A Survey of Applications of Algebraic Geometry

Juncheng Wan, 120033910148, jorgenwan@gamil.com

June 9th, 2022

1 Introduction

As a student of the Department of Computer Science and Engineering, I am interested in the various applications of algebraic geometry. During my survey, I find that there are preliminary attempts to apply theorems in algebraic geometry into application fields, such as combinatorics [1], graph theory [2], and even machine learning [9, 3]. In this survey, I mainly focus on the applications of combinatorics and graph theory. Most of the contents are from works [1, 2]. My work is more like a research porters, as some theorems and proofs are paraphrased from others' work.

Combinatorics and graph theory deal with arbitrary finite collections of objects such as points, lines, or graphs. Recently, people try to use tools of algebraic geometry to tackle many problems in these directions that have been solved by algebraic means, which is called the polynomial method. The work I appreciate the most is Alon's Combinatorial Nullstellensatz. From my understanding, these methods have the following three steps:

1. Treat the problem about some points in a vector space;
2. Find a polynomial of lowest possible degree (or other measure of complexity) that vanishes on the points;
3. Use tools from algebraic geometry to understand the structure of this zero set and solve the problem.

I also find that some inequalities in algebraic geometry are useful for probability estimation. For example, the Lang-Weil Bound [6] can be used to improve the probabilistic proof in extremal graph theory problem. I will also give a brief introduction to it.

2 Alon's Combinatorial Nullstellensatz and Applications

Alon's work on his combinatorial Nullstellensatz [1] can be seen as the root of the polynomial method.

First, I will recall Hilbert's Nullstellensatz learned in the class.

Theorem 1 (Hilbert's Nullstellensatz)**.** *Given an arbitrary set of n-variate polynomials g_i over an algebraically closed field F, if some other n-variate polynomial f vanishes over the common zeros of the g_i 's,*

then f raised to some power is contained in the ideal generated by the g_i 's. In other words, for such an $f, \exists k \in \mathbb{N}$ such that

$$f^k = \sum_{i=1}^{m} h_i g_i$$

where m is the number of g_i.

The proof is given by the teacher in the class. Thus, here I skip the proof.

To be different, Alon gives combinatorial nullstellensatsz with two restrictions for better applications:

1. There are as many g_i 's as there are dimensions;
2. Make a specific choice of g_i 's that allows concluding that they form a basis for all polynomials vanishing on their common zeros

The second property leads in particular to a corollary about the existence of non vanishing elements in sets that are much larger than the degree of a polynomial.

2.1 Alon's Combinatorial Nullstellensatz

In the original paper, there are two forms of Alon's Combinatorial Nullstellensatz. I list both of them below.

Theorem 2 (Alon's Nullstellensatz 1st form)**.** *Let F be an arbitrary field, $f \in F[x_1, \ldots, x_n]$. Let $S_1, \ldots, S_n$ be nonempty subsets of F and define $g_i(x_i) = \prod_{s \in S_i}(x_i - s)$. If f vanishes over all common zeros of $g_1, \ldots, g_n$, then there are polynomials $h_1, \ldots, h_n \in F[x_1, \ldots, x_n]$ satisfying $\deg(h_i) \leq \deg(f) - \deg(g_i)$ such that*

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

Moreover, if $f, g_1, \ldots, g_n$ lie in $R[x_1, \ldots, x_n]$ for some subring R of F then there are polynomials $h_i \in R[x_1, \ldots, x_n]$

In a sense, it states that the g_i 's form a basis for any polynomial vanishing on the entirety of $S_1 \times \ldots \times S_n$. The second form is as follows.

Theorem 3 (Alon's Nullstellensatz 2nd form)**.** *Let F be an arbitrary field, and let $f \in F[x_1, \ldots, x_n]$. Suppose the degree $\deg(f)$ of f is $\sum_{i=1}^{n} t_i$, where each t_i is a nonnegative integer, and suppose the coefficient of the term $\prod_{i=1}^{n} x_i^{t_i}$ in f is nonzero. Then if $S_1, \ldots, S_n$ are subsets of F with $|S_i| > t_i$, there is a point $(s_1, \ldots, s_n) \in S_1 \times \ldots \times S_n$ so that*

$$f(s_1, \ldots, s_n) \neq 0$$

This theorem can be used to find some polynomial that admits a non root in some product of subsets of $\mathbb{F}$ if and only if some property that we want holds. Then by the conditions of the problem, one shows that the desired coefficient is nonzero and that the polynomial has small degree, allowing the application of the Nullstellensatz.

