($  $, G_v\Big)\Big) = p\Big($  $, G\Big).$

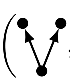
Applying Jensen's Inequality on (3.16) for the concave function $\sqrt{\cdot}$, we obtain

$p\Big($  $, G\Big) \geq c - \sqrt{\frac{a}{3} \cdot}$  $+ o(1).$

Since it is out-regular with degree $cn + o(n)$, the graph G has $cn^2 + o(n^2)$ edges. It follows that $p\Big($  $, G\Big) = 1 - (cn^2 + o(n^2))/\binom{n}{2} = 1 - 2c + o(1)$. Consequently,

$3c - 1 \leq \sqrt{\frac{a}{3} \cdot p\Big(}$  $\overline{, G\Big)} + o(1).$

Since $3c - 1 \geq 0$, we can take the square

$(3c-1)^2 \leq \frac{a}{3} \cdot p\Big($  $, G\Big) + o(1),$

which proves the theorem. $\square$

Let σ be a type of size k. A σ-source is a σ-flag of size $k+1$ where the only unlabeled node has out-degree zero. Let $\mathcal{F}^{\sigma,\rightarrow}$ be the set of all σ-sources and F_0^σ the particular σ-source where all vertices of the type have an edge to the unlabeled vertex. Fix a constant c_1 for which we later assume that $c_1 \geq c_0$. Let us define

$$f(\sigma) = \sum_{F \in \mathcal{F}^{\sigma,\rightarrow}} F + (c_1 - 1)F_0^\sigma - c.$$

Note that f depends on c and c_1.

Theorem 3.14 ([27], Lemma 4.4). *Assume $c \in [c, c_1]$ and let σ be a type with one vertex with in-degree $|\sigma| - 1$. For every $G \in \mathbf{H}_c$ and every choice of rooted G^σ, the following holds.*

$$p(f(\sigma), G^\sigma) \geq o(1).$$

Proof. Let G^σ be a σ-flag rooted on the vertex $S = \{v_1, \ldots, v_k\}$ (so $G[S] = \sigma$). Assume that v_1 is the vertex with in-degree $k-1$ in σ.

The set $V_{F_0^\sigma} := \bigcap_{i=1}^k N^+(v_i)$, is the set of vertices $v \in V \setminus S$ such that the graph $G[S \cup \{v\}]$ rooted on $v_1, \ldots, v_k$ is isomorphic to the rooted flag F_0^σ, so its size can be expressed as $|V_{F_0^\sigma}| = p(F_0^\sigma, G^\sigma)n + o(n)$.

More generally, let V_F be the set of vertices $v \in V \setminus S$ such that the graph $G[S \cup \{v\}]$ rooted on $v_1, \ldots, v_k$ is isomorphic to F, for $F \in \mathcal{F}^{\sigma,\rightarrow}$. We have $|V_F| = p(F, G^\sigma)n + o(n)$. We distinguish two cases.

Assume that $|V_{F_0^\sigma}| = 0$, so $p(F_0^\sigma, G^\sigma) = 0$. For every $v \in N^+(v_1)$, the rooted flag $G[S \cup \{v\}]$ is a σ-source: indeed, if there is an edge from v to a vertex v_i with $2 \leq i \leq k$, then $v_1 v v_i$ is a directed triangle. As a consequence,

$$\sum_{F \in \mathcal{F}^{\sigma,\rightarrow}} |V_F| \geq d^+(v_1) = cn + o(n)$$

i.e.

$$\sum_{F \in \mathcal{F}^{\sigma,\rightarrow}} p\,(F, G^\sigma) \geq c + o(1).$$

This proves the theorem in the case where $|V_{F_0^\sigma}| = 0$.

Assume now that $|V_{F_0^\sigma}| > 0$. As an induced subgraph of G, the directed graph $G[V_{F_0^\sigma}]$ has no directed triangle. By the definition of c_0, there further is a vertex $v \in V_{F_0^\sigma}$ with out-degree at most $c_0|V_{F_0^\sigma}| + o(n) \leq c_1 p(F_0^\sigma, G^\sigma)n + o(n)$ in $G[V_{F_0^\sigma}]$. Since moreover, $d_G^+(v) = cn + o(n)$, it follows that

$$|N^+(v) \setminus V_{F_0^\sigma}| \geq cn - c_1 p(F_0^\sigma, G^\sigma) + o(n).$$

Note that every $w \in N^+(v) \setminus V_{F_0^\sigma}$ form a σ-source with $v_1, \ldots, v_k$. Indeed, if there is an edge from w to v_i for some $i \in [k]$, then $v w v_i$ is a directed triangle. It follows that

$$\sum_{F \in \mathcal{F}^{\sigma,\rightarrow}} |V_F| - |F_{F_0^\sigma}| \geq |N^+(v) \setminus V_{F_0^\sigma}|.$$

and further

$$\sum_{F \in \mathcal{F}^{\sigma,\rightarrow}} n \cdot p(F, G^\sigma) - n \cdot p(F_0^\sigma, G^\sigma) \geq cn - c_1 n \cdot p(F_0^\sigma, G^\sigma) + o(n)$$

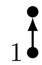
$$\sum_{F \in \mathcal{F}^{\sigma,\rightarrow}} p(F, G^\sigma) + (c_1 - 1)p(F_0^\sigma, G^\sigma) - c \geq o(1)$$

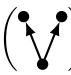
This finishes the proof of the theorem. ∎

3.4.3 Connection with graph limits

We pass Lemma 3.10 and Theorems 3.13 and 3.14 to the limit to deduce constraint on some homomorphism. Let $\lambda = 1\bullet$ be the unique type made of only one rooted vertex.

