Algebraic Phenomena in Combinatorics: Shattering-Extremal Families and the Combinatorial Nullstellensatz

by

Tamás Mészáros

Submitted to

Central European University

Department of Mathematics and its Applications

In partial fulfillment of the requirements for the degree of

Doctor of Philosophy

Supervisor: Professor Lajos Rónyai

Budapest, Hungary

2015

Acknowledgments

First and foremost I would like to thank my supervisor, Lajos Rónyai, for initiating me into this beautiful topic on the border of extremal combinatorics and algebra, and for proposing me this instructive project. His ideas, advice and guiding in the research during the past few years had a fundamental contribution to this dissertation itself as well as to my professional development in general.

I also would like to thank the Central European University for supporting me during my PhD studies and the whole staff for the assistance in various issues.

Finally let me express my gratitude to my family, in particular to my wife Franciska and to my mother for their support and for undertaking difficulties that arise naturally during such a project.

Abstract

This PhD dissertation is based on a research that originates from extremal combinatorics. In the first part we consider the problem of characterizing shattering-extremal set systems and extremal vector systems. We propose two different approaches, an algebraic and a graph theoretical one, and prove several characterisations of these extremal structures. The algebraic approach uses the standard monomials and Gröbner bases of vanishing ideals of finite point sets, while the key elements of the graph theoretical approach are the inclusion graphs of set systems. The second part of the dissertation is devoted to Noga Alon's famous Combinatorial Nullstellesatz and Non-vanishing Theorem. We prove generalizations of these results in different directions. First we introduce a version for multisets, then we consider the problem over arbitrary commutative rings instead of fields. At the end we investigate the problem of determining which finite sets X, beside discrete boxes, admit a version of the Combinatorial Nullstellensatz and the Non-vanishing Theorem.

Table of Contents

Acknowledgments **ii**

Abstract **iii**

List of Figures **v**

1 Introduction **1**

1.1 Historical background 1

1.2 Introduction to the theory 3

1.3 Overview of the thesis 5

2 Preliminaries **8**

2.1 Standard monomials and Gröbner bases 8

2.2 Shattering and strong shattering 17

2.3 Some set system operations 19

I Shattering-extremal set systems **24**

3 An algebraic approach to shattering **25**

3.1 Order shattering 25

3.2 Shattering and standard monomials 26

4 Extremality in the general case **31**

5 Graph theoretical characterization of extremality **39**

5.1 Extremal families of VC dimension 1 42

5.2 Extremal families of VC dimension 2 50

5.3 The eliminability conjecture 61

II Alon's Combinatorial Nullstellensatz **66**

6 The Combinatorial Nullstellensatz and the Non-vanishing Theorem **67**

6.1 The Non-vanishing Theorem for multisets 70

6.2 The Combinatorial Nullstellensatz and the Non-vanishing Theorem over commutative rings 72

6.3 The Combinatorial Nullstellensatz and the Non-vanishing Theorem for balanced systems 77

Bibliography **87**

List of Figures

5.1 The inclusion graph of $\{\{2\},\{1,5\},\{2,5\},\{1,2,5\},\{2,4,5\},\{2,3,4,5\}\}$ . . 40

5.2 Step A 52

5.3 Step B 53

5.4 Case of Step B 54

5.5 Example of the building process in $\mathcal{E}$ 55

5.6 Path of 4 cycles 55

5.7 Case $l > 1$ 57

5.8 Case $\ell = 1$ 58

5.9 Building up extremal set systems 60

Chapter 1

Introduction

1.1 Historical background

Extremal combinatorics is one of the central branches of discrete mathematics. It is not just an independent mathematical discipline, but appears also in several other mathematical areas. It deals with the problem of estimating the size of a combinatorial structure satisfying certain requirements by providing a lower or upper bound. In case this bound is sharp, it also tries to characterize somehow the extremal examples. It the particular case when the structures considered are set systems over a finite ground set, we talk about extremal set theory. A good survey on the topic in general is provided by [21] and [27].

In the past many results in this field were obtained mainly by ingenuity and detailed reasoning, however during the last few decades this field has experienced an impressive growth and some very efficient tools were developed. One of the main general techniques that played an important role in this development was the application of algebraic methods. Among these methods one of the first and probably the best-known are the linear algebra methods, see [9] for a good survey. Here we deal with a relatively new collection belonging to the family of polynomial methods, see [48] for a recent survey. It borrows some of the philosophy of algebraic geometry. Given a combinatorial object $\mathcal{C}$, we fix a field $\mathbb{F}$, represent $\mathcal{C}$ as a finite set of points in the affine space $\mathbb{F}^n$, and study its vanishing ideal $I(\mathcal{C})$ instead of the combinatorial object itself.

When studying such polynomial ideals, it turned out, that standard monomials and Gröbner bases can be very useful. While standard monomials form a nice basis of the $\mathbb{F}$-vector space $\mathbb{F}[x_1, \dots, x_n]/I(\mathcal{C})$, Gröbner bases are special systems of generators of the ideal. They were introduced in 1965 by Austrian mathematician Bruno Buchberger in his Ph.D. thesis, and named after his supervisor, Wolfgang Gröbner. He was motivated by questions from commutative algebra and algebraic geometry, but since then, Gröbner bases have been applied in various fields of mathematics e.g. coding theory, symbolic computation, automatic theorem proving, integer programming, statistics, partial differential equations and numerical computations. A good survey is provided by [1] and [13].

The combinatorial structures we will examine using the above technique are shattering-extremal set systems. They are set systems that are extremal with respect to the well known Sauer inequality that was first proved in the 70's. Since then many interesting results have been obtained in connection with these combinatorial objects, among others by Bollobás, Leader and Radcliffe in [11], by Bollobás and Radcliffe in [12], by Frankl in [21]. Füredi and Quinn in [23], and recently Kozma and Moran in [33] provided interesting examples of shattering-extremal set systems. However for the time being shattering-extremal set systems do not have a good structural description, ... *a structural description of extremal systems is still sorely lacking.*" (Bollobáas, Radcliffe 1995).

Probably the best-known member of the family of polynomial methods is Noga Alon's famous Combinatorial Nullstellensatz and the resulting Non-vanishing Theorem. Since its publication in 1999 it became one of the most powerful algebraic tools in combinatorics. It has several beautiful and strong applications, see [3] for some classical ones and [14], [18],

[25], [28], [29], [30], [41], [46] for some recent ones.

1.2 Introduction to the theory

To get a better overview of what is contained in this thesis, let us shortly (and sometimes informally) introduce the basic concepts.

Extremal set and vector systems

We say that a set system $\mathcal{F} \subseteq 2^{[n]}$ shatters a given set $S \subseteq [n]$ if $2^S = \{F \cap S : F \in \mathcal{F}\}$. The size of the largest set shattered by $\mathcal{F}$ is called the Vapnik-Chervonenkis dimension of $\mathcal{F}$. The Sauer inequality states that in general, a set system $\mathcal{F}$ shatters at least $|\mathcal{F}|$ sets. Here we concentrate on the case of equality. A set system is called shattering-extremal if it shatters exactly $|\mathcal{F}|$ sets. Our aim is to characterize somehow shattering-extremal families.

Let $\mathbb{F}$ be an arbitrary field. When representing sets by their characteristic vectors, then a set system $\mathcal{F} \subseteq 2^{[n]}$ can also be viewed as a set of $0-1$ vectors in the affine space $\mathbb{F}^n$. In this way one can study the vanishing ideal $I(\mathcal{F})$ instead of the set system itself, which consists of all polynomials from $\mathbb{F}[\mathbf{x}] = \mathbb{F}[x_1, \dots, x_n]$ that vanish at all of the characteristic vectors of elements of $\mathcal{F}$.

In the study of these ideals, our key tools are standard monomials and Gröbner bases. To define them, one first needs a term order - a complete ordering of the monomials of $\mathbb{F}[\mathbf{x}]$, with some additional properties. As an example one can take the standard lexicographic ordering of monomials. By reordering the variables, we can get another lexicographic order, so we will talk about a lexicographic term order based on some permutation of the variables $x_1, x_2, \dots, x_n$. The greatest monomial of a polynomial with respect to some term order is called its leading monomial. In the univariate case, the leading monomial is simply the highest degree term of the polynomial. Given a polynomial ideal $I \lhd \mathbb{F}[\mathbf{x}]$, a monomial is called standard, if it is not the leading monomial of a any polynomial $f \in I$. We write $Sm(I)$ for the collection of all standard monomials for the ideal I. Standard monomials possess some very nice properties, among others they form an $\mathbb{F}$-vector space basis of $\mathbb{F}[\mathbf{x}]/I$. A finite subset G of I is a Gröbner basis of I, if the leading monomial of any nonzero $f \in I$ is divisible by the leading monomial of some $g \in G$. Then G also generates the ideal in a 'nice' manner. For more details on Gröbner bases see [1].

Given a family $\mathcal{F} \subseteq 2^{[n]}$, by investigating the standard monomials of the vanishing ideal $I(\mathcal{F})$ for different lexicographic term orders, one can prove a very useful characterization of shattering-extremality, namely that a set system is shattering-extremal if and only if the family of standard monomials of its vanishing ideal is the same for every lexicographic term order. Based on this characterization one can fully describe the reduced Gröbner bases of shattering-extremal set systems, it results an efficient algorithm for testing extremality and also allows us to generalize the notion of extremality for general finite sets of vectors, not merely set systems. We define a finite system $\mathcal{V} \subseteq \mathbb{F}^n$ to be extremal if the family of standard monomials of the vanishing ideal $I(\mathcal{V})$ is the same for every lexicographic term order.

Beside the previously outlined algebraic one, we can take a graph theoretical approach to shattering-extremal set systems as well. The inclusion graph of a set system is a directed edge-labelled graph whose vertices are the elements of the set system, and there is a directed edge going from G to F with label j exactly if $F = G \cup \{j\}$. Inclusion graphs offer a good framework to study different properties of set systems. In case of shattering-extremal set systems they possess some very nice properties, for example they are always connected.

Alon's Combinatorial Nullstellensatz and Non-vanishing Theorem

Let $S_1, S_2, \dots, S_n$ be finite nonempty subsets of an arbitrary field $\mathbb{F}$ and let

$$\mathbf{S} = S_1 \times S_2 \times \cdots \times S_n \subseteq \mathbb{F}^n.$$

For $i = 1, \dots, n$ put $g_i = g_i(x_i) = \prod_{s \in S_i}(x_i - s) \in \mathbb{F}[\mathbf{x}]$. Alon's Combinatorial Nullstellensatz (Theorem 1.1 in [3]) is a specialized and strengthened version of the Hilbertsche Nullstellensatz for the vanishing ideal $I(\mathbf{S})$. It states that if a polynomial $f(\mathbf{x}) \in \mathbb{F}[\mathbf{x}]$ vanishes over all the common zeros of $g_1, \dots, g_n$ (i.e. $f \in I(\mathbf{S})$), then there are polynomials $h_1, \dots, h_n \in \mathbb{F}[\mathbf{x}]$, satisfying some degree bounds, so that $f = \sum_{i=1}^{n} h_i g_i$.

This result can be easily rephrased using Gröbner bases. In Gröbner terminology it states that the polynomials $g_1, \dots, g_n$ form a universal Gröbner basis of $I(\mathbf{S})$.

From this, a simple and widely applicable non-vanishing criterion has been deduced. It provides a sufficient condition for a polynomial $f \in \mathbb{F}[\mathbf{x}]$ for not vanishing everywhere on the discrete box $\mathbf{S}$. Let $t_1, \dots, t_n$ be nonnegative integers such that $|S_i| > t_i$ for every i. Given a polynomial $p \in \mathbb{F}[\mathbf{x}]$ of degree $\sum_{i=1}^{n} t_i$, suppose that the coefficient of the monomial $x_1^{t_1} \cdots x_n^{t_n}$ in p is not 0. Then there is some vector $\mathbf{s} \in \mathbf{S}$ such that $p(\mathbf{s}) \neq 0$.

1.3 Overview of the thesis

After the Introduction, in Chapter 2, all the necessary concepts and theorems needed to understand the subsequent parts are collected. We discuss standard monomials, Gröbner bases and shattering-extremal set systems in detail. Right away comes the essential portion of the thesis divided into two separate parts.

Part I deals with shattering-extremal set systems and extremal vector systems, and contains 3 chapters. The first one, Chapter 3, summarizes results concerning the algebraic approach to shattering-extremal set systems. We present a characterization using standard monomials and describe the reduced Gröbner bases of the vanishing ideals of such systems. The chapter also contains an efficient algorithm for determining the extremality of a set system. Next, Chapter 4 follows, where we prove the generalizations of the previously mentioned results for arbitrary finite sets of vectors instead of set systems. As a corollary of these results we also give a new proof for a result of Li, Zhang and Dong from [34], where they investigated the standard monomials of zero dimensional polynomial ideals. In the last chapter of this part, Chapter 5, we take a different approach and investigate the inclusion graphs of shattering-extremal set systems. Although this method is not at all an algebraic one, for the sake of completeness it is still included in this thesis and the results provide a better understanding of the structure of these combinatorial objects. The general task of giving a good description of the inclusion graphs of shattering-extremal set systems seems to be too complex at this point. We restrict therefore our attention to the simplest cases, where the Vapnik-Chervonenkis dimension is bounded by some fixed natural number t. First we characterize in a nice way the inclusion graphs of shattering-extremal set systems of Vapnik-Chervonenkis dimension 1. This allows us to relate things to earlier chapters and to compute a Gröbner basis of the vanishing ideal of such set systems. Next, shattering-extremal families of Vapnik-Chervonenkis dimension 2 follow. We give an algorithmic procedure for constructing the inclusion graphs of all such set systems. To finish this chapter we study a conjecture about the eliminability of elements from a shattering-extremal set system in such a way that the resulting system is still shattering-extremal. There are several examples where the conjecture holds, e.g. our results provide a positive answer for every shattering-extremal set system of Vapnik-Chervonenkis dimension at most 2.

Part II consists of one single chapter, Chapter 6, and is devoted to generalizations of the Combinatorial Nullstellensatz and the Non-vanishing Theorem. Both theorems offer various directions to look for generalizations. After introducing the original theorems of Alon, for the sake of completeness, in Section 6.1 we present a generalization of the Non-vanishing Theorem for multisets - discrete boxes where we add multiplicities to the elements in some natural way. Next, in Section 6.2 we consider variants over arbitrary commutative rings, not merely fields, and as an application we present a generalization of Theorem 6.3 from [3]. The last section of this chapter, Section 6.3, contains a possible generalization of the Combinatorial Nullstellensatz for balanced systems - vector systems with the property that independently of how we fix the last some coordinates, we get the same number of vectors.

The present thesis is mostly based on our papers [31], [36], [37] and [38]. We tried to make the thesis itself as coherent as possible and to select topics in an order that is the easiest to understand. Some parts, mostly from Chapter 4, have not been published yet.

Chapter 2

Preliminaries

Before getting started with the main definitions, we introduce some notation. Throughout the thesis $\mathbb{F}$ will stand for an ordinary field, and n will be a positive integer. The set $\{1, 2, \ldots, n\}$ will be referred to shortly as $[n]$, its powerset as $2^{[n]}$ and the collection of k-sets by $\binom{[n]}{k}$. Vectors of length n will be denoted by boldface letters, and we denote their coordinates by the same letter indexed by respective numbers, for example $\mathbf{y} = (y_1, \ldots, y_n) \in \mathbb{F}^n$. For the ring of polynomials in n variables over $\mathbb{F}$ we will use the usual notation $\mathbb{F}[x_1, \ldots, x_n] = \mathbb{F}[\mathbf{x}]$. To shorten our notation, for a polynomial $f(x_1, \ldots, x_n)$ we will write $f(\mathbf{x})$. If $\mathbf{w} \in \mathbb{N}^n$, we write $\mathbf{x}^{\mathbf{w}}$ for the monomial $x_1^{w_1} \ldots x_n^{w_n} \in \mathbb{F}[\mathbf{x}]$. For a subset $M \subseteq [n]$, the monomial x_M will be $\prod_{i \in M} x_i$ (and $x_\emptyset = 1$).

2.1 Standard monomials and Gröbner bases

Term orders

To start with, we will need the notion of term orders.

Definition 2.1. *A relation $\prec$ on the monomials from $\mathbb{F}[x_1, \ldots, x_n]$ is called a term order if it is a linear order with 1 as minimal element and it is monotone with respect to multiplication.*

An example of a term order is the standard lexicographic (lex for short) ordering of monomials. We say that $\mathbf{x}^w$ is smaller than $\mathbf{x}^u \neq \mathbf{x}^w$ according to the standard lexicographic order if for the first index i such that $w_i \neq u_i$, we have $w_i < u_i$. This is clearly a term order, and we have $x_1 \leq x_2 \leq \cdots \leq x_n$.

For example for $n = 2$ the lexicographic ordering of the first few monomials is the following:

$$1 \prec x_2 \prec x_2^2 \prec x_2^3 \prec \cdots \prec x_1 \prec x_1x_2 \prec x_1x_2^2 \prec \cdots \prec x_1^2 \prec x_1^2x_2 \prec x_1^2x_2^2 \prec \ldots$$

By reordering the variables, we can get another lexicographic term order. Let $\pi \in S_n$ be an arbitrary permutation, where S_n denotes the symmetric group of order n. $\mathbf{x}^w$ is smaller than $\mathbf{x}^u \neq \mathbf{x}^w$ according to the lex order based on π if for the first index i such that $w_{\pi(i)} \neq u_{\pi(i)}$, we have $w_{\pi(i)} < u_{\pi(i)}$. In this way one can obtain $n!$ different lex orders. As a special case, choosing π to be the identity permutation, we get the standard lex order back.

Term orders in general posses two fundamental properties:

Proposition 2.2. *([1, Proposition 1.4.5 and Theorem 1.4.6]) Every term order $\prec$ is the refinement of the divisibility between monomials and is a well-ordering.*

Proof. For the first part, let us suppose that $\mathbf{x}^u | \mathbf{x}^v$. Then $\frac{\mathbf{x}^v}{\mathbf{x}^u}$ is a monomial as well, so $1 \preceq \frac{\mathbf{x}^v}{\mathbf{x}^u}$. And then when multiplying by $\mathbf{x}^u$ we get the desired inequality.

For the proof of the second part see [1, Theorem 1.4.6]. $\square$

The second part of Proposition 2.2 about term orders being a well ordering is known in the literature as Dickson's lemma.

Leading and standard monomials

From now on, if not stated otherwise, we will always assume that we are given a fixed term order $\prec$. The leading monomial $lm(f(\mathbf{x}))$ of a nonzero polynomial $f(\mathbf{x}) \in \mathbb{F}[\mathbf{x}]$ is the greatest monomial according to $\prec$ appearing in $f(\mathbf{x})$ with nonzero coefficient, when written as the usual linear combination of monomials. The leading monomial of a polynomial $f(\mathbf{x})$ together with its coefficient is called the leading term of $f(\mathbf{x})$ and is denoted by $lt(f(\mathbf{x}))$. For an ideal of polynomials $I \triangleleft \mathbb{F}[\mathbf{x}]$ we denote by $Lm(I)$ the set of leading monomials of the polynomials in I:

$$Lm(I) = \{lm(f(\mathbf{x})) \mid f(\mathbf{x}) \in I, f \neq 0\}.$$

A monomial which is not a leading monomial of any polynomial in I is called a standard monomial. The set of standard monomials is denoted by $Sm(I)$:

$$Sm(I) = \{\mathbf{x}^{\mathbf{w}} \in \mathbb{F}[\mathbf{x}]\} \backslash Lm(I) = \{\mathbf{x}^{\mathbf{w}} \mid \nexists f(\mathbf{x}) \in I, \text{ for which } lm(f) = \mathbf{x}^{\mathbf{w}}\}.$$

The standard monomials will be important tools in our arguments.

Definition 2.3. *A set $S \subseteq \{\mathbf{x}^{\mathbf{w}} \in \mathbb{F}[\mathbf{x}]\}$ is downward (upward) closed with respect to divisibility or shortly a down-set (up-set) if $\mathbf{x}^{\mathbf{w}} \in S$ and $\mathbf{x}^{\mathbf{u}} | \mathbf{x}^{\mathbf{w}}$ ($\mathbf{x}^{\mathbf{w}} | \mathbf{x}^{\mathbf{u}}$) imply that $\mathbf{x}^{\mathbf{u}} \in S$.*

It is easy to see by definition that $Lm(I)$ is an up-set and hence $Sm(I)$ is a down-set. The set of standard monomials has a very important property:

Proposition 2.4. *([35, Proposition 2.3]) The canonical image of $Sm(I)$ is a basis of $\mathbb{F}(\mathbf{x})/I$ as an $\mathbb{F}$ vector space.*

Corollary 2.5. *Each leading monomial* $\mathbf{x^u}$ *has a representation by standard monomials, i.e. there are standard monomials* $x^{\mathbf{u}_1}, \ldots, x^{\mathbf{u}_\ell}$ *and coefficients* $\alpha_1, \ldots, \alpha_\ell \in \mathbb{F}$ *such that* $\mathbf{x^u} + \sum\limits_{i=1}^{\ell} \alpha_i x^{\mathbf{u}_i} \in I$*.*

Proof. If we consider $\mathbf{x^u}$ as an element of $\mathbb{F}(\mathbf{x})/I$, Proposition 2.4 implies that it is an $\mathbb{F}$-linear combination of the $x^{\mathbf{u}_1}$'s. Hence there are coefficients $\beta_1, \ldots, \beta_\ell \in \mathbb{F}$ such that $\mathbf{x^u} = \sum\limits_{i=1}^{\ell} \beta_i x^{\mathbf{u}_i}$ in $\mathbb{F}(\mathbf{x})/I$. However this happens if and only if $\mathbf{x^u} - \sum\limits_{i=1}^{\ell} \beta_i x^{\mathbf{u}_i} \in I$. Putting $\alpha_i = -\beta_i$ we get the desired result. □

Gröbner bases

When working with a polynomial ideal, a nice ideal basis can facilitate things. An important example of such nice bases are Gröbner bases.

Definition 2.6. *Let* I *be a nonzero ideal of* $\mathbb{F}[\mathbf{x}]$*. For a fixed term order* $\prec$*, a finite subset* $G \subseteq I$ *is a Gröbner basis of* I *if for every nonzero* $f \in I$ *there exists* $g \in G$ *such that* $lm(g)$ *divides* $lm(f)$*.*

Consider two polynomials $f, g \in \mathbb{F}[\mathbf{x}]$, and suppose that there is one monomial $\mathbf{x^w}$ in f with nonzero coefficient c_f that is divisible by $lm(g)$. Let the coefficient of $lm(g)$ in g be c_g and let

$$\widehat{f}(\mathbf{x}) = f(\mathbf{x}) - \frac{c_f \cdot \mathbf{x^w}}{c_g \cdot lm(g)} g(\mathbf{x}).$$

Since the leading monomial of $\frac{\mathbf{x^w}}{lm(g)} g(\mathbf{x})$ is $\mathbf{x^w}$, in the above operation the term $c_f \mathbf{x^w}$ in f gets eliminated and is replaced by a linear combination of monomials strictly less than $\mathbf{x^w}$. This operation is called reduction.

If G is a finite set of polynomials and $f(\mathbf{x}) \in \mathbb{F}[\mathbf{x}]$ is an arbitrary polynomial, we say that f is reduced with respect to G if there is no monomial in f with nonzero coefficient that is divisible by $lm(g)$ for some $g \in G$. Now take an arbitrary f and reduce it with the elements of G, at each step replacing the greatest monomial with smaller ones, until we get a reduced polynomial with respect to G. Since there is no infinite downward chain of monomials starting with $lm(f)$, this process terminates in finitely many steps ending up with a factorization

$$f(\mathbf{x}) = \sum_{i=1}^{m} g_i(\mathbf{x})h_i(\mathbf{x}) + \widehat{f}(\mathbf{x}),$$

where $G = \{g_1, \ldots, g_m\}$, $h_1, \ldots, h_m \in \mathbb{F}[\mathbf{x}]$, $\widehat{f}$ is reduced with respect to G and $lm(g_i h_i) \preceq lm(f)$ for every index i. We say that $\widehat{f}$ is a remainder of f with respect to G if such polynomials $h_1, h_2, \ldots, h_m$ exist.

Example 2.7. Let $g_1(x_1, x_2) = x_1^2x_2^2 + x_1$, $g_2(x_1, x_2) = x_1^2x_2^2 + x_2$, $G = \{g_1, g_2\}$, $\prec$ the standard lexicographic order and consider $f(x_1, x_2) = x_1^2x_2^2$. If we reduce f with g_1 we get $-x_1$, if with g_2 we get $-x_2$. It is easy to see that both, $-x_1$ and $-x_2$ are reduced with respect to G.

This example shows that in general the remainder of a polynomial $f(\mathbf{x})$ is not necessary unique with respect to some fixed family of polynomials G.

Gröbner bases have several characterizations, the next proposition contains two of them.

Proposition 2.8. *([1, Theorem 1.6.2 and Theorem 1.6.7]) Let $I \lhd \mathbb{F}[\mathbf{x}]$ be a nonzero ideal. Then the following statements are equivalent for a finite family $G \subseteq I$ of nonzero polynomials.*

(i) G is a Gröbner basis of I.

(ii) A polynomial $f \in \mathbb{F}[\mathbf{x}]$ belongs to I if and only if it can be reduced to 0 using G.

(iii) Every polynomial $f \in \mathbb{F}[\mathbf{x}]$ has a unique remainder with respect to G.