During my survey, I find that the second formulation is more commonly used in applications than the first one.

Before proving the two forms of Alon's Combinatorial Nullstellensatz, I first list the lemma below and prove it.

Lemma 1. *Let $f = f(x_1, \ldots, x_n)$ be a polynomial in n variables over some field $\mathbb{F}$. Suppose that the degree of f in the ith variable is at most t_i, let $S_i \subset \mathbb{F}$ be a set of at least $t_i + 1$ distict elements of $\mathbb{F}$. If $f(x_1, \ldots, x_n) = 0$ for all n-tuples $(x_1, \ldots, x_n) \in S_1 \times \ldots \times S_n$, then f is the zero polynomial.*

Proof. By induction on n. $n = 1$. This is the well-known statement that a polynomial of degree at most d cannot have more than d roots. $n > 1$. Rewrite f as a polynomial in x_n :

$$f(x_1, \ldots, x_{n-1}, x_n) = \sum_{i=1}^{t_n} f_i(x_1, \ldots, x_{n-1}) x_n^i$$

Fixing some tuple $(s_1, \ldots, s_{n-1}) \in S_1 \times \ldots \times S_{n-1}$, we see that as a single variable polynomial (in the nth coordinate), by hypothesis, f is identically zero on S_n. Thus $\forall 1 \leq i \leq n-1, f_i(s_1, \ldots, s_{n-1}) = 0$. Hence each f_i is a polynomial in $n-1$ variables such that its degree in the j th variable is at most t_j that vanishes on all points $(s_1, \ldots, s_{n-1}) \in S_1 \times \ldots \times S_{n-1}$, with $|S_j| \geq t_j + 1$

Then, induction hypothesis can be used to get that f_i are zero polynomials, and conclude that f must also be the zero polynomial. □

It is hard to prove the second theorem directly from this lemma. The reason is that the condition that the degree of f in the i th variable is at most t_i is not necessarily satisfied. Indeed, we could have a term that has degree in x_i which is $t_i + 1$ while the total degree of f is still exactly $\sum_{i=1}^{n} t_i$.

Proof of Alon's Nullstellensatz 1st form. Define $t_i = |S_i| - 1$ and $g_i(x_i) = \prod_{s \in S_i} (x_i - s)$. Isolating the highest degree term of g_i we write

$$g_i(x_i) = x_i^{t_i+1} - \sum_{j=0}^{t_i} g_{ij} x_i^j$$

When $x_i \in S_i$ we have $g_i(x_i) = 0$ and thus the equality $x_i^{t_i+1} = \sum_{j=0}^{t_i} g_{ij} x_i^j$. Using this equality, we can modify f by repeatedly replacing every instance of x_i^k where $k > t_i$ by $x_i^{k-(t_i+1)} \sum_{j=0}^{t_i} g_{ij} x_i^j$.

Let the newly obtained polynomial be $\widetilde{f}$, we can check that we can rearrange $f - \widetilde{f}$ so that it is of the form $\sum_{i=1}^{n} h_i g_i$, where each polynomial h_i has degree at most $\deg(f) - \deg(g_i)$. Looking at a single substitution, we see that

$$\begin{aligned} f - \widetilde{f} &= h_i(x) x_i^{t_i+1} - h_i(x) \sum_{j=0}^{t_i} g_{ij} x_i^j \\ &= h_i g_i \end{aligned}$$

One can then generalize this after a notation heavy calculation. Note further that $\widetilde{f}$ now has degree at most t_i in the i th variable.

Moreover, as we replaced terms of f by term that evaluate to the same values on the cross product of the S_i 's, we have equality between f and $\widetilde{f}$ for all $x \in S_1 \times \ldots \times S_n$. Thus $\widetilde{f}$ is zero on $S_1 \times \ldots \times S_n$. Applying the lemma, we conclude that f_i is the zero polynomial. Thus

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

□

I think the proof is rather similar to the proof of Hilbert's Nullstellensatz in the class. The key point is to use the equality $x_i^{t_i+1} = \sum_{j=0}^{t_i} g_{ij} x_i^j$ to reduce the degree of monomials iteratively.