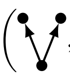
Theorem 3.15. *If $c \in [c_1, c_0]$, then there is a limit ℓ of directed graphs without directed triangles that satisfies*

1. $\ell\left(\left[\!\left[$  $\cdot F - cF\right]\!\right]_\lambda\right) = 0$ *for every $F \in \mathcal{F}^\lambda$;*

2. $\ell\left(\right.$  $\left.\right) \geq \frac{3(3c-1)^2}{a}$*; and*

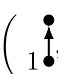
3. $\ell([\![f(\sigma)F]\!]_\sigma) \geq 0$ *for every type σ with a sink and every $F \in \mathcal{F}^\sigma$.*

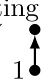
Proof. Let $(G_n)_{n\in\mathbf{N}}$ be a sequence of graphs of $\mathbf{H}_c$ whose size tends to infinity. Up to extracting a subsequence, we may assume that $(G_n)_{n\in\mathbf{N}}$ converges to a limit $\ell \in \mathrm{Hom}^+(\mathcal{A}, \mathbf{R})$. We know from Theorem 3.13 that

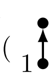
$p\left(\right.$  $\left., G_n\right) \geq \dfrac{3(3c-1)^2}{a} + o(1).$

Letting n tends to infinity gives Item 2.

Since G_n is out-regular with degree $cn + O(1)$, we know that the random rooted flag G^λ rooted on a random vertex satisfies

$p\left(\right.$  $\left., G_n\right) = c + o(1).$

Letting n tends to infinity, it follows that the random homomorphism ℓ^λ satisfies $\ell^\lambda\left(\right.$  $\left. - c\right) = 0$ with probability 1. For every $F \in \mathcal{F}^\sigma$, we know that $\ell^\sigma(F) \geq 0$, so

$\ell^\lambda\left(\left(\right.\right.$  $\left.\left. - c\right) \cdot F\right) = 0.$

Item 1 follows from Corollary 3.3.

Item 3 is proved similarly. Taking the limit in Theorem 3.14 applied to G_n, we deduce that the random homomorphism ℓ^σ satisfies

$$\ell^\sigma(f(\sigma)) \geq 0$$

with probability 1. Since $\ell^\sigma(F) \geq 0$ for every $F \in \mathcal{F}^\sigma$, it follows that $\ell^\sigma(f(\sigma)F) \geq 0$. Item 3 then follows from Corollary 3.3. $\square$

3.4.4 Results of the algorithm

Let $P(c, c_1)$ be the problem described by Theorem 3.15, that is the question of whether there is a limit ℓ of directed graphs without triangle satisfying the constraints 3.15.1, 3.15.2 and 3.15.3.

This problem can be solved using the semi-definite method with the program presented in the next section. This program uses as input the inequalities in Theorem 3.15 with the standard flag algebra inequalities described in the s.d.p. section (i.e. every flag is non-negative, the sum of flags of size n is 1 and Cauchy-Schwarz inequalities).

The problem of monotonicity

The problem $P(c, c_1)$ is monotone in c_1 in the sense that if $c_1 \geq c_1'$ and $P(c, c_1)$ has a solution, then $P(c, c_1')$ also admits a solution. We have however no evidence that this problem is monotone in c. There may be c, c' and c'' with $c \leq c' \leq c''$ such that $P(c, c_1)$ and $P(c'', c_1)$ have a solution but not $P(c', c_1)$.

If the problem $c \mapsto P(c, c)$ happen to be monotonous, it would be enough to run the program on $P(c, c)$ to directly prove the bound $c_0 \leq c$ (in case there is no solution).

Results

For $n = 6$, that is for flags on 6 vertices, $c = 0.3392$ and $c_1 = 0.3465$, the program concludes that the system has no solution, which further contradicts the hypothesis of Theorem 3.15, showing that $c_0 \notin [c, c_1]$. Since $c_0 \leq 0.3465 = c_1$ by the result of Hladký, Král and Norin [27], we conclude that $c_0 < c = 0.3392$.

Further iterations can be made starting with $c_1 = 0.3392$. The following table shows results of other runs of the algorithm.

c	c_1	n	CSDP output
0.3392	0.3465	6	Dual infeasible
0.3391	0.3392	6	Dual infeasible*
0.339	0.3391	6	Dual infeasible*
0.3386	0.3386	6	SDP solved
0.342	0.3465	5	Dual infeasible
0.341	0.341	5	SDP solved
0.3475	0.37	4	Dual infeasible
0.3465	0.3475	4	Dual infeasible
0.3464	0.3464	4	SDP solved

The star (*) indicates that the solver did not converge properly but still gave a certificate that allows us to witness a contradiction.

As previously mentioned, a result of dual infeasibility means that no assignment of density on graphs of size n satisfies the constraints, which contradicts the hypothesis that $c_0 \in [c, c_1]$. On the other hand, the existence of a solution means that such an assignment exists, and that therefore suggests that it is not possible to find a better solution without increasing n.