As a corollary one can prove the following important property of Gröbner bases.

Proposition 2.9. *([1, Corollary 1.6.3]) If a finite family G of nonzero polynomials is a Gröbner basis of the ideal $I \lhd \mathbb{F}[\mathbf{x}]$, then $I = \langle G \rangle$.*

The question naturally arises, whether every nonzero ideal $I \lhd \mathbb{F}[\mathbf{x}]$ has a Gröbner basis with respect to a fixed term order $\prec$. The answer is fortunately yes. A possible way to prove this is Buchberger's algorithm, which explicitly constructs a Gröbner basis of an ideal $I \lhd \mathbb{F}[\mathbf{x}]$.

Definition 2.10. *The S-polynomial of nonzero polynomials $f, g \in \mathbb{F}[\mathbf{x}]$ is*

$$S(f,g) = \frac{L}{lt(f)} f - \frac{L}{lt(g)} g,$$

where L is the least common multiple of the monomials $lm(f)$ and $lm(g)$.

Buchberger's theorem gives us another characterization of Gröbner bases.

Proposition 2.11. *([1, Theorem 1.7.4]) A finite family G of nonzero polynomials $g_1, \dots, g_t$ is the Gröbner basis of the ideal $I = \langle g_1, \dots, g_t \rangle$ if and only if for all $1 \le i < j \le t$ the S-polynomial $S(g_i, g_j)$ can be reduced to 0 using G.*

Accordingly, Buchberger's algorithm starts from a generating set G of I. In each step it reduces using G the S-polynomial of two polynomials from G and if the remainder is not 0, then adds it to G. The main point is to prove that this process finishes after finitely many steps.

Proposition 2.12. *([1, Theorem 1.7.8]) Given a finite family of nonzero polynomials $\{f_1, \dots, f_s\}$, Buchberger's algorithm will produce in finitely many steps a Gröbner basis G of the ideal $I = \langle f_1, \dots, f_s \rangle$.*

We remark that although Buchberger's algorithm computes a Gröbner basis of the ideal I, the size of G may be exponentially large in s and so from a practical point of view the algorithm is not really efficient. However, a recent improved version from [17] could produce results for surprisingly large examples as well.

The Gröbner basis G of a nonzero ideal I is not unique. For example by adding finitely many polynomials from I to G, the resulting set of polynomials will also be a Gröbner basis of I. For uniqueness we need a bit more.

Definition 2.13. *If G is a Gröbner basis of some nonzero ideal $I \lhd \mathbb{F}[\mathbf{x}]$, and every polynomial $g \in G$ has leading coefficient 1 and is reduced with respect to $G \setminus \{g\}$, then G is called a reduced Gröbner basis of I.*

Reformulating this, we get that a Gröbner basis G is reduced if and only if every polynomial $g \in G$, apart from $lm(g)$, consists only of standard monomials and has 1 as leading coefficient.

Proposition 2.14. *([1, Theorem 1.8.7]) Every nonzero ideal I has a unique reduced Gröbner basis with respect to any fixed term order.*

Standard monomials of vanishing ideals of finite point sets

Let $\mathcal{V} \subseteq \mathbb{F}^n$ be a finite set of vectors, and denote by $I(\mathcal{V})$ the set of polynomials vanishing on $\mathcal{V}$, i.e.:

$$I(\mathcal{V}) = \{f(\mathbf{x}) \in \mathbb{F}[\mathbf{x}] \mid f(\mathbf{y}) = 0 \text{ for all } \mathbf{y} \in \mathcal{V}\}.$$

It is easy to see that $I(\mathcal{V})$ is an ideal in $\mathbb{F}[\mathbf{x}]$, it is called the vanishing ideal of $\mathcal{V}$. According to Proposition 2.4 for any fixed term order $Sm(I(\mathcal{V}))$ is a basis of $\mathbb{F}(\mathbf{x})/I(\mathcal{V})$ as an $\mathbb{F}$ vector space. On the other hand as $\mathcal{V}$ is finite, by interpolation we get that this factor space is isomorphic to the space of functions from $\mathcal{V}$ to $\mathbb{F}$, which clearly has dimension $|\mathcal{V}|$. In this way we obtain the following:

Proposition 2.15. *Let $\mathcal{V} \subseteq \mathbb{F}^n$ be a finite set of vectors. Then $|Sm(I(\mathcal{V}))| = |\mathcal{V}|$ for every term order.* $\square$

For a finite set of vectors $\mathcal{V} \subseteq \mathbb{F}^n$ let $A \subseteq \mathbb{F}$ be the collection of all elements from $\mathbb{F}$ that appear as coordinates, i.e. A is the smallest subset of $\mathbb{F}$ for which $\mathcal{V} \subseteq A^n$. Denote the size of A by k.

Felszeghy, Ráth and Rónyai in [19] gave a combinatorial description of the standard monomials of vanishing ideals of finite point sets by introducing a two player game, called the Lex game. This combinatorial description has many interesting corollaries. Using it, one can prove for example, that the degree of any variable in a lexicographic standard monomial of $I(\mathcal{V})$ can be at most $k-1$, i.e.

$$Sm(I(\mathcal{V})) \subseteq \{\mathbf{x}^{\mathbf{w}} : \mathbf{w} \in \{0, 1, \ldots, k-1\}^n\}.$$

Further, the Lex game also shows that given the lexicographic standard and leading monomials of the vanishing ideal $I(\mathcal{V})$ one can compute them for $I(\mathcal{V}^c)$ as well, where $\mathcal{V}^c = A^n \setminus \mathcal{V}$.

Proposition 2.16. *([19, Corollary 9]) For every exponent vector* $\mathbf{w} \in \{0, 1, \ldots, k-1\}^n$ *we have* $x_1^{w_1} \ldots x_n^{w_n} \in Sm(I(\mathcal{V}))$ *if and only if* $x_1^{k-1-w_1} \ldots x_n^{k-1-w_n} \in Lm(I(\mathcal{V}^c))$.

Now suppose that $\widehat{\mathbb{F}}$ is an arbitrary field, and for $j = 1, 2, \ldots, n$ take injective functions $\varphi_j : A \to \widehat{\mathbb{F}}$. Let $\widehat{\mathcal{F}}$ be the image of $\mathcal{V}$ under the action of $(\varphi_1, \ldots, \varphi_n)$, i.e.

$$\widehat{\mathcal{V}} = \{(\varphi_1(v_1), \ldots, \varphi_n(v_n)) \mid (v_1, \ldots, v_n) \in \mathcal{V}\}.$$

The following proposition is also a direct consequence of the Lex game:

Proposition 2.17. *([19, Corollary 7]) The standard monomials of* $I(\mathcal{V}) \lhd \mathbb{F}[\mathbf{x}]$ *are the same as those of* $I(\widehat{\mathcal{V}}) \lhd \widehat{\mathbb{F}}[\mathbf{x}]$ *for any lexicographic term order.*

However the most important corollary of the Lex game is an efficient algorithm that calculates the standard monomials with respect to any lex order in essentially linear time.

Proposition 2.18. *([19, Theorem 14]) Let $\mathcal{V} \subseteq \mathbb{F}^n$ be a finite set of vectors and $\prec$ an arbitrary lex order. Also let A be the smallest subset of $\mathbb{F}$ for which $\mathcal{V} \subseteq A^n$ and denote its size by k. Then there is an algorithm that computes $Sm(I(\mathcal{V}))$ in $O(n|\mathcal{V}|k)$ time. If we assume that there exists an ordering on the coordinate set A, which can be tested in constant time then the algorithm makes $O(n|\mathcal{V}|\log k)$ steps.*

Note that the natural size of the input of the algorithm is $n|\mathcal{V}|\log k$. This result shows, that from an algorithmic point of view, standard monomials are a promising tool, since they are as efficiently computable as possible.

Standard monomials of set systems

The characteristic vector $\mathbf{v}_F$ of a set $F \subseteq [n]$ is a $0-1$ vector of length n, whose ith entry is 1 if and only if $i \in F$. Let now $\mathcal{F} \subseteq 2^{[n]}$ be a set system. We will identify $\mathcal{F}$ with the collection of characteristic vectors of its elements. In this way it makes sense to consider the vanishing ideal $I(\mathcal{F})$ of $\mathcal{F} \subseteq \{0,1\}^n \subseteq \mathbb{F}^n$. The study of this ideal, in particular the study of the standard monomials and Gröbner bases of $I(\mathcal{F})$, turned out to be very useful when investigating combinatorial properties of $\mathcal{F}$.

The polynomial $x_i^2 - x_i$ is trivially a member of the ideal $I(\mathcal{F})$ for every index $i \in [n]$ and for every term order, hence $\{x_1^2, \ldots, x_n^2\} \subseteq Lm(I(\mathcal{F}))$. This latter fact implies that $Sm(I(\mathcal{F}))$ contains only square-free monomials. Note that any square-free monomial $\mathbf{m}$ can be uniquely written as $\mathbf{x}_H = \prod_{i\in H} x_i$ for an appropriate set $H \subseteq [n]$, and so, via the bijection $\mathbf{m} = \mathbf{x}_H \leftrightarrow H \subseteq [n]$, the family of standard monomials of a set system has also a set system representation, and conversely any set system can also be considered as a family of square-free monomials. It will always be clear from the context which representation is considered.

From the algorithmic point of view one can note that when computing standard monomials of finite set systems with respect to some lex order, then, as the size of the coordinate set is 2, the running time of the algorithm from Proposition 2.18 is $O(n|\mathcal{F}|)$, hence in this case it is linear.

2.2 Shattering and strong shattering

Shattering

Definition 2.19. *A set system $\mathcal{F} \subseteq 2^{[n]}$ shatters a given set $S \subseteq [n]$ if*

$$2^S = \{F \cap S : F \in \mathcal{F}\}.$$

The family of subsets of $[n]$ shattered by $\mathcal{F}$ is denoted by $Sh(\mathcal{F})$. The notion of shattering occurs in various fields of mathematics, such as combinatorics, statistics, computer science and logic. As an example, one can mention the Vapnik-Chervonenkis dimension of a set system $\mathcal{F}$.

Definition 2.20. *The Vapnik-Chervonenkis dimension of a set system $\mathcal{F} \subseteq 2^{[n]}$, VC dimension for short and denoted by $dim_{VC}(\mathcal{F})$, is the maximum cardinality of a set shattered by $\mathcal{F}$.*

The Vapnik-Chervonenkis dimension is a widely known and used notion, appearing in several areas of mathematics, among others in machine learning (see e.g. [10], [39], [20]) and probability theory (see e.g. [47]).

Definition 2.21. *A set system $\mathcal{F} \subseteq 2^{[n]}$ is downward (upward) closed with respect to inclusion or shortly a down-set (up-set) if $F \in \mathcal{F}$ and $H \subseteq F$ ($F \subseteq H$) imply that $H \in \mathcal{F}$.*

Clearly, by definition $Sh(\mathcal{F})$ is always a down-set, and if $\mathcal{F} \subseteq \mathcal{F}'$ are set systems then $Sh(\mathcal{F}) \subseteq Sh(\mathcal{F}')$. The following inequality deals with the size of $Sh(\mathcal{F})$ and states that in general, a set system $\mathcal{F}$ shatters at least $|\mathcal{F}|$ sets.

Proposition 2.22. *(See e.g. [7, Theorem 1.1].)* $|Sh(\mathcal{F})| \geq |\mathcal{F}|$ *for every set system* $\mathcal{F} \subseteq 2^{[n]}$.

This result, known as Sauer inequality, was proved by various authors (Aharoni and Holzman [2], Pajor [40], Sauer [43], Shelah [44]), and studied by many others.

We are interested in the case of equality, when a set system shatters exactly $|\mathcal{F}|$ sets.

Definition 2.23. *The set system* $\mathcal{F}$ *is called shattering-extremal if* $|Sh(\mathcal{F})| = |\mathcal{F}|$.

As an example let $\mathcal{F}$ be an arbitrary down-set. It is easy to see that in this case $Sh(\mathcal{F}) = \mathcal{F}$, and so by definition every down-set is shattering-extremal.

The main aim here is to characterize somehow shattering-extremal set systems. Before getting started with this, we first present an interesting results in connection with shattering.

Proposition 2.24. *([43, Theorem 1]) Let* $\mathcal{F}$ *be a family of subsets of* $[n]$ *with no shattered set of size* k*. Then*

$$|\mathcal{F}| \leq \sum_{i=0}^{k-1} \binom{n}{i},$$

and this inequality is best possible.

Proposition 2.24, also known as Sauer lemma, has found applications in a variety of contexts, including applied probability.

Strong shattering

In [11] and [12] a different version of shattering, strong shattering was introduced .

Definition 2.25. *A set system* $\mathcal{F} \subseteq 2^{[n]}$ *strongly shatters the set* $F \subseteq [n]$*, if there exists* $I \subseteq [n] \backslash F$ *such that*

$$2^F + I = \{H \cup I \mid H \subseteq F\} \subseteq \mathcal{F}.$$

The family of all sets strongly shattered by some set system $\mathcal{F}$ is denoted by $st(\mathcal{F})$. Clearly $st(\mathcal{F}) \subseteq Sh(\mathcal{F})$, $st(\mathcal{F})$ is also a down set, and similarly to $Sh(\mathcal{F})$ if $\mathcal{F} \subseteq \mathcal{F}'$ are set systems then $st(\mathcal{F}) \subseteq st(\mathcal{F}')$. Also note that $S \in st(\mathcal{F})$ exactly if $[n] \backslash S \notin Sh(2^{[n]} \backslash \mathcal{F})$. Using this duality for the size of $st(\mathcal{F})$ one can prove the so called reverse Sauer inequality:

Proposition 2.26. *([11, Theorem 1.1])* $|st(\mathcal{F})| \leq |\mathcal{F}|$ *for every set system* $\mathcal{F} \subseteq 2^{[n]}$.

In [11] the authors actually prove the statement for so called "order-convex" sets, but the same proof yields this general form of the reverse Sauer inequality as well.

This inequality would enable us to define a new type of extremality, however according to [12] this is not necessary.

Proposition 2.27. *([12, Theorem 2])* $\mathcal{F} \subseteq 2^{[n]}$ *is extremal with respect to the Sauer inequality (shattering-extremal) if and only if it is extremal with respect to the reverse Sauer inequality i.e.* $|st(\mathcal{F})| = |\mathcal{F}| \iff |Sh(\mathcal{F})| = |\mathcal{F}|$.

Since the two extremal cases coincide, we will call such set systems shortly just extremal. As a consequence of the above facts, we obtain that for extremal set systems we have $st(\mathcal{F}) = Sh(\mathcal{F})$.

2.3 Some set system operations

Definition 2.28. *The standard subdivision of a set system* $\mathcal{F} \subseteq 2^{[n]}$ *with respect to an element* $i \in [n]$ *consists of the following two set systems:*

$$\mathcal{F}_0^{(i)} = \{F \mid F \in \mathcal{F} \text{ and } i \notin F\} \subseteq 2^{[n] \backslash \{i\}},$$
$$\mathcal{F}_1^{(i)} = \{F \backslash \{i\} \mid F \in \mathcal{F} \text{ and } i \in F\} \subseteq 2^{[n] \backslash \{i\}}.$$

For a pair $A \subseteq B \subseteq [n]$ of sets let

$$\mathcal{F}_{A,B} = \{F \backslash A \mid F \in \mathcal{F},\ A \subseteq F \subseteq B\}$$

and

$$\widehat{\mathcal{F}}_{A,B} = \{F \mid F \in \mathcal{F},\ A \subseteq F \subseteq B\} \subseteq \mathcal{F}.$$

With the above definitions in mind, note that

$$\mathcal{F}_0^{(i)} = \mathcal{F}_{\emptyset,[n]\backslash\{i\}}, \qquad \mathcal{F}_1^{(i)} = \mathcal{F}_{\{i\},[n]},$$

and if $A = \{i_1, \ldots, i_k\} \subseteq B \subseteq [n]$ and $[n]\backslash B = \{j_1, \ldots, j_\ell\}$ then

$$\mathcal{F}_{A,B} = (\cdots(((\cdots(\mathcal{F}_1^{(i_1)})_1^{(i_2)}\cdots)_1^{(i_k)})_0^{(j_1)})\cdots)_0^{(j_\ell)}.$$

The standard subdivision of a set system can be used to prove Proposition 2.22. For the sake of completeness we provide a possible proof, whose main idea will be useful later on.

Proof. (of Proposition 2.22) We will prove the statement by induction on n. For $n = 1$ the result is trivial. Now suppose that $n > 1$, and consider the standard subdivision of $\mathcal{F}$ with respect to the element n. As $\mathcal{F}_0^{(n)}, \mathcal{F}_1^{(n)} \subseteq 2^{[n-1]}$, by the induction hypothesis we have $|Sh(\mathcal{F}_0^{(n)})| \geq |\mathcal{F}_0^{(n)}|$ and $|Sh(\mathcal{F}_1^{(n)})| \geq |\mathcal{F}_1^{(n)}|$. Moreover $|\mathcal{F}| = |\mathcal{F}_0^{(n)}| + |\mathcal{F}_1^{(n)}|$, $Sh(\mathcal{F}_0^{(n)}) \cup Sh(\mathcal{F}_1^{(n)}) \subseteq Sh(\mathcal{F})$ and if $S \in Sh(\mathcal{F}_0^{(n)}) \cap Sh(\mathcal{F}_1^{(n)})$, then according to the definition of $\mathcal{F}_0^{(n)}$ and $\mathcal{F}_1^{(n)}$ we have $S \cup \{n\} \in Sh(\mathcal{F})$. Summarizing

$$|Sh(\mathcal{F})| \geq |Sh(\mathcal{F}_0^{(n)})| + |Sh(\mathcal{F}_1^{(n)})| \geq |\mathcal{F}_0^{(n)}| + |\mathcal{F}_1^{(n)}| = |\mathcal{F}|.$$

$\square$

From the proof of Proposition 2.22 it is immediate to see, that if $\mathcal{F}$ is extremal, then so are the systems $\mathcal{F}_0^{(i)}$ and $\mathcal{F}_1^{(i)}$ in the standard subdivision with respect to any element $i \in [n]$, and hence, by a previous observation, so is $\mathcal{F}_{A,B}$ for all pairs of sets $A \subseteq B \subseteq [n]$. On the other hand $\widehat{\mathcal{F}}_{A,B}$ can be obtained from $\mathcal{F}_{A,B}$ by adding A to every set in it, and as this does not change neither the size of the family, nor the family of shattered sets, we also get that the subsystem $\widehat{\mathcal{F}}_{A,B}$ of $\mathcal{F}$ is also extremal.

Definition 2.29. *For $i \in [n]$ let φ_i be the ith bit flip operation, i.e. for $F \in 2^{[n]}$ we have*

$$\varphi_i(F) = F \triangle \{i\} = \begin{cases} F \backslash \{i\} & \text{if } i \in F \\ F \cup \{i\} & \text{if } i \notin F \end{cases}$$

and for $\mathcal{F} \subseteq 2^{[n]}$ put $\varphi_i(\mathcal{F}) = \{\varphi_i(F) \mid F \in \mathcal{F}\}$.

If $\mathcal{F} \subseteq 2^{[n]}$, then the family of shattered sets is trivially invariant under the bit flip operation, i.e. $Sh(\mathcal{F}) = Sh(\varphi_i(\mathcal{F}))$ for all $i \in [n]$, and hence so is extremality. This means that when dealing with a nonempty set system $\mathcal{F}$, and examining its extremality, we can without loss of generality assume that $\emptyset \in \mathcal{F}$, otherwise we could apply bit flips to it, to bring $\emptyset$ inside.

Definition 2.30. *The downshift operation for a set system $\mathcal{F} \subseteq 2^{[n]}$ by the element $i \in [n]$ is defined as*

$$\begin{aligned} D_i(\mathcal{F}) &= \{F \mid F \in \mathcal{F},\ i \notin F\} \ \cup\ \{F \mid F \in \mathcal{F},\ i \in F,\ F\backslash\{i\} \in \mathcal{F}\} \\ &\quad \cup \{F\backslash\{i\} \mid F \in \mathcal{F},\ i \in F,\ F\backslash\{i\} \notin \mathcal{F}\} \\ &= \{F\backslash\{i\} \mid F \in \mathcal{F}\} \ \cup\ \{F \mid F \in \mathcal{F},\ i \in F,\ F\backslash\{i\} \in \mathcal{F}\}. \end{aligned}$$

It is not hard to see that $|D_i(\mathcal{F})| = |\mathcal{F}|$ and $Sh(D_i(\mathcal{F})) \subseteq Sh(\mathcal{F})$, hence D_i also preserves extremality (e.g. [12, Lemma 1]).

The downshift operation is an important tool in the study of set systems, in particular downshifts can be used to give a possible combinatorial description of the family of standard monomials of the vanishing ideal $I(\mathcal{F})$ for lexicographic term orders. For indices $i_1, i_2, \dots, i_\ell$ put

$$D_{i_1,i_2,\dots,i_\ell}(\mathcal{F}) := D_{i_1}(D_{i_2}(\dots(D_{i_\ell}(\mathcal{F})))).$$

Proposition 2.31. *([35, Theorem 6.1]) Let $\mathcal{F} \subseteq 2^{[n]}$ and $\prec$ be a lexicographic term order for which $x_{i_1} \succ x_{i_2} \succ \dots \succ x_{i_n}$. Then*

$$Sm(I(\mathcal{F})) = D_{i_n,i_{n-1},\dots,i_1}(\mathcal{F}).$$

Note that in the above equality the set system representation of $Sm(I(\mathcal{F}))$ is considered.

Definition 2.32. *For a set systems $\mathcal{F} \subseteq 2^{[n]}$ and $i \in [n]$ let*

$$M_i(\mathcal{F}) = \mathcal{F}_0^{(i)} \cap \mathcal{F}_1^{(i)},$$
$$U_i(\mathcal{F}) = \mathcal{F}_0^{(i)} \cup \mathcal{F}_1^{(i)}.$$

The following equalities follow easily from the definitions:

$$M_i(\mathcal{F}) = \mathcal{F}_0^{(i)} \cap \mathcal{F}_1^{(i)} = (D_i(\mathcal{F}))_1^{(i)},$$
$$U_i(\mathcal{F}) = \mathcal{F}_0^{(i)} \cup \mathcal{F}_1^{(i)} = (D_i(\mathcal{F}))_0^{(i)}.$$

From these it follows that if $\mathcal{F}$ is extremal then so are $M_i(\mathcal{F})$ and $U_i(\mathcal{F})$, since we can obtain them from $\mathcal{F}$ using operations preserving extremality.

Definition 2.33. *For a set system $\mathcal{F} \subseteq 2^{[n]}$ and a set $B \subseteq [n]$ let*

$$\mathcal{F}(B) = \{I \subseteq [n]\backslash B \mid I + 2^B \subseteq \mathcal{F}\}.$$

We remark that if $B = \{i_1, \dots, i_m\} \subseteq [n]$, then

$$\mathcal{F}(B) = M_{i_1}(M_{i_2}(\dots M_{i_m}(\mathcal{F}) \dots)),$$

and hence $\mathcal{F}(B)$ is extremal if $\mathcal{F}$ is extremal.

Definition 2.34. *The projection of a set system* $\mathcal{F} \subseteq 2^{[n]}$ *to a set of indices* $X \subseteq [n]$ *is*

$$\mathcal{F}|_X = \{F \cap X \mid F \in \mathcal{F}\}.$$

Note that $X \in Sh(\mathcal{F})$ if and only if $\mathcal{F}|_X = 2^X$. Also if $X = \{x_1, \ldots, x_m\}$ then $\mathcal{F}|_X$ is just $U_{x_1}(U_{x_2}(\ldots U_{x_m}(\mathcal{F})\ldots))$, thus if the original set system is extremal, then so is its projected version. Moreover, if $Y \subseteq X$ then we have that $Y \cap (F \cap X) = Y \cap F$ for all $F \subseteq [n]$, meaning that $Y \subseteq X$ is in $Sh(\mathcal{F}|_X)$ if and only if it is in $Sh(\mathcal{F})$, in particular $Sh(\mathcal{F}|_X) = Sh(\mathcal{F})|_X$.

Part I

Shattering-extremal set systems

Chapter 3

An algebraic approach to shattering

3.1 Order shattering

Anstee, Rónyai and Sali in [7] were the first who related standard monomials to shattering. They defined the concept of order shattered in an inductive way.

Definition 3.1. *We say that the set $S = \{s_1, s_2, \ldots, s_d\} \subseteq [n]$ is order shattered by a given family $\mathcal{F} \subseteq 2^{[n]}$ if the following holds: in the case $S = \emptyset$ the family $\mathcal{F}$ has to contain a set; when $|S| > 0$ and $s_1 < s_2 < \cdots < s_d$, then there are 2^d sets in $\mathcal{F}$ that can be divided into two families $\mathcal{F}_0$ an $\mathcal{F}_1$ such that $s_d \notin F_0$ for all $F_0 \in \mathcal{F}_0$, $s_d \in F_1$ for all $F_1 \in \mathcal{F}_1$, and both $\mathcal{F}_0, \mathcal{F}_1$ order shatter the set $S \backslash \{s_d\}$, furthermore $T \cap F_0 = T \cap F_1$ holds for $T = \{s_d + 1, s_d + 2, \ldots, n\}$ and for all $F_0 \in \mathcal{F}_0$, $F_1 \in \mathcal{F}_1$.*

Let $osh(\mathcal{F})$ be the family of sets order shattered by $\mathcal{F}$. It is easy to see that $osh(\mathcal{F})$ is a down-set for every $\mathcal{F} \subseteq 2^{[n]}$ and $osh(\mathcal{F}) \subseteq Sh(\mathcal{F})$. For the size of $osh(\mathcal{F})$ Anstee, Rónyai and Sali proved in [7] the following:

Proposition 3.2. *([7, Theorem 1.4]) Let $\mathcal{F}$ be a family of subsets of $[n]$. Then*

$$|osh(\mathcal{F})| = |\mathcal{F}|.$$

In fact they proved, that $osh(\mathcal{F})$ is equal (with the usual bijection between sets and square-free monomials in mind) to the family of standard monomials of $I(\mathcal{F})$ with respect to the standard lex order, and so the above definition also gives a combinatorial description of standard monomials. For further definitions and properties of $osh(\mathcal{F})$ see [7] and as an example of its application see [22].