The second form of Alon's Nullstellensatz can be derived from the first form.

Proof of Alon's Nullstellensatz 2nd form. Without loss of generality, let $|S_i| = t_i + 1$. If we assume that f is zero on the whole $S_1 \times \ldots \times S_n$. Then we may write it as $\sum_{i=1}^n h_i g_i$ with g_i defined as previously. But then we clearly have a contradiction as if we look at the coefficient of the term $\prod_{j=1}^n x^{t_j}$, it must come from at least one of the summands $g_i h_i$. This means that h_i must have degree at least $\sum_{j \neq i} t_j$. But then the term $g_i h_i$ has degree $1 + \sum_{j=1}^n t_j$, which is larger than the degree of f. This is a ontradiction and we complete the proof. □

2.2 Applications in Algebra Problems

Theorem 2.1 are essentially useful for tackling algebra problems. Here, I list a classical problem about finite field which are solved by Theorem 2.1.

Theorem 4 (Chevalley-Warning)**.** *Let p be a prime, let $P_1, \ldots, P_m \in \mathbb{F}_p[x_1, \ldots, x_n]$. If $n > \sum_{i=1}^m \deg(P_i)$ and the polynomials P_i have a common zero $(c_1, \ldots, c_n)$, then they have another common zero.*

Proof. Suppose by contradiction that the point $(c_1, \ldots, c_n)$ is the only common zero. First, we note that for a fixed point $(s_1, \ldots, s_n) \in \mathbb{F}_p^n$, the following holds:

$$\prod_{i=1}^m \left(1 - P_i(s_1, \ldots, s_n)^{p-1}\right) = \begin{cases} 1, & \text{if } \forall i, P_i(s_1, \ldots, s_n) = 0 \\ 0, & \text{if } \exists \text{ i s.t. } P_i(s_1, \ldots, s_n) \neq 0 \end{cases}$$

In other words, this product is an indicator of the common zeros of the P_i 's. We can further note that its degree is at most $(p-1)\sum_{i=1}^m \deg(P_i)$. Next, we can define an indicator of the point $(c_1, \ldots, c_n)$:

$$\prod_{j=1}^n \prod_{c \in \mathbb{Z}_p, c \neq c_j} (s_j - c) = \begin{cases} 1/\delta, & \text{if } (s_1, \ldots, s_n) = (c_1, \ldots, c_n) \\ 0, & \text{otherwise} \end{cases}$$

Where δ is some non zero constant.

Now if we define the following polynomial:

$$f(x_1, \ldots, x_n) = \prod_{i=1}^m \left(1 - P_i(x_1, \ldots, x_n)^{p-1}\right) - \delta \prod_{j=1}^n \prod_{c \in \mathbb{Z}_p, c \neq c_j} (x_j - c)$$

We have by the above observations that f vanishes both at $(c_1, \ldots, c_n)$ and at every other point of $\mathbb{F}_p^n$ (since by assumption $(c_1, \ldots, c_n)$ is the only common zero). We can also see the coefficient of the term $\prod_{i=1}^n x_i^{p-1}$ comes only from the second sum since by an above remark, the degree of the first term is $(p-1)\sum_{i=1}^m \deg(P_i) < n(p-1)$. This means that the coefficient must be δ, thus nonzero.

Applying the second Nullstellensatsz to the set $\mathbb{F}_p \times \ldots \times \mathbb{F}_p$, we conclude that f must have a non root in $\mathbb{F}_p^n$, which is a contradiction. □

There exists a stronger version of this theorem, which states that the number of common zeros needs in fact to be a multiple of the characteristic of the field. However, it seems hard to derive this result from the Nullstellensatsz.

One is tempted approach this stronger version using a modification of the above polynomial:

$$f(x_1, \ldots, x_n) = \prod_{i=1}^{m} \left(1 - P_i(x_1, \ldots, x_n)^{p-1}\right) - \sum_{i=1}^{k} \delta_k \prod_{j=1}^{n} \prod_{c \in \mathbb{Z}_p, c \neq c_{ij}} (x_j - c)$$

where k is the number of common roots and $c_i = (c_{i1}, \ldots, c_{in})$ the common roots. But this approach fails as $\sum_{i=1}^{k} \delta_k$ might be zero.

2.3 Applications in Combinatorics Problems

The second applications is in additive combinatorics problems.