The method has been run in an independent program written in C by Jan Volec and obtained the bound $c < 0.3388$.

Deducing a proof

To prove that $c_0 \notin [c, c_1]$, we need to recover the certificate of dual infeasibility provided by the solver. This certificate consist of a (huge if $n = 6$) (approximately) semi-definite matrix X of floating point numbers, that when reinjected in our inequality, gives a computation that

$$\sum_{i}^{m} \epsilon_i F_i \geq \alpha \tag{3.17}$$

where $\alpha \approx 1$ and each $|\epsilon_i|$ is close to 0. Recall that $(F_i)_i^m$ is the family of flags of size n and that $\sum_i^m F_i = 1$. If true, such an inequality indeed gives a contradiction.

To properly establish (3.17), we first approximate X by a matrix $\tilde{X}$ with rational coefficients that is semi-definite positive. To do so, we compute the (approximated) eigenvectors $(e_i)_{i=1}^k$ of X with respective eigenvalues $(\lambda_i)_{i=1}^k$, then we round each e_i to a rational vector $\tilde{e}_i$ and each λ_i to a rational number $\tilde{\lambda}_i$ and we set $\tilde{X} := \sum_{i=1}^k \tilde{\lambda} \tilde{e}_i^T \tilde{e}_i$. Up to checking that $\tilde{\lambda}_i \geq 0$ for every i, the matrix $\tilde{X}$ is indeed s.d.p. To complete the proof, we recompute the constraints of the s.d.p. problem in rational numbers and we compute the inequality of type (3.17) given by $\tilde{X}$. It then only remains to check that the parameters found satisfy $\max_{i=1}^m \epsilon_i - \min_{i=1}^m \epsilon_i < \alpha$ to ensures that the contradiction holds.

This method confirmed the results given in the table above, hence proving that the extremal value c_0 of the triangle case in the Caccetta-Häggkvist conjecture is at most 0.339.

3.5 A generic flag algebra program

My program is written in OCaml (Objective Caml) and is available at `https://github.com/avangogo/flag`. This program takes advantage of the module and functor features of OCaml in order to deal with any type of flag algebras. Given a module that describes the theory on which we want to work, the program builds the corresponding flag algebra and provides features to create and manipulate inequalities in flag algebras and generate s.d.p. problems. Those s.d.p. problems are written in the sdpa format that can be used as an input by the s.d.p. solver Csdp[1].

In this section, I describe the general architecture of my program and the main ideas behind it.

[1] https://projects.coin-or.org/Csdp/

3.5.1 For non-OCaml readers

In this section we use the notation of the Ocaml type system to describe the type of the objects implemented. We gives a few information the reader has to know to understand the end of this chapter. In OCaml, every object (including function) has a type. The type `a -> b` is the type of functions `f` with input of type `a` and output of type `b`. The application of `f` to `x` (of type `a`) is written `f x`. The type `a -> b -> c` has to be read `a -> (b -> c)` and refers to a function (in Curry form) with two inputs of types `a` and `b` and an output of type `c`. An object of type `a array` is an array of objects of type `a`. An object of type `a list` is a linked list of objects of type `a`. Types can be constructed inductively. For instance, `int list array ->int` is the type for functions that take an array of linked lists of integers and return an integer. In a signature, `type t` requires the definition of a type and `val foo : t` requires the definition of an object of type `t` named `foo`.

3.5.2 The flags

The program permits to implement flag algebras on every kind of flags. A type of flag is implemented in a module that essentially gives the elements described in Definition 3.1, together with some additional tools. A flag module has the following signature.

```
module type S =
sig
  type t
  val size : t -> int
  val induce : int array -> t -> t
  val apply_morphism : int array -> t -> t

  val iso_invariant : t -> int -> int list array
  val invariant_size : int

  val superflags : t -> t list
  val span : int -> t list

  val name : string
  val draw : ?root:int -> t -> Graphic.drawable
  val print : t -> string
end
```

This signature lists the elements that are needed for a flag module. Let us see these elements in more detail.

```
  type t
```

The type of the flags. This is an abstract type, which means that an element outside the module does not know (and does not need to know) the implementation of an object of type `t`. For instance, if the flags are graphs represented by their adjacency matrix, `t` can be a type of matrices (e.g. `int array array`).

This implies that the rest of the program manipulates flags only by the mean of the few function described in this module (plus the built-in equality and comparison provided by OCaml).

```
val size : t -> int
```

The function `size` returns the number s of vertices of the flag F as input. For the next functions, the vertices of F are identified by the integers $\{0, \ldots, s-1\}$.

```
val induce : int array -> t -> t
```

The function `induce` takes as input a set of vertices S (represented by an array of integers) and a flag F and returns the subflag $F[S]$ of F induced by S.

```
val apply_morphism : int array -> t -> t
```

The function `apply_morphism` takes in input a permutation ϕ of $\{0, \ldots, s-1\}$ and a flag F of size s and returns the flag F^{ϕ} obtained by reordering the vertices of F according to ϕ.