3.2 Shattering and standard monomials

In [35] and [42] we developed a new algebraic technique for the investigation of extremal set systems. Most results in this section were already part of my master's thesis at Budapest University of Technology and Economics, see [35], however they are included in this thesis as well as they are essential to get a more complete picture of extremality.

Lemma 3.3. *([42, Lemma 1.]) Let $\mathcal{F} \subseteq 2^{[n]}$.*

(a) If $x_H \in Sm(I(\mathcal{F}))$ for some term order, then $H \in Sh(\mathcal{F})$.

(b) If $H \in Sh(\mathcal{F})$, then there is a lex order for which we have $x_H \in Sm(I(\mathcal{F}))$.

Combining the two parts of Lemma 3.3, we obtain the following:

Proposition 3.4. *([42, Equation 1])*

$$Sh(\mathcal{F}) = \bigcup_{\text{term orders}} Sm(I(\mathcal{F})) = \bigcup_{\text{lex term orders}} Sm(I(\mathcal{F})),$$

where any set $H \subseteq Sh(\mathcal{F})$ is identified with the square-free monomial $\mathbf{x}_H$.

Note that besides of giving an algebraic description of the family of shattered sets, the second equality of Proposition 3.4 is interesting on its own as well, only considering the algebraic setting.

To obtain a similar result for $st(\mathcal{F})$ first recall that

$$S \in st(\mathcal{F}) \iff [n] \backslash S \notin Sh(2^{[n]} \backslash \mathcal{F}),$$

and so by Proposition 3.4

$$\iff \mathbf{x}_{[n] \backslash S} \in \bigcap_{\text{lex term orders}} Lm(I(2^{[n]} \backslash \mathcal{F})).$$

However by Proposition 2.16 for any lex term order

$$x_{[n] \backslash S} \in Lm(I(2^{[n]} \backslash \mathcal{F})) \iff \mathbf{x}_S \in Sm(I(\mathcal{F}).$$

Putting things together we get the following:

Proposition 3.5.

$$st(\mathcal{F}) = \bigcap_{\text{term orders}} Sm(I(\mathcal{F})),$$

where any set $H \subseteq st(\mathcal{F})$ is identified with the square-free monomial $\mathbf{x}_H$.

As noted previously, for set systems one can compute $Sm(I(\mathcal{F}))$ efficiently for any lex term order. However, as the number of lex orders is $n!$, Proposition 3.4 does not immediately provide an efficient way to calculate $Sh(\mathcal{F})$, nevertheless, when comparing cardinalities, it results at once a simple algebraic characterization of extremal set systems:

Theorem 3.6. *([42, Theorem 18]) For a set system $\mathcal{F} \subseteq 2^{[n]}$ the following are equivalent:*

(i) $\mathcal{F}$ is extremal.

(ii) $Sm(I(\mathcal{F}))$ is the same for all term orders.

(iii) $Sm(I(\mathcal{F}))$ is the same for all lex term orders.

Theorem 3.6 leads to an algebraic characterization of extremal set systems, involving the Gröbner bases of $I(\mathcal{F})$.

Definition 3.7. *A Gröbner basis $\mathcal{G}$ of and ideal $I \lhd \mathbb{F}[\mathbf{x}]$ is called a universal Gröbner basis of I if it is a Gröbner basis for every term order.*

Definition 3.8. *For a pair of sets $H \subseteq S \subseteq [n]$ let*

$$f_{S,H}(\mathbf{x}) = \left(\prod_{i \in H} x_i\right)\left(\prod_{j \in S \setminus H} (1 - x_j)\right).$$

Note that $f_{S,H}(\mathbf{v}_F) \neq 0$ exactly if $S \cap F = H$, and the leading monomial of $f_{S,H}$ is $\mathbf{x}_S$ for every term order.

Theorem 3.9. *([42, Theorem 19]) $\mathcal{F} \subseteq 2^{[n]}$ is extremal if and only if there are polynomials of the form $f_{S,H}$, which together with $\{x_i^2 - x_i, i \in [n]\}$ form a universal Gröbner basis of $I(\mathcal{F})$.*

The sets S in the above theorem are actually the minimal sets outside $Sh(\mathcal{F})$, implying that the above collection of polynomials is reduced with respect to any term order. For one such fixed S, H is the unique(!) subset of S for which there is no $F \in \mathcal{F}$ for which $F \cap S = H$. In this way one can assign to every extremal family $\mathcal{F}$ a collection of pairs of sets

$$\mathcal{P}_{\mathcal{F}} = \{(H, S) \mid f_{S,H} \text{ is in the reduced Gröbner basis of } I(\mathcal{F})\}.$$

It may be interesting to obtain insight into the structure of $\mathcal{P}_{\mathcal{F}}$.

Open problem 1. *Given a finite collection of pairs of sets*

$$\mathcal{P} = \{(H_i, S_i) \mid H_i \subseteq S_i \subseteq [n],\ i \in I\},$$

under what condition is there an extremal family $\mathcal{F} \subseteq 2^{[n]}$ such that $\mathcal{P} = \mathcal{P}_{\mathcal{F}}$?

We remark also that in Theorem 3.9 it would be enough to require that $I(\mathcal{F})$ has a suitable Gröbner basis for some term order. Indeed suppose that $I(\mathcal{F})$ has an appropriate Gröbner basis for some fixed term order $\prec$, and take some monomial $\mathbf{x}^{\mathbf{u}} \in Lm(I(\mathcal{F}))$. Since $\mathcal{G}$ is a Gröbner basis, there is some $g \in \mathcal{G} \subseteq I(\mathcal{F})$ such that $lm(g)|\mathbf{x}^{\mathbf{u}}$. From this we get that the polynomial $\frac{\mathbf{x}^{\mathbf{u}}}{lm(g)}g(\mathbf{x})$ is a member of $I(\mathcal{F})$. g is either $x_i^2 - x_i$ for some i or is of the form $f_{S,H}$. Since term orders are monotone with respect to multiplication, in both cases the leading monomial of $\frac{\mathbf{x}^{\mathbf{u}}}{lm(g)}g(\mathbf{x})$ is $\mathbf{x}^{\mathbf{u}}$ for every term order. This implies that $\mathbf{x}^{\mathbf{u}}$ is a leading monomial for every term order. However by Proposition 2.15 the number of standard monomials, i.e. the number of non-leading monomials, is the same for every term order, namely $|\mathcal{F}|$. Accordingly the previous observation also means that the family of standard monomials is the same for every term order, and hence by Theorem 3.6 $\mathcal{F}$ is extremal.

In addition to this characterization, Theorem 3.6 leads also to an efficient algorithm for testing the extremality of a set system. The test is based on the theorem below.

Theorem 3.10. *([42, Theorem 20]) Take n orderings of the variables such that for every index i there is one, in which x_i is the greatest element, and take the corresponding lex term orders. If $\mathcal{F}$ is not extremal, then among these we can find two term orders for which the standard monomials of $I(\mathcal{F})$ differ.*

Accordingly to decide whether $\mathcal{F}$ is extremal or not, it is enough to compute and compare the standard monomials for n lex orders. Using the lex game one can compute the standard monomials for one fixed lex order in linear, i.e. $O(n|\mathcal{F}|)$ time, and so the total running time of this algorithm is $O(n^2|\mathcal{F}|)$. This (in typical cases, when $n < |\mathcal{F}|^2$) improves the algorithm given in [24] by Greco, where the time bound is $O(n|\mathcal{F}|^3)$. But it is still open whether one can do better.

Open problem 2. *Given a family $\mathcal{F} \subseteq 2^{[n]}$, can the extremality of $\mathcal{F}$ be tested in linear, i.e. $O(n|\mathcal{F}|)$ time?*

Chapter 4

Extremality in the general case

Set systems, when representing their elements by their characteristic vectors, can be considered as special types of finite point sets. Some of the previous results concerning extremality remain true in a more general setting as well.

There is a usual way of generalizing the notion of shattering (see e.g. [45]) for collections of vectors from $\{0, 1, ..., k-1\}^n$. Here the vectors are considered as $[n] \to \{0, 1, ..., k-1\}$ functions.

Definition 4.1. *Let $\mathcal{V}$ be a class of $[n] \to \{0, 1, ..., k-1\}$ functions. We say that $\mathcal{V}$ shatters a set $S \subseteq [n]$ if for every function $\mathbf{g} : S \to \{0, 1, ..., k-1\}$ there exists a function $\mathbf{f} \in \mathcal{V}$ such that $\mathbf{f}|_S = \mathbf{g}$.*

As previously, for a finite set $\mathcal{V} \subseteq \{0, 1, ..., k-1\}^n$ let $Sh(\mathcal{V})$ denote the family of shattered sets. In the definition of extremality the Sauer inequality played a key role, however in this case we cannot expect a similar inequality to hold. Indeed, as $Sh(\mathcal{V}) \subseteq 2^{[n]}$, there are at most 2^n sets shattered, but at the same time the size of $\mathcal{V}$ can be much larger, up to k^n.

This lack of a Sauer-like inequality suggests to forget about shattering, and define extremality according to Theorem 3.6. Before this we first prove the equivalence $(ii) \Leftrightarrow (iii)$ from Theorem 3.6 in this general setting.

Definition 4.2. *A polynomial $f(\mathbf{x}) \in \mathbb{F}[\mathbf{x}]$ is called degree dominated if it is of the form $f(\mathbf{x}) = \mathbf{x}^{\mathbf{w}} + \sum_{i=1}^{\ell} \alpha_i \mathbf{x}^{\mathbf{v}_i}$, where $\mathbf{x}^{\mathbf{v}_i} | \mathbf{x}^{\mathbf{w}}$ for every i. $\mathbf{x}^{\mathbf{w}}$ is called the dominating term of f.*

As an example of a degree dominated polynomial one can consider any polynomial of the form $f_{S,H}$ or for $i = 1, \dots, n$ the polynomial $x_i^2 - x_i$, all of them appearing in Theorem 3.9.

Note that any monomial appearing in a degree dominated polynomial divides its dominating term. Accordingly, since any term order is the refinement of the divisibility of monomials (Proposition 2.2), the dominating term of such a polynomial is also its leading term for every term order.

Proposition 4.3. *If $\mathcal{V} \subseteq \{0, 1, ..., k-1\}^n$ is a finite set, then $Sm(I(\mathcal{V}))$ is the same for every lexicographic term order if and only if $Sm(I(\mathcal{V}))$ is the same for every term order.*

Proof. One direction is just trivial. For the other direction suppose that the standard monomials of $I(\mathcal{V})$ are the same for every lex order, and denote this collection of monomials by $\mathcal{S}$. Take an arbitrary monomial $\mathbf{x}^{\mathbf{u}} \notin \mathcal{S}$. $\mathbf{x}^{\mathbf{u}}$ is a leading monomial with respect to every lex order. Fix one lex order, and take the standard representation of $\mathbf{x}^{\mathbf{u}}$ according to Corollary 2.5.

$$f(\mathbf{x}) = \mathbf{x}^{\mathbf{u}} + \sum_{x^{\mathbf{v}} \in \mathcal{S}} \alpha_{\mathbf{v}} x^{\mathbf{v}} \in I(\mathcal{V}).$$

As $\mathcal{S}$ is the family of standard monomials for every other lex order as well, the leading monomial of f can be only $\mathbf{x}^{\mathbf{u}}$ for them as well. This is possible only $f(\mathbf{x})$ is a degree dominated polynomial with dominating term $\mathbf{x}^{\mathbf{u}}$. Indeed suppose this is not the case, and there is some monomial $x^{\mathbf{v}} \in \mathcal{S}$ that appears with a nonzero coefficient in f and $x^{\mathbf{v}} \nmid x^{\mathbf{u}}$.

For this there has to be an index i for which $v_i > u_i$, but then for any lex order where x_i is the largest variable we would have that $x^{\mathbf{u}} \prec x^{\mathbf{v}}$, contradicting our assumption that $lm(f) = x^{\mathbf{u}}$.

On the other hand, by our earlier remark, the leading monomial of f is $x^{\mathbf{u}}$ for every term order. This results, that every monomial $\mathbf{x}^{\mathbf{u}} \notin \mathcal{S}$ is a leading monomial for every term order. Adding the fact that the number of standard monomials, i.e. the number of non-leading monomials, is $|\mathcal{V}|$ for every term order, we get that the family of leading monomials, and hence the family of standard monomials of $I(\mathcal{V})$ is the same for every term order as desired. $\square$

Definition 4.4. *A finite set of vectors $\mathcal{V} \subseteq \{0, 1, ..., k-1\}^n \subseteq \mathbb{R}^n$ is extremal if $Sm(I(\mathcal{V}))$ is the same for every lexicographic term order, or equivalently if $Sm(I(\mathcal{V}))$ is the same for every term order.*

Proposition 4.3 was needed to guarantee that the definition of extremality in this general setting is compatible with the special case of set systems.

In the above definition $I(\mathcal{V})$ is considered inside $\mathbb{R}[\mathbf{x}]$. At first sight we restrict our attention only to special types of finite vector systems, however Proposition 2.17 justifies that in fact it is enough to deal with the special vector systems from Definition 4.4.

As mentioned earlier, one can use the Lex game to describe combinatorially the family of standard monomials of the vanishing ideal of some finite point set for lexicographic orders. In the case of set systems the downshift operation provided us another method to get some insight into the structure of the family of standard monomials. This method, especially Proposition 2.31, can be generalized to the present setting as well.

For $1 \le i \le n$, the i-section of $\mathcal{V} \subseteq \{0, 1, \ldots, k-1\}^n$ for $n-1$ arbitrary elements

$$\alpha_1, \ldots, \alpha_{i-1}, \alpha_{i+1}, \ldots, \alpha_n \in \{0, 1, \ldots, k-1\}$$

is defined as

$$\mathcal{V}_i(\alpha_1,\ldots,\alpha_{i-1},\alpha_{i+1},\ldots,\alpha_n)=\{\alpha \mid (\alpha_1,\ldots,\alpha_{i-1},\alpha,\alpha_{i+1},\ldots,\alpha_n)\in\mathcal{V}\}.$$

Using i-sections one can define D_i, the downshift operation at coordinate i in the general case. For any finite point set $\mathcal{V}\subseteq\{0,1,\ldots,k-1\}^n$, $D_i(\mathcal{V})$ is the unique point set in $\{0,1,\ldots,k-1\}^n$, for which

$$(D_i(\mathcal{V}))_i(\alpha_1,..,\alpha_{i-1},\alpha_{i+1},..,\alpha_n)=\{0,1,..,|\mathcal{V}_i(\alpha_1,..,\alpha_{i-1},\alpha_{i+1},..,\alpha_n)|-1\}$$

whenever $\mathcal{V}_i(\alpha_1,\ldots,\alpha_{i-1},\alpha_{i+1},\ldots,\alpha_n)$ is nonempty, otherwise it is empty as well.

As before, for indices $i_1,i_2,\ldots,i_\ell$ let

$$D_{i_1,i_2,\ldots,i_\ell}(\mathcal{V}):=D_{i_1}(D_{i_2}(\ldots(D_{i_\ell}(\mathcal{V})))).$$

Proposition 4.5. *([35, Theorem 10.1]) Let $\mathcal{V}\subseteq\{0,1,\ldots,k-1\}^n$ be a finite point set and $\prec$ the lexicographic term order order for which $x_{i_1}\succ x_{i_2}\succ\cdots\succ x_{i_n}$. Then*

$$Sm(I(\mathcal{V}))=D_{i_n,i_{n-1},\ldots,i_1}(\mathcal{V}).$$

In [35], beside Proposition 4.5, several other results concerning this general setting were proved, however the general versions of the two main results from Chapter 3, Theorem 3.9 and Theorem 3.10, were missing. Now we eliminate this shortcoming.

Theorem 4.6. *A finite set of vectors $\mathcal{V}\subseteq\{0,1,...,k-1\}^n\subseteq\mathbb{R}^n$ is extremal if and only if there is a finite family $\mathcal{G}\subseteq\mathbb{R}[\mathbf{x}]$ of degree dominated polynomials that form a universal Gröbner basis of $I(\mathcal{V})$.*

Proof. First suppose that $\mathcal{V}\subseteq\{0,1,...,k-1\}^n\subseteq\mathbb{R}^n$ is extremal. By definition $Sm(I(\mathcal{V}))$

is the same for every term order. Denote the family of all minimal (with respect to division) monomials outside of it by $\mathcal{S}$. From this point we follow the line of thinking from the proof of Proposition 4.3. Each monomial $\mathbf{x^u} \in \mathcal{S}$ is a leading monomial for every term order, in particular for the standard lex order as well. By Corollary 2.5, $\mathbf{x^u}$ has a representation by standard monomials with respect to the standard lex order, i.e. there are standard monomials $x^{\mathbf{u}_1}, \ldots, x^{\mathbf{u}_\ell}$, and coefficients $\alpha_1, \ldots, \alpha_\ell \in \mathbb{R}$ such that $f_{\mathbf{u}}(\mathbf{x}) = \mathbf{x^u} + \sum_{i=1}^{\ell} \alpha_i x^{\mathbf{u}_i} \in I(\mathcal{V})$. As by assumption $Sm(I(\mathcal{V}))$ is the same for every term order, and except of $x^{\mathbf{u}}$ every monomial in $f_{\mathbf{u}}(\mathbf{x})$ is a standard one, we necessarily have that $lm(f_{\mathbf{u}}) = x^{\mathbf{u}}$ for every term order. However, in the same way as in the proof of Proposition 4.3, this is possible only if $f_{\mathbf{u}}$ is degree dominated with dominating term $x^{\mathbf{u}}$. Put now

$$\mathcal{G} = \{f_{\mathbf{u}} \mid \mathbf{x^u} \in \mathcal{S}\}.$$

By the definition of $\mathcal{S}$ the family $\mathcal{G}$ is clearly a Gröbner basis for every term order, i.e. it is a universal Gröbner basis.

For the other direction suppose that we are given a finite family $\mathcal{G} \subseteq \mathbb{R}[\mathbf{x}]$ of degree dominated polynomials that form a universal Gröbner basis of $I(\mathcal{V})$. We prove that the fact that there is a common Gröbner basis for every term order, without knowing anything about the members of the Gröbner basis, already guarantees the extremality of $\mathcal{V}$.

If we are given a Gröbner basis $\mathcal{G}$ of some ideal I for a fixed term order, it determines $Lm(I)$, and hence $Sm(I)$, namely

$$Lm(I) = \{\mathbf{x^u} \mid \exists g \in \mathcal{G} \text{ such that } lm(g) | \mathbf{x^u}\}.$$

Indeed the containment in one direction follows from the definition of Gröbner bases. For the other direction note that if for some $g \in \mathcal{G}$ we have that $lm(g)|\mathbf{x^u}$, then the polynomial $\frac{\mathbf{x^u}}{lm(g)} g(\mathbf{x}) \in I$ shows that $\mathbf{x^u} \in Lm(I)$.

However as $\mathcal{G}$ is a common Gröbner basis for every term order, it gives us the same family of standard monomials for every term order, and so the extremality of $\mathcal{V}$ follows. □

We remark that similarly as in the case of Theorem 3.9, in Theorem 4.6 it is also enough to require that $I(\mathcal{V})$ has a suitable Gröbner basis for some term order.

Theorem 4.7. *Take n orderings of the variables such that for every index i there is one in which x_i is the greatest element, and take the corresponding lex orders. If a finite set of vectors $\mathcal{V} \subseteq \{0, 1, ..., k-1\}^n \subseteq \mathbb{R}^n$ is not extremal, then among these we can find two term orders for which the standard monomials of $I(\mathcal{V})$ differ.*

Proof. By contraposition it is enough to prove that if the standard monomials of $I(\mathcal{V})$ are the same for the above term orders, then $\mathcal{V}$ is extremal. Accordingly suppose the condition holds, and denote the collection of standard monomials for the above term orders by $\mathcal{S}$. From now on we again follow the proof of Proposition 4.3. Take an arbitrary monomial $\mathbf{x}^{\mathbf{u}} \notin \mathcal{S}$. In this case $\mathbf{x}^{\mathbf{u}}$ is a leading monomial with respect to all of the n lex orders considered. Fix one of these lex orders, and take the standard representation of $\mathbf{x}^{\mathbf{u}}$ according to Corollary 2.5.

$$f(\mathbf{x}) = \mathbf{x}^{\mathbf{u}} + \sum_{x^{\mathbf{v}} \in \mathcal{S}} \alpha_{\mathbf{v}} x^{\mathbf{v}} \in I(\mathcal{V}).$$

As $\mathcal{S}$ is the family of standard monomials for the other $n-1$ lex orders as well, the leading monomial of f can be only $\mathbf{x}^{\mathbf{u}}$ for them as well. This is possible only if f is degree dominated with dominating term $x^{\mathbf{u}}$. Indeed suppose this is not the case, and there is some monomial $x^{\mathbf{v}} \in \mathcal{S}$ that appears with a nonzero coefficient in f and $x^{\mathbf{v}} \nmid x^{\mathbf{u}}$. Now there has to be an index i for which $v_i > u_i$, but then for any lex order where x_i is the largest variable, in particular for the one in our collection of lex orders, we would have that $x^{\mathbf{u}} \prec x^{\mathbf{v}}$, contradicting our assumption that $lm(f) = x^{\mathbf{u}}$.

From this the extremality of $\mathcal{V}$ follows exactly as in Proposition 4.3. $\square$

Theorem 4.7 has several interesting consequences. First of all it means that in the definition of extremality it would have been enough to require that the family of standard monomials is the same for a particular family of lex orders of size n.

Next, Theorem 4.7 also results an efficient algorithm for deciding whether a finite set of vectors $\mathcal{V} \subseteq \{0, 1, ..., k-1\}^n \subseteq \mathbb{R}^n$ is extremal or not. As in the special case of set systems, it is enough to compute and compare the standard monomials for n lex orders. According to Proposition 2.18 that can be done in $O(n^2|\mathcal{V}|k)$ time.

To finish, we remark that Theorem 4.7 also proves a result of Li, Zhang and Dong from [34], where they investigated the standard monomials of zero dimensional polynomial ideals.

Definition 4.8. *Let $\mathbb{F}$ be a field and $I \lhd \mathbb{F}[\mathbf{x}]$ a polynomial ideal. I is called zero dimensional if the factor space $\mathbb{F}[\mathbf{x}]/I$ is a finite dimensional $\mathbb{F}$-vector space.*

It is easy to see that vanishing ideals of finite point sets are special types of zero dimensional ideals.

A term order $\prec$ is called an elimination order with respect to the variable x_i if x_i is larger than any monomial from $\mathbb{F}[x_1, \ldots, x_{i-1}, x_{i+1}, \ldots, x_n]$. As an example one can consider any lex order where x_i is the largest variable.

Now for $1 \leq i \leq n$ let $\prec_i$ be an elimination order with respect to x_i. Part (2) $\Leftrightarrow$ (3) of Theorem 3.1 in [34] states that if $\mathbb{F}$ has characteristic zero, then the standard monomials of any zero dimensional ideal $I \lhd \mathbb{F}[\mathbf{x}]$ are the same for every term order if and only if they are the same for $\prec_1, \ldots, \prec_n$. We claim that Theorem 4.7 together with Proposition 2.17 prove the same result for arbitrary fields. For this first note, that in Theorem 4.7 the lex orders can be substituted by arbitrary elimination term orders with respect to the variables, since in the proof only the elimination property is used. Similarly also observe that about the ideal considered we only needed that the number of standard monomials is the same for every term order. However as the standard monomials form a linear basis of the $\mathbb{F}$-vector space $\mathbb{F}[\mathbf{x}]/I$ this is by definition true in general for zero dimensional ideals, not merely vanishing ideals of finite point sets.

With these observations in mind one gets the following form of Theorem 4.7, which generalizes part (2) $\Leftrightarrow$ (3) of Theorem 3.1 from [34] to arbitrary fields instead of algebraically closed ones.

Theorem 4.9. *Let $\mathbb{F}$ be an arbitrary field and for $1 \leq i \leq n$ let $\prec_i$ be an elimination order with respect to x_i. Then the standard monomials of any zero dimensional ideal $I \lhd \mathbb{F}[\mathbf{x}]$ are the same for every term order if and only if they are the same for $\prec_1, \ldots, \prec_n$.* $\square$

Chapter 5

Graph theoretical characterization of extremality

In this chapter we introduce a simple graph theoretical interpretation of the topic of the original set theoretic setting and develop an effective method for the study of extremal set systems.

