# Special case algorithms for Nullstellensatz and transcendence degree

Abhibhav Garg
Supervisor: Nitin Saxena

June 14, 2020

## Abstract

Checking for the existence of a common root of a set of polynomials is a fundamental problem in computer science, primarily since it affords reduction from many other problems. It is easy to see for example that SAT, and therefore all NP complete problems reduce to checking the existence of common roots. It is therefore unlikely that this problem has an efficient algorithm.

In this thesis, we study this problem, and some related problems, in the special case when the input polynomials have low transcendence degree. The transcendence degree of a set of polynomials is the size of the largest subset of them that do not satisfy any polynomial relationship. The case of low transcendence degree generalizes the case of having fewer polynomials than variables. In particular, the three problems we study are radical membership, effective Nullstellensatz and transcendence degree computation. The radical membership problem is to check, given polynomials $f_1, \ldots, f_m$ and a polynomial $g$, whether some power of $g$ belongs to the ideal generated by the polynomials $f_i$. By Hilbert's Nullstellensatz, taking $g = 1$ in the above is equivalent to checking for the existence of a common root of $f_1, \ldots, f_m$. In the case when 1 is in the ideal generated by $f_1, \ldots, f_m$, it is natural to compute witnesses $h_1, \ldots, h_m$ that satisfy $1 = \sum f_i h_i$. The effective Nullstellensatz gives degree upper bounds on the $h_i$ that depend on $m$ and the degree of the polynomials $f_i$. The transcendence degree problem is to compute, given a set of polynomials $f_1, \ldots, f_m$ what their transcendence degree is. For each of these, we give bounds and algorithms that depend on the transcendence degree of the polynomials $f_1, \ldots, f_m$.

We also provide exposition of Hilbert's Nullstellensatz and the Effective Nullstellensatz. We also study the algebraic independence problem, and provide alternative proofs for many known results about the problem.

# Acknowledgements

I am indebted to Nitin, not only for supervising this thesis, but for teaching me nearly everything I know about theoretical computer science. The amount I have learnt from him is immeasurable, both over the course of working on this thesis, and in the five courses that I did with him.

I am grateful to all of the other professors I have had over the years, including but not limited to Satyadev, Somenath, Manindra, Rajat, Puru, Amit, and Nandini. I thank Ramprasad for hosting me during the summer of 2019 at TIFR.

Neel, Tushant, Sahay, Gopa, Anurag, and Amur have played large roles in my life over the past four years, and I am more than grateful to all of them. I have learnt a lot from the discussions I have had with Amit, Abhiroop, Prerona, Anamay, and Subhayan, and this thesis would not have been possible without them.

I am also grateful to all the folk back at home; my parents, grandparents, Yash, Ayush and Prithvi.

# Contents

# Chapter 1

# Introduction

Given a set of multivariate polynomials $f_1, \dots, f_m$, there is a natural certificate for the existence of a common root of these polynomials, namely the root itself. Hilbert's Nullstellensatz [Kru50] is a fundamental theorem in algebraic geometry that states that there exists a certificate for the non existence of a common root, in the form of polynomials $h_1, \dots, h_m$ such that $\sum f_i h_i = 1$. These certificates are not efficient: every root can have exponential bit complexity and every set of polynomials $h_1, \dots, h_m$ with the above property can have exponential degree. It is therefore natural to ask if there are efficient certificates for one or both of the above. In this thesis, we study the above problems in the special case when the polynomials $f_1, \dots, f_m$ have low transcendence degree.

The thesis is divided into seven chapters, including this one. All chapters except chapter 6 are expositional. Chapter 6 is the main contribution of this thesis, and is based on [GS20].

In chapter 2, we establish some notation that we will use for the rest of the thesis, and also list some basic facts from field theory and algebraic geometry.

In chapter 3, we discuss effective versions of the hyperplane intersection theorem and Noether normalization. We show that random hyperplanes intersect varieties—both projective and affine—property with high probability. We also provide a proof of the Noether normalization theorem, and what it means in the algebraic geometric setting.

In chapter 4, we discuss the Nullstellensatz. We first provide a proof of the classical Nullstellensatz. We then present a proof of an effective version of the Nullstellensatz, which gives degree bounds on the certificates.

In chapter 5, we discuss the notion of transcendence degree and algebraic independence. We state the algebraic independence problem, and provide alternative proofs of some well known properties and theorems regarding algebraic independence.