Dealing with isomorphisms

The program needs to reduce the flags modulo isomorphisms. To handle isomorphismic flags recognition, I implemented a simplified and adapted version of an algorithm of Brendan McKay [37, 38] for graph isomorphism (this is implemented in `algebra.ml`). This algorithm finds for every flag F a "normal form" F_{nf} isomorphic to F with the property that F and G are isomorphic if and only if $F_{nf} = G_{nf}$.

The two following functions are used as hints for the normal form algorithm.

```
val iso_invariant : t -> int -> int list array
val invariant_size : int
```

Given a flag F and a vertex of v of F, the function `iso_invariant` can provide one or more lists of vertices $\{u_1, \ldots, u_k\}$ that are invariant upon isomorphisms in the sense that if F is reordered by the permutation $\pi : \{0, \ldots, s-1\} \to \{0, \ldots, s-1\}$, then the list corresponding to $\pi(v)$ is $\{\pi(u_1), \ldots, \pi(u_k)\}$ in the new flag.

Typically, for graphs, this function can return the neighborhood of v. For directed graphs, this function can return two sets, the in-neighborhood and the out-neighborhood of v. The constant integer `invariant_size` specifies the number of lists given by `iso_invariant`. The `invariant_size` can be 0, and, consistently, `iso_invariant` should always return an empty array in this case.

This invariant is in theory optional : it is only used to help with the normal form algorithm. If `invariant_size` is 0 (so `iso_invariant` gives no information), then the algorithm boils down to compute every possible permutation of the input flag F and to return the minimal one for some order (whose specification does not matter, I use the built-in structural comparison of OCaml).

This therefore may quickly be too long in practice. If the flags implemented are graphs and `iso_invariant g v` gives the neighborhood of v in the flag g, then the normal form algorithm works as the original algorithm of McKay.

This algorithm is very efficient for graphs and can be more or less adapted to other structures. The normal form algorithm is a central element of the program since it is used to compute the set of flags and all the densities.

Generation of flags

The two following functions are used to generate all the flags up to a given size.

```
val superflags : t -> t list
```

Given a flag F of size s, `superflags` returns the list of flags of size $s+1$ that extend F, i.e. the flags containing F as induced sub-flag. For graphs, it can return the 2^s graphs corresponding to all the ways to adding a vertex and its incident vertices to a graph of size s.

The set of flags of size $s+1$ is then constructed inductively from the set of flags of size s (reduced by isomorphisms) by applying `superflags` to each of them and then reducing the union of the obtained sets by isomorphism (using the algorithm presented in the previous section).

```
val span : int -> t list
```

The function `span` directly generates a list of all flags of the size given as input, typically with a naive algorithm that is not supposed to be efficient. This list can have redundancy and is a priori not reduced by isomorphism. It is used firstly to initialize the induction above by providing the list of flags of size 1 and secondly for testing the correctness of the optimized inductive algorithm.

Pretty-printing

```
val name : string
val draw : ?root:int -> t -> Graphic.drawable
val print : t -> string
```

These three last values have no algorithmic use and are mainly used for debugging and pretty-printing.

3.5.3 Storing operators

A *basis* of flags refers to the set of flags of a given size of a (possibly rooted) flag algebra. A flag is identified by a basis and a unique identifier in this basis. A basis is identified by three numbers : the size n of the flags considered, the size s of the type and the identifier of the type in the basis of unrooted flags of size s. This implies that the type is reduced by isomorphism, and avoids to compute equivalent flag algebras.

The different parameters of the flag algebra may be long to compute. The list of representative of flags of a given basis, densities matrices $p(.,.)$, multiplications tables (i.e. lists of values $(p(F_1, F_2; F))$ where F_1, F_2 and F are in some bases), and unlabeling operator (given for each graph of a given basis, by the coefficient $q_\sigma(.)$ and the identifier of the unrooted graph in the corresponding basis) are computed the first time a part of the program asks for them and are then stored in files. On further runs of the program, these files are loaded.

As an optimization the *multiply and unlabel* operator $(F^\sigma, G^\sigma) \to [\![F^\sigma \cdot G^\sigma]\!]_\sigma$ is also stored on a file. This avoids to directly manipulate products $F^\sigma \cdot G^\sigma$ that live in a space of larger dimension. This operator is also the operator used in the Cauchy-Schwarz inequalities.

Except when computing (once for all) the above values, the program in particular almost never manipulates the combinatorial object (i.e. the graphs, hypergraphs, ...) but only vectors and matrices of numbers.

As a consequence, this program can take a very long time the first time a computation on a new type of flags is launched. After that, the program is typically faster than the solver launched on its output.

3.5.4 Examples

Computing in the flag algebra

Let us use the program to check the computations in Example 3.5. The following program performs the computation in this example and draws the intermediate steps.