Theorem 5 (Cauchy-Davenport)**.** *Given A, B non-empty subsets of $\mathbb{F}_p$, for some prime p, the following holds:*

$$|A + B| \geq \min\{p, |A| + |B| - 1\}$$

This is a tight bound. If A and B are two singletons respectively, $|A + B| = 1$, we have the bound.

Proof. When $|A| + |B| > p$ then taking any $g \in \mathbb{F}_p$, the sets A and $g - B$ must intersect, thus we can write $g = a + b$ for some $a \in A, b \in B$.

Otherwise, assume by contradiction that $|A + B| < |A| + |B| - 1$. Take a subset C of $\mathbb{F}_p$ such that $A + B \subset C$ and $|C| = |A| + |B| - 2$. Then if we define

$$f(x, y) = \prod_{c \in C} (x + y - c)$$

we have that $f(a, b) = 0$ for $a \in A, b \in B$, since $A + B \subset C$. Furthermore, the coefficient of the monomial $x^{|A|-1}y^{|B|-1}$ is $\begin{pmatrix} |C| \\ |A| - 1 \end{pmatrix}$. Since $|C| = |A| + |B| - 2$ and $|A| + |B| \leq p$, this coefficient is nonzero in $\mathbb{F}_p$. Finally, this monomial is a maximum degree term in $\mathbb{F}$. We can thus apply the Nullstellensatz to $A \times B$ and derive a contradiction. □

From the paper, I also notice that the statement is not true for general finite fields. If the characteristic of the field can divide the coefficient of the $x^{|A|-1}y^{|B|-1}$, the above proof can not go through.

2.4 Applications in Graph Problems

I think it is very interesting to apply the theorem into graph problems. Berge and Sauer conject that any simple 4-regular graph contains a 3-regular subgraph. This is proved by Táskinov [7] in 1982.

Theorem 6. *In any 4-regular simple graph, there exists a 3-regular subgraph.*

This proposition is almost a special case (when $p = 3$) of the following theorem. But it does not quite work with a perfectly 4-regular graph. That being said, if we allow ourselves to slightly raise the average degree (eg. by adding an extra edge), we get that the newly obtained graph has a 3-regular subgraph.

Theorem 7. *For any prime p, any loopless graph $G = (V, E)$ with average degree bigger than $2p - 2$ and maximum degree at most $2p - 1$ contains a p-regular subgraph.*

Proof. For each edge $e \in E$ we define a variable x_e. The idea is to use these variables as selectors for the edges: $x_e = 1$ means e is in the subgraph, otherwise it is not. Let $a_{v,e} = 1$ if v is incident to $e, 0$ otherwise.

Since the maximum degree is $2p - 1$, for a fixed $v \in V, \sum_{e \in E} a_{v,e} x_e = 0 (\mathrm{mod} p)$ is equivalent to saying that either exactly p of the x_e are 1 or they are all 0 (i.e. p or no edges have been chosen, respectively). Now define the following polynomial:

$$f = \prod_{v \in V} \left(1 - \left(\sum_{e \in E} a_{v,e} x_e \right)^{p-1} \right) - \prod_{e \in E} (1 - x_e)$$

By the above observation, the first product is nonzero if and only if each individual vertex has either degree 0 or p. But the second product is 1 if and only if all x_e are zero (i.e. no edges have been picked), and otherwise zero. Thus f is zero for any assignment of x_e unless at least one x_e is nonzero and all vertices have degree p or zero, i.e. the subgraph induced by the selected edges is p-regular.

The degree of the first product is $(p - 1)|V|$, and by our initial hypothesis on the average degree we $2|E| > |V|(2p - 2)$. Hence the coefficient of $\prod_{e \in E} x_e$ comes only from the second product. We can see that it is nonzero.

Applying the Nullstellensatsz to the set $\{0, 1\}^{|E|}$, we obtain that f has a non root in this set, i.e. G has a p-regular subgraph. □

The above proof takes each edge as the variable and let the selected edges to be the subgraph. In fact, we can also take vertices as variables. I want to share the application of Theorem 2.1 in another graph problem which takes vertices as variables.

