Algebraic Independence

UGP : CS498A Report
Advisor : Dr. Nitin Saxena

Tushant Mittal
Indian Institute of Technology, Kanpur
Contents

1 Introduction 2
  1.1 The Problem ................................................. 2
  1.2 Motivation ................................................. 2
  1.3 Preliminary Definitions .................................... 3

2 Previous Work 4
  2.1 Computability ............................................. 4
    2.1.1 The “Brute force” Algorithm .......................... 4
  2.2 Characteristic 0 (or large) fields ........................... 5
  2.3 Witt-Jacobian Criterion .................................... 5
  2.4 Generalizing the Jacobian .................................. 5

3 Dimension Reduction 6
  3.1 Notation ................................................. 6
  3.2 The first approach ....................................... 6
  3.3 The k-gap .............................................. 7
    3.3.1 Bivariate Case ..................................... 7

4 New Criterion 9
  4.1 Ideal Shrink ........................................... 9
  4.2 Criterion ............................................. 10

5 Conclusion and Future Directions 12

6 Acknowledgements 13
Chapter 1

Introduction

1.1 The Problem

The concept of algebraic independence is a natural generalization of the familiar notion of linear dependence. More formally,

**Definition 1.1.**

A subset \( S \) of a field \( L \) is algebraically dependent over a subfield \( K \) if the elements of \( S \) satisfy a non-trivial polynomial equation with coefficients in \( K \).

A few concrete examples are:

- **Algebraic/Transcendental Numbers**: \( L = \mathbb{C}, K = \mathbb{Q}, S = \{\alpha\} \)
- **Polynomials**: \( L = \mathbb{F}(x_1, \cdots, x_n), K = \mathbb{F}, S = \{f_1, \cdots, f_n\} \)

The problem of testing algebraic independence is then,

Given a set of polynomials \( \{f_1, \cdots, f_n\} \) determine if they are algebraically dependent i.e does there exist \( A \in \mathbb{F}[y_1, \cdots, y_n] \) such that \( A(f_1, \cdots, f_n) = 0 \). (\( A \) is called its annihilating polynomial).

**Examples**

1. The set \( f = \{x_1, x_2, \cdots, x_k\} \) is always algebraically independent.

2. Algebraic dependence depends on the underlying field, \( \{x_1 + x_2, x_1^p + x_2^p\} \) is independent over \( \mathbb{Q} \) but is dependent over \( \mathbb{F}_p \) with \( y_2 - y_1^p \) as the annihilating polynomial.

1.2 Motivation

It’s a natural algebraic question connections to many fields of mathematics like Algebraic Geometry, dimension theory, field theory etc. It has also many applications in theoretical computer science.
especially in arithmetic circuit complexity. A few classical ones are

- Schönhage’s simplified proof of Strassen’s lower bounds [Sch76]
- Kalorkoti’s lower bounds on determinant computation [Kal82].
- More recently, Dvir, Gabizon and Wigderson’s construction of explicit rank extractors used the idea of algebraic independence [DGW09]
- Beecken, Mittmann, Saxena defined a notion of rank for arithmetic circuits and gave its applications to the long-standing problem of Polynomial Identity Testing [BMS11].

1.3 Preliminary Definitions

Before we begin our exploration let us first define a few important terms.

**Definition 1.2** (Minimal Polynomial). If $L/K$, then $\alpha \in L$ is said to be algebraic over $K$ if $\exists f \in K[x]$ such that $f(\alpha) = 0$. Of all such $f$’s, the one with the lowest degree is called the minimal polynomial of $\alpha$.

**Definition 1.3** (Transcendence degree). The transcendence degree of a set of polynomials $f = \{f_1, \cdots, f_n\} f_i \in F[x]$ is the size of its maximal subset that is algebraically independent.

**Definition 1.4** (Separable Polynomial). A polynomial $f \in F[x]$ is separable if $f$ has no repeated roots in its splitting field (i.e the smallest field extension over $F$ containing all its roots).

**Definition 1.5** (Inseparable degree). The inseparable degree of a set of polynomials $f = \{f_1, \cdots, f_n\} f_i \in F[x]$ (i.e of the field extension $F(x)/F(f)$) is the least integer $d$ such that the minimal polynomial of $x_i^d$ is separable $\forall i \in [n]$.