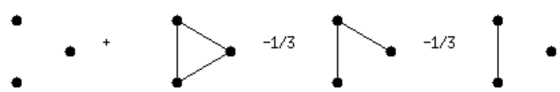
```
(* Loading Modules for rational valued quantum graphs *)
module S = Storage.Make (Graph)
module Vect = Vectors.Vect (Rational) (Graph)
open Vect

(* Defining the rooted edes and non-edge *)
let rootedBasis = S.basis_id ~typeSize:1 ~typeId:0 2
let rootedEdge = flag rootedBasis (Graph.make 2 [(0,1)])
let rootedNonedge = flag rootedBasis (Graph.make 2 [])

(* Operations in the flag algebra *)
let diff = rootedEdge -~ rootedNonedge
let square = diff *~ diff
let average_square = untype square
let total_sum = one (S.basis_id 3)
let result = average_square +~
               (scalar_mul (Rational.make 1 3) total_sum );;

(* Displaying the results *)
draw diff;;
draw square;;
draw average_square;;
draw total_sum;;
```

```
draw result
```

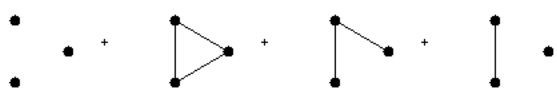
Let us look at the pictures given by this program. In the program, `diff` is the rooted edge minus the rooted non-edge in the flag algebra of graphs rooted on one vertex. It is displayed as follows, where the red vertex is the root.

Its square `square` is displayed as follows, where, again, the red vertex is the root.

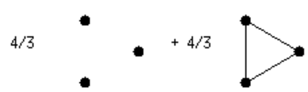
Applying the operator $[\![.]\!]_\sigma$ gives `average_square`.



We just computed (3.8). Consider the sum `total_sum` of all graphs of size 3.

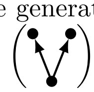


Adding one third of this sum with `average_square` gives the following.



We know from the construction that this flag has values at least $\frac{1}{3}$, which proves the asymptotic Goodman bound (3.10).

Maximal density of forks in digraphs

The following code generates a semi-definite program looking for the maximal density of the fork  in a directed graph.

```
(* Maximal density of forks in a Digraph *)

(* Building and loading modules *)
module S = Storage.Make (Digraph)
module I = Inequality.Make (Rational) (Digraph)
```

```
module Vect = Vectors.Vect (Rational) (Digraph)
module Problem = Problem.Make (Rational) (Digraph)
open Storage
open I
open Vect

(* We compute on flags of size 4 *)
let basis = S.basis_id 4

(* Constructing the vector corresponding to the fork *)
let fork_flag =
  let fork_digraph = Digraph.make 3 [|(0,1);(0,2)|] in
  flag ~name:"fork" (S.basis_id 3) fork_digraph

(* We want to maximize minus the density of forks *)
let objective = expand basis (opposite fork_flag)

(* Inequalities on flags we use *)
let inequalities =
  List.concat [
     all_flags_nonnegative basis; (* A flag is nonnegative *)
     equality (totalsum basis) (* Flags of size 4 sum to 1 *)
    ];;

(* Writing the corresponding s.d.p. program *)
Problem.write "forks_density" inequalities objective
```

This program gives the following output.

```
Building sdp problem (forks_density)
on basis : digraph/basis4_type0_id0 (42 flags)
Maximizing : -fork
Building Cauchy-Schwartz blocks (3 bases)
Building Inequalities block (44 inequalities)
Inequalities list :
  Total sum is 1 (Equality) (2)
  Flag is non-negative (42)
Writing problem in file "forks_density.sdpa"
```

Running `csdp fork_density.csdp` then gives the following.

```
Success: SDP solved
Primal objective value: -4.6410162e-01
Dual objective value: -4.6410161e-01
Relative primal infeasibility: 5.94e-13
Relative dual infeasibility: 1.82e-10
Real Relative Gap: 5.56e-10
XZ Relative Gap: 9.36e-10
DIMACS error measures: 8.27e-13 0.00e+00 3.86e-10 0.00e+00 5.
Elements time: 0.004356
Factor time: 0.000590
Other time: 0.011753
```

```
Total time: 0.016699
```

The solver found an upper bound (on minus the density of forks) that is around -0.46101, which approximately corresponds to $-(2\sqrt{3}-3)$. It can be indeed proved that the maximum density of forks is $-(2\sqrt{3}-3)$. The lower bound was found using the semi-definite method [17, Theorem 27].

Edge density in graphs without C_5

We give a similar example on graphs. The purpose of the following program is to find an upper bound on the edge density of graphs without cycle of size 5.

```
(* Maximal density of graphs without cycle of size 5 *)
module S = Storage.Make (Graph)
module I = Inequality.Make (Field.Float) (Graph)
module Problem = Problem.Make (Field.Float) (Graph)
module Vect = Vectors.Vect (Field.Float) (Graph)
open I
open Vect

(* We compute on flags of size 5 *)
let basis = S.basis_id 5

(* This function decides whether a graph is hamiltonian *)
let is_hamiltonian g =
  let rec aux u path = function
    | 0 -> u == 0
    | n -> List.exists
             ( fun v -> not (List.mem v path)
                        && aux v (v::path) (n-1) )
             ( Graph.neibrs g u ) in
  aux 0 [] (Graph.size g)

(* Computes the sum of flags of size 5 containing C5 *)
let c5_subgraphs =
  let indicator _ g = if is_hamiltonian g then 1. else 0. in
  Vect.make basis indicator;;

(* Inequalities used *)
let inequalities =
  List.concat [
      (* The density of every flag is non-negative *)
      all_flags_nonnegative basis;
      (* The sum of the flags of size 5 is 1 *)
      equality (totalsum basis);
      (* The density of graphs with a C5 is at most 0 *)
      [ at_most c5_subgraphs 0. ]
    ]

(* The objective is the density of non-edges *)
```

```
(* ( expressed with flags of size 5 ) *)
let nonedge =
  Vect.flag ~name:"E2" (S.basis_id 2) (Graph.make 2 [])

let objective = expand basis nonedge;;

(* Maximum number of non-edges *)
Problem.write "C5-free" inequalities objective
```