Definition 5.1. *The inclusion graph of a set system $\mathcal{F} \subseteq 2^{[n]}$, denoted by $\mathbb{G}_{\mathcal{F}}$, is the simple directed edge-labelled graph whose vertices are the elements of $\mathcal{F}$, and there is a directed edge with label $j \in [n]$ going from G to F exactly when $F = G \cup \{j\}$.*

See Figure 5.1 for an example. The inclusion graph of the complete set system $2^{[n]}$ will be denoted by $\mathbb{H}_n$. The undirected version of $\mathbb{H}_n$ is often referred to as the Hamming graph $H(n, 2)$, or as the hypercube of dimension n, whose vertices are all $0-1$ vectors of length n, and two vertices are adjacent if and only if they differ in exactly one coordinate. When computing distances between vertices in the inclusion graph $\mathbb{G}_{\mathcal{F}}$, we forget about the direction of edges. We define the distance between vertices $F, G \in \mathcal{F}$, denoted by

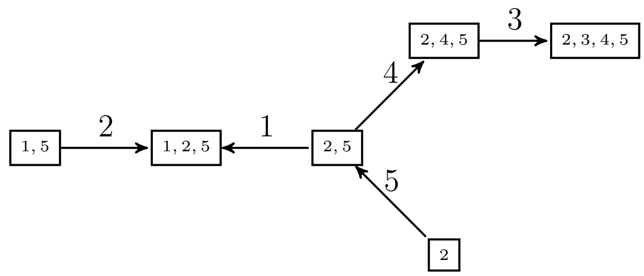


Figure 5.1: The inclusion graph of $\{\{2\},\{1,5\},\{2,5\},\{1,2,5\},\{2,4,5\},\{2,3,4,5\}\}$

$d_{\mathbb{G}_{\mathcal{F}}}(F,G)$, as their graph distance in the undirected version of $\mathbb{G}_{\mathcal{F}}$, i.e. the length of the shortest path between them in the undirected version of $\mathbb{G}_{\mathcal{F}}$. Similarly, some edges in $\mathbb{G}_{\mathcal{F}}$ form a path between two vertices if they do so in the undirected version of $\mathbb{G}_{\mathcal{F}}$. For example, the distance between two vertices $F,G \subseteq [n]$ in $\mathbb{H}_n$ is just the size of the symmetric difference $F \triangle G$, i.e. $d_{\mathbb{H}_n}(F,G) = |F \triangle G|$. As a consequence, when only distances of vertices will be considered, and the context will allow, we omit the directions of edges to avoid unnecessary case analysis, and will specify edges by merely listing their endpoints.

In terms of the inclusion graph, the ith bit flip operation φ_i (see Definition 2.29) flips the directions of edges with label i, i.e. there is a bijection between the vertices of $\mathbb{G}_{\mathcal{F}}$ and $\mathbb{G}_{\varphi_i(\mathcal{F})}$ that preserves all edges with label different from i, and reverses edges with label i. This bijection is simply given by the reflection with respect to the hyperplane $x_i = \frac{1}{2}$ in the Hamming graph, when viewed as a subset of $\mathbb{R}^n$.

This graph theoretical point of view has already appeared in the literature several times. As an example consider a result of Greco. To be able to state the result, first note that for any set system $\mathcal{F} \subseteq 2^{[n]}$, the identity map naturally embeds the inclusion graph $\mathbb{G}_{\mathcal{F}}$ into $\mathbb{H}_n$. We say that the inclusion graph $\mathbb{G}_{\mathcal{F}}$ is isometrically embedded (into $\mathbb{H}_n$), if this embedding is an isometry, meaning that for arbitrary $F,G \in \mathcal{F}$ we have $d_{\mathbb{G}_{\mathcal{F}}}(F,G) = d_{\mathbb{H}_n}(F,G)$, i.e. there is a path of length $d_{\mathbb{H}_n}(F,G) = |F \triangle G|$ between F and G inside the undirected version of $\mathbb{G}_{\mathcal{F}}$. Greco in [24] proved the following:

Proposition 5.2. *([24, Lemma 8]) If $\mathcal{F} \subseteq 2^{[n]}$ is extremal, then $\mathbb{G}_{\mathcal{F}}$ is isometrically embedded.*

As this fact will be used several times, we provide the reader with our simple proof from [35]:

Proof. Suppose the contrary, namely that $\mathbb{G}_{\mathcal{F}}$ is not isometrically embedded. Then there exist sets $A, B \in \mathcal{F}$ such that $d_{\mathbb{H}_n}(A, B) = k < d_{\mathbb{G}_{\mathcal{F}}}(A, B)$. Suppose that A and B are such that k is minimal. Clearly $k \geq 2$. Without loss of generality we may suppose that $A = \emptyset$ and $|B| = k$, otherwise one could apply bit flips to the set system to achieve this. Note that distances both in $\mathbb{G}_{\mathcal{F}}$ and in $\mathbb{H}_n$ are invariant under bit flips.

We claim that there is no set $C \in \mathcal{F}$ different from A with $C \subsetneq B$. Indeed suppose such C exists, then

$$d_{\mathbb{H}_n}(A, C) + d_{\mathbb{H}_n}(C, B) = d_{\mathbb{H}_n}(A, B) = k < d_{\mathbb{G}_{\mathcal{F}}}(A, B) \leq d_{\mathbb{G}_{\mathcal{F}}}(A, C) + d_{\mathbb{G}_{\mathcal{F}}}(C, B).$$

From this we have either $d_{\mathbb{H}_n}(A, C) < d_{\mathbb{G}_{\mathcal{F}}}(A, C)$ or $d_{\mathbb{H}_n}(C, B) < d_{\mathbb{G}_{\mathcal{F}}}(C, B)$. Since $d_{\mathbb{H}_n}(A, C)$ and $d_{\mathbb{H}_n}(C, B)$ are less then k, we get a contradiction in both cases with the minimality of k.

Now, since $\mathcal{F}$ is extremal, so must be $\widehat{\mathcal{F}}_{\emptyset,B}$ (see Definition 2.28). However, in our case $\widehat{\mathcal{F}}_{\emptyset,B} = \{\emptyset, B\}$, and so if $B = \{b_1, \ldots, b_k\}$, then $Sh(\widehat{\mathcal{F}}_{\emptyset,B}) = \{\emptyset, \{b_1\}, \ldots, \{b_k\}\}$. Counting cardinalities we get that $|Sh(\widehat{\mathcal{F}}_{\emptyset,B})| = |B| + 1 = k + 1 \geq 3 > 2 = |\widehat{\mathcal{F}}_{\emptyset,B}|$, implying that $\widehat{\mathcal{F}}_{\emptyset,B}$ cannot be extremal. This contradiction finishes the proof. $\square$

Another earlier appearance of inclusion graphs was in [11], where Bollobás, Leader and Radcliffe characterized extremality in terms of the inclusion graphs of the systems $\mathcal{F}(B)$ (see Definition 2.33):

Proposition 5.3. *([11, Theorem 2.3]) $\mathcal{F} \subseteq 2^{[n]}$ is extremal if and only if $\mathbb{G}_{\mathcal{F}(B)}$ is connected for every $B \subseteq [n]$.*

Proposition 5.3 in this form appears first in [12], the authors in [11] prove the statement for so called "order convex" sets, but the same proof yields this form of the proposition as well.

The "only if" direction follows easily by earlier results. Indeed, previously we already noted that if $\mathcal{F}$ is extremal then so is $\mathcal{F}(B)$ for every $B \subseteq [n]$. However if $\mathcal{F}(B)$ is extremal, then $\mathbb{G}_{\mathcal{F}(B)}$ is isometrically embedded, in particular connected.

It is easy to see that $S \in st(\mathcal{F})$ (and so in the extremal case $S \in Sh(\mathcal{F})$) is just equivalent to the fact that $\mathbb{G}_{2^S}$ is isomorphic to a subgraph of $\mathbb{G}_{\mathcal{F}}$ as a directed edge-labelled graph, i.e. there exists a bijection between the vertices of $\mathbb{G}_{2^S}$ and $2^{|S|}$ vertices of $\mathbb{G}_{\mathcal{F}}$ preserving edges, edge labels and edge directions. If this happens, then we will say, that there is a copy of $\mathbb{G}_{2^S}$ in $\mathbb{G}_{\mathcal{F}}$.

Suppose that for a set $S \subseteq [n]$ there are 2 different copies of $\mathbb{G}_{2^S}$ in $\mathbb{G}_{\mathcal{F}}$, i.e. there are two different sets $I_1, I_2 \subseteq [n] \backslash S$ such that $2^S + I_1, 2^S + I_2 \subseteq \mathcal{F}$. Since $I_1 \neq I_2$, there must be an element $\alpha \notin S$ such that $\alpha \in I_1 \triangle I_2$. For this element α we clearly have that $\mathcal{F}$ shatters $S \cup \{\alpha\}$.

Observation 5.4. *If $\mathcal{F} \subseteq 2^{[n]}$ is extremal and the set $S \subseteq [n]$ is a maximal element in $st(\mathcal{F}) = Sh(\mathcal{F})$, in the sense that $S \in st(\mathcal{F}) = Sh(\mathcal{F})$ and for all $S' \supsetneq S$ we have $S' \notin st(\mathcal{F}) = Sh(\mathcal{F})$, then S is uniquely strongly shattered, i.e. there is one unique copy of $\mathbb{G}_{2^S}$ in $\mathbb{G}_{\mathcal{F}}$.*

5.1 Extremal families of VC dimension 1

When examining the inclusion graphs of some specific extremal set systems, one can come up with the observation, that in some sense the complexity of the inclusion graph depends on the VC dimension of the set system considered. Accordingly, to get some insight we first restrict our attention to simple cases, where the VC dimension of $\mathcal{F}$ is bounded by some small fixed natural number t. This kind of relaxation of problems is a usual method in extremal combinatorics, see [8].

To start with, in [36] we considered the case $t = 1$.

Proposition 5.5. *([36, Proposition 2]) A set system $\mathcal{F} \subseteq 2^{[n]}$ is extremal and of VC dimension at most 1 if and only if $\mathbb{G}_{\mathcal{F}}$ is a tree and all labels on the edges are different.*

Proof. For the 'only if' direction suppose that $\mathcal{F}$ is extremal and $dim_{VC}(\mathcal{F}) \leq 1$. According to Proposition 5.2 we know that $\mathbb{G}_{\mathcal{F}}$ must be isometrically embedded into $\mathbb{H}_n$, in particular $\mathbb{G}_{\mathcal{F}}$ is connected. Next we prove that all labels on the edges of $\mathbb{G}_{\mathcal{F}}$ are different. Suppose for contradiction, that there are two edges with the same label. Without loss of generality we may assume that this label is 1. Since there are no two edges going out from a set with the same label, there are sets $A, B, C, D \in \mathcal{F}$, all different, such that $1 \in A \cap B$, $C = A \backslash \{1\}$ and $D = B \backslash \{1\}$. Since $A \neq B$, $A \bigtriangleup B$ is nonempty, so there is an element $a \neq 1 \in A \bigtriangleup B$. Without loss of generality we may assume that $a \in A \backslash B$. Now

$$\{1, a\} \cap A = \{1, a\} \quad \{1, a\} \cap B = \{1\} \quad \{1, a\} \cap C = \{a\} \quad \{1, a\} \cap D = \emptyset.$$

So $\{1, a\}$ is shattered by $\{A, B, C, D\}$, consequently $\{1, a\} \in Sh(\mathcal{F})$, contradicting the assumption $dim_{VC}(\mathcal{F}) \leq 1$.

To finish with this direction note that the fact that all labels are different implies that $\mathbb{G}_{\mathcal{F}}$ is acyclic. Indeed suppose for contradiction that it is not the case, and $\mathbb{G}_{\mathcal{F}}$ contains a cycle. Pick one edge from this cycle and let a be its label. On the remaining part of the cycle there must be another edge-labelled with a, since it connects a set containing a with one not containing a. However this is impossible, since all labels are different. Adding the connectedness of $\mathbb{G}_{\mathcal{F}}$, we obtain that it is actually a tree as wanted.

For the reverse direction suppose that $\mathbb{G}_{\mathcal{F}}$ is a tree and all labels on the edges are different. It is easily seen that this implies that $\mathbb{G}_{\mathcal{F}}$ is isometrically embedded into $\mathbb{H}_n$. Otherwise a path from a set A to B in $\mathbb{G}_{\mathcal{F}}$ which is not a shortest in $\mathbb{H}_n$ would contain 2 edges with the same label, corresponding to the addition and deletion of the same element of $[n]$.

Now we prove that $dim_{VC}(\mathcal{F}) \leq 1$. Suppose the contrary, namely that $\mathcal{F}$ shatters a set of size 2, e.g. $\{1,2\}$. This means that there are sets $A, B, C, D \in \mathcal{F}$ such that

$$\{1,2\} \cap A = \{1,2\} \quad \{1,2\} \cap B = \{1\} \quad \{1,2\} \cap C = \{2\} \quad \{1,2\} \cap D = \emptyset.$$

Consider a shortest path in $\mathbb{G}_{\mathcal{F}}$ from A to B. Since $2 \in A \backslash B$, this shortest path has to contain an edge labelled with 2. Repeating this argument for C and D one gets another, different (since on a shortest path between A and B every set contains the element 1, on the other hand on a shortest path between C and D none of the sets does) edge with label 2, what contradicts the assumption that all labels are different.

Now we calculate $Sh(\mathcal{F})$. If $i \in [n]$ is not an edge label, then either all sets from $\mathcal{F}$ contain i or none of them does. In particular $\{i\}$ is not shattered by $\mathcal{F}$. Thus $Sh(\mathcal{F})$ consists of $\emptyset$ and the sets $\{i\}$, where i is an edge label. However all edge labels are different, so we get that $|Sh(\mathcal{F})| = |E(\mathbb{G}_{\mathcal{F}})| + 1 = |\mathcal{F}|$ (since $\mathbb{G}_{\mathcal{F}}$ is a tree), i.e. $\mathcal{F}$ is extremal. $\square$

Let now $\mathcal{F} \subseteq 2^{[n]}$ be an extremal family such that $supp(\mathcal{F}) = \cup_{F \in \mathcal{F}} F = [n]$ and $\cap_{F \in \mathcal{F}} F = \emptyset$. By Proposition 5.5 to every extremal family of VC dimension at most 1 we can associate a directed edge-labelled tree $\mathbb{G}_{\mathcal{F}}$, all edges having distinct labels. We have seen that $Sh(\mathcal{F})$ consists of $\emptyset$ and the sets $\{i\}$, where i is an edge label. On the other hand, since $\cap_{F \in \mathcal{F}} F = \emptyset$, we also have that $Sh(\mathcal{F}) = \{\emptyset\} \cup \{\{j\} \mid j \in supp(\mathcal{F}) = [n]\}$. As a consequence the tree must have n edges and thus $n+1$ vertices, i.e. such an extremal family has $n+1$ elements.

Now conversely suppose that we are given a directed edge-labelled tree T on $n+1$ vertices with n edges, all having a different label from $[n]$. This tree at the same time also defines a set system $\mathcal{T} = \{F_v \mid v \in T\}$. Take the edges one by one. When considering an edge with label s going from u to v, then for all vertices w closer to v than to u in the undirected tree put s into F_w. Clearly $T = \mathbb{G}_{\mathcal{T}}$, and by the previous proposition $\mathcal{F}$ must be extremal. Figure 5.1 illustrates such an example with $n = 5$.

This gives a bijection between the set of all extremal families of VC dimension at most 1 and directed edge-labelled trees.

Theorem 5.6. *([36, Theorem 3]) Let $n \geq 1$ be an integer. There is a one-to-one correspondence between extremal families $\mathcal{F} \subseteq 2^{[n]}$ of Vapnik-Chervonenkis dimension 1 with supp($\mathcal{F}$) $= [n]$, $\cap_{F\in\mathcal{F}} F = \emptyset$ and directed edge labelled trees on $n+1$ vertices, all edges having a different label from $[n]$.*

As a corollary one can prove the following statement.

Corollary 5.7. *([36, Corollary 4]) There are $2^n(n+1)^{n-2}$ different extremal families $\mathcal{F} \subseteq 2^{[n]}$ of Vapnik-Chervonenkis dimension at most 1 with supp($\mathcal{F}$) $= [n]$ and $\cap_{F\in\mathcal{F}} F = \emptyset$.*

Proof. There are $(n+1)^{n-2}$ different edge-labelled undirected trees on $n+1$ vertices (see e.g. [15, Proposition 2.1]), all edges having a different label from $[n]$ and each of these trees can be directed in 2^n ways. □

Extremal families from consecutive layers

For an uniform family $\mathcal{F}$ the inclusion graph $\mathbb{G}_{\mathcal{F}}$ is not connected, hence $\mathcal{F}$ cannot be extremal. As a relaxation of uniformity we consider families which belong to two consecutive layers of $2^{[n]}$. The next proposition shows that extremal families among them are actually special cases of the previously studied ones.

Proposition 5.8. *([36, Proposition 5]) Let* $\mathcal{F} \subseteq \binom{[n]}{k} \cup \binom{[n]}{k-1}$ *be an extremal family of subsets of* $[n]$ *with* $n \geq k \geq 1$*. Then we have* $dim_{VC}(\mathcal{F}) \leq 1$*.*

Proof. We will do induction on k. For $k = 1$ the statement is just trivial for every possible value of n. Now suppose that $k > 1$ and the result holds for all values smaller than k (for all possible values of n). Note that for $n \leq 2$ the statement can be verified by an easy case analysis, hence we may suppose that $n > 2$.

We prove that such an extremal family cannot shatter a subset of size 2. Suppose the contrary, namely that $\mathcal{F}$ shatters for example $\{1, 2\}$. Consider the standard subdivision of $\mathcal{F}$ with respect to the element n:

$$\mathcal{F}_0^{(n)} = \{F \mid F \in \mathcal{F} \text{ and } n \notin F\} \subseteq \binom{[n-1]}{k} \cup \binom{[n-1]}{k-1},$$

$$\mathcal{F}_1^{(n)} = \{F \backslash \{n\} \mid F \in \mathcal{F} \text{ and } n \in F\} \subseteq \binom{[n-1]}{k-1} \cup \binom{[n-1]}{k-2}.$$

Since $\mathcal{F}$ is extremal, both $\mathcal{F}_0^{(n)}$ and $\mathcal{F}_1^{(n)}$ must be extremal and for the shattered sets we have the general formula

$$Sh(\mathcal{F}) = Sh(\mathcal{F}_0^{(n)}) \cup Sh(\mathcal{F}_1^{(n)}) \cup \{F \cup \{n\} \mid F \in Sh(\mathcal{F}_0^{(n)}) \cap Sh(\mathcal{F}_1^{(n)})\}.$$

Since $n > 2$, by the induction hypothesis $\{1, 2\} \in Sh(\mathcal{F}_1^{(n)})$ cannot hold, thus we have $\{1, 2\} \in Sh(\mathcal{F}_0^{(n)})$. In this way we constructed an extremal family with the same properties but on a smaller ground set. Continuing this we arrive to an extremal family $\mathcal{F} \subseteq \binom{[k]}{k} \cup \binom{[k]}{k-1}$ that shatters $\{1, 2\}$. However this is easily seen to be impossible, because for any $F \in \mathcal{F}$ we have $|F \cap \{1, 2\}| \geq 1$. This finishes the proof. $\square$

Using essentially the same argument one can prove the following:

Proposition 5.9. *([36, Proposition 6]) Let* $\mathcal{F} \subseteq \binom{[n]}{k} \cup \binom{[n]}{k-1} \cup \cdots \cup \binom{[n]}{k-t+1}$ *be an extremal family of subsets of* $[n]$ *with* $n \geq k \geq t-1 \geq 1$*. Then we have* $dim_{VC}(\mathcal{F}) \leq t-1$*.* $\square$

We return now to the situation when $\mathcal{F} \subseteq \binom{[n]}{k} \cup \binom{[n]}{k-1}$ for some $n \geq k \geq 1$ and $supp(\mathcal{F}) = [n]$, $\cap_{F\in\mathcal{F}} F = \emptyset$. Proposition 5.5 states in this case that $\mathcal{F}$ is extremal if and only if $\mathbb{G}_{\mathcal{F}}$ (the undirected version) is a tree and all labels on the edges are different. As before, we also have that this tree has $n+1$ vertices and n edges.

Now suppose that we are given a tree T on $n+1$ vertices having n edges labelled with elements of $[n]$, all edges having a different label. T can also be viewed as a bipartite graph (since it is acyclic, and so contains no odd cycles) with color classes $\mathcal{A}, \mathcal{B}$. Direct all edges from $\mathcal{A}$ to $\mathcal{B}$, and let $\mathcal{T}$ be as before the set system this directed tree just defines. It is easily seen that we have $\mathcal{T} \subseteq \binom{[n]}{k} \cup \binom{[n]}{k-1}$, where $k = |\mathcal{A}|$ and using the characterization of extremal families we also get that $\mathcal{T}$ is extremal. If we swap the role of $\mathcal{A}$ and $\mathcal{B}$ we get the "dual" set system

$$\mathcal{T}' = \{[n] \backslash F \mid F \in \mathcal{T}\} \subseteq \binom{[n]}{n-k+1} \cup \binom{[n]}{n-k},$$

which is clearly also extremal using the same reasoning.

Permuting the labels on the edges corresponds just to a permutation of the ground set $[n]$, so if we want to characterize extremal set systems up to isomorphism, we can freely omit the labels from the edges.

Summarizing the preceding discussion, we have the following:

Theorem 5.10. *([36, Theorem 7]) Up to isomorphism and the operation of taking the "dual" of a set system, there is a one to one correspondence between extremal set systems $\mathcal{F}$ from two consecutive layers on the ground set $[n]$ ($supp(\mathcal{F}) = [n]$ and $\cap_{F\in\mathcal{F}} F = \emptyset$) and trees on $n+1$ vertices. The bijection is realized via the map $\mathcal{F} \to \mathbb{G}_{\mathcal{F}}$.* $\square$

Ideal bases of extremal families of VC dimension 1

As an application of Proposition 5.5, we determine the Gröbner bases of extremal set systems of VC dimension 1.

Suppose that $\mathcal{F} \subseteq 2^{[n]}$ is an extremal system such that $dim_{VC}(\mathcal{F}) = 1$, $supp(\mathcal{F}) = [n]$ and $\cap_{F \in \mathcal{F}} F = \emptyset$. By Theorem 3.9 there are polynomials of the form $f_{S,H}$, which together with $\{x_i^2 - x_i, i \in [n]\}$ form a universal Gröbner basis of $I(\mathcal{F})$. Moreover any set S in the above description is a minimal set that is not shattered by $\mathcal{F}$, and H is the unique subset of S for which there is no $F \in \mathcal{F}$ for which $F \cap S = H$. In our case we know that $Sh(\mathcal{F})$ is the collection of all sets of size at most 1, so the minimal sets outside $Sh(\mathcal{F})$ are exactly the sets of size 2. Fix one such set $S = \{\alpha, \beta\}$, and consider the 2 edges in the inclusion graph $\mathbb{G}_{\mathcal{F}}$ labelled by α and β. As $\mathbb{G}_{\mathcal{F}}$ is a tree, one can consider the unique path connecting the 2 edges. There are 4 possibilities:

- The edges are directed towards each other on this path:

  $$\boxed{\alpha \notin, \beta \in} \xrightarrow{\quad\alpha\quad} \boxed{\alpha \in, \beta \in} \text{ - - - - - } \boxed{\alpha \in, \beta \in} \xleftarrow{\quad\beta\quad} \boxed{\alpha \in, \beta \notin}$$

  In this case the corresponding set H is $\emptyset$, so $f_{S,H} = (x_\alpha - 1)(x_\beta - 1)$. Indeed then every $F \in \mathcal{F}$ contains either α or β.

- The edges are directed away from each other each other on this path:

  $$\boxed{\alpha \in, \beta \notin} \xleftarrow{\quad\alpha\quad} \boxed{\alpha \notin, \beta \notin} \text{ - - - - - } \boxed{\alpha \notin, \beta \notin} \xrightarrow{\quad\beta\quad} \boxed{\alpha \notin, \beta \in}$$

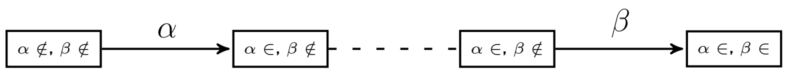
  In this case the corresponding set H is $\{\alpha, \beta\}$, so $f_{S,H} = x_\alpha x_\beta$. No $F \in \mathcal{F}$ contains $\{\alpha, \beta\}$.

- The edges are directed in the same direction towards the edge with label α on this path:

  $$\boxed{\alpha \in, \beta \in} \xleftarrow{\quad\alpha\quad} \boxed{\alpha \notin, \beta \in} \text{ - - - - - } \boxed{\alpha \notin, \beta \in} \xleftarrow{\quad\beta\quad} \boxed{\alpha \notin, \beta \notin}$$

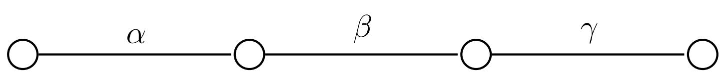
In this case the corresponding set H is $\{\alpha\}$, so $f_{S,H} = x_\alpha(x_\beta - 1)$. If $\alpha \in F$ for some $F \in \mathcal{F}$ then $\beta \in F$ as well.

- The edges are directed in the same direction towards the edge with label β on this path:



Similarly to the previous case $H = \{\beta\}$, so $f_{S,H} = (x_\alpha - 1)x_\beta$.