**Definition 1.6** (Separating Transcendence Basis). A subset $g \subset f = \{f_1, \cdots, f_n\} f_i \in F[x]$ is a separating transcendence basis if there exists a separable annihilating polynomial of $\bar{g} \cup \{f_i\} \forall i \in [n]$. 


Chapter 2

Previous Work

2.1 Computability

It is not evident from the problem statement that the problem is even computable. Oskar Perron in 1927 gave a degree bound for the annihilating polynomial which enables computability via a natural algorithm.

**Theorem 2.1 (Perron ’27).** Let \( f_i \in K[x_1, \cdots, x_n] \) be a set of \( n+1 \) non-constant polynomials and let \( \delta_i := \text{deg}(f_i) \). Then \( \exists A \in K[y_1, \cdots, y_{n+1}] \) such that \( A(f_1, \cdots, f_{n+1}) = 0 \) and

\[
\text{deg}(A) \leq \frac{\delta_1 \cdots \delta_{n+1}}{\min\{\delta_1, \cdots, \delta_{n+1}\}} \leq (\max\{\delta_1, \cdots, \delta_{n+1}\})^n
\]

A detailed proof can be found in [Plo05]. Kayal [Kay09] generalized it to sets with arbitrary number of polynomials over fields of zero characteristic. His result depends on the transcendence degree and is independent of the number of variables. Mittman [Mit12] generalised Kayals result to fields of arbitrary characteristic.

2.1.1 The “Brute force” Algorithm

Since, the annihilating polynomial’s degree is bounded we can consider a general equation of the polynomial

\[
F = \sum_{\sum_i w_i < d^n} a_w \prod_i y_i^{w_i}, \quad a_w \in K
\]

Substituting the \( f_i \)s in \( y_i \)s and setting coefficient of each monomial to 0 leads to a system of linear equations. If no solution exists then the polynomials are independent. But since the degree \((d^n)\) is high, the system is exponential sized and its complexity is in \( \text{PSPACE} \). Moreover, Kayal has showed that this bound is tight and that computing even the constant of the annihilating polynomial is \#P hard [Kay09].
2.2 Characteristic 0 (or large) fields

The earliest criterion was due to Carl Gustav Jacob Jacobi in 1841 which naturally leads to a randomized poly-time algorithm.

Theorem 2.2 (Jacobian Criterion). Let $f_i \in K[x]$ be a set of non-constant polynomials with $\text{deg}(f_i) < d$ and let $\text{char}(K) = 0$ or $> d$

$$rk_{K[x]}(J_x(f)) = \text{trdeg}(f) \quad \text{where} \quad J_x(f) = (\partial_j f_i)_{i,j}$$

In particular, $f$ is algebraically dependent iff its Jacobian is 0.

The reader may refer to [BMS11] for a proof. Using the DeMillo-Lipton-Schwartz-Zippel lemma [Zip79], we can check whether the $\det(J_x) = 0$ by evaluating it at a random set of points in polynomial time.

2.3 Witt-Jacobian Criterion

Mittmann, Saxena, Scheiblechner [MSS12] gave the first non-trivial algorithm to test independence. The idea is to lift the problem to a $\text{char} 0$ field namely, the p-adic field ($\hat{\mathbb{Z}}_p$). The algorithm reduced the complexity from PSPACE to NP$\#P$ which is where the problem is currently placed.

2.4 Generalizing the Jacobian

- Pandey, Saxena, Sinhababu (2016) [PSS16] gave a new criterion that relates algebraic dependence to approximate functional dependence
- It identifies the inseparable degree as a crucial parameter and shows that if a set of polynomials are independent then they can’t be approximately functionally dependent up to any precision greater than this inseparable degree.