This program outputs the following.

```
Building sdp problem (C5-free)
on basis : graphs/basis5_type0_id0 (34 flags)
Maximizing : E2
Building Cauchy-Schwartz blocks (5 bases)
Building Inequalities block (37 inequalities)
Inequalities list :
  Total sum is 1 (Equality) (2)
  Flag is non-negative (34)
  No label (1)
Writing problem in file "C5-free.sdpa"
```

Processing the `.sdpa` file with the command `csdp fork_density.csdp` then gives the following.

```
Success: SDP solved
Primal objective value: 5.0000000e-01
Dual objective value: 5.0000000e-01
Relative primal infeasibility: 3.42e-14
Relative dual infeasibility: 1.41e-09
Real Relative Gap: 2.63e-09
XZ Relative Gap: 6.01e-09
DIMACS error measures: 7.14e-14 0.00e+00 3.89e-09 0.00e+00 2.
Elements time: 0.008245
Factor time: 0.000726
Other time: 0.034204
Total time: 0.043175
```

The found bound is close to $\frac{1}{2}$, which is indeed the maximum edge density of graphs without cycle of size 5. This is a consequence of the Erdős-Stone-Simonovits Theorem [15], which states that the asymptotic maximum edge density of H-free graphs is $1 - \frac{1}{1-\chi(H)}$, where $\chi(H)$ stands for the chromatic number of H. Since $\chi(C_5) = 3$, this indeed gives $1 - \frac{1}{1-3} = \frac{1}{2}$. This density is in particular reached by the limit of complete bipartite graphs $(K_{n,n})_{n\in\mathbf{N}}$.

Caccetta-Häggkvist

The following code (together with the common files of my program) encodes the constraints of Theorem 3.15. It has three parameters : the maximal size n of the flags used, a constant c_1 for which it is known that $c_0 \leq c_1$ and the constant c for which we assume for a contradiction that $c \leq c_0$.

```
(* Caccetta-Haggkvist *)

open Digraph
open Storage
module S = Storage.Make (Trianglefree)
module I = Inequality.Make (Field.Float) (Trianglefree)
module V = Vectors.Vect (Field.Float) (Trianglefree)
module Problem = Problem.Make (Field.Float) (Trianglefree)
open V
open I

(* **** Parameters **** *)
let flagSize = 6 (* Size of the flags used *)
let c  = 0.3392  (* Constant for which we prove the result *)
let c1 = 0.3465  (* Constant for which we know it holds *)

(* Basis on which we build the sdp problem *)
let b = S.basis_id flagSize

(* **** Computing Inequalities of Theorem 3.15 **** *)

(* ==== 1. Degree is c ==== *)

let outedge =
  let basis2_1 = S.basis_id ~typeSize:1 ~typeId:0 2 in
  V.flag ~name:"outedge" basis2_1 (Digraph.make 2 [|(0,1)|])

let outdegree_geq_c = { ( at_least outedge c )
                      with boundName = Some "c" }

let outdegree_is_c =
  let ineqs = multiply_and_unlabel b (outdegree_geq_c) in
  name_list "1.␣Outdegree␣is␣c"
    (List.concat (List.map (equality ~epsilon:1e-11) ineqs))

(* ==== 2. Density of forks is at least 3(3c-1)^2/a ==== *)

(* Constant for the Chudnovsky-Seymour-Sullivan conjecture *)
let a = 0.88

(* basis of flags of size 3 *)
let b3 = S.basis_id 3

(* fork *)
let fork =
  V.flag ~name:"fork" b3 (Digraph.make 3 [|(0,1);(0,2)|])

let fork_ineq =
  let x = 3. *. c -. 1. in
  let y = (3. *. x *. x) /. a in
```

```
  { ( at_least fork y )
  with name = Some "2.␣Fork" ;
       boundName = Some "3*(3c-1)^2/a" }

(* ==== 3. [|f(sigma)F|] >= 0 for every sigma-source ==== *)

(* Determine if the flag in input is a sigma-source *)
let is_a_sigma_source typeSize flag =
  let edge_condition (i, j) =
    not (i >= typeSize && j < typeSize) in
  Common.array_for_all edge_condition flag.e

(* Return the sum of sigma sources of the basis in input *)
let sum_of_sigma_sources basis =
  let indicator typeSize f =
    if is_a_sigma_source typeSize f then 1. else 0. in
  V.make ~name:"Sigma␣sources" basis indicator

(* Build the graph f0 obtained by adding a sink to sigma *)
let f0_flag sigma_flag =
  let n = sigma_flag.n in
  let new_edges =  Array.init n (fun i -> (i, n)) in
  let edges = Array.concat [sigma_flag.e; new_edges] in
  Digraph.make (n+1) edges

(* Recover the type of this basis in input (as a graph) *)
let get_type basis =
  ( S.get_basis (S.basis_id basis.typeSize) ).(basis.typeId)

(* Return the vector corresponding to f0 *)
let f0 basis =
  let f0 = f0_flag (get_type basis) in
  V.flag ~name:"F0" basis f0

(* Compute f(sigma) = sum of sigma-sources + (c1-1)F0 - c *)
let f basis =
  let x_f0 = scalar_mul ~name:"(c1-1)" (c1-.1.) (f0 basis) in
  let c_one = scalar_mul ~name:"c" c (one basis) in
  let sum = sum_of_sigma_sources basis in
  sum +~ x_f0 -~ c_one   (* f(sigma) *)

(* f(sigma) >= 0 *)
let f_inequality basis = at_least (f basis) 0.

let has_dominated_vertex g =
  Common.array_exists ( (==) (g.n - 1) ) (in_degrees g)

(* Build the inequalities of the type f(sigma) >= 0 *)
(* for the types sigma with a dominated vertex *)
let f_rooted_ineqs =
```

```
  let res = ref [] in
  for n = 2 to b.flagSize - 1 do
    let types = S.get_basis (S.basis_id n) in
    for i = 0 to (Array.length types) - 1 do
      if has_dominated_vertex types.(i) then begin
        let basis = S.basis_id ~typeSize:n ~typeId:i (n+1) in
        res := ( f_inequality basis ) :: !res
      end
    done
  done;
  !res

(* Build the inequalities of the type [|f(sigma)*F|] >= 0 *)
let f_inequalities =
  name_list "3.␣Sigma-sources"
    (List.concat
      (List.map (multiply_and_unlabel b) f_rooted_ineqs))

(* **** Construction of the problem **** *)

(* List of inequalties used *)
let inequalities =
  List.concat
    [
      all_flags_nonnegative b;  (* A flag is non-negative *)
      [ at_least (one b) 1. ];  (* Sum of flags at least 1 *)
      outdegree_is_c;           (* Constraint 1. *)
      [ expand b (fork_ineq) ]; (* Constraint 2. *)
      f_inequalities;           (* Constraint 3. *)
    ]

(* Print the sdp program in CH.sdpa *)
let _ = Problem.write "CH" inequalities (one b)
```