Now given $\mathbb{G}_\mathcal{F}$, if we do the above analysis for every pair $\alpha, \beta \in [n]$, we obtain a Gröbner basis for $I(\mathcal{F})$. This Gröbner basis will have $\binom{n}{2} + n$ elements. If we want just an ideal basis of $I(\mathcal{F})$ and not necessarily a Gröbner basis, things become easier as we do not need to consider all pairs. Indeed consider 3 consecutive edges in $\mathbb{G}_\mathcal{F}$, i.e. a path of length 3 with labels α, β, γ.



They define 3 pairs and hence 3 polynomials,

$$f_{\alpha,\beta} = (x_\alpha - \varepsilon_\alpha)(x_\beta - \varepsilon_\beta),\ \ f_{\alpha,\gamma} = (x_\alpha - \varepsilon_\alpha)(x_\gamma - \varepsilon_\gamma),\ \ f_{\beta,\gamma} = (x_\beta - 1 + \varepsilon_\beta)(x_\gamma - \varepsilon_\gamma),$$

where ε_α, ε_β and ε_γ are 0 or 1 depending on the orientations of the edges. By checking all 2^3 possibilities for the values of ε_α, ε_β, ε_γ one can easily verify that

$$(x_\gamma - \varepsilon_\gamma)f_{\alpha,\beta} - (x_\alpha - \varepsilon_\alpha)f_{\beta,\gamma} = (1 - 2\varepsilon_\beta)f_{\alpha,\gamma}.$$

Here $1 - 2\varepsilon_\beta$ is either 1 or -1, so $f_{\alpha,\gamma}$ is superfluous in the ideal basis, since it can be obtained from $f_{\alpha,\beta}$ and $f_{\beta,\gamma}$. This means that when constructing an ideal basis of $I(\mathcal{F})$ it is enough to consider only adjacent pairs of edges in $\mathbb{G}_\mathcal{F}$. In this way, depending on the structure of $\mathbb{G}_{\mathcal{F}}$ one can substantially decrease the number of polynomials. For example if $\mathbb{G}_{\mathcal{F}}$ is one single path of length n, we get only $n-1+n$ polynomials. However, if $\mathbb{G}_{\mathcal{F}}$ is a star, one has to take all $\binom{n}{2}+n$ polynomials as in the Gröbner basis.

5.2 Extremal families of VC dimension 2

In [38] we increased t and considered families of VC dimension 2. We characterized extremal systems in this case by providing an algorithmic procedure for constructing the inclusion graphs of all such set systems. In the following first we describe a building process for the set system and then study how the inclusion graph evolves in the meantime.

Let Step 0 be the initialization, after which we are given the set system $\{\emptyset\}$. Now suppose we are given a set system $\mathcal{F}$ and consider the following two types of operations to enlarge $\mathcal{F}$:

- **Step A** - If such exists, take an element $\alpha \in [n]\backslash supp(\mathcal{F})$ together with a set $W \in \mathcal{F}$ and add the set $V = \{W, \alpha\}$ to $\mathcal{F}$.

  Note that the singleton $\{\alpha\}$ is strongly shattered by $\mathcal{F} \cup \{V\}$, as shown by the sets W and V, but is not by $\mathcal{F}$, by the assumption $\alpha \notin supp(\mathcal{F})$.

- **Step B** - If there exist, take two elements $\alpha, \beta \in supp(\mathcal{F})$ such that $\{\alpha, \beta\} \notin st(\mathcal{F})$, together with sets $P, W, Q \in \mathcal{F}$ such that $Q \triangle W = \{\alpha\}$ and $P \triangle W = \{\beta\}$. Let $V = W \triangle \{\alpha, \beta\}$. V is also the unique set satisfying $P \triangle V = \{\alpha\}$ and $Q \triangle V = \{\beta\}$. For these sets we have that $\{P, W, Q, V\} = W \cap V + 2^{\{\alpha,\beta\}} = P \cap Q + 2^{\{\alpha,\beta\}}$, and hence V cannot belong to $\mathcal{F}$, otherwise the sets P, W, Q, V would strongly shatter $\{\alpha, \beta\}$, contradicting our assumption. Therefore, it is reasonable to add V to $\mathcal{F}$.

  Note that the set $\{\alpha, \beta\}$ is strongly shattered by $\mathcal{F} \cup \{V\}$, as shown by the sets P, W, Q and V, but is not by $\mathcal{F}$ by assumption.

Let $\mathcal{E}$ be the collection of all set systems $\mathcal{F}$ that can be built up starting with Step 0 and then using steps of type A and B in an arbitrary but valid order.

Lemma 5.11. *([38, Lemma 14]) Any set system $\mathcal{F} \in \mathcal{E}$ is extremal and $dim_{VC}(\mathcal{F}) \leq 2$.*

Proof. We will use induction on the size of $\mathcal{F}$. If $|\mathcal{F}| = 1$ then necessarily $\mathcal{F} = \{\emptyset\}$, which is clearly extremal and $dim_{VC}(\mathcal{F}) = 0$. Now suppose we know the result for all members of $\mathcal{E}$ of size at most $m \geq 1$, and consider a system $\mathcal{F} \in \mathcal{E}$ of size $m + 1$. As $\mathcal{F} \in \mathcal{E}$ it can be built up starting from $\{\emptyset\}$ using Steps A and B. Fix one such building process, and let $\mathcal{F}'$ be the set system before the last building step. As noted previously, independently of the type of the last step there is a set S that is strongly shattered by $\mathcal{F}$ but is not strongly shattered by $\mathcal{F}'$. S is either a singleton or a set of size 2, depending on the type of the last step. By the induction hypothesis $\mathcal{F}'$ is extremal and $dim_{VC}(\mathcal{F}') \leq 2$. Using the reverse Sauer inequality we get that

$$|\mathcal{F}'| = |st(\mathcal{F}')| < |st(\mathcal{F})| \leq |\mathcal{F}| = |\mathcal{F}'| + 1,$$

what is possible only if $|st(\mathcal{F})| = |st(\mathcal{F}')| + 1$ and $|st(\mathcal{F})| = |\mathcal{F}|$, in particular $\mathcal{F}$ is extremal.

However in the extremal case the family of shattered sets is the same as the family of strongly shattered sets, and so the above reasoning also gives that there is exactly one set that is shattered by $\mathcal{F}$ and is not shattered by $\mathcal{F}'$, namely S, and so

$$dim_{VC}(\mathcal{F}) \leq max(dim_{VC}(\mathcal{F}'), |S|) \leq 2.$$

□

The proof of Lemma 5.11 also describes how the family of shattered/strongly shattered sets grows during a building process. After each step it grows by exactly one new set, namely by $\{\alpha\}$, if the step considered was Step A with the label α, and by $\{\alpha, \beta\}$, if the

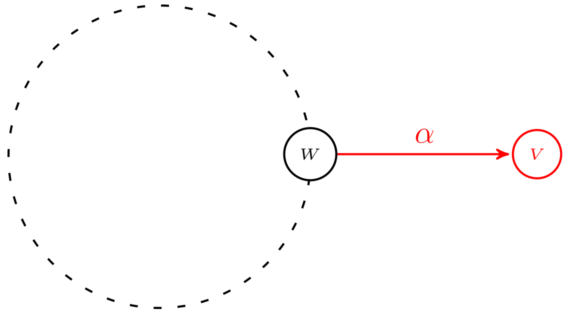


Figure 5.2: Step A

step considered was Step B with labels α, β. By our assumptions on the steps it also follows that a valid building process for a set system $\mathcal{F} \in \mathcal{E}$ cannot involve twice Step A with the same label α, neither twice Step B with the same pair of labels α, β. Moreover we also have that

$$Sh(\mathcal{F}) = st(\mathcal{F}) = \begin{array}{c} \Big\{\emptyset\Big\} \bigcup \Big\{\{\alpha\} \mid \text{Step A is used with label } \alpha\Big\} \bigcup \\ \Big\{\{\alpha, \beta\} \mid \text{Step B is used with labels } \alpha \text{ and } \beta\Big\} \end{array}.$$

Now consider a valid building process from $\mathcal{E}$, and let us examine, how the inclusion graph evolves. We use the notation from the definitions of Steps A and B. Suppose we have already built up a set system $\mathcal{F}$, and we are given its inclusion graph $\mathbb{G}_{\mathcal{F}}$.

In Step A we add a new vertex, namely V to $\mathbb{G}_{\mathcal{F}}$, together with one new directed edge with label α going from W to V. As $\alpha \notin supp(\mathcal{F})$, V has no other neighbors in $\mathbb{G}_{\mathcal{F}}$. Figure 5.2 shows Step A in terms of the inclusion graph.

In Step B we also add one new vertex to $\mathbb{G}_{\mathcal{F}}$, namely V. As the distance of V from both P and Q is 1, and $P \triangle V = \{\alpha\}$ and $Q \triangle V = \{\beta\}$, we have to add at least 2 new edges, one between P and V with label α and one between Q and V with label β. The direction of these edges is predetermined by the vertices P, W and Q. Figure 5.3 shows all possible cases for the directions of these edges. We claim that no other edges need to be added, i.e. V has no other neighbors in $\mathbb{G}_{\mathcal{F}}$. Indeed suppose that the new vertex V has

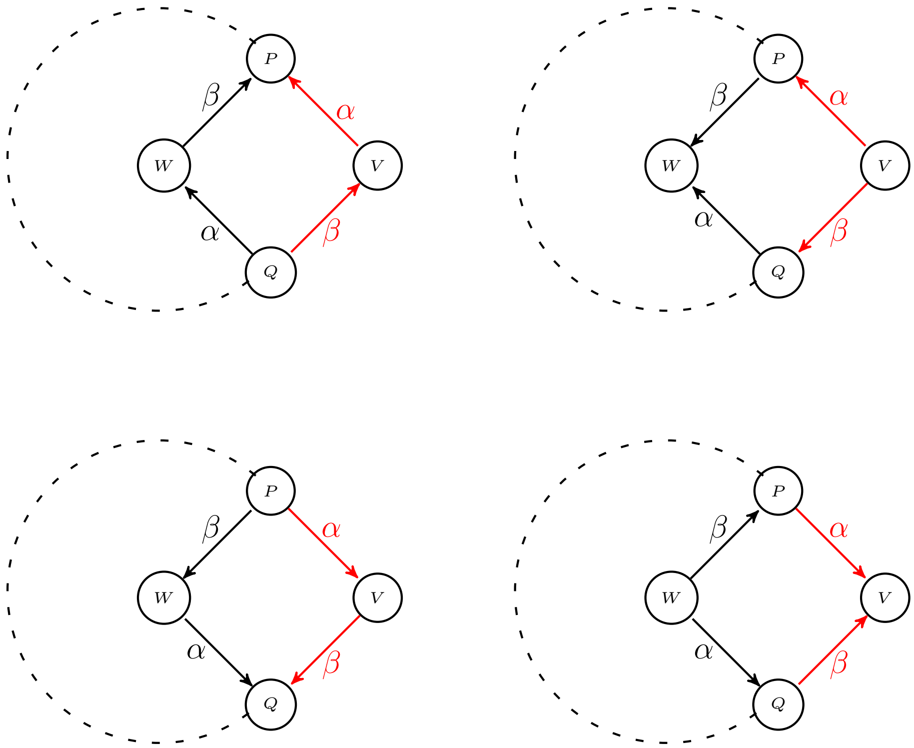


Figure 5.3: Step B

another neighbor X in $\mathbb{G}_{\mathcal{F}}$, different from P and Q, that should be connected to it with some label γ different from α and β. See Figure 5.4, where edge directions are ignored, only edge labels are shown.

Here $d_{\mathbb{H}_n}(P, X) = |P \triangle X| = |\{\alpha, \gamma\}| = 2$. On the other hand as $\mathcal{F}$ was built using Steps A and B starting from $\{\emptyset\}$, it is a member of $\mathcal{E}$, and so by Lemma 5.9 it is extremal. According to Proposition 5.2 this implies that $\mathcal{G}_{\mathcal{F}}$ is isometrically embedded. This means that there should be a vertex Y in $\mathbb{G}_{\mathcal{F}}$ connected to both P and X with edges with labels γ and α respectively. The same reasoning applies for Q and V with some intermediate vertex Z and edge labels β, γ. However in this case, independently of the directions of the edges, we have $\{X \cap \{\alpha, \beta\}, Y \cap \{\alpha, \beta\}, Z \cap \{\alpha, \beta\}, W \cap \{\alpha, \beta\}\} = 2^{\{\alpha, \beta\}}$, i.e. the sets X, Y, Z, W shatter the set $\{\alpha, \beta\}$, and so by the extremality of $\mathcal{F}$ we have that $\{\alpha, \beta\}$ is also strongly shattered, what contradicts the assumptions of Step B.

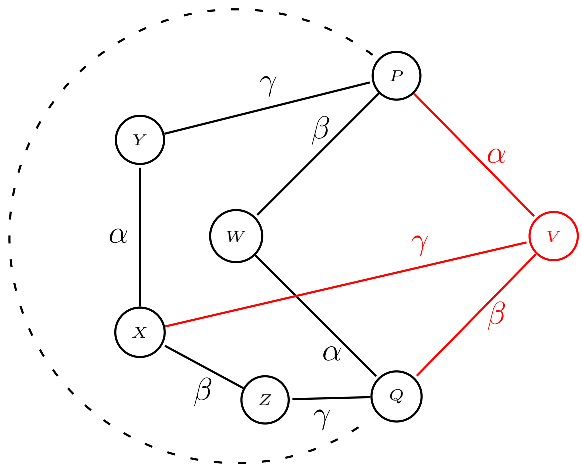


Figure 5.4: Case of Step B

From now on it will depend on the context whether we regard Steps A and B as building steps for extremal set systems of VC dimension at most 2 or as building steps for their inclusion graphs.

Figure 5.5 shows a possible building process in $\mathcal{E}$ for the set system

$$\mathcal{F} = \{\emptyset, \{1\}, \{2\}, \{3\}, \{2, 3\}\}$$

in terms of the inclusion graph.

Take an element of $\mathcal{E}$ and fix a valid building process for it. The above observations also imply, that when observing the evolution of the inclusion graph, after the first occurrence of an edge with some fixed label α, new edges with the same label can come up only when using Step B always with a different label next to α. By easy induction on the number of building steps, this results that between any two edges with the same label α there is a "path of 4-cycles". See Figure 5.6. Note that by assumptions all the β_i's in Figure 5.6 must be different. Along this path of 4-cycles we also obtain a shortest path between X_1 and X_2, and similarly between Y_1 and Y_2.

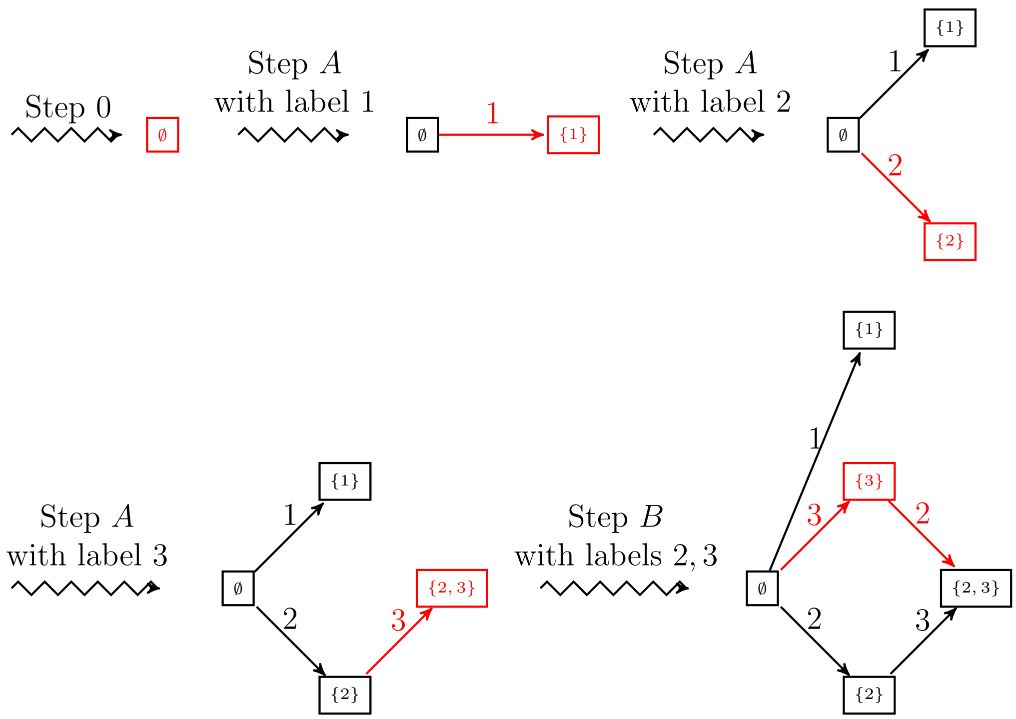


Figure 5.5: Example of the building process in $\mathcal{E}$

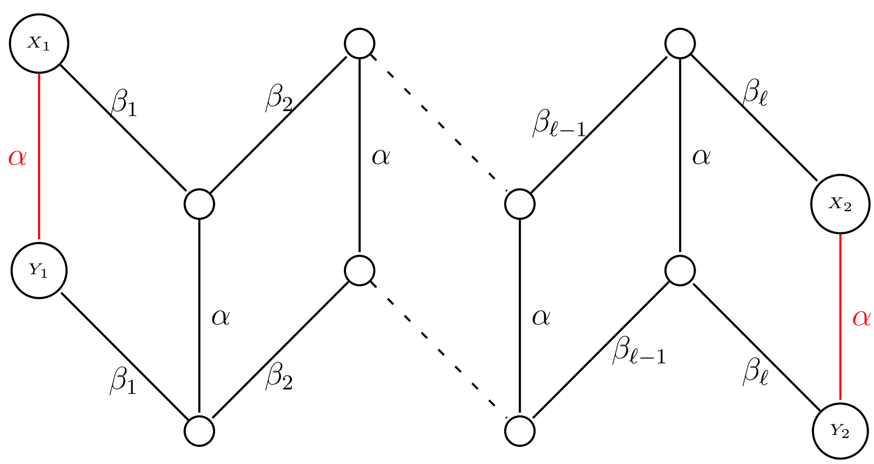


Figure 5.6: Path of 4 cycles

The first of the main results of this section is that the set systems in $\mathcal{E}$, described above, are actually all the extremal set systems of VC dimension at most 2 and containing $\emptyset$.

Theorem 5.12. *([38, Theorem 15]) $\mathcal{F} \subseteq 2^{[n]}$ is an extremal set system with $\emptyset \in \mathcal{F}$ and $dim_{VC}(\mathcal{F}) \leq 2$ if and only if $\mathcal{F} \in \mathcal{E}$.*

Before turning to the proof of Theorem 5.12, we first prove a lemma about the building processes in $\mathcal{E}$, that will play a key role further on.

Lemma 5.13. *([38, Lemma 16]) Suppose that $\mathcal{F}', \mathcal{F}$ are elements of $\mathcal{E}$ such that $\mathcal{F}' \subseteq \mathcal{F}$. Then $\mathcal{F}'$ can be extended with valid building process to build up $\mathcal{F}$.*

Proof. Suppose this is not the case, and consider a counterexample. Without loss of generality we may suppose that the counterexample is such that $\mathcal{F}'$ cannot be continued with any valid step towards $\mathcal{F}$. $\mathcal{F}'$ and $\mathcal{F}$ are both extremal and so $\mathbb{G}_{\mathcal{F}'}$ and $\mathbb{G}_{\mathcal{F}}$ are both isometrically embedded, in particular connected, hence the neighborhood of $\mathbb{G}_{\mathcal{F}'}$ inside $\mathbb{G}_{\mathcal{F}}$ is nonempty. Now take a closer look at the edges on the boundary of $\mathbb{G}_{\mathcal{F}'}$.

If there would be an edge going out from $\mathbb{G}_{\mathcal{F}'}$ with a label $\alpha \in supp(\mathcal{F}) \backslash supp(\mathcal{F}')$, then Step A would apply with this label α. On the other hand there cannot be an edge going into $\mathbb{G}_{\mathcal{F}'}$ with a label $\alpha \notin supp(\mathcal{F}')$, otherwise the endpoint of this edge inside $\mathbb{G}_{\mathcal{F}'}$ would contain α, what would be a contradiction.

We can therefore assume that the label of any edge on the boundary of $\mathbb{G}_{\mathcal{F}'}$, independently of the direction of the edge, is an element of $supp(\mathcal{F}')$. However as $\emptyset \in \mathcal{F}'$ and $\mathbb{G}_{\mathcal{F}'}$ is isometrically embedded, an element belongs to $supp(\mathcal{F}')$ only if it appears as an edge label in $\mathbb{G}_{\mathcal{F}'}$. Now take an edge (W, V) on the boundary of $\mathbb{G}_{\mathcal{F}'}$ with $W \in \mathcal{F}'$, $V \in \mathcal{F} \backslash \mathcal{F}'$ and with some label α, together with an edge (X, Y) with the same label inside $\mathbb{G}_{\mathcal{F}'}$. Denote the distance of the edges (W, V) and (X, Y) by ℓ, i.e. $d_{\mathbb{H}_n}(W, X) = d_{\mathbb{H}_n}(V, Y) = \ell$. The latter equality means, that depending on the direction of the edges, W and X both do

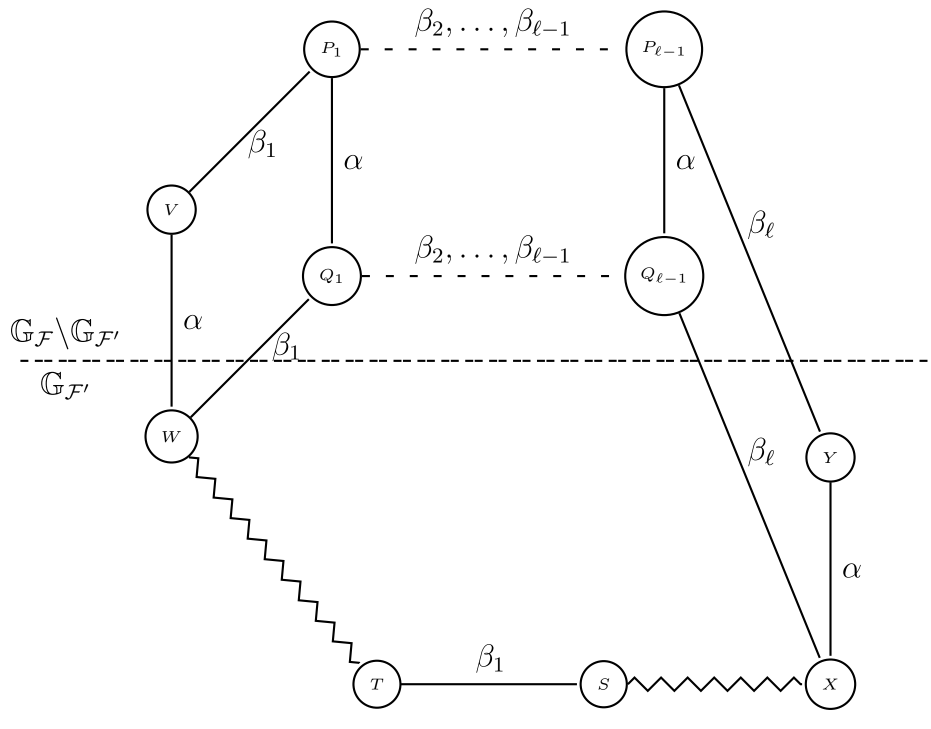


Figure 5.7: Case $l > 1$

contain the element α, or neither of them does. Suppose that the triple α, (W, V), (X, Y) is such that the distance ℓ is minimal.

First suppose that $\ell > 1$. Since the edges $(W, V), (X, Y)$ have the same label and $\mathcal{F} \in \mathcal{E}$, there is a path of 4-cycles of length ℓ between them inside $\mathbb{G}_{\mathcal{F}}$. This path of 4-cycles also provides shortest paths between the endpoints of the edges $(W, V), (X, Y)$. By the minimality of our choice, in this path, except the edges at the ends, there cannot be an edge with label α neither totally inside $\mathbb{G}_{\mathcal{F}'}$, neither on the boundary of it, meaning that this path of 4-cycles is essentially going outside $\mathbb{G}_{\mathcal{F}'}$. See Figure 5.7.

Since $\mathbb{G}_{\mathcal{F}'}$ is isometrically embedded and $d_{\mathbb{H}_n}(W, X) = \ell$, there must be a path of length ℓ between W and X inside $\mathbb{G}_{\mathcal{F}'}$. As this path runs inside $\mathbb{G}_{\mathcal{F}'}$, it has to be disjoint from the path of 4-cycles. Along the path of 4-cycles all the β_i's are different, so for each i exactly one of the sets W and X contains the element β_i. In particular for $i = 1$, the shortest path between W and X inside $\mathbb{G}_{\mathcal{F}'}$ also has to contain an edge (T, S) with label β_1 with

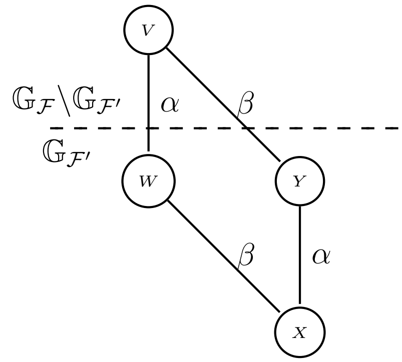


Figure 5.8: Case $\ell = 1$

direction determined by the sets W and X. However the distance between W and T is at most $\ell - 1$, and hence the triple $\beta_1, (W, Q_1), (T, S)$ contradicts with the minimality of the initial triple $\alpha, (W, V), (X, Y)$ where the distance was ℓ.