Theorem 2.3. Denote $f = \{f_1, \ldots, f_n\}$. If $\text{trdeg} f = k$, then there exist algebraically independent $\{g_1, \ldots, g_k\} \subset f$ such that for random $a \in \overline{\mathbb{F}}^n$, there are polynomials $h_i \in \overline{\mathbb{F}}[Y_1, \ldots, Y_k]$ satisfying,

$\forall \, i \in [m], \quad f_i^{\leq t}(x + a) = h_i^{\leq t}(g_1(x + a), \ldots, g_k(x + a))$

Theorem 2.4. If $f$ are algebraically independent with inseparable degree $p^i$. Then,

- $\forall \, 1 \leq t \leq p^i$ for random $a \in \overline{\mathbb{F}}^n \ \exists h_j \in \overline{\mathbb{F}}[Y_1, \ldots, Y_{n-1}], \forall \, j \in [n], \quad f_j^{\leq t}(x + a) = h_j^{\leq t}(f_1(x + a), \ldots, f_{j-1}(x + a), f_{j+1}(x + a), \cdots, f_n(x + a))$
- $\forall \, t > p^i$ for random $a \in \overline{\mathbb{F}}^n \ \nexists h, \quad f_n^{\leq t}(x + a) = h^{\leq t}(f_1(x + a), \cdots, f_{n-1}(x + a))$

This gives an algorithm to check if $f$ is algebraically independent by checking approximate functional dependence up to the inseparable degree
Chapter 3

Dimension Reduction

The idea is to map each variable \( x_i \) to a random polynomial in just one variable \( t \). Clearly this will lead to an algebraically dependent set of polynomials but we investigate whether the functional (in)dependence of these one-dimensional polynomials is related to the algebraic (in)dependency of the original polynomials.

3.1 Notation

The map \( \phi_i : x_i \rightarrow \mathbb{F}_p[t] \) and denote by \( \phi(f) := f(\phi_1(x_1), \cdots, \phi_n(x_n)) \). Also, \( \bar{x} \) is used to succinctly represent \( x_1, x_2, \cdots, x_n \).

3.2 The first approach

The first “natural” idea we had was to map each variable to a random univariate of appropriately enough high degree. The following observation, however, shows that such a naive dimension reduction can’t work. \( \phi_i(x_i) = a_{i0} + a_{i1}t + a_{i2}t^2 + \cdots + a_{iN}t^N \), where \( a_{ij} \) are random elements \( \in \mathbb{F}_p \).

**Theorem 3.1.** Given \( f, g \in \mathbb{F}_p[t] \) such that \( \phi(g) \) has non-zero \( t \) coefficient \( (a_1) \), then, \( \forall d, \exists h_d \) such that \( f = h_d(g) \mod <t^{d+1}> \)

**Proof.** We will prove it using induction. For \( d=0 \), it is trivial as we set \( h_0 = f(0) \) Assume it is true \( \forall d < D, \implies f = h_D(g) + bDt^D < t^{D+1} > \) If \( b_D = 0 \), we are done. Else, choose \( h_{D+1} = h_D - \frac{b_D}{a_1^D} (g - g(0))^D \) \( \square \)
3.3 The k-gap

Since it is the coefficient of $t$ that is the cause, we make it 0. We, thus, modify the earlier map by multiplying it by a $t^{k_i}$ factor. $\phi_i(x_i) = t^{k_i}(a_{ik_1} + a_{ik_1+1}t + \cdots a_{ik_i+N}t^N)$, where $a_{ij}$ are random elements $\in \mathbb{F}_p$. But before we begin let us prove a simple but necessary lemma that let’s us translate equations between polynomial rings.

**Lemma 3.2.** Let $g(\bar{x}) = 0 \mod <\bar{x}^d>$, then $\phi(g(\bar{x})) = 0 \mod <t^{k_i}>$ where $K = \min k_i$ where the minimum is over those $i$ such that $g \notin \mathbb{F}_p[\bar{x} \setminus x_i]$.

**Proof.** Given a d degree monomial $\prod x_{i}^{\alpha_{i}}$, $\phi(\prod x_{i}^{\alpha_{i}})$ has the least degree = $\sum \alpha_{i}k_{i}$ where $\sum \alpha_{i} = d$. Therefore, $\sum \alpha_{i}k_{i} \geq dK$. We can easily construct cases where equality occurs and thus this choice of $K$ is the least that can be chosen in general. □

3.3.1 Bivariate Case

We now show that if the original set of polynomials were algebraically dependent, then the reduced polynomials are functionally dependent.