The program then construct the s.d.p. problem and output the following summary (in verbose mode).

```
Building sdp problem (CH)
on basis : trianglefree/basis6_type0_id0 (6583 flags)
Maximizing : 1
Building Cauchy-Schwartz blocks (35 bases)
Building Inequalities block (12563 inequalities)
Inequalities list :
  1. Outdegree is c (2858)
    [| (outedge - c)*x |] >= -1e-11 (1429)
    -[| (outedge - c)*x |] >= -1e-11 (1429)
  2. Fork (expanded) (1)
    fork >= 3*(3c-1)^2/a (1)
  3. Sigma-sources (3120)
    [| (Sigma sources + (c1-1).F0 - c)*x |] >= 0. (3120)
  Flag is non-negative (6583)
```

```
    x >= 0. (6583)
  No label (1)
    1 >= 1. (1)
Writing problem in file "CH.sdpa"
```

Bibliography

[1] Oswin Aichholzer, Franz Aurenhammer, and Hannes Krasser. Enumerating order types for small point sets with applications. *Order*, 19(3):265–281, 2002.

[2] Oswin Aichholzer and Hannes Krasser. The point set order type data base: A collection of applications and results. In *Proceedings of the 13th Canadian Conference on Computational Geometry, University of Waterloo, Ontario, Canada, August 13-15, 2001*, pages 17–20, 2001.

[3] Noga Alon. The number of polytopes, configurations and real matroids. *Mathematika*, 33(1):62–71, 1986.

[4] Noga Alon and Jacob Fox. Easily testable graph properties. *Combinatorics, Probability & Computing*, 24(4):646–657, 2015.

[5] Noga Alon and Michael Krivelevich. Testing k-colorability. *SIAM J. Discrete Math.*, 15(2):211–227, 2002.

[6] Noga Alon and Asaf Shapira. A characterization of easily testable induced subgraphs. *Combinatorics, Probability & Computing*, 15(6):791–805, 2006.

[7] Noga Alon and Asaf Shapira. A characterization of the (natural) graph properties testable with one-sided error. *SIAM J. Comput.*, 37(6):1703–1727, 2008.

[8] John Adrian Bondy. Counting subgraphs a new approach to the caccetta-häggkvist conjecture. *Discrete Mathematics*, 165-166:71–80, 1997.

[9] Kellogg S. Booth and George S. Lueker. Testing for the consecutive ones property, interval graphs, and graph planarity using pq-tree algorithms. *J. Comput. Syst. Sci.*, 13(3):335–379, 1976.

[10] Louis Caccetta and Roland Häggkvist. On minimal digraphs with given girth. *Congressus Numerantium*, 21:181–187, 1978.

[11] Maria Chudnovsky, Paul D. Seymour, and Blair Sullivan. Cycles in dense digraphs. *Combinatorica*, 28(1):1–18, 2008.

[12] David Conlon and Jacob Fox. Bounds for graph regularity and removal lemmas. *Geometric and Functional Analysis*, pages 1–66, 2012.

[13] Roy O Davies. Subsets of finite measure in analytic sets. In *Indagationes Mathematicae (Proceedings)*, volume 55, pages 488–489. Elsevier, 1952.