In chapter 6, we show the existence of improved Nullstellensatz certificates, radical membership algorithms, and transcendence degree computing algorithms for polynomials with low transcendence degree. This chapter is the main contribution of this thesis.

Finally, in chapter 7 we provide a conclusion followed by some potential next steps.

Given the range of topics discussed, we have deferred motivations and literature surveys of the above topics to the respective chapters.

# Chapter 2

# Preliminaries

We first establish some notation that we will use throughout this thesis. We also state some basic facts from field theory and algebraic geometry that will be useful.

## 2.1 Notation

- We use $k$ to denote the underlying field of constants. This will generally be the algebraic closure of $\mathbb{F}_p$ for some prime $p$ that is either arbitrary or clear from context.
- We use $k[x]$ and $k(x)$ respectively to denote the ring of polynomials with coefficients from $k$ with indeterminate $x$, and its field of fractions.
- We use vector notation to denote indexed sets of objects when the indexing set is clear: for example, we use $\mathbf{x}$ to denote variables $x_1, \ldots, x_n$ if the number of variables is clear. We extend this vector notation greatly. For example, if $f_1, \ldots, f_m$ are polynomials each in the same $n$ variables and $a_1, \ldots, a_n \in k$ then $\mathbf{f}(\mathbf{a})$ denotes the evaluations

  $$(f_1(a_1, \ldots, a_n), f_2(a_1, \ldots, a_n), \ldots, f_m(a_1, \ldots, a_n)).$$

  If $\mathbf{x}$ is a set of $n$ variables and $\mathbf{m} \in \mathbb{N}^n$ is a vector of natural numbers then $\mathbf{x}^{\mathbf{m}}$ denotes the monomial $\prod x_i^{m_i}$.
- We use $\mathbb{A}^n$ and $\mathbb{P}^n$ respectively to denote the $n$ dimensional affine and projective spaces. We use $\mathbb{P}^n_\infty$ to denote the hyperplane at infinity.
- Given a variety $X$, we use $k[X]$ to denote its coordinate ring.

## 2.2 Algebra preliminaries

### 2.2.1 The polynomial identity lemma

We make extensive use of the following theorem that controls the number of roots of multivariate polynomials. We refer to it as the polynomial identity lemma.

**Lemma 2.2.1** (Polynomial identity lemma, [Sch80, Zip79, Ore22, DL78]). *Let $f(x_1, \dots, x_n)$ be a polynomial of degree $d$ over the field $k$. Let $S$ be a subset of the field $k$. Then the number of roots of $f$ in $S \times \cdots \times S$ is at most $d|S|^{n-1}$. Equivalently, if points $a_1, \dots, a_n$ are sampled uniformly and independently from $S$ then the probability that $f(a_1, \dots, a_n)$ is 0 is bounded above by $d/|S|$.*

### 2.2.2 Field theory and commutative algebra preliminaries

We state some basic definitions from the theory of field extensions and commutative algebra that we will use throughout this thesis.

Suppose $K$ is a field extension of $k$. Given a subset $S$ of $K$, we say that $S$ is algebraically independent if the elements of $S$ do not satisfy any polynomial equation with coefficients in $k$. The *transcendence degree* of the extension $K/k$ is the cardinality of the largest algebraically independent subset of $K$. If $T$ is such a subset, then by definition the extension $K/k(T)$ is algebraic. All maximal algebraically independent subsets of $K$ have the same cardinality. This is akin to the notion of linear independence of vectors, and in fact the proof is similar too. We refer the reader to [Lan02, Theorem 1.1, Chapter 8] for an elementary proof in the case of interest here, which is that of extensions with finite transcendental degree.

Suppose now that $K/k$ is an algebraic extension. An element $a \in K$ is called separable if the minimal polynomial of $a$ over $k$ does not have repeated roots. The extension is called separable if every element is separable. Suppose $K_s$ is the subfield of separable elements of $K$ over $k$. Then every element $a \in K$ and $a \notin K_s$ is such that $a^{p^n} \in K_s$. Here $p$ is the characteristic of the field. Inseparability is a property that can only arise in fields of finite characteristic (more specifically, only in fields that are not perfect).

All of the rings we consider will be unital and commutative. Suppose $S/R$ is an extension of rings. The ring $S$ is called an integral extension of $R$ if every $s \in S$ satisfies a monic polynomial with coefficients in $R$. If $S/R$ is an integral extension and $S'/S$ is another integral extension, then $S'/R$ is also an integral extension.