**Theorem 3.3.** Given $f(\bar{x}) \exists \phi_i$ such that $\phi(f)$ are algebraically dependent if $f$ are functionally dependent.

**Proof.** Using the result from [PSS16], we have that $\exists g \subset f$ such that $\forall i$, $\exists h_i$ the equation $f_i^{\leq d}(x + a) = h_i(g_1(x + a), \cdots , g_k(x + a))$ holds. From the above lemma, we get, $\phi(f_i) = a + h_i(\phi(g_1)) \mod <t^{dK} > a \in \mathbb{F}_p$, $K = \min_{i \in [n]}k_i$ □

We prove the converse only for the bivariate case.

**Theorem 3.4.** Given $f(\bar{x}) \exists \phi_i$ such that $\phi(f)$ are algebraically independent if $f$ are functionally independent.

**Proof.** Let $p_i$ be the inseparable degree.

An equivalent criteria for $f$ being algebraically independent is that each $x_j^{p_i}$ depends on it separably. Thus we have that,

$x_i^{p_i} = F_i(f) \mod <x^{p_i+1}> \forall i \in [2]$

Applying the $\phi$ map,

$(a_{ik_i}t^{k_i} + a_{ik_i+1}t^{k_i+1} + \cdots )^{p_i} = F_i(\phi(f)) \mod <t^{K(p_i+1)} >$

Assume that $\phi(f)$ are functionally dependent thus we can discard one, say $f_2$ from the set and still have these 2 equations. Thus,
$$a_{1k_1}t^{k_1p^e} + a_{1k_1+1}t^{(k_1+1)p^e} + \cdots = \sum_j c_j \phi(f_1)^j \mod t^{K(p^e+1)}$$

Least degree (non-zero) of LHS is $t^{p^e k_1}$ and thus least degree of $\phi(f_1)$ say $t^l|t^{p^e k_1}$ i.e. $l|p^e k_1$ and similarly, $l|p^e k_2$.

And this clearly gives us the required contradiction as $k_1, k_2$ can be chosen to be coprime. \qed

This proof doesn’t generalize because the divisibility criteria holds only for $n = 2$. It is thus, not clear whether an efficient reduction in a general case is possible. Such a reduction will however, lead to a significant improvement in the time complexity of the problem. Thus, we ask the following question,

**Open Problem 3.1**

*Does there exist a polynomial map $\phi_i : x_i \rightarrow F[x_1, \cdots, x_c] \forall i \in [n]$ where $c = O(1)$, such that for any $f \in F[x_1, \cdots, x_n]$ $f$ is algebraically dependent $\iff \phi(f)$ is?*
Chapter 4

New Criterion

We now introduce a new criterion that is equivalent to algebraic independence but is in the form of linear dependence of shifted polynomials modulo the square of the ideal generated by these polynomials.

4.1 Ideal Shrink

Lemma 4.1. Let $f_i \in \langle \bar{x} \rangle \subset \mathbb{F}[x]$. If $f_n = \sum_{i}^{n-1} c_i f_i \mod < f_1, \cdots, f_n >^2$, $c_i \in \mathbb{F}$, then $< f_1, \cdots, f_n >^2 = < f_1, \cdots, f_{n-1} >^2$

Proof. Let $I = < f_1, \cdots, f_{n-1} >^2$. We will show that each of the generators of $< f_1, \cdots, f_n >^2$ lie in $I$. The only non-overlapping generators are $\{f_n f_i | i \in [n] \}$. By the hypothesis we have,

$$f_n = \sum_{i}^{n-1} c_i f_i + f_n \left( \sum_{i}^{n-1} g_i f_i \right) + f_n^2 G \mod I$$

$$f_j f_n = \sum_{i}^{n-1} c_i (f_i f_j) + f_n \left( \sum_{i}^{n-1} g_i (f_i f_j) \right) + (f_j f_n) f_n G \mod I \forall j \in [n-1]$$