[14] Molly Dunkum, Peter Hamburger, and Attila Pór. Destroying cycles in digraphs. *Combinatorica*, 31(1):55–66, 2011.

[15] Paul Erdős and Miklós Simonovits. A limit theorem in graph theory. In *Studia Sci. Math. Hung.* Citeseer, 1965.

[16] Kenneth Falconer and Pertti Mattila. Strong marstrand theorems and dimensions of sets formed by subsets of hyperplanes, 2015.

[17] Victor Falgas-Ravry and Emil R. Vaughan. Turán h-densities for 3-graphs. *Electr. J. Comb.*, 19(3):40, 2012.

[18] Jacob Fox, János Pach, and Andrew Suk. A polynomial regularity lemma for semi-algebraic hypergraphs and its applications in geometry and property testing. *CoRR*, abs/1502.01730, 2015.

[19] Lior Gishboliner and Asaf Shapira. Efficient removal without efficient regularity. *ArXiv e-prints*, 2017.

[20] Lior Gishboliner and Asaf Shapira. Removal lemmas with polynomial bounds. In *Proceedings of the 49th Annual ACM SIGACT Symposium on Theory of Computing, STOC 2017, Montreal, QC, Canada, June 19-23, 2017*, pages 510–522, 2017.

[21] Xavier Goaoc, Alfredo Hubard, Rémi de Joannis de Verclos, Jean-Sébastien Sereni, and Jan Volec. Limits of order types. In *31st International Symposium on Computational Geometry, SoCG 2015, June 22-25, 2015, Eindhoven, The Netherlands*, pages 300–314, 2015.

[22] Oded Goldreich, Shari Goldwasser, and Dana Ron. Property testing and its connection to learning and approximation. *J. ACM*, 45(4):653–750, July 1998.

[23] Oded Goldreich and Luca Trevisan. Three theorems regarding testing graph properties. *Random Struct. Algorithms*, 23(1):23–57, 2003.

[24] Timothy Gowers. Lower bounds of tower type for szemerédi's uniformity lemma. *GAFA*, 7(2):322–337, May 1997.

[25] Peter Hamburger, Penny E. Haxell, and Alexandr V. Kostochka. On directed triangles in digraphs. *Electr. J. Comb.*, 14(1), 2007.

[26] Yahya Ould Hamidoune. An application of connectivity theory in graphs to factorizations of elements in groups. *Eur. J. Comb.*, 2(4):349–355, 1981.

[27] Jan Hladký, Daniel Král, and Sergey Norin. Counting flags in triangle-free digraphs. *Electronic Notes in Discrete Mathematics*, 34:621–625, 2009.

[28] Krishnam Raju Jampani and Anna Lubiw. Simultaneous interval graphs. In *Algorithms and Computation - 21st International Symposium, ISAAC 2010, Jeju Island, Korea, December 15-17, 2010, Proceedings, Part I*, pages 206–217, 2010.

[29] Svante Janson. Standard representation of multivariate functions on a general probability space. *Elect. Comm. Probab.*, 14:343–346, 2009.

[30] Svante Janson. Graphons, cut norm and distance, couplings and rearrangements. *ArXiv e-prints*, September 2010.

[31] Henri Lebesgue. *Leçons Sur L'integration Et la Recherche Des Fonctions Primitives.* Collection de monographies sur la théorie des fonctions. Gauthier-Villars, 1904.

[32] László Lovász. *Large networks and graph limits*, volume 60 of *American Mathematical Society Colloquium Publications.* American Mathematical Society, Providence, RI, 2012.

[33] László Lovász and Balázs Szegedy. Limits of dense graph sequences. *J. Comb. Theory, Ser. B*, 96(6):933–957, 2006.

[34] László Lovász and Balázs Szegedy. Limits of dense graph sequences. *Journal of Combinatorial Theory, Series B*, 96(6):933 – 957, 2006.

[35] John M. Marstrand. Some fundamental geometrical properties of plane sets of fractional dimensions. *Proc. London Math. Soc. (3)*, 4:257–302, 1954.

[36] Pertti Mattila. Hausdorff dimension, orthogonal projections and intersections with planes. *Ann. Acad. Sci. Fenn. Ser. A I Math.*, 1(2):227–244, 1975.

[37] Brendan D. McKay. Practical graph isomorphism. *Congressus Numerantium*, 30:45–87, 1981.

[38] Brendan D. McKay and Adolfo Piperno. Practical graph isomorphism, ii. *J. Symb. Comput.*, 60:94–112, 2014.

[39] Nicolai E. Mnëv. Varieties of combinatorial types of projective configurations and convex polytopes. *Doklady Akademii Nauk SSSR*, 283(6):1312–1314, 1985.

[40] Alexander A. Razborov. Flag algebras. *J. Symb. Log.*, 72(4):1239–1282, 2007.

[41] Jian Shen. Directed triangles in digraphs. *J. Comb. Theory, Ser. B*, 74(2):405–407, 1998.

[42] Roman Sikorski. Funkcje rzeczywiste I. *Państwowe Wydawnictwo Naukowe, Warszawa*, 1958.

[43] Elias M. Stein and Rami Shakarchi. *Real Analysis: Measure Theory, Integration, and Hilbert Spaces (Princeton Lectures in Analysis) (Bk. 3)*. Princeton University Press, 2005.