Theorem 8. *Let p be a prime, and let $G = (V, E)$ be a graph on a set of $|V| > d(p-1)$ vertices. Then there is a nonempty subset U of vertices of G such that the number of cliques of d vertices of G that intersect U is 0 modulo p.*

Proof. For each subset I of vertices of G, let $K(I)$ denote the number of copies of K_d in G that contain I. Associate each vertex $v \in V$ with a variable x_v, and consider the polynomial

$$F = \prod_{v \in V} (1 - x_v) - 1 + G,$$

where

$$G = \left[\sum_{\emptyset \neq I \subset V} (-1)^{|I|+1} K(I) \prod_{i \in I} x_i \right]^{p-1}$$

over $GF(p)$. Since $K(I)$ is obviously zero for all I of cardinality bigger than d, the degree of this polynomial is $|V|$, as the degree of G is at most $d(p-1) < |V|$. Moreover, the coefficient of $\prod_{v \in V} x_v$ in F is $(-1)^{|V|} \neq 0$. Therefore, by Theorem 1.2, there are $x_v \in \{0, 1\}$ for which $F(x_v : v \in V) \neq 0$. Since F vanishes on the all 0

vector, it follows that not all numbers x_v are zero, and hence that $G(x_v : v \in V) \neq 1$, implying, by Fermat's little Theorem that

$$\sum_{\emptyset \neq I \subset V} (-1)^{|I|+1} K(I) \prod_{i \in I} x_i \equiv 0 (\mathrm{mod} p).$$

However, the left hand side of the last congruence is precisely the number of copies of K_d that intersect the set $U = \{v : x_v = 1\}$, by the Inclusion-Exclusion formula. Since U is nonempty, the desired result follows. □

3 Lang-Weil Bound and Applications

3.1 Lang-Weil Bound

In the survey, I find that algebraic geometry method can be used to construct probability measure with certain property. For example, the probability measure can not be so "continuous" and make the approximation tighter.

One excellent work is by Boris Bukh [2] in 2014, which utilizes Lang-Weil Bound [6] to guarantee such property of probability measure and applies it to the graph problem.

The following is Lang-Weil Bound theorem.

Theorem 9 (Lang-Weil Bound, 1954 [6])**.** *For every s and d there exists a constant C such the following holds: Suppose $f_1(Y), \ldots, f_s(Y)$ are s polynomials on $\mathbb{F}_q^s$ of degree at most d, and consider the set*

$$W = \left\{ y \in \mathbb{F}_q^s : f_1(y) = \cdots = f_s(y) = 0 \right\}.$$

Then exactly one of the following holds:

1. *(Zero-dimensional case) $|W| \leq C$,*
2. *(Higher-dimensional case) $|W| \geq q - C\sqrt{q}$. The constant C depends only on s and the degrees of f 's.*

The proof of this theorem is complicated for me. Some concepts like the degree of variety, and some results like Bezout's inequality seem fresh for me. If the reader is interested, I recommend the references [6, 2].

3.2 Turán's Question for Complete Bipartite

In extremal graph theory, Turán question is famous: how many edges can a graph have if it does not contain H as a subgraph?

Let $\mathrm{ex}(n, H)$ be the maximum number of edges in any n-vertex H-free graph. Turán [8] determined $\mathrm{ex}(n, H)$ when H is a clique.

Theorem 10. *If $H = K_r$, then the maximum $\mathrm{ex}(n, H)$ is attained by a complete $(r-1)$-partite graph whose parts are as equal as possible.*

The proof of this theorem can be found in the elementary book of graph theory.

In 1965, Erdős-Stone and Simonovits [4] showed that for every graph H the largest H-free graph is appropriately close to a complete multipartite graph on $\chi(H)-1$ parts, where $\chi(H)$ is the chromatic number of H.

Theorem 11.

$$\mathrm{ex}(n,H)=\left(1-\frac{1}{\chi(H)-1}\right)\binom{n}{2}+o\left(n^2\right) \tag{1}$$

We can interpret $1-\frac{1}{\chi(H)-1}$ as the fraction of the total number of edges in the complete graph. When $\chi(H)\geq 3$, the fraction is positive and Equation (1) is a satisfactory asymptotics.

However, if we want to determine $\mathrm{ex}(n,H)$ for bipartite graph H, the above formula fails. Because we have $\chi(H)=2$ and the main term in Equation (1) vanishes, leaving a notoriously hard open problem of finding an asymptotics for $\mathrm{ex}(n,H)$ when H is bipartite.