$$f_j f_n = f_j f_n^2 G \mod I$$

$$f_j f_n = f_j f_n (f_n G)^k \mod I \forall k > 0$$

$$\Rightarrow f_j f_n \in I + < f_n^k > \forall k > 0$$

$$f_n = \sum_{i}^{n-1} c_i (f_i) + \left( \sum_{i}^{n-1} g_i (f_i f_n) \right) + f_n^2 G \mod I$$

$$= \sum_{i}^{n-1} c_i (f_i) + f_n^2 G' \mod I$$

$$\Rightarrow f_n^2 = f_n^2 \sum_{i}^{n-1} c_i (f_i) + f_n^4 G \mod I$$
\[ f_n^k H + + f_n^4 G \mod I \]
\[ \Rightarrow f_n^2 \in I + < f_n^4 > \]
Continuing this we get that,
\[ \Rightarrow f_n^2 \in I + < f_n^l > \forall l > 1 \]

Let \( l \) be larger than \( \max_i \{ \deg(f_i^2) \} \). Now since, \( f_n^2 = \sum_{i=1}^{n-1} f_i^2 g_i + f'_i g_n \) we can remove the dependence by looking at \( g_i \mod x^{\deg(f_i^2)} \) and thus \( f_n^2 \in I \).

4.2 Criterion

We pick a point randomly from \( \mathbb{F}^n \) say, \( \bar{\alpha} \in_r \mathbb{F}^n \), and define the constant free shifted polynomial \( Hf_i = f_i(\bar{x} + \bar{\alpha}) - f_i(\bar{\alpha}) \).

**Theorem 4.2.** \( f_1, \cdots, f_n \) are algebraically dependent iff \( \exists c \in \mathbb{F}^n \setminus 0^n \) such that \( \sum_{i=1}^{n} c_i Hf_i \in < Hf_1, \cdots, Hf_n >_{\mathbb{F}[x]}^2 \).

**Proof.** Case 1 \( f \) are algebraically dependent.

This part of the proof is very similar to that of theorem 10 in [PSS16].

Let \( g = \{g_1, g_2, \cdots, g_k\} \subset f \) be its separating transcendence basis. For any \( i \), let \( g_0 := f_i \), then, \( \{g_0\} \cup g \) has a minimal separable annihilating polynomial say \( A_i(y) = \sum e_i a_{e_i} y^{e_i} \). Now, \( A_i(g) = \sum e_i g^{e_i} = 0 \) and replacing \( \bar{x} \) by \( \bar{x} + \bar{\alpha} \) where \( \bar{\alpha} \) is a randomly chosen element of \( \mathbb{F}^n \). Writing \( g_i(\bar{x} + \bar{\alpha}) = Hg_i + g(\bar{\alpha}) \) and expanding the entire sum using Taylor’s series, we get,

\[
A_i(g) = \sum e_i a_{e_i} \prod_{j=0}^{k} (Hg_j + g_j(\bar{\alpha}))^{e_{ij}}
\]

\[ 0 = A(g(\bar{\alpha})) + \sum_{j=0}^{k} \frac{\partial A_i}{\partial y_j} \bigg|_{g(\bar{\alpha})} Hg_j \mod < Hg_0, Hg_1, \cdots, Hg_k >_{\mathbb{F}[x]}^2 \]

\[ 0 = \sum_{j=0}^{k} c_j Hg_j \mod < Hg_0, Hg_1, \cdots, Hg_k >_{\mathbb{F}[x]}^2 \]

Now we need to check that \( c_0 \neq 0 \). Since, \( A \) is separable \( A'(\bar{y}) \) is not identically 0 defines a polynomial which has to be non-zero due to the minimality of \( A \) and thus \( \exists \bar{\alpha} \) such \( c_0 = A'(\bar{\alpha}) \neq 0 \).