## 2.3 Algebraic geometry preliminaries

Here we state some basic facts from algebraic geometry. An excellent reference for this material is [SR13]. Another excellent reference with emphasis on computational aspects is [CLO07]. We use [Eis13] for the commutative algebra facts.

### 2.3.1 Basic definitions

Let $k$ be an algebraically closed field, and let $A$ be the ring $k[x_1, \dots, x_n]$. The ring $k$ is Noetherian since it is a field, and by repeated applications of the

Hilbert Basis Theorem [Eis13, Theorem 1.2, Chapter 1] we obtain that $A$ is Noetherian. This implies that every ideal of $A$ is finitely generated.

Let $f_1, \ldots, f_m$ be a set of polynomials from $A$. We denote by $V(\mathbf{f})$ the set of common zeroes of $\mathbf{f}$. More explicitly

$$V(\mathbf{f}) = \{(c_1, \ldots, c_n) \,|\, \forall i, f_i(\mathbf{c}) = 0\}.$$

Any simultaneous root of $\mathbf{f}$ is a root of every polynomial in the ideal generated by the $\mathbf{f}$, and therefore $V(\mathbf{f}) = V(\langle \mathbf{f} \rangle)$. Since every ideal is finitely generated, every $V(I)$ for an arbitrary ideal $I$ will be of the form $V(\mathbf{g})$ where $\mathbf{g}$ is a generating set for $I$. We will call $V(\mathbf{f})$ the affine variety defined by $\mathbf{f}$.

We think of $V(\cdot)$ as a map from the set of ideals of $A$ to subsets of $\mathbb{A}^n$. We also have a natural map in the opposite direction: Given a subset $U$ of $\mathbb{A}^n$, we define $I(U)$ to be the set of all polynomials that vanish on every element of $u$. More explicitly,

$$I(U) = \{f \in A \,|\, \forall u \in U, f(u) = 0\}.$$

The maps $I(\cdot)$ and $V(\cdot)$ are inclusion reversing. For any variety $V$ we have $V(I(V)) = V$. For every ideal $I$ we have $I(V(I)) = \sqrt{I}$. This second statement is one version of the Nullstellensatz, and will be proved in chapter 4.

### Zariski topology

The $n$ dimensional affine space $\mathbb{A}^n$ can be given a topology where the closed sets are exactly the zerosets of a finite number of polynomials. This topology is called the Zariski topology. Given a variety $V = V(I)$, it is irreducible in the Zariski topology if and only if $I$ is a prime ideal. The Zariski topology in a Noetherian topological space, which means that every descending chain of closed subsets stabilizes. This follows easily by noting that a descending chain of closed subsets corresponds to an ascending chain of ideals. As a consequence of this, every closed subset can be written uniquely as the union of irreducible closed sets, none containing another.

Suppose $f$ is a polynomial. The set of points given by $f \neq 0$ is an open set. The polynomial identity lemma stated above essentially states that given an open set of this form, a point chosen at random lies in this open set with high probability.

### Coordinate rings

Given a variety $V \subseteq \mathbb{A}^n$ corresponding to ideal $I$, the ring $k[x_1, \ldots, x_n]/I$ is called the coordinate ring of $I$. We denote it by $k[V]$. It consists of polynomial functions on $V$, that is, functions $V \to k$ that are given by polynomials. We have $k[\mathbb{A}^n] = k[x_1, \ldots, x_n]$. In the general case, $x_1, \ldots, x_n$ generate $k[X]$ as an algebra, and we call these the coordinate functions on $X$.

Suppose $V$ is irreducible. Then $I$ is prime, and $k[x_1, \ldots, x_n]/I$ is a domain. The field of fractions of this domain is called the function field of $V$, and is denoted by $k(V)$. It consists of functions on $V$ that are defined on some open subset of $V$.

### Polynomial maps

Suppose $X, Y$ are varieties in $\mathbb{A}^n, \mathbb{A}^m$ respectively. A map $\phi : X \to Y$ is called a regular map (or polynomial map) if each coordinate function $\phi_i$ of $\phi$ is an element of $k[X]$. Rephrasing, the map is called regular if there exists $\phi_1, \dots, \phi_m \in k[X]$ such that for every point $(a_1, \dots, a_n)$ in $X$ we have

$$\phi(a_1, \dots, a_n) = (\phi_1(a_1, \dots, a_n), \dots, \phi_m(a_1, \dots, a_n)).$$

Polynomial maps are continuous in the Zariski topology. Given a point $y \in Y$, the set $\phi^{-1}(y)$ is called the fibre of $y$ under $\phi$, or just the fibre of $y$ if $\phi$ is clear from context. Since points are closed in the Zariski topology, and polynomial maps are continuous, the fibre of any point is a variety.