In 1954, people determined a bound of $\mathrm{ex}(n,H)$ for bipartite graph as follows. I also give the original proof.

Theorem 12 (Kovári-Sós-Turán, 1954 [5])**.** *For each s and t there is a constant C such that* $\mathrm{ex}\left(n,K_{s,t}\right)\leq Cn^{2-1/s}$.

The following proof is not necessarily important in this paper. The reader can skip it without any worry.

Proof. We let C be a large constant (to be specified later). Suppose $G=(V,E)$ is a $K_{s,t}$-free graph. It suffices to prove that G contains a vertex of degree less than $Cn^{1-1/s}$, for then we may remove it, and apply the induction on the number of vertices since $C(n-1)^{2-1/s}+Cn^{1-1/s}\leq Cn^{2-1/s}$. Assume, for contradiction's sake, that $\deg(v)\geq Cn^{1-1/s}$ for all $v\in V$.

Let N denote the number of copies $K_{1,s}$ in G. We count N in two different ways. On one hand, denoting by $\deg(v)$ the degree of $v\in V$, we obtain

$$N=\sum_{v\in V}\binom{\deg(v)}{s}$$

the summand being the number of copies of $K_{1,s}$ with the apex v. Since $\deg(v)\geq Cn^{1-1/s}$ for all v and C is sufficiently large in terms of s, we have $\binom{\deg(v)}{s}\geq\left(\frac{1}{2}Cn^{1-1/s}\right)^s/s!=2^{-s}C^sn^{s-1}/s!$ and hence

$$N\geq 2^{-s}C^sn^s/s! \tag{2}$$

On the other hand, if $\{u_1,\ldots,u_s\}$ is any set of s vertices, then no more than $t-1$ vertices can be adjacent to all of these s vertices, as G is $K_{s,t}$-free. Thus

$$N\leq(t-1)\binom{n}{s}. \tag{3}$$

Combining (2) and (3) together with the simple bound $(t-1)\binom{n}{s}\leq(t-1)n^s/s!$ yields a contradiction unless $C\leq 2(t-1)^{1/s}$. □

3.3 Random Algebraic Construction

The probabilistic method for this problem is best $\mathrm{ex}(n, K_{s,t}) = \Omega\left(n^{2-\frac{2}{s+1}}\right)$, which is not as good as the expected lower bound.

The general procedure for obtaining the lower bound of a formula using a probabilistic method is to consider a random graph with an expected number of edges, then count the approximate number of edges in the random graph, destroy them all, and finally count the remaining edges of the graph.

Why doesn't this approach have a good lower bound? I think the problem lies in the number of numbers $K_{s,t}$. For example, for a set of points s, we can only calculate the expected number of U by estimating the number of common neighborhood points $\mathbb{P}(|N(U)| \geqslant t)$. The problem here is that the probability given by the random graph model is continuous. This results in a smooth probability distribution with a long, de-decaying tail, even though $\mathbb{P}(|N(U)| \geqslant t)$ has a low expectation. This makes it extremely difficult to control the number of formulas, dragging down the value of the formula in the random graph so that there are not as many edges to be obtained in this way.

We discussed earlier that one of the reasons why pure probability methods are so bad is that we define random graph models that are smooth/continuous in their probability distributions. Algebraic methods tend to imply discrete/non-smooth. There's a naive intuition that if we can embed something that's not smooth on top of the probabilistic approach, maybe we'll get better results.

In the work [2], Bukh focus on the best-understood class of bipartite graphs, the complete bipartite graphs. I think the questions the paper want to deal are as follows:

1. The paper need to find a random graph substitute, polynomial on finite field seems to be a good choice, we can specify the upper bound of polynomial degree,
2. How do we define the points of our constructed graph and the rules for linking points to points?
3. How to characterize $K_{s,t}$ from the view of polynomial?
4. What conclusions from algebraic geometry can be applied?