By the above reasoning only $\ell = 1$ is possible. In this case the endpoints of the edges (W, V), (X, Y) are connected by edges with the same label. Let this label be β. See Figure 5.8. The direction of these edges is predetermined by $\mathbb{G}_{\mathcal{F}'}$. $\{\alpha, \beta\} \notin st(\mathcal{F}')$, otherwise there would be already a copy of $\mathbb{G}_{2\{\alpha,\beta\}}$ in $\mathbb{G}_{\mathcal{F}'}$, which together with the vertices W, V, X, Y would give us two different copies of it inside $\mathbb{G}_{\mathcal{F}}$, which is impossible by Observation 5.4, as $\{\alpha, \beta\}$ is a maximal set strongly shattered by the extremal family $\mathcal{F}$. Hence Step B applies with new vertex V, edges $(W, V), (V, Y)$ and labels α, β respectively, contradicting with the fact, that we started with a counterexample. $\square$

Now we are ready to prove Theorem 5.12.

Proof. One direction of the theorem is just Lemma 5.11. For the other direction we use induction on the number of sets in $\mathcal{F}$. If $|\mathcal{F}| = 1$, then $\mathcal{F}$ is necessarily $\{\emptyset\}$, and so it belongs trivially to $\mathcal{E}$. Now suppose we proved the statement for all set systems with at most $m-1$ members, and let $\mathcal{F}$ be an extremal family of size m, of VC dimension at most 2 and containing $\emptyset$. Take an arbitrary element α appearing as a label of an edge going out from $\emptyset$ in $\mathbb{G}_{\mathcal{F}}$, i.e. an element α such that $\{\alpha\} \in \mathcal{F}$. Consider the standard subdivision of

$\mathcal{F}$ with respect to the element α with parts $\mathcal{F}_0$ and $\mathcal{F}_1$ (see Definition 2.28), and let

$$\widehat{\mathcal{F}}_1 = \{F \cup \{\alpha\} \ : \ F \in \mathcal{F}_1\}.$$

Note that with respect to shattering and strong shattering $\mathcal{F}_1$ and $\widehat{\mathcal{F}}_1$ behave in the same way. Since $\mathcal{F}$ is extremal, so are $\mathcal{F}_0$, $\mathcal{F}_1$ and hence $\widehat{\mathcal{F}}_1$ as well, and clearly their VC dimension is at most 2. The collection of all edges with label α in the inclusion graph $\mathbb{G}_{\mathcal{F}}$ forms a cut. This cut divides $\mathbb{G}_{\mathcal{F}}$ into two parts, that are actually the inclusion graphs $\mathbb{G}_{\mathcal{F}_0}$ and $\mathbb{G}_{\widehat{\mathcal{F}}_1}$. Note that $\mathbb{G}_{\mathcal{F}_1}$ and $\mathbb{G}_{\widehat{\mathcal{F}}_1}$ are isomorphic as directed edge-labelled graphs. Let T_0 and T_1 be the induced subgraphs on the endpoints of the cut edges in $\mathbb{G}_{\mathcal{F}_0}$ and $\mathbb{G}_{\widehat{\mathcal{F}}_1}$, respectively. See Figure 5.9. T_0 and T_1 are isomorphic, and they are actually the inclusion graphs of the set systems $\mathcal{T}_0 = \mathcal{F}_0 \cap \mathcal{F}_1 = M_\alpha(\mathcal{F})$ (see Definition 2.32) and $\mathcal{T}_1 = \{F \cup \{\alpha\}, F \in \mathcal{T}_0\}$. Similarly to the pair $\mathcal{F}_1$, $\widehat{\mathcal{F}}_1$, the set systems $\mathcal{T}_0$ and $\mathcal{T}_1$ also behave in the same way with respect to shattering and strong shattering. By assumption $\mathcal{F}$ is extremal, and hence so are $\mathcal{T}_0$ (since M_α preserves extremality) and $\mathcal{T}_1$. For every set S in $Sh(\mathcal{T}_0) = Sh(\mathcal{F}_0 \cap \mathcal{F}_1) \subseteq 2^{[n]\setminus\{\alpha\}}$ the set $S \cup \{\alpha\}$ is shattered by $\mathcal{F}$, implying that $dim_{VC}(\mathcal{T}_0) \leq dim_{VC}(\mathcal{F}) - 1 \leq 1$. Therefore $\mathcal{T}_0$ is an extremal family of VC dimension at most 1, and so by Proposition 5.5 we get that T_0 (and hence T_1) is a directed edge-labelled tree having all edge labels different. Note that for any edge label β appearing in T_0 (and hence in T_1), there is a copy of $\mathbb{G}_{2^{\{\alpha,\beta\}}}$ along the cut, implying that $\{\alpha, \beta\} \in st(\mathcal{F}) = Sh(\mathcal{F})$. By the VCdimension constraint on $\mathcal{F}$ the set $\{\alpha, \beta\}$ is a maximal element of $st(\mathcal{F}) = Sh(\mathcal{F})$, and so by Observation 5.4 there cannot be another copy of $\mathbb{G}_{2^{\{\alpha,\beta\}}}$ in $\mathbb{G}_{\mathcal{F}}$, neither in $\mathbb{G}_{\mathcal{F}_0}$ nor in $\mathbb{G}_{\widehat{\mathcal{F}}_1}$, in particular $\{\alpha, \beta\} \notin st(\mathcal{F}_0)$.

Let's now turn to the building process of $\mathcal{F}$. Our choice of α guarantees that $\emptyset \in \mathcal{F}_0, \mathcal{F}_1$ and so by the induction hypothesis both of them belong to $\mathcal{E}$. In particular we can build up $\mathcal{F}_0$, and in the meantime $\mathbb{G}_{\mathcal{F}_0}$, according to the building rules in $\mathcal{E}$. $\alpha \notin supp(\mathcal{F}_0)$ and so we

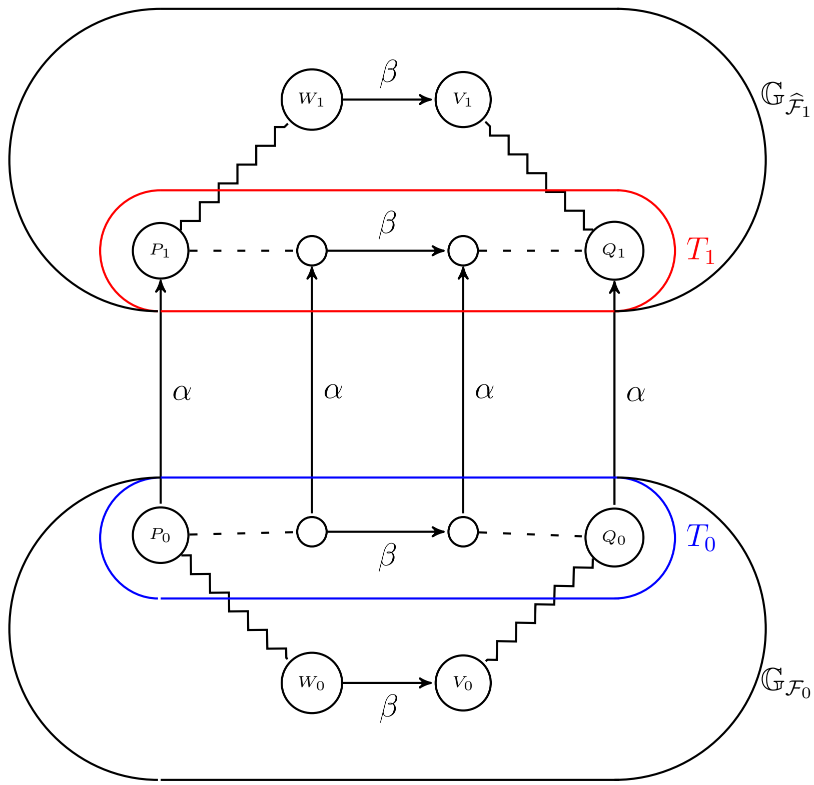


Figure 5.9: Building up extremal set systems

can apply Step A with α to add one fixed cut edge to $\mathbb{G}_{\mathcal{F}_0}$. Then we apply Step B several times to add the whole of T_1 to $\mathbb{G}_{\mathcal{F}_0}$ and simultaneously $\mathcal{T}_1$ to $\mathcal{F}_0$. By earlier observations all edge labels of T_1 are different, and if β is such a label, then $\{\alpha, \beta\} \notin st(\mathcal{F}_0)$, and hence all these applications of Step B will be valid ones. The building process so far shows that $\mathcal{F}_0 \cup \mathcal{T}_1$ is also a member of $\mathcal{E}$. $\mathbb{G}_{\mathcal{F}_0 \cup \mathcal{T}_1}$ is just $\mathbb{G}_{\mathcal{F}_0}$ and T_1 glued together along the cut in the way described above.

T_0 shows that $\mathcal{T}_0$ can be built up using only Step A, and hence it belongs to $\mathcal{E}$. The inclusion $\mathcal{T}_1 \subseteq \widehat{\mathcal{F}}_1$ shows that $\mathcal{T}_0 \subseteq \mathcal{F}_1$, therefore by Lemma 5.13 $\mathcal{T}_0$ can be extended with a valid building process to build up $\mathcal{F}_1$. This extension can also be considered as building up $\widehat{\mathcal{F}}_1$ from $\mathcal{T}_1$. $\emptyset \notin \mathcal{T}_1, \widehat{\mathcal{F}}_1$ and so neither of the two systems is a member of $\mathcal{E}$, however this causes no problems, as the pairs $\mathcal{T}_0$, $\mathcal{T}_1$ and $\mathcal{F}_1$, $\widehat{\mathcal{F}}_1$ behave in the same way with respect to shattering and strong shattering, and so all building steps remain valid.

We claim, that this last building procedure remains valid, and so completes a desired building process for $\mathcal{F}$, if we start from $\mathcal{F}_0 \cup \mathcal{T}_1$ instead of $\mathcal{T}_1$. First note that if there is a label appearing both in $\mathbb{G}_{\mathcal{F}_0}$ and $\mathbb{G}_{\widehat{\mathcal{F}}_1}$, then it appears also in T_0, and hence in T_1. Indeed let β be such a label, and consider 2 edges with this label, one going from W_0 to V_0 in $\mathbb{G}_{\mathcal{F}_0}$ and the other going from W_1 to V_1 in $\mathbb{G}_{\widehat{\mathcal{F}}_1}$. See Figure 5.9. $\mathbb{G}_{\mathcal{F}}$ is isometrically embedded, therefor there is a shortest path both between W_0 and W_1 and between V_0 and V_1 in $\mathbb{G}_{\mathcal{F}}$. Thanks to β these two paths have to be disjoint. Both of these paths must have a common edge with the cut, say (P_0, P_1) and (Q_0, Q_1), with P_0 and Q_0 in $\mathbb{G}_{\mathcal{F}_0}$. Since $\beta \in P_0 \triangle Q_0$, along the shortest path between P_0 and Q_0 in the isometrically embedded inclusion graph T_0 of the extremal family $\mathcal{T}_0$ there must be an edge with label β. According to this, when applying Step A in the extension process, then the used element will be new not just when we start from $\mathcal{T}_1$, but also when starting from $\mathcal{F}_0 \cup \mathcal{T}_1$.

Finally suppose that an application of Step B with some labels β, γ in the extension process turns invalid when we start from $\mathcal{F}_0 \cup \mathcal{T}_1$ instead of $\mathcal{T}_1$. This is possible only if $\{\beta, \gamma\} \in st(\mathcal{F}_0 \cup \mathcal{T}_1) \setminus st(\mathcal{T}_0)$, i.e. there is a copy of $\mathbb{G}_{2^{\{\beta,\gamma\}}}$ already in $\mathbb{G}_{\mathcal{F}_0 \cup \mathcal{T}_1}$. However this copy together with the copy, that the invalid use of Step B results, gives two different occurrences of $\mathbb{G}_{2^{\{\beta,\gamma\}}}$ inside $\mathbb{G}_{\mathcal{F}}$, which is impossible by Observation 5.4, as $\{\beta, \gamma\}$ is a maximal set strongly shattered by the extremal family $\mathcal{F}$. $\square$

5.3 The eliminability conjecture

Concerning the structure of extremal set systems the question naturally arises whether an extremal family can be built up from the empty system by adding sets to it one-by-one in such a way that at each step we have an extremal family. Accordingly in [36] we posed the following question:

Open problem 3. *For a nonempty extremal family $\mathcal{F} \subseteq 2^{[n]}$ does there always exist a set*

$F \in \mathcal{F}$ such that $\mathcal{F} \backslash \{F\}$ is still extremal?

From Theorem 2 of [12] we know that $\mathcal{F}$ is extremal if and only if $2^{[n]} \backslash \mathcal{F}$ is extremal, thus the above question has an equivalent form:

Open problem 4. *For an extremal family $\mathcal{F} \subsetneq 2^{[n]}$ does there always exist a set $F \in \mathcal{F}$ such that $\mathcal{F} \cup \{F\}$ is still extremal?*

There are several special cases when the answer appears to be true for Open problem 3.

Example 5.14. As previously noted, if $\mathcal{F}$ is a nonempty down-set then $\mathcal{F}$ is extremal. Moreover in this case if we omit any maximal element from $\mathcal{F}$ then it remains still a down set and so it will be still extremal.

Example 5.15. If $\mathcal{F}$ is an extremal family of VC dimension 1, then according to Proposition 5.5, if we omit a set corresponding to a leaf, i.e. to a vertex of degree 1 in $G_{\mathcal{F}}$, then the resulting set system will still be extremal.

Example 5.16. Take $\mathcal{F} \subseteq 2^{[n]}$ to be a nonempty extremal family of VC dimension at most 2. Let $F \in \mathcal{F}$ be an arbitrary set from the set system and let $\varphi = \prod_{i \in F} \varphi_i$. Since bit flips preserve extremality, $\varphi(\mathcal{F})$ is extremal as well. Moreover $\varphi(F) = \emptyset \in \varphi(\mathcal{F})$, and hence by Theorem 5.12 we have $\varphi(\mathcal{F}) \in \mathcal{E}$, and we can consider a building process for it. Let $V \in \varphi(\mathcal{F})$ be the set added in the last step of this building process. The same building process shows that $\varphi(\mathcal{F}) \backslash \{V\} \in \mathcal{E}$, and hence by Theorem 5.12 we have that $\varphi(\mathcal{F}) \backslash \{V\}$ is an extremal family of VC dimension at most 2 and containing $\emptyset$. However $\varphi(\mathcal{F}) \backslash \{V\}$ is clearly $\varphi(\mathcal{F} \backslash \{\varphi(V)\})$, and since bit flips preserve extremality, we get that $\varphi(\varphi(\mathcal{F} \backslash \{\varphi(V)\})) = \mathcal{F} \backslash \{\varphi(V)\}$ is also extremal, meaning that the set $\varphi(V) \in \mathcal{F}$ can be removed from the extremal system $\mathcal{F}$ so that the result is still extremal.

Example 5.17. Anstee in [6] considered maximal set systems $\mathcal{F} \subseteq 2^{[n]}$, $|\mathcal{F}| = \binom{n}{0}+\binom{n}{1}+\binom{n}{2}$ without triangles, i.e. set systems with the property, that for all 3-element subsets F we have that $\mathcal{F}|_F$ (see Definition 2.34) does not contain all 2-element subsets of F. Note that in particular the VC dimension of $\mathcal{F}$ is bounded from above by 2, hence we have that $Sh(\mathcal{F}) \subseteq \binom{[n]}{0} \cup \binom{[n]}{1} \cup \binom{[n]}{2}$, implying that $|Sh(\mathcal{F})| \leq \binom{n}{0} + \binom{n}{1} + \binom{n}{2}$. Comparing the sizes of $\mathcal{F}$ and $Sh(\mathcal{F})$ we obtain that such set systems are extremal.

Clearly any such maximal set system $\mathcal{F}$ contains the extremal subsystem $\binom{[n]}{0} \cup \binom{[n]}{1}$. For the remaining part of these set systems Anstee's construction can be interpreted in an inductive way as follows:

- $\mathcal{F}_1 := \binom{[n]}{0} \cup \binom{[n]}{1}$
- For $k = 2, 3, \ldots, n$ suppose we already constructed $\mathcal{F}_{k-1}$. Let $\mathcal{G}_{k-1}$ be the collection of all $k-1$-element sets in $\mathcal{F}_{k-1}$. Define G_{k-1} to be a graph, whose vertex set is $\mathcal{G}_{k-1}$ and there is an edge between $A, B \in \mathcal{G}_{k-1}$ exactly when $|A \triangle B| = 2$. Take a spanning tree T_{k-1} of G_{k-1}.

$$\mathcal{F}_k := \mathcal{F}_{k-1} \cup \{A \cup B \mid (A, B) \text{ is an edge of } T_{k-1}\}$$

- $\mathcal{F} := \mathcal{F}_n$

It is not hard to prove that when we add $A \cup B$, there will be a unique new element that gets into the family of shattered sets, namely $A \triangle B$, hence the resulting system after each step will be extremal. Reversing it, if $\mathcal{F}$ is such an example, then its elements can be deleted one-by-one in such a way that the remaining set system is extremal after each step.

Example 5.18. More generally one can consider set systems $\mathcal{F} \subseteq 2^{[n]}$ with the property, that for all t-element subsets F we have that $\mathcal{F}|_F$ does not contain all l-element subsets of F, for some l with $n \geq t \geq l \geq 0$. Füredi and Quinn in [23] constructed for all values $n \geq t \geq l \geq 0$ a set system $\mathcal{F}(n, t, l)$ with the desired property and of size $\sum_{i=0}^{t-1} \binom{n}{i}$. The same argument as above shows that $Sh(\mathcal{F}(n, t, l))$ consists of all sets of size at most $t - 1$ and hence $\mathcal{F}(n, t, l)$ is extremal for all possible values. Their construction is as follows.

For $x_1, \ldots, x_i \in [n]$, $x_1 < \cdots < x_i$ let

$$E(x_1, \ldots, x_i) = \{x \in [n] \mid x = x_j \text{ for } j \leq l\}$$
$$\cup \{x \in [n] \mid x > x_l \text{ but } x \neq x_j \text{ for any } j > l\},$$

in particular $E(\emptyset) = \emptyset$. Let $\mathcal{F}(n, t, l)$ consist of all $E(x_1, \ldots, x_i)$ where $i \leq t - 1$. Order the sets of $\mathcal{F}(n, t, l)$ as follows: $E(X) \succ E(Y)$ if either $|X| > |Y|$, or $|X| = |Y|$ and $X \succ Y$ with respect to the standard lexicographic ordering. It is not hard to see, that if we remove the elements of $\mathcal{F}(n, t, l)$ with respect to this ordering one-by-one, starting from the largest one, then each time when we remove some $E(X)$, then X is eliminated from the family of shattered sets, hence after each step the resulting family will be still extremal.

To finish this chapter we remark that the building process from Section 5.2 can be generalized to the case when the VC dimension bound is some fixed natural number $t > 2$ as well. We can define a building step for every set $S \subseteq [n]$ with $|S| \leq t$. Let Step($\emptyset$) be the initialization, after which we are given the set system $\{\emptyset\}$. For some set $S \subseteq [n]$ with $|S| \leq t$, Step(S) can be applied to a set system $\mathcal{F}$, if there exists some set $F \subseteq [n]$, $F \notin \mathcal{F}$, such that $S \in st(\mathcal{F} \cup \{F\}) \backslash st(\mathcal{F})$. If such set F exists, choose one, and let the resulting system be $\mathcal{F} \cup \{F\}$. In terms of the inclusion graph $S \in st(\mathcal{F} \cup \{F\}) \backslash st(\mathcal{F})$ means, that by adding the set F there arises a copy of $\mathbb{G}_{2^S}$ inside $\mathbb{G}_{\mathcal{F} \cup \{\mathcal{F}\}}$ containing the vertex F. Similarly as previously, one can prove that F's only neighbors are the ones contained in this copy of $\mathbb{G}_{2^S}$. Using this observation Step(S) could have been defined in terms of the inclusion graph as well (as it was done in the case $t = 2$).

Restrict our attention to those set systems, that can be built up starting with Step($\emptyset$), and then using always new building steps, i.e. not using a building step with the same set S twice. Along the same lines of thinking as in Lemma 5.11, one can prove that every such set system is extremal and of VC dimension at most t. We think, that these set systems are actually all the extremal families of VC dimension at most t. Unfortunately, for the time being we were unable to prove a suitable generalization of Lemma 5.13. Once it is done, the generalization of Theorem 5.12 would follow easily for general t. Although this general version would not give such a transparent structural description of extremal systems as in the case $t = 1$, but still it would imply an affirmative answer for Open problem 3.

Part II

Alon's Combinatorial Nullstellensatz

Chapter 6

The Combinatorial Nullstellensatz and the Non-vanishing Theorem

Alon's famous Combinatorial Nullstellensatz is a specialized and strengthened version of the Hilbertsche Nullstellensatz, a fundamental theorem of algebraic geometry.

Let, as before, $\mathbb{F}$ be a field and for a finite set of points $V \subseteq \mathbb{F}^n$ let $I(V)$ be the vanishing ideal of V. On the other hand, for an ideal $I \lhd \mathbb{F}[\mathbf{x}]$ let its vanishing set be defined as

$$V(I) = \{\mathbf{v} \in \mathbb{F}^n \mid f(\mathbf{v}) = 0 \text{ for every } f \in I\}.$$

At first sight one would think that these two operations are in some sense inverses of each other, i.e. for an ideal $I \lhd \mathbb{F}[\mathbf{x}]$ one has $I(V(I)) = I$. However this is not necessarily the case as it is shown by the ideal $I = (x^2) \lhd \mathbb{F}[x]$, where we have that $I(V(I)) = (x) \supsetneq (x^2)$. Before resolving this problem, we will need one more definition. The radical of an ideal $I \lhd \mathbb{F}[\mathbf{x}]$ is

$$\sqrt{I} = \{f \in \mathbb{F}[\mathbf{x}] \mid \exists t \in \mathbb{N} \text{ such that } f^t \in I\}.$$

Note that in the above example we have that $\sqrt{(x^2)} = (x)$, i.e. $I(V((x^2))) = \sqrt{(x^2)}$. With an extra condition on the ground field this is true in general.

Theorem 6.1. *(Hilbert's Nullstellensatz - see e.g. [16, Theorem 1.6]) If $\mathbb{K}$ is an algebraically closed field and $I \lhd \mathbb{K}[\mathbf{x}]$ a polynomial ideal, then $I(V(I)) = \sqrt{I}$.*

By Hilbert's Basis Theorem I is finitely generated as an ideal, i.e. there are polynomials $f_1(\mathbf{x}), \ldots, f_N(\mathbf{x}) \in \mathbb{K}[\mathbf{x}]$ such that $I = (f_1, \ldots, f_N)$. This means that some polynomial $f(\mathbf{x}) \in \mathbb{K}[\mathbf{x}]$ belongs to I if and only if there are polynomials $h_1(\mathbf{x}), \ldots, h_N(\mathbf{x}) \in \mathbb{K}[\mathbf{x}]$ such that

$$f = h_1 f_1 + \cdots + h_N f_N,$$

and also that $V(I)$ is just the set of common zeros of the f_i's. With this in mind Hilbert's Nullstellensatz can be reformulated. Indeed it states that over an algebraically closed field $\mathbb{K}$ a polynomial $f(\mathbf{x}) \in \mathbb{K}[\mathbf{x}]$ vanishes over all the common zeros of some polynomials $f_1(\mathbf{x}), \ldots, f_N(\mathbf{x}) \in \mathbb{K}[\mathbf{x}]$ if and only if there are polynomials $h_1(\mathbf{x}), \ldots, h_N(\mathbf{x}) \in \mathbb{K}[\mathbf{x}]$ and a positive integer t such that

$$f^t = h_1 f_1 + \cdots + h_N f_N.$$

Let's now turn to the Combinatorial Nullstellensatz. Henceforth we won't assume anymore that the field we are working with is algebraically closed. We return to the assumption that $\mathbb{F}$ is an arbitrary field, however we restrict ourselves to polynomials of some special type.

Theorem 6.2. *(Alon's Combinatorial Nullstellensatz - [3, Theorem 1.1]) Let $\mathbb{F}$ be an arbitrary field and let $f(\mathbf{x})$ be a polynomial in $\mathbb{F}[\mathbf{x}]$. Further let $S_1, \ldots, S_n$ be finite, nonempty* *subsets of* $\mathbb{F}$ *and for* $1 \le i \le n$ *define*

$$g_i(x_i) = \prod_{s \in S_i} (x_i - s).$$

If f *vanishes over all the common zeros of the polynomials* $g_1, \ldots, g_n$ *(that is* $f(\mathbf{s}) = 0$ *for all* $\mathbf{s} \in \mathbf{S} = S_1 \times S_2 \times \cdots \times S_n$*), then there are polynomials* $h_1(\mathbf{x}), \ldots, h_n(\mathbf{x}) \in \mathbb{F}[\mathbf{x}]$ *satisfying* $\deg(h_i) \le \deg(f) - \deg(g_i)$ *so that*

$$f = \sum_{i=1}^{n} h_i g_i.$$

The point set $\mathbf{S}$ is called a discrete box. Note that the vanishing set of the ideal $(g_1, \ldots, g_n)$ is just $\mathbf{S}$ and by the above theorem $I(\mathbf{S}) = (g_1, \ldots, g_n)$ also holds.