Given a polynomial map $\phi : X \to Y$, there is an induced map $\phi^* : k[Y] \to k[X]$ that takes $f \in k[Y]$ to $f \circ \phi \in K[X]$. This map is a ring homomorphism. Alternatively, given a ring homomorphism $\psi : k[Y] \to k[X]$, there exists a map $\psi^* : X \to Y$ such that $\psi = (\psi^*)^*$. This map is defined by the coordinate functions $\psi_i^* = \psi(y_i)$, where $y_i$ is the $i^{th}$ coordinate function of the ambient space $\mathbb{A}^m$ of $Y$ (strictly speaking, the image of the coordinate function in $k[Y]$).

Suppose $\phi : X \to Y$ is a polynomial map. The image $\phi(X)$ might not be a variety. We use $\overline{\phi(X)}$ to denote the Zariski closure of the image. This is the intersection of all closed sets containing $\phi(X)$.

If the map is such that $\overline{\phi(X)} = Y$, then the map is called dense. Although a dense map is not surjective, every polynomial that vanishes on the image of a dense map will vanish on the entire codomain. If the map $\phi$ is dense, then the induced map $\phi^*$ is an injection. To see this, suppose we have $\phi^*(f) = 0$ for some $f \in k[Y]$. Then $\phi^*(f)$ vanishes on every point in $X$, and therefore $f$ vanishes on every point in $\phi(X)$. Since the map is dense, $f$ vanishes on every point in $Y$ whence $f = 0$ in $k[Y]$.

### Projective varieties

Let $\mathbb{P}^n$ be the quotient of the space $\mathbb{A}^{n+1} \setminus \{0\}$, where every line through the origin is identified. In other words, elements of $\mathbb{P}^n$ are $n+1$ tuples $(x_0, \dots, x_n)$ of elements of $k$, with at least one $x_i$ nonzero, and the equivalence $(x_0, \dots, x_n) = (\lambda x_0, \dots, \lambda x_n)$ for every nonzero $\lambda \in k$. Given a polynomial in $n+1$ variables, we do not get a well defined map on $\mathbb{P}^n$ since polynomial functions are not scale invariant. However, given a homogeneous polynomial, the zeroset of the polynomial is a well defined subset of $\mathbb{P}^n$. We can give $\mathbb{P}^n$ a topology where the closed sets are exactly the zerosets of a finite number of homogeneous polynomials.

Given the $n$-dimensional projective space $\mathbb{P}^n$, there is a copy of $\mathbb{A}^n$ embedded in it, namely the set of points of $\mathbb{P}^n$ of the form $(1, x_1, \dots, x_n)$. We call this an affine chart. The complement of this set, namely the set of points of the form $(0, x_1, \dots, x_n)$ is a closed linear subspace of $\mathbb{P}^n$ given by $x_0 = 0$, and is called the hyperplane an infinity. Suppose we have an affine variety $X$. This set $X$ is subset of $\mathbb{P}^n$ via the above inclusion of $\mathbb{A}^n$ into $\mathbb{P}^n$. Let $X^p$ be the closure of this set in the Zariski topology of $\mathbb{P}^n$. This set is called the projective closure of $X$.

There exists an ideal variety correspondence in the projective case which is similar to the one discussed above. There also exists notions of function fields for projective varieties. These constructions are more involved than their affine counterparts, and therefore we do not present them here.

The advantage of projective varieties is that they behave better with respect to intersections. Many of our arguments will therefore involve starting with an affine variety, considering its projective closure, studying intersections in the projective closure and deducing properties of the original affine variety.

### 2.3.2 The dimension and degree of a variety

**Dimension**

To every affine variety we can assign a dimension. It seems natural to want the dimension of $\mathbb{A}^n$ to be $n$, and the dimension of an affine subspace [1] to be equal to the linear algebraic dimension of the subspace. There are more motivations for the following definition of dimension which we do not provide here.