Let q be a prime power, and let $\mathbb{F}_q$ be the finite field of order q. We shall assume that $s \geq 4$ is fixed, and that q is sufficiently large as a function of s. Let $d = s^2 - s + 2, n = q^s$. The graph G that we will construct in this section will be bipartite. Each of the two parts, L and R, will be identified with $\mathbb{F}_q^s$

Suppose f is a polynomial in $2s$ variables over $\mathbb{F}_q$. We write the polynomial as $f(X, Y)$ where $X = (X_1, \ldots, X_s)$ and $Y = (Y_1, \ldots, Y_s)$ are the first and the last s variables respectively. Such a polynomial induces a bipartite graph in the natural way: pair $(x, y) \in L \times R$ is an edge if $f(x, y) = 0$. Let $\mathcal{P} \subset \mathbb{F}_q[X, Y]$ be the set of all polynomials of degree at most d in each of X and Y. Pick a polynomial f uniformly from $\mathcal{P}$ and let G be the associated graph. We shall show that G, on average, contains many edges but hardly any copies of $K_{s,t}$ for $t = s^d + 1$. We will then remove few vertices from G to render G completely free of $K_{s,t}$'s while still leaving many edges left.

We show that G behaves very similarly to the random graph that we constructed in the previous section with $p = 1/q$. We begin by counting the number of edges in G.

Lemma 2. *For every $u, v \in \mathbb{F}_q^s$, we have $\Pr[f(u, v) = 0] = 1/q$. In particular, the expected number of edges in G is n^2/q.*

Proof. Fix $u, v \in \mathbb{F}_q^s$. Let $\mathcal{P}_0 = \{f \in \mathcal{P} : f(0,0) = 0\}$ be the set of polynomials with zero constant term. Every $f \in \mathcal{P}$ can be written uniquely as $f = g + h$, where $g \in \mathcal{P}_0$ and h is a constant. So, a way to sample $f \in \mathcal{P}$ uniformly is to first sample g from $\mathcal{P}_0$, and then sample h from $\mathbb{F}_q$. It is clear that having chosen g, out of q possible choices for h exactly one choice results in $f(u, v) = 0$. □

To count the copies of $K_{s,t}$ we shall look at the distribution of $|N(U)|$, where U is an arbitrary set of s vertices in the same part. We shall focus on the case $U \subset L$, the other case being symmetric. Computing the distribution of $|N(U)|$ directly is hard. Instead, we will compute moments of $|N(U)|$ with aid of the following two lemmas:

Lemma 3. *Suppose $u, u' \in \mathbb{F}_q^s$ are two distinct points, and L is a linear function chosen uniformly among all linear functions $\mathbb{F}_q^s \to \mathbb{F}_q$. Then* $\Pr[\ Lu = Lu\ '] = 1/q$.

Proof. Since u and u' are distinct, there is a coordinate in which they differ. Without loss of generality, it is the first coordinate. A linear function is uniquely determined by its action on the basis vectors $e_1, \ldots, e_s$. Sample L by first sampling $Le_2, \ldots, Le_s$ and then sampling Le_1. Having chosen $Le_2, \ldots, Le_s$ there is precisely one choice for Le_1 such that $Lu = Lu'$. □

Lemma 4. *Suppose $r, s \leq \min(\sqrt{q}, d)$. Let $U \subset \mathbb{F}_q^s$ and $V \subset \mathbb{F}_q^s$ be sets of size s and r respectively. Then*

$$\Pr[f(u, v) = 0 \text{ for all } u \in U, v \in V] = q^{-sr}.$$

I think the proof of this lemma is similar to the above lemmas, but with more dedicated operations. I omit it for clarity.

The following operations is common in graph techniques by calculating the moments of a randan variable. For a fixed set $U \subset \mathbb{F}_q^s$ of size s. For $v \in \mathbb{F}_q^s$, put $I(v) = 1$ if $f(u, v) = 0$ for all $u \in U$, and $I(v) = 0$ if $f(u, v) \neq 0$ for some $u \in U$. The d 'th moment of $|N(U)|$ is easily computed by writing $|N(U)|$ as a sum of $I(v)$'s and expanding:

$$\begin{aligned}\mathbb{E}\left[|N(U)|^d\right] &= \mathbb{E}\left[\left(\sum_{v \in \mathbb{F}_q^s} I(v)\right)^d\right] = \mathbb{E}\left[\sum_{v_1, \ldots, v_d \in \mathbb{F}_q^s} I(v_1) I(v_2) \cdots I(v_d)\right] \\ &= \sum_{v_1, \ldots, v_d \in \mathbb{F}_q^s} \mathbb{E}\left[I(v_1) I(v_2) \cdots I(v_d)\right]\end{aligned}$$