**Case 2 -** \( f \) are algebraically independent.

Let \( p^e \) be the inseparable degree. An equivalent criteria for \( f \) being algebraically independent is that each \( x_j^{p^e} \) depends on it separably. Thus we have from the above proved statement that,

\[ x_j^{p^e} = \sum_{i=1}^{n} c_{ij} Hf_j \mod I \forall i \in [n] \]

Assume also that the \( Hf_i \) are \( \mathbb{F} \) linearly dependent.
mod $< H f_1, \cdots, H f_{n-1} >^2$. Thus, via the ideal shrink lemma, we can eliminate one say $H f_n$ from the equations. We will now reach a contradiction.

$$x_i^{p^e} = \sum_{j=1}^{n-1} c_{ij} H f_j \mod < H f_1, \cdots, H f_{n-1} >^2 \quad \forall i \in [n]$$

We rewrite $H f_j = D_j + H f_j^{p^e}$ where $D_j$ is the part of $H f_j$ with total degree $< p^e$. This gives us that,

$$0 = \sum_{j=1}^{n-1} c_{ij} D_j \mod < H f_1, \cdots, H f_{n-1} >^2$$

$$\implies 0 = \sum_{j=1}^{n-1} c_{ij} D_j \mod < D_1, \cdots, D_{n-1} >^2$$

Let $k$ be the number of linearly independent $D_j$. Using the ideal shrink lemma, we can shrink the ideal to $< D_1, \cdots, D_k >$. Thus, we have at most $n - 1 - k$ linearly independent linear relations. Now without loss of generality, assume that the last $n - 1 - k$ equations are independent. For $i \in [1, k + 1]$, $\exists L_i(x_i^{p^e}, x_{k+2}^{p^e}, x_{k+3}^{p^e}, \cdots, x_n^{p^e})$ such that $L_i = 0 \mod < D_1, \cdots, D_k >^2$. Now, let’s look at the zero set of the ideal generated by the $L_i$, $L = < L_1, L_2, \cdots L_n >$.

$$L \subset < D_1, \cdots, D_k >^2$$

$$Z(L) \supset Z(< D_1, \cdots, D_k >^2)$$

$$Z(L) \supset Z(< D_1, \cdots, D_k >)$$

$$dim(Z(L)) \geq dim(Z(< D_1, \cdots, D_k >))$$

Clearly, $Z(L) = \{(f_1(a_1, \cdots, a_{n-k-1}), \cdots, f_{k+1}(a_1, \cdots, a_{n-k-1}), a_1, \cdots, a_{n-k-1}) | a \in \mathbb{F}^{n-k-1}\}$ and thus, $dim(Z(L)) = n - k - 1$. That this gives us a contradiction follows from the following result from dimension theory,

**Theorem 4.3** (Krull’s dimension Theorem). Let $R$ be a Noetherian ring and $a \subset R$ an ideal generated by elements $a_1, \cdots, a_r$. Then $ht(p) \leq r$ for every minimal prime divisor $p$ of $a$.

**Corollary 4.4** (Exercise 1.9 from [Har77]). Let $a = k[x_1, \cdots, x_n]$ be an ideal which can be generated by $r$ elements. Then every irreducible component of $Z(a)$ has dimension $\geq n - r$. 


Chapter 5

Conclusion and Future Directions

We have thus explored a couple of possible approaches to attacking the problem of algebraic dependence over finite characteristic fields. While the computational utility of such a criteria is not currently clear it will hopefully provide some geometric insight into the problem. Due to to the great disparity between the problem’s current known complexity $NP^{\#P}$ and the conjectured one ($RP$), there could be numerous successful lines of attack. A few of them could be

- Studying the algorithmic consequences of the criterion and see if it can be harnessed to create an efficient algorithm.

- Trying to deduce a dimension reduction for the general n-variate case.

- Looking at special cases like supersparse polynomials or n bivariates

- Gaining a better understanding of the relations between the different but associated notions of algebraic, analytic and functional dependence
Chapter 6

Acknowledgements

I am indebted to Prof. Nitin Saxena for investing his time and effort to advise me. The weekly meetings I have had with him were as enjoyable as they were enlightening and his keen eye made sure I didn’t push anything under the rug. The project wouldn’t have been half as interesting without Amit Sinhababu who spent a significant amount of time answering my doubts, discussing ideas and pointing me to interesting bodies of work. Thanks also to the members of SIGTACS, IITK and in particular Prof. Rajat Mittal and Sumanta Ghosh who attended both my talks and provided constructive criticism. I would also like to thank Zeyu Guo for interesting discussions.
Bibliography