Theorem 6.2 can be easily reformulated using Gröbner bases. It states that the polynomials $\{g_1, \ldots, g_n\}$ form a universal Gröbner basis of the vanishing ideal $I(\mathbf{S})$.

As a corollary, a simple and widely applicable non-vanishing criterion has been deduced. It provides a sufficient condition for a polynomial $f(\mathbf{x}) \in \mathbb{F}[\mathbf{x}]$ for not vanishing everywhere on the discrete box $\mathbf{S}$.

Theorem 6.3. *(Alon's Non-vanishing Theorem - [3, Theorem 1.2]) Let* $\mathbb{F}$ *be a field,* $S_1, \ldots, S_n \subseteq \mathbb{F}$, $|S_i| > t_i$*, where each* t_i *is a nonnegative integer. Put* $\mathbf{S} = S_1 \times S_2 \times \cdots \times S_n$ *and let* $p(\mathbf{x}) \in \mathbb{F}[\mathbf{x}]$ *be a polynomial. Suppose that* $\deg p = \sum_{i=1}^{n} t_i$ *and the coefficient of the monomial*

$$m = x_1^{t_1} x_2^{t_2} \cdots x_n^{t_n}$$

in p *is not 0. Then there exists* $\beta \in \mathbf{S}$*, such that* $p(\beta) \neq 0$.

In [31], [32] and [37] generalizations of these results are considered in various settings.

6.1 The Non-vanishing Theorem for multisets

One possible direction when trying to generalize Theorem 6.3 is to allow multiple points. Results in this section were formulated and proven by Géza Kós and Lajos Rónyai in [32], and so are included in this thesis only for the sake of completeness. The proof of the main result will be omitted.

It is well known that if $\mathbb{F}$ is a field then for an arbitrary $\mathbf{s} \in \mathbb{F}^n$ we can express a polynomial $f(\mathbf{x}) \in \mathbb{F}[\mathbf{x}]$ as

$$f(\mathbf{x}) = \sum_{\mathbf{u}} f_{\mathbf{u}}(\mathbf{s})(\mathbf{x}-\mathbf{s})^{\mathbf{u}},$$

where the coefficients $f_{\mathbf{u}}(\mathbf{s}) \in \mathbb{F}$ are uniquely determined by f, $\mathbf{u}$ and $\mathbf{s}$. In particular we have $f_{\mathbf{0}}(\mathbf{s}) = f(\mathbf{s})$ for all $\mathbf{s} \in \mathbb{F}^n$. Observe that if $u_1 + \cdots + u_n \geq \deg f$, then $f_{\mathbf{u}} = f_{\mathbf{u}}(\mathbf{s})$ does not depend on $\mathbf{s}$.

Suppose now that $S_1, S_2, \ldots, S_n$ are nonempty finite subsets of $\mathbb{F}$, and assume further that for $i = 1, \ldots, n$ we have a positive integer multiplicity $m_i(s)$ attached to every element $s \in S_i$. This way we can view the pair (S_i, m_i) as a multiset which contains the element $s \in S_i$ precisely $m_i(s)$ times. We shall consider the sum $d_i = d(S_i) := \sum_{s \in S_i} m_i(s)$ as the size of the multiset (S_i, m_i). As before, we put $\mathbf{S} = S_1 \times S_2 \times \cdots S_n$, for an element $\mathbf{s} = (s_1, \ldots, s_n) \in \mathbf{S}$ we set the multiplicity vector $m(\mathbf{s}) = (m_1(s_1), \ldots, m_n(s_n))$ and write $|m(\mathbf{s})| = m_1(s_1) + \cdots + m_n(s_n)$.

Now we are able to formulate a version of the Non-vanishing Theorem for multiple points over fields. From this one can obtain Alon's result by setting $m_i(s) = 1$ identically.

Theorem 6.4. *([32, Theorem 6]) Let $\mathbb{F}$ be a field, $f = f(\mathbf{x}) \in \mathbb{F}[\mathbf{x}]$ be a polynomial of degree $\sum_{i=1}^{n} t_i$, where each t_i is a nonnegative integer. Assume, that the coefficient in f of the monomial $x_1^{t_1} x_2^{t_2} \cdots x_n^{t_n}$ is nonzero. Suppose further that $(S_1, m_1), (S_2, m_2), \ldots, (S_n, m_n)$ are multisets of $\mathbb{F}$ such that for the size d_i of (S_i, m_i) we have $d_i > t_i$ $(i = 1, \ldots, n)$. Then there exists a point $\mathbf{s} = (s_1, \ldots, s_n) \in \mathbf{S} = S_1 \times \cdots \times S_n$ and an exponent vector $\mathbf{u} = (u_1, \ldots, u_n)$*

with $u_i < m_i(s_i)$ for each i, such that $f_{\mathbf{u}}(\mathbf{s}) \neq 0$.

As an application of Theorem 6.4 we present an extension of [3, Theorem 6.3]. The original result, which is the special case when every multiplicity is 1, was obtained by Alon and Füredi, see [4, Theorem 1], and later reproved by Alon using the original non-vanishing argument. By a hyperplane H in $\mathbb{F}^n$ we understand the set of zeros of a linear polynomial of the form

$$\ell(\mathbf{x}) = a_1x_1 + \cdots + a_nx_n - b = (\mathbf{a}, \mathbf{x}) - b,$$

where $a_1, \ldots, a_n, b \in \mathbb{F}$.

Theorem 6.5. *([32, Theorem 12]) Let $(S_1, m_1), \ldots, (S_n, m_n)$ be finite multisets from the field $\mathbb{F}$ and put $\mathbf{S} = S_1 \times \cdots \times S_n$. Suppose that $0 \in S_i$, with $m_i(0) = 1$ for every i, and $H_1, \ldots, H_k$ are hyperplanes in $\mathbb{F}^n$ such that every point $\mathbf{s} \in \mathbf{S} \setminus \{\mathbf{0}\}$ is covered by at least $|m(\mathbf{s})| - n + 1$ hyperplanes and the point $\mathbf{0}$ is not covered by any of the hyperplanes. Then $k \geq d(S_1) + d(S_2) + \cdots + d(S_n) - n$.*

Proof. For $j = 1, \ldots, k$ let $\ell_j(\mathbf{x})$ be the linear polynomial defining the hyperplane H_j, set $f(\mathbf{x}) = \prod_{j=1}^{k} \ell_j(\mathbf{x})$ and $t_i = d(S_i) - 1$. Let

$$P(\mathbf{x}) = \prod_{i=1}^{n} \prod_{s \in S_i \setminus \{0\}} (x_i - s)^{m_i(s)}$$

and

$$F(\mathbf{x}) = P(\mathbf{x}) - \frac{P(\mathbf{0})}{f(\mathbf{0})} f(\mathbf{x}).$$

Note that we have $f(\mathbf{0}) \neq 0$, because the hyperplanes do not cover $\mathbf{0}$. If the statement is false, then the degree of F is $t_1 + t_2 + \cdots + t_n$ and the coefficient of $x_1^{t_1} \cdots x_n^{t_n}$ is 1. Theorem 6.4 applies with $(S_1, m_1), \ldots, (S_n, m_n)$ and $t_1, \ldots, t_n$: there exists a vector $\mathbf{s} \in S$, and an exponent vector $\mathbf{u}$ with $u_i < m_i(s_i)$ for each i, such that $F_{\mathbf{u}}(\mathbf{s}) \neq 0$. We observe that $\mathbf{s}$ can not be $\mathbf{0}$, because $F(\mathbf{0}) = 0$. Thus $\mathbf{s}$ must have at least one nonzero coordinate, implying that $P_{\mathbf{u}}(\mathbf{s}) = 0$.

Moreover, as $\mathbf{s}$ is a nonzero vector, $f(\mathbf{x})$ must vanish at $\mathbf{s}$ at least $|m(\mathbf{s})| - n + 1$ times, implying that $f_{\mathbf{u}}(\mathbf{s}) = 0$ (expand the product at $\mathbf{s}$; for every term $(\mathbf{x} - \mathbf{s})^{\mathbf{v}}$ obtained there will be an index j such that $v_j \geq m_j(s_j)$). These facts imply that $F_{\mathbf{u}}(\mathbf{s}) = 0$, a contradiction. This finishes the proof. □

6.2 The Combinatorial Nullstellensatz and the Non-vanishing Theorem over commutative rings

Another possible direction to generalize Theorems 6.2 and 6.3 is to consider the problem over commutative rings instead of fields.

First we remark that if in Theorem 6.2 all the polynomials $f, g_1, \dots, g_n$ lie in $R[\mathbf{x}]$ for some subring R of $\mathbb{F}$, then the same can be required for the polynomials $h_1, \dots, h_n$, see [3]. Accordingly, both theorems remain true if we replace $\mathbb{F}$ by some of its subrings. However if R is an arbitrary commutative ring, then some additional assumption is needed.

Example 6.6. Let $R = \mathbb{Z}_6$, the ring of integers modulo 6 and consider polynomials in two variables. Further let $S_1 = S_2 = \{2, 4\}$, i.e. $g_1(x) = (x-2)(x-4)$, $g_2 = (y-2)(y-4)$ and consider the polynomial $f(x, y) = 3x^2 + 3xy + 3y^2$. It is easy to check that f vanishes everywhere on $\mathbf{S} = S_1 \times S_2 \subseteq \mathbb{Z}_6^2$. If Theorem 6.2 would hold, then according to the degree bounds h_1 and h_2 should be constant polynomials. However a linear combination of the form $c_1 g_1(x) + c_2 g_2(y)$ does not contain the monomial xy, and so cannot be equal to f. Also, if we put $t_1 = t_2 = 1$, then the coefficient of xy in f is nonzero, which shows that Theorem 6.3 does not hold either.

When we examine the proof of the Combinatorial Nullstellensatz from an algebraic point of view, then we can come up with a natural extra assumption on the S_i's.

Theorem 6.7. *([31, Theorem 3]) Let R be a commutative ring and $f(\mathbf{x})$ a polynomial from $R[\mathbf{x}]$. Further let $S_1, \ldots, S_n$ be nonempty finite subsets of R with the property that if $s \neq s^* \in S_i$, then $s - s^*$ is a unit in R, and for $1 \leq i \leq n$ define*

$$g_i(x_i) = \prod_{s \in S_i} (x_i - s).$$

Then for every polynomial $f(\mathbf{x}) \in R[\mathbf{x}]$ there are polynomials $h_1(\mathbf{x}), \ldots, h_n(\mathbf{x}), r(\mathbf{x}) \in R[\mathbf{x}]$ such that $\deg(h_i) \leq \deg(f) - \deg(g_i)$ for all i and the degree of r is less than $\deg(g_i) = |S_i|$ in every x_i, for which

$$f(\mathbf{x}) = r(\mathbf{x}) + \sum_{i=1}^{n} h_i(\mathbf{x}) g_i(x_i).$$

Moreover, if we put $\mathbf{S} = S_1 \times \cdots \times S_n$, then $f \in I(\mathbf{S})$ if and only if r is identically zero, hence

$$I(\mathbf{S}) = (g_1, \ldots, g_n).$$

Proof. Let us denote by V the R module of all functions from $\mathbf{S}$ to R. V is a free R module of rank

$$\text{rank}_R V = |\mathbf{S}| = \prod_{i=1}^{n} d_i,$$

where $d_i = |S_i| = \deg(g_i)$. In fact, for $\mathbf{s} \in \mathbf{S}$ we denote by $f_{(\mathbf{s})}$ the $\mathbf{S} \longrightarrow R$ function taking value 1 on $\mathbf{s}$ and 0 everywhere else in $\mathbf{S}$. Then the set $F = \{f_{(\mathbf{s})} | \mathbf{s} \in \mathbf{S}\}$ is a free generating set of V over R, and $|F| = |\mathbf{S}| = \prod_{i=1}^{n} d_i$. Next observe, that every $f_{(\mathbf{s})}$ can be written as a polynomial from $R[\mathbf{x}]$, using interpolation. For $\mathbf{s} = (s_1, \ldots, s_n) \in S$ we have

$$f_{(\mathbf{s})}(\mathbf{x}) = \prod_{i=1}^{n} (\prod_{\alpha \in S_i, \alpha \neq s_i} (x_i - \alpha)(s_i - \alpha)^{-1}).$$

Since $s_i \neq \alpha \in S_i$, the element $s_i - \alpha$ is a unit in R by assumption, hence the definition of $f_{(\mathbf{s})}(\mathbf{x})$ makes sense.

Consider the following set of monomials:

$$M = \{\mathbf{x}^{\mathbf{w}};\ w_i \leq d_i - 1,\ i = 1, \dots, n\}.$$

Take an arbitrary polynomial f from $R[\mathbf{x}]$, reduce it with $G = \{g_1, \dots, g_n\}$ and denote the remainder by r. Note that during this reduction process no divisions with ring elements have to be made, hence the process is well defined. The fact that r is reduced with respect to G just means that the degree of r is less than d_i in every x_i, meaning that r is an R-linear combination of monomials from M. During this reduction process we also obtain polynomials $h_1(\mathbf{x}), \dots, h_n(\mathbf{x}) \in R[\mathbf{x}]$ such that $\deg(h_i) \leq \deg(f) - d_i$ for all i and

$$f(\mathbf{x}) = r(\mathbf{x}) + \sum_{i=1}^{n} h_i(\mathbf{x}) g_i(x_i).$$

To finish the proof first note that f and r are equal as functions on $\mathbf{S}$. Accordingly, if we reduce elements of F with G, we obtain a collection of $|\mathbf{S}|$ polynomials which, as functions, are independent over R, and each of them is a linear combination of monomials from M. Using also that $|M| = |\mathbf{S}|$, we infer that M, as a set of functions from $\mathbf{S}$ to R, is also linearly independent over R, meaning that $f \in I(S)$ if and only if r is the all zero linear combination. □

We remark that the proof actually gives that the polynomials $g_1, \dots, g_n$ form a universal Gröbner basis of $I(S)$. When developing Gröbner theory over commutative rings instead of fields one has to be cautious, for details we refer the reader to [1, Chapter 4].

From Theorem 6.7, using the original argument of Alon, one can easily deduce a version of the Non-vanishing Theorem over commutative rings.

Theorem 6.8. *([31, Theorem 2]) Let R be a commutative ring, and let $f(\mathbf{x})$ be a polyno-*

mial in $R[\mathbf{x}]$. Suppose the degree $\deg(f)$ of f is $\sum_{i=1}^n t_i$, where t_i is a nonnegative integer, and suppose that the coefficient of $\prod_{i=1}^n x_i^{t_i}$ in f is nonzero. Suppose further that $S_1, \dots, S_n$ are subsets of R with $|S_i| > t_i$, and with the property that if $s \neq s^* \in S_i$, then $s - s^*$ is a unit in R. Then there exists a vector $\mathbf{s} \in \mathbf{S} = S_1 \times \cdots \times S_n$, such that $f(\mathbf{s}) \neq 0$.

Proof. The proof is essentially the same as the one in [3]. Clearly we may assume that $|S_i| = t_i + 1$ for all i. Suppose that the result is false, i.e. $f \in I(\mathbf{S})$, and for all i define $g_i(x_i) = \prod_{s \in S_i}(x_i - s)$. By Theorem 6.7 there are polynomials $h_1(\mathbf{x}), \dots, h_n(\mathbf{x}) \in R[\mathbf{x}]$ such that $\deg(h_j) \leq \deg(f) - \deg(g_j) = \sum_{i=1}^n t_i - (t_j + 1)$ for all j, for which

$$f(\mathbf{x}) = \sum_{i=1}^n h_i(\mathbf{x}) g_i(\mathbf{x}).$$

Here the degree of $h_i(\mathbf{x})g_i(\mathbf{x})$ is at most $\deg(f) = \sum_{i=1}^n t_i$, and if there are any monomials of degree $\deg(f)$ in it, then they are divisible by $x_i^{|S_i|} = x_i^{t_i+1}$ for a suitable i. It follows that the coefficient of $\prod_{i=1}^n x_i^{t_i}$ on the right hand side is zero. However by our assumption the coefficient of $\prod_{i=1}^n x_i^{t_i}$ on the left hand side is nonzero, and this contradiction completes the proof. □

In the case $R = \mathbb{F}$, when we work over a field, the preceding results specialize to Theorems 6.2 and 6.3. To see this, only note that if R is a field, then $s - s^*$ is always a unit, whenever s and s^* are different.

Next, as an application of Theorem 6.8, we present another generalization of [3, Theorem 6.3] to the Boolean cube over a commutative ring R.

Theorem 6.9. *([31, Theorem 10]) Let R be a commutative ring, and let $H_1, \dots, H_m$ be hyperplanes in R^n such that $H_1, \dots, H_m$ cover all the vertices of the unit cube $\{0,1\}^n \subseteq R^n$, with the exception of $\mathbf{0}$. For $i = 1, \dots, m$ let $(\mathbf{a}^i, \mathbf{x}) - b_i$ be the polynomial defining H_i. If $\prod_{i=1}^m b_i \neq 0$, then $m \geq n$.*

Proof. The proof is essentially the same as the one in [3]. Assume that the assertion is false, i.e $m < n$, and consider the polynomial

$$P(\mathbf{x}) = (-1)^{n+m+1} \prod_{j=1}^{m} b_j \prod_{i=1}^{n} (x_i - 1) + \prod_{i=1}^{m} [(\mathbf{a}^i, \mathbf{x}) - b_i].$$

The degree of this polynomial is clearly n, and the coefficient of $\prod_{i=1}^{n} x_i$ in P is

$$(-1)^{n+m+1} \prod_{j=1}^{m} b_j,$$

which is nonzero by assumption. By applying Theorem 6.8 with $S_i = \{0, 1\}$, $t_i = 1$, we obtain a point $\mathbf{s} \in \{0, 1\}^n$ for which $P(\mathbf{s}) \neq 0$. This point is not the all zero vector, as P vanishes on $\mathbf{0}$. But otherwise, if $\mathbf{s} \neq \mathbf{0}$ then $s_j \neq 0$ for some j and $(\mathbf{a}_i, \mathbf{s}) - b_i = 0$ for some i (as $\mathbf{s}$ is covered by some hyperplane H_i), implying that P does vanish on $\mathbf{s}$, and so resulting a contradiction. $\square$

If we put $R = \mathbb{Z}_n$ for some square-free integer $n \in \mathbb{N}$, then in this particular case an application of the original Non-vanishing Theorem over $\mathbb{F}_p$, where p is a prime factor of n for which $\prod_{i=1}^{m} b_i \neq 0$ in $\mathbb{F}_p$, proves the statement. However, if n has square factors, then Theorem 6.8 appears to give a new result.

To finish this section, we remark that [31] also contains a common generalization of Theorems 6.4 and 6.8. Similarly, as Theorem 6.4, this result is also due to Géza Kós and Lajos Rónyai, and is included only for the sake of completeness.

Theorem 6.10. *([31, Theorem 4]) Let R be a commutative ring, $f = f(\mathbf{x}) \in R[\mathbf{x}]$ be a polynomial of degree $\sum_{i=1}^{n} t_i$, where each t_i is a nonnegative integer. Assume, that the coefficient in f of the monomial $x_1^{t_1} x_2^{t_2} \cdots x_n^{t_n}$ is nonzero. Suppose further that (S_1, m_1), (S_2, m_2), … ,(S_n, m_n) are multisets of R such that for the size d_i of (S_i, m_i) we have $d_i > t_i$ ($i = 1, \ldots, n$), and for each i if $s \neq s^* \in S_i$, then $s - s^*$ is a unit in R. Then there exists* *a point* $\mathbf{s} = (s_1, \ldots, s_n) \in \mathbf{S} = S_1 \times \cdots \times S_n$ *and an exponent vector* $\mathbf{u} = (u_1, \ldots, u_n)$ *with* $u_i < m_i(s_i)$ *for each* i*, such that* $f_{\mathbf{u}}(\mathbf{s}) \neq 0$.

6.3 The Combinatorial Nullstellensatz and the Non-vanishing Theorem for balanced systems

In [37] we examined the possibility of extending the Combinatorial Nullstellensatz and the Non-vanishing Theorem for a wider class of point sets, not merely discrete boxes.

Let again $\mathbb{F}$ be an arbitrary field and $X \subseteq \mathbb{F}^n$ a finite point set. For $1 \leq k \leq n$ define the projection of X to the last $n-k+1$ coordinates as

$$X_k = \{(s_k, \ldots, s_n) \mid \exists\ s_1, \ldots, s_{k-1} \in \mathbb{F} \text{ such that } (s_1, \ldots, s_n) \in X)\} \subseteq \mathbb{F}^{n-k+1}.$$

Theorem 6.11. *([37, Theorem 3.1]) For a nonempty finite set* $X \subseteq \mathbb{F}^n$ *and for positive integers* $d_1, \ldots, d_n$ *the following are equivalent:*

(i) $Sm(I(X)) = \{\mathbf{x}^{\mathbf{u}} \mid u_i < d_i$ *for all* $1 \leq i \leq n\}$ *with respect to the standard lex order.*

(ii) The reduced Gröbner basis of $I(X)$ *with respect to the standard lex order is of the form* $\{F_1, \ldots, F_n\}$*, where for all* $1 \leq i \leq n$ *we have* $lm(F_i) = x_i^{d_i}$.

(iii) For all $k = 1, \ldots, n-1$ *the size of*

$$\{s \in \mathbb{F} \mid (s, s_{k+1}, \ldots, s_n) \in X_k\}$$

is d_k *for all* $(s_{k+1}, \ldots, s_n) \in X_{k+1}$*, and* $|X_n| = d_n$.

Proof. First we prove that (i) and (ii) in Theorem 6.11 are actually equivalent for every zero dimensional ideal and every term order.

Lemma 6.12. *Let $I \lhd \mathbb{F}[\mathbf{x}]$ be a zero dimensional ideal, $\prec$ an arbitrary term order and $d_1, \ldots, d_n$ positive integers. Then with respect to $\prec$*

$$Sm(I) = \{\mathbf{x}^{\mathbf{u}} \mid u_i < d_i \textit{ for all } 1 \leq i \leq n\}$$

if and only if the reduced Gröbner basis of I is of the form $\{F_1, \ldots, F_n\}$, where for all $1 \leq i \leq n$ we have $lm(F_i) = x_i^{d_i}$.

Proof. Fix an term order $\prec$ and suppose that for this $Sm(I) = \{\mathbf{x}^{\mathbf{u}} \mid u_i < d_i \text{ for all } i\}$. By assumption $x_i^{d_i}$ is a leading monomial for all i, hence by Corollary 2.5 it has a representation by standard monomials. Denote the corresponding polynomial by F_i. The leading monomial of F_i is clearly $x_i^{d_i}$, since any other monomial appearing in F_i is a standard one. Now if we take any non-standard, i.e. leading monomial $\mathbf{x}^{\mathbf{u}}$, then by the structure of $Sm(I)$ there will be an index j such that $u_j \geq d_j$. This implies that $lm(F_j)|\mathbf{x}^{\mathbf{u}}$, in particular $\{F_1, F_2, \ldots, F_n\}$ is a Gröbner basis of I. Moreover, $lm(F_i) = x_i^{d_i}$ cannot divide a standard monomial (again by the structure of $Sm(I)$) neither $lm(F_j)$ for $j \neq i$, meaning that the family of polynomials $\{F_1, F_2, \ldots, F_n\}$ is a reduced Gröbner basis.

For the other direction, suppose that the reduced Gröbner basis of $I \lhd \mathbb{F}[\mathbf{x}]$ is of the form $\{F_1, F_2, \ldots, F_n\}$, where for all $1 \leq i \leq n$ we have that $lm(F_i) = x_i^{d_i}$. By the properties of Gröbner bases, for any leading monomial $\mathbf{x}^{\mathbf{u}} \in Lm(I)$ there is an index i such that $lm(F_i) = x_i^{d_i}|\mathbf{x}^{\mathbf{u}}$. On the other hand, if for some monomial $\mathbf{x}^{\mathbf{u}}$ there is an index i such that $x_i^{d_i}|\mathbf{x}^{\mathbf{u}}$ (i.e. $d_i \leq u_i$), then $\mathbf{x}^{\mathbf{u}}$ is the leading monomial of the polynomial $\frac{\mathbf{x}^{\mathbf{u}}}{x_i^{d_i}} F_i \in I$. These facts together imply that $Sm(I) = \{\mathbf{x}^{\mathbf{u}} \mid u_i < d_i \text{ for all } 1 \leq i \leq n\}$. $\square$

For $(ii) \implies (iii)$ suppose that $X \subseteq \mathbb{F}^n$ is such that the reduced Gröbner basis of $I(X)$ with respect to the standard lex order is of the form $G = \{F_1, \ldots, F_n\}$, where for all $1 \leq i \leq n$ we have that $lm(F_i) = x_i^{d_i}$. The fact $lm(F_i) = x_i^{d_i}$ implies that x_j with $j < i$

does not occur in F_i, i.e.

$$F_i \in \mathbb{F}[x_i, \dots, x_n] \subseteq \mathbb{F}[x_1, \dots, x_n].$$

For $k = 1, 2, \dots, n$ put $G_k = \{F_k, F_{k+1}, \dots, F_n\}$ and $I_k = \langle G_k \rangle \lhd \mathbb{F}[x_k, \dots, x_n]$. As a special case we have that $G = G_1$ and $I(X) = I_1$.