The preceding lemma tells us that the summand is equal to q^{-rs} if there are exactly r distinct points among $v_1, \ldots, v_d$. Let M_r be the number of surjective functions from a d-element set onto an r element set, and let $M = \sum_{r \leq d} M_r$. Breaking the sum according to the number of distinct elements among $v_1, \ldots, v_d$, we see that

$$\mathbb{E}\left[|N(U)|^d\right] \leq \sum_{r \leq d} \begin{pmatrix} q^s \\ r \end{pmatrix} M_r q^{-rs} \leq \sum_{r \leq d} M_r = M.$$

We can use the moments to bound the probability that $|N(U)|$ is large:

$$\Pr[|N(U)| \geq \lambda] = \Pr\left[|N(U)|^d \geq \lambda^d\right] \leq \frac{\mathbb{E}\left[|N(U)|^d\right]}{\lambda^d} \leq \frac{M}{\lambda^d}.$$

The following is important, where the Lang-Weil Bound limits the choice of $N(U)$. For s polynomials $f(u, \cdot)$ as u ranges over U. The preceding lemma then says that either $|N(U)| \leq C$ or $|N(U)| \geq q/2$ if q is

sufficiently large in terms of s. Notice that when q is large, $q - C\sqrt{q} > q/2$. From Lang-Weil Bound (9), we have

$$\Pr[|N(U)| > C] = \Pr[|N(U)| \geq q/2] \leq \frac{M}{(q/2)^d}.$$

Call a set of s vertices of G bad if their common neighborhood has more than C vertices. Let B the number of bad sets. The above shows that

$$\mathbb{E}[B] \leq 2 \begin{pmatrix} n \\ s \end{pmatrix} \frac{M}{(q/2)^d} = O\left(q^{s-2}\right). \tag{4}$$

Remove a vertex from each bad set counted by B from G to obtain graph G'. Since no vertex has degree more than q^s, the number of edges in G' is at most Bq^s fewer than in G. Hence, the expected number of edges in G' is at least

$$n^2/q - \mathbb{E}[B]q^s = \Omega\left(n^{2-1/s}\right)$$

where n^2/q comes from Lemma 2 , and the estimation of $\mathbb{E}[B]$ comes from Equation (4). Therefore, there exists a graph with at most $2n$ vertices and $\Omega\left(n^{2-1/s}\right)$ edges, but without $K_{s,C+1}$.

4 Conclusion

In this work, I survey preliminary attempts to apply theorems in algebraic geometry into application fields, such as combinatorics and graph theory.

I learn Alon's Combinatorial Nullstellensatz and its versatility. These methods need dedicated construction of the polynomials to make the contradiction.

I also learn the use of Lang-Weil Bound to guarantee the property of specific probability measures. However, I do have difficulties understanding the proof of Lang-Weil Bound.

From my perspective, though an engineer use theorems most of the time, it is not enough to just utilize these theorems in algebraic geometry. It is important to understand why we get these theorems and why these theorems capture the essence of the problems in such a subtle way.

References

[1] Noga Alon. "Combinatorial nullstellensatz". In: *Combinatorics, Probability and Computing* 8.1-2 (1999), pp. 7–29.

[2] Boris Bukh. "Random algebraic construction of extremal graphs". In: *arXiv preprint arXiv:1409.3856* (2014).

[3] Michael R Douglas. "From algebraic geometry to machine learning". In: *Pure and Applied Mathematics Quarterly* 17.2 (2021), pp. 605–617.

[4] Paul Erdős and Miklós Simonovits. "A limit theorem in graph theory". In: *Studia Sci. Math. Hung.* Citeseer. 1965.

[5] P Kővári, Vera T Sós, and Pál Turán. "On a problem of Zarankiewicz". In: *Colloquium Mathematicum*. Vol. 3. Polska Akademia Nauk. 1954, pp. 50–57.

[6] Serge Lang and André Weil. "Number of points of varieties in finite fields". In: *American Journal of Mathematics* 76.4 (1954), pp. 819–827.

[7] VA Taskinov. “Regular subgraphs of regular graphs”. In: *Soviet Math. Dokl.* Vol. 26. 1. 1982, pp. 37–38.

[8] Paul Turán. “On an external problem in graph theory”. In: *Mat. Fiz. Lapok* 48 (1941), pp. 436–452.

[9] Sumio Watanabe. *Algebraic geometry and statistical learning theory*. 25. Cambridge university press, 2009.