Lemma 6.13. *If a polynomial $f \in \mathbb{F}[x_k, \dots, x_n]$ reduces to 0 using G inside $\mathbb{F}[x_1, \dots, x_n]$, then it reduces to 0 using G_k inside $\mathbb{F}[x_k, \dots, x_n]$.*

Proof. Take a reduction process for f inside $\mathbb{F}[x_1, \dots, x_n]$ that results 0. We claim that this reduction process takes place actually inside $\mathbb{F}[x_k, \dots, x_n]$. The first step in the reduction of f by G can only be by a polynomial $g \in G_k \subseteq G$, as only these have their leading term in $\mathbb{F}[x_k, \dots, x_n]$. For the polynomial $\widetilde{f}$, obtained after the first reduction step we again have $\widetilde{f} \in \mathbb{F}[x_k, \dots, x_n]$ as $G_k \subseteq \mathbb{F}[x_k, \dots, x_n]$. The claim now follows by induction on the length of the reduction process. □

Lemma 6.14. *G_k is the reduced Gröbner basis of I_k for $1 \leq k \leq n$.*

Proof. Recall that by Buchberger's theorem (see Proposition 2.10) G_k is a Gröbner basis of I_k if and only if the S-polynomial (see Definition 2.11) of any two polynomials in G_k can be reduced to 0 using G_k inside $\mathbb{F}[x_k, \dots, x_n]$. Now take F_i, F_j, $k \leq i < j \leq n$, and let $S(F_i, F_j)$ be their S-polynomial. Since G is a Gröbner basis of $I(X)$, $S(F_i, F_j) \in \mathbb{F}[x_i, \dots, x_n]$ can be reduced to 0 using G inside $\mathbb{F}[x_1, \dots, x_n]$, and so by Lemma 6.13 it can be reduced to 0 using $G_i \subseteq G_k$ inside $\mathbb{F}[x_i, \dots, x_n] \subseteq \mathbb{F}[x_k, \dots, x_n]$.

The fact that G_k is reduced easily follows as it is a subset of G which is a reduced basis. □

It is easily seen that I_k is a zero dimensional ideal, however a bit more is true.

Lemma 6.15. $I(X_k) = I_k$

Proof. $I_k \subseteq I(X_k)$ follows directly from the definitions. For the other direction let f be an arbitrary polynomial from $I(X_k) \lhd \mathbb{F}[x_k, \ldots, x_n]$. Since G_k is a Gröbner basis of I_k, to prove that $f \in I_k$ it suffices to show that it can be reduced to 0 using G_k. By the definition of X_k the fact that $f \in I(X_k) \lhd \mathbb{F}[x_k, \ldots, x_n]$ implies that $f \in I(X)$, and hence it can be reduced to 0 using the Gröbner basis G inside $\mathbb{F}[x_1, \ldots, x_n]$. Again by Lemma 6.13 this means that it can be reduced to 0 using G_k inside $\mathbb{F}[x_k, \ldots, x_n]$ as well. $\square$

Lemma 6.14 and 6.15 together imply that $G_k = \{F_k, \ldots, F_n\}$ is the reduced Gröbner basis of the vanishing ideal $I(X_k) \lhd \mathbb{F}[x_k, \ldots, x_n]$. Now Lemma 6.12 implies that

$$Sm(I(X_k)) = \{x_k^{u_k} \cdots x_n^{u_n} \mid u_i < d_i \text{ for all } k \leq i \leq n\},$$

and hence by the properties of standard monomials of vanishing ideals of finite point sets we get that $|X_k| = |Sm(I(X_k))| = \prod_{i=k}^{n} d_i$, in particular $|X_n| = d_n$.

Independently from the proof, we remark that from the general properties of elimination term orders (see [1, Theorem 2.3.4]) we know that G_k is a Gröbner basis (and hence an ideal basis) of the elimination ideal $I(X) \cap \mathbb{F}[x_k, \ldots, x_n]$ as well, and hence

$$I(X) \cap \mathbb{F}[x_k, \ldots, x_n] = I(X_k).$$

Now fix $1 \leq k \leq n-1$, let $(s_{k+1}, \ldots, s_n) \in X_{k+1}$ and put $h(x_k) = F_k(x_k, s_{k+1}, \ldots, s_n)$. h is a polynomial in $\mathbb{F}[x_k]$ of degree d_k. If $s \in \mathbb{F}$ is such that $(s, s_{k+1}, \ldots, s_n) \in X_k$, then $h(s) = F_k(s, s_{k+1}, \ldots, s_k) = 0$, i.e. s is a root of h. By the degree bound on h, the number of such elements s is at most d_k. However $|X_k| = d_k \cdot |X_{k+1}|$, what is possible only if for all fixed $(s_{k+1}, \ldots, s_n) \in X_{k+1}$ the number of suitable elements s is exactly d_k. This finishes the $(ii) \implies (iii)$ part of the proof.

To complete the proof of Theorem 6.11, suppose that the finite set $X \subseteq \mathbb{F}^n$ satisfies the given combinatorial condition, i.e. for all $k = 1, \ldots, n-1$ the size of

$$\{s \in \mathbb{F} \mid (s, s_{k+1}, \ldots, s_n) \in X_k\}$$

is d_k for all $(s_{k+1}, \ldots, s_n) \in X_{k+1}$, and $|X_n| = d_n$. Let A be the set of all field elements occurring as a coordinate in X and put $m = |A| - 1$. For $i = 1, \ldots, n$ fix some injective functions $\varphi_i : A \longrightarrow \{0, 1, \ldots, m\} \subseteq \mathbb{R}$, and using them, define $\widehat{X} \subseteq \{0, 1, \ldots, m\}^n \subseteq \mathbb{R}^n$ as in Proposition 2.17. By the injectivity of the φ_i's $\widehat{X}$ inherits from X its structural property, i.e. for all $1 \leq k \leq n-1$ the number of elements α for which $(\alpha, \alpha_{k+1}, \ldots, \alpha_n) \in \widehat{X}_k$ is d_k for all $(\alpha_{k+1}, \ldots, \alpha_n) \in \widehat{X}_{k+1}$, and $|\widehat{X}_n| = d_n$. However this strict structure immediately implies that after applying downshifts in a suitable order we get that

$$D_n(D_{n-1}(\ldots D_1(\widehat{X}) \ldots)) = \{\mathbf{u} \in \mathbb{N}^n \mid u_i < d_i \text{ for all } i\},$$

and hence using Proposition 4.5 and Proposition 2.17

$$\{\mathbf{x}^{\mathbf{u}} \mid u_i < d_i \text{ for all } 1 \leq i \leq n\} = Sm(I(\widehat{X})) = Sm(I(X)).$$

This finishes the proof of Theorem 6.11. $\square$

We remark that Theorem 6.11 remains true if we replace the standard lex order with another lex order based on some permutation $\sigma \in S_n$, only that projections have to be defined with respect to the appropriate ordering of the variables, i.e.

$$X_k^\sigma = \{(s_k, \ldots, s_n) \mid \exists\, s_1, \ldots, s_{k-1} \in \mathbb{F} \text{ such that } (s_{\sigma^{-1}(1)}, \ldots, s_{\sigma^{-1}(n)}) \in X)\} \subseteq \mathbb{F}^{n-k+1},$$

and accordingly part (iii) of the theorem also has to be modified:

$(iii)^{\sigma}$ For all $k = 1, \dots, n-1$ the size of

$$\{s \in \mathbb{F} \mid (s, s_{k+1}, \dots, s_n) \in X_k^{\sigma}\}$$

is $d_{\sigma(k)}$ for all $(s_{k+1}, \dots, s_n) \in X_{k+1}^{\sigma}$, and $|X_n^{\sigma}| = d_{\sigma(n)}$.

There are several examples of point sets that satisfy the structural property in part (iii) of Theorem 6.11.

Example 6.16. Let $\mathbf{S} = S_1 \times \cdots \times S_n$ be a discrete box as in Alon's original Nullstellensatz. Here we have $d_i = |S_i|$ and $F_i(x_i, \dots, x_n) = F_i(x_i) = \prod_{s \in S_i}(x_i - s)$, $i = 1, \dots, n$. This example shows that Theorem 6.11 is indeed in some sense a generalization of the Combinatorial Nullstellensatz.

Example 6.17. Let $a_1, \dots, a_n$ be different elements from $\mathbb{F}$, and consider all possible permutations of these elements as vectors in $\mathbb{F}^n$.

$$P_n(a_1, \dots, a_n) = \{(a_{\pi(1)}, a_{\pi(2)}, \dots, a_{\pi(n)}) \mid \pi \in S_n\},$$

where S_n is the symmetric group of degree n. The reduced Gröbner basis of the vanishing ideal $I(P_n(a_1, \dots, a_n))$ with respect to the lex order was determined in [26]. There we have $d_i = i$ for $1 \le i \le n$. For the precise polynomials and proofs see [26].

Example 6.18. Let A be an $n \times n$ matrix with entries $a_{i,j}$, $1 \le i, j \le n$ from the field $\mathbb{F}$, and suppose that each column contains n different elements, i.e. $a_{i_1 j} \ne a_{i_2 j}$ for all j and $i_1 \ne i_2$. Put

$$\mathcal{P}(A) = \{(a_{1\pi(1)}, a_{2\pi(2)}, \dots, a_{n\pi(n)}) \mid \pi \in S_n\},$$

where S_n is the symmetric group of degree n. Sets of the form $\mathcal{P}(A)$ are the generalizations of permutations, and clearly satisfy the combinatorial condition (iii) from Theorem 6.11.

In connection with norm graphs (see [5]), the polynomials

$$f_i(x_1,\dots,x_n)=\prod_{j=1}^{n}(x_j-a_{ij}),\ i=1,\dots,n$$

turn up, where the field elements a_{ij} satisfy the same condition as above. Their set of common zeros is exactly $\mathcal{P}(A)$.

For simplicity, above we considered only the standard lex order, however, because of their symmetry, all 3 examples behave similarly for other lex orders as well.

Example 6.19. For this example let $\mathbb{F}=\mathbb{C}$ and for $n>1$ different nonzero complex numbers $z_1,z_2,\dots,z_n$ put

$$f(x,y)=x^n-y,$$

$$g(y)=(y-z_1)(y-z_2)\cdots(y-z_n).$$

For $1\le i\le n$ let $w_i\in\mathbb{C}$ be one n^{th} root of z_i, and let $\varepsilon\in\mathbb{C}$ be a primitive n^{th} root of unity. The vanishing set of $I=\langle f,g\rangle\lhd\mathbb{C}[x,y]$ is

$$X=\{(\varepsilon^k w_i,z_i)\mid 1\le i,k\le n\}\subseteq\mathbb{C}^2.$$

X clearly possesses the desired combinatorial property with $d_1=d_2=n$, and hence by Theorem 6.11 for the lex order we have $Sm(I(X))=\{x^\alpha y^\beta\mid\alpha,\beta<n\}$. $f,g\in I(X)$ by definition, moreover, using f and g any polynomial $h\in\mathbb{C}[x,y]$ can be reduced to some polynomial $\widetilde{h}$ whose degree is smaller than n both in x and in y, and so $\widetilde{h}$ is a linear combination of standard monomials. This implies that f and g form a reduced Gröbner basis of $I(X)$ with respect to the lex order, in particular $I(X)=\langle f,g\rangle$.

Similar examples can be given in higher dimensions as well.

A Gröbner basis G is called *degree reducing*, if for every element $g\in G$ the leading monomial $lm(g)$ is the unique monomial of maximal degree, i.e. $deg(lm(g)) = deg(g)$ and for any other monomial $\mathbf{x^u}$ occurring in g with a nonzero coefficient we have that $deg(\mathbf{x^u}) < deg(lm(g))$.

If $X \subseteq \mathbb{F}^n$ is such that $I(X)$ has a degree reducing Gröbner basis, then the original proof of the Non-vanishing Theorem from [3] applies to obtain:

Proposition 6.20. *([37, Proposition 3.7]) Let $X \subseteq \mathbb{F}^n$ be a nonempty set such that $I(X)$ has a degree reducing Gröbner basis G for some term order. If a polynomial $f \in \mathbb{F}[\mathbf{x}]$ of degree d contains a standard monomial for $I(X)$ of degree d with nonzero coefficient, then there is some point $\mathbf{s} \in X$ where f does not vanish, i.e. $f(\mathbf{s}) \neq 0$.*

Proof. Denote the elements of G by $g_1, \dots, g_t$ and let $\mathbf{w}$ be the exponent vector of the standard monomial of degree d appearing in f with a nonzero coefficient. Suppose by contradiction that the statement is false, i.e. $f \in I(X)$. As G is a degree reducing Gröbner basis of $I(X)$, the polynomial f can be reduced to 0 using G, and during this reduction process no terms of degree more the d can appear, meaning that at the end we obtain a representation

$$f(\mathbf{x}) = \sum_{i=1}^{t} h_i(\mathbf{x}) g_i(\mathbf{x}),$$

where $h_1, \dots, h_t \in \mathbb{F}[\mathbf{x}]$ and $deg(h_i g_i) \leq deg(f)$ for every i. The coefficient of $\mathbf{x^w}$ on the left side is nonzero by assumption, hence so it must be on the right side. Because of the degree bounds on the summands this is possible only if $\mathbf{x^w} = lm(h_i) \cdot lm(g_i)$ for some i. However in this case $lm(g_i) | \mathbf{x^w}$ and so the polynomial $\frac{\mathbf{x^w}}{lm(g_i)} \cdot g_i(\mathbf{x})$ shows that $\mathbf{x^w}$ cannot be a standard monomial, which is a contradiction. $\square$

Note that in the original case of the Non-vanishing Theorem in [3] the polynomials $g_1, \dots, g_n$ formed a universal degree reducing Gröbner basis. An interesting feature of Example 6.19 is that it provides an example of a finite point set that is not a discrete box, but we still have a degree reducing Gröbner basis and hence a Non-vanishing Theorem.

Moreover, in this case, by Theorem 6.11, the condition from Proposition 6.20, that there is a standard monomial of maximal degree, reduces to a simple degree bound as in the original Non-vanishing Theorem.

We also remark, as pointed out by Gábor Hegedűs, that, according to Theorem 3.6 and Theorem 4.6, shattering-extremal families and extremal vector systems from Part I provide us other examples of finite point sets with degree reducing Gröbner bases. For this only note that $x_i^2 - x_i$ for $i \in [n]$, polynomials of the form $f_{S,H}$ for $H \subseteq S \subseteq [n]$ (see Definition 3.8) and degree dominated polynomials (see Definition 4.2) are all special types of polynomials with a leading monomial being the unique monomial of maximal degree. In this case however the condition from Proposition 6.20 may be harder to check. It would be interesting to find combinatorial problems where the Non-vanishing Theorem, applied to extremal systems, would work.

Example 6.21. For our last example suppose that for $1 \leq i \leq N$ we are given positive integers $d_{i1}, \ldots, d_{in_i}$ and a point set $X^{(i)} \subseteq \mathbb{F}^{n_i}$ satisfying property (iii) from Theorem 6.11. By Theorem 6.11 the vanishing ideal $I(X^{(i)}) \lhd \mathbb{F}[x_{i1}, \ldots, x_{in_i}]$ has a reduced Gröbner basis $G_i = \{F_{i1}, \ldots, F_{in_i}\}$ such that $lm(F_{ij}) = x_{ij}^{d_{ij}}$, $1 \leq j \leq n_i$ with respect to the "standard" lex order $\prec_i$ (for which $x_{in_i} \prec_i x_{i(n_i-1)} \prec_i \cdots \prec_i x_{i1}$). Now let

$$X = X^{(1)} \times X^{(2)} \times \cdots \times X^{(N)} \subseteq \mathbb{F}^{\sum_{i=1}^{N} n_i},$$

$$G = \bigcup_{i=1}^{N} G_i$$

and $\prec$ be the "standard" lex order for which $x_{Nn_N} \prec \cdots \prec x_{N1} \prec x_{(N-1)n_{N-1}} \prec \cdots \prec x_{11}$. From the construction it follows that X satisfies the given combinatorial property as well, so by Theorem 6.11 for the lex order $\prec$ we have

$$Sm(X) = \{\mathbf{x}^{\mathbf{u}} \mid u_{ij} < d_{ij} \text{ for all } 1 \leq i \leq N \text{ and } 1 \leq j \leq n_i\}.$$

On the other hand, using G, any polynomial f in the variables x_{ij}, $1 \leq i \leq N$, $1 \leq j \leq n_i$ can be reduced to a form $\widetilde{f}$ where the degree in each variable x_{ij} is less than d_{ij}, and so $\widetilde{f}$ is the linear combination of standard monomials. These imply that G is a reduced Gröbner basis of $I(X)$.

This direct product construction allows us to combine the earlier examples and to obtain more complicated ones.

In Example 6.19 we introduced a wider class of point sets, not merely discrete boxes, where the Non-vanishig Theorem holds in its full generality. The conditions of Theorem 6.11 are in general not sufficient for the Non-vanishing Theorem to hold. For example in the case of permutations (Example 6.17), the polynomial

$$f(x_1, \ldots, x_n) = \sum_{i=1}^{n} x_i - \sum_{i=1}^{n} a_i$$

has standard monomials in its maximal degree part ($x_2, x_3, \ldots, x_n$ are all standard monomials), but it vanishes on the whole set of permutations, it is actually a member of the reduced Gröbner basis. It would be interesting to develop an understanding of the finite sets $X \subseteq \mathbb{F}^n$ for which a version of the Non-vanishing Theorem holds.

Bibliography

[1] W. W. Adams and P. Loustaunau. *An introduction to Gröbner bases*, volume 3 of *Graduate Studies in Mathematics*. American Mathematical Society, 1996.

[2] R. Aharoni and R. Holzman. personal communication, cited in [24].

[3] N. Alon. Combinatorial Nullstellensatz. *Combinatorics, Probability and Computing*, 8:7–29, 1999.

[4] N. Alon and Z. Füredi. Covering the cube by affine hyperplanes. *European Journal of Combinatorics*, 14:79–83, 1993.

[5] N. Alon, L. Rónyai, and T. Szabó. Norm-graphs: variations and applications. *Journal of Combinatorial Theory, Series B*, 76:280–290, 1999.

[6] R.P. Anstee. Properties of $(0-1)$-matrices with no triangles. *Journal of Combinatorial Theory, Series A*, 29:186–198, 1980.

[7] R.P. Anstee, L. Rónyai, and A. Sali. Shattering news. *Graphs and Combinatorics*, 18:59–73, 2002.

[8] R.P. Anstee and A. Sali. Sperner families of bounded VC-dimension. *Discrete Mathematics*, 175:12–21, 1997.

[9] L. Babai and P. Frankl. *Linear algebra methods in combinatorics with applications to geometry and computer science*. Preliminary Version 2 (September 1992).

[10] A. Blumer, A. Ehrenfeucht, D. Haussler, and M.K. Warmuth. Learnability and the Vapnik-Chervonenkis dimension. *Journal of the Association for Computing Machinery*, 36(4):929–965, 1989.

[11] B. Bollobás, I. Leader, and A.J. Radcliffe. Reverse Kleitman inequalities. *Proceedings of the London Mathematical Society*, s3-58:153–168, 1989.

[12] B. Bollobás and A.J. Radcliffe. Defect Sauer results. *Journal of Combinatorial Theory, Series A*, 72:189–208, 1995.

[13] B. Buchberger and F. Winkler, editors. *Gröbner bases and applications*, volume 251 of *London Mathematical Society Lecture Note Series*. Cambridge University Press, 1998.

[14] M. Cámara, A. Lladó, and J. Moragas. On a conjecture of Graham and Häggkvist with the polynomial method. *European Journal of Combinatorics*, 30:1585–1592, 2009.

[15] P. J. Cameron. Counting two-graphs related to trees. *The Electronic Journal of Combinatorics*, 2:R4, 1995.

[16] D. Eisenbud. *Commutative algebra with a view toward algebraic geometry*, volume 150 of *Graduate Texts in Mathematics*. Springer-Verlag, 1995.

[17] J.-C. Faugère. A new efficient algorithm for computing Gröbner bases without reduction to zero (F5). In *Proceedings of the 2002 international symposium on Symbolic and algebraic computation*, pages 75–83. ACM New York, 2002.

[18] B. Felszeghy. On the solvability of some special equations over finite fields. *Publicationes Mathematicae Debrecen*, 68(1-2):15–23, 2006.

[19] B. Felszeghy, B. Ráth, and L. Rónyai. The lex game and some applications. *Journal of Symbolic Computation*, 41:663–681, 2006.

[20] S. Floyd and M.K. Warmuth. Sample compression, learnability, and the Vapnik-Chervonenkis dimension. *Machine Learning*, 21:269–304, 1995.

[21] P. Frankl. *Extremal set systems*, volume 2 of *Handbook of combinatorics*, chapter 24. Elsevier Science B.V/The MIT Press, 1995.

[22] K. Friedl and L. Rónyai. Order shattering and Wilson's theorem. *Discrete Mathematics*, 270:127–136, 2003.

[23] Z. Füredi and F. Quinn. Traces of finite sets. *Ars Combinatoria*, 18:195–200, 1983.

[24] G. Greco. Embeddings and the trace of finite sets. *Information Processing Letters*, 67:199–203, 1998.

[25] B. Green and T. Tao. The distribution of polynomials over finite fields, with applications to the Gowers norms. *Contributions to Discrete Mathematics*, 4(2):1–36, 2009.

[26] G. Hegedűs, A. Nagy, and L. Rónyai. Gröbner bases for permutations and oriented trees. *Annales Universitatis Scientiarum Budapestinensis de Rolando Eötvös Nominatae Sectio Computatorica*, 23:137–148, 2004.

[27] S. Jukna. *Extremal combinatorics with applications in computer science*. Texts in Theoretical Computer Science, An EATCS Series. Springer, 2nd edition, 2011.

[28] Gy. Károlyi. Restricted set addition: the exceptional case of the Erdős-Heilbronn conjecture. *Journal of Combinatorial Theory, Series A*, 116:741–746, 2009.

[29] Gy. Károlyi and Z. L. Nagy. A simple proof of the Zeilberger-Bressoud q-Dyson theorem. *Proceedings of the American Mathematical Society*, 142:3007–3011, 2014.

[30] Gy. Károlyi, Z. L. Nagy, F. V. Petrov, and V. Volkov. A new approach to constant term identities and Selberg-type integrals. accepted to Advances in Mathematics, arXiv:1312.6369.

[31] G. Kós, T. Mészáros, and L. Rónyai. Some extensions of Alon's Nullstellensatz. *Publicationes Mathematicae Debrecen*, 79(3-4):507–519, 2011.

[32] G. Kós and L. Rónyai. Alon's Nullstellensatz for multisets. *Combinatorica*, 32(5):589–605, 2012.

[33] L. Kozma and S. Moran. Shattering, graph orientations and connectivity. *The Electronic Journal of Combintaorics*, 20(3):P44, 2013.

[34] Z. Li, S. Zhang, and T. Dong. Finite sets of affine points with unique associated monomial order quotient bases. *Journal of Algebra and its Applications*, 11(2), Article ID: 1250025, 2012.

[35] T. Mészáros. S-extremal set systems and Gröbner bases. Master's thesis, Budapest University of Technology and Economics, 2010. `http://www.math.bme.hu/~slovi/thesiswork.pdf`.

[36] T. Mészáros and L. Rónyai. Shattering-extremal set systems of small VC-dimension. *ISRN Combinatorics*, Volume 2013, Article ID: 126214, 2013.

[37] T. Mészáros and L. Rónyai. A note on Alon's Combinatorial Nullstellensatz. *Annales Universitatis Scientiarum Budapestinensis de Rolando Eötvös Nominatae Sectio Computatorica*, 42:249–260, 2014.

[38] T. Mészáros and L. Rónyai. Shattering-extremal set systems of VC dimension at most 2. *The Electronic Journal of Combintorics*, 21(4):P4.30, 2014.

[39] B.K. Natarajan. *Machine learning: a theoretical approach*. Morgan Kaufmann, 1991.

[40] A. Pajor. *Sous-spaces LN/L des espaces de Banach*. Travaux en cours. Hermann, Paris, 1985.

[41] H. Pan and Z-W. Sun. A new extension of the Erdős-Heilbronn conjecture. *Journal of Combinatorial Theory, Series A*, 116:1374–1381, 2009.

[42] L. Rónyai and T. Mészáros. Some combinatorial applications of Gröbner bases. In F. Winkler, editor, *Algebraic Informatics, Proc. CAI 2011*, volume 6742 of *Lecture Notes in Computer Science*, pages 65–83. Springer, 2011.

[43] N. Sauer. On the density of families of sets. *Journal of Combinatorial Theory, Series A*, 13:145–147, 1972.

[44] S. Shelah. A combinatorial problem: stability and order for models and theories in infinitary languages. *Pacific Journal of Mathematics*, 41:247–261, 1972.

[45] A. Shinohara. Complexity of computing Vapnik-Chervonekis dimension and some generalized dimensions. *Theoretical Computer Science*, 137:129–144, 1995.

[46] Z-W. Sun. On value sets of polynomials over a field. *Finite Fields and Their Applications*, 14:470–481, 2008.

[47] M. Talagrand. Vapnik-Chervonenkis type conditions and uniform Donsker classes of functions. *The Annals of Probability*, 31(3):1565–1582, 2003.

[48] T. Tao. Algebraic combinatorial geometry: the polynomial method in arithmetic combinatorics, incidence combinatorics and number theory. *EMS Surveys in Mathematical Sciences*, 1:1–46, 2014.