Suppose $X$ is an irreducible affine variety. As discussed above, $k[X]$ is a domain, and therefore $k(X)$ is a field extension of $X$. We define the dimension of $X$ to be equal to the transcendence degree of the extension $k(X)/k$. Suppose $Y$ is an arbitrary affine variety. As discussed above, we can write $Y$ uniquely as a union of irreducible varieties. We define the dimension of $Y$ to be the maximum of the dimensions of these irreducible varieties.

We can alternatively define the dimension as follows. Let $X$ be an irreducible variety. Consider chains of irreducible varieties of the form

$$\emptyset \neq X_0 \subsetneq X_1 \subsetneq \cdots \subsetneq X_n = X.$$

We define the dimension of $X$ to be $n$ if the above is a maximal chain of irreducible varieties. Given any chain of shorter length, it can always be refined to a chain of length $n$. This definition is very similar to the definition of linear algebraic dimension of affine spaces. In that setting, given a vector space of dimension $n$, the only vector spaces that it strictly contains are those of dimension $n-1$. The refinement statement also holds.

Note that the second definition applies directly to projective varieties, while the first definition depends on the function field, which we have not defined here. For a projective variety $W$ the intersection $W \cap \mathbb{A}^n$—where $\mathbb{A}^n$ is considered as a subset of $\mathbb{P}^n$ as discussed above—is an affine variety, and the dimension of $W$ matches the dimension of this affine variety. It also holds that the

[1] Here, by affine subspace we mean a translate of an linear subspace of $\mathbb{A}^n$ when it is treated as a $k$-vector space with origin $\mathbf{0}$.

dimension of the projective closure of an affine variety is equal to the dimension of the original variety.

There is a third definition of dimension based on the Hilbert polynomial of the coordinate ring of the variety, which we do not discuss here. A fundamental theorem in dimension theory is that these definitions are all equivalent. The proof of this fact is beyond the scope of the thesis, and can be found in the mentioned references.

Given a variety $Y \subseteq X$ with $\dim X = n$ and $\dim Y = m$, we define the codimension of $Y$ in $X$ to be $n - m$. This is denoted by $\operatorname{codim}_X Y$. When $X$ is not explicitly mentioned, we assume that $X$ is the ambient space in which $Y$ lies.

A hypersurface in $\mathbb{A}^n$ is a variety defined by a single polynomial. The variety is irreducible if and only if the polynomial is. Every irreducible component of a hypersurface has codimension 1. This is similar to the notion of hyperplanes.

We now study how the dimension behaves with the intersection of varieties. Consider the linear algebra case. Suppose we had a linear subspace $L$, and a hyperplane $H$. Then the dimension of $L \cap H$ can be either $\dim L$ or $\dim L - 1$. The first case only occurs when $L$ is contained in $H$. If not, then the second case occurs. If we consider affine subspaces, then there is a third possibility, namely that $H \cap L = \emptyset$. The linear algebraic dimension therefore either remains the same, or drops by exactly 1 when the intersection is nonempty. In our setting, affine varieties behave like affine subspaces: intersection of an affine variety with a hypersurface reduces the dimension by at most 1, as long as the intersection is nonempty. Projective varieties behave like linear subspaces: intersections of projective varieties and hypersurfaces are always nonempty (unless the projective variety has dimension 0), and therefore the intersection reduces the dimension by at most 1. We state the above, and some corollaries of the above as theorems. These will be used repeatedly throughout this thesis.

**Theorem 2.3.1** ([SR13, Cor 1.13, Section 6, Chapter 1])**.** *Suppose $W$ is an irreducible projective variety, and $H$ is a hypersurface that does not contain $W$. Then every component of $W \cap H$ has dimension $\dim W - 1$. In particular, if $\dim W \geqslant 1$, then the intersection is nonempty.*

*Suppose $X$ is an irreducible affine variety and $H$ is a hypersurface that does not contain $X$. Then every nonempty component of $X \cap H$ has dimension $\dim X - 1$.*

We say that a hyperplane intersects a variety properly if the dimension drops by exactly 1. We can repeatedly apply the above theorem to obtain the following corollary.

**Corollary 2.3.2** ([SR13, Cor 1.14, Section 6, Chapter 1])**.** *Suppose $W$ is an irreducible projective variety, and $Z \subset W$ is the set of zeroes of $m$ homogeneous polynomials on $W$. Then every component of $Z$ has dimension at least $\dim W - m$.*

*Suppose $X$ is an irreducible affine variety and $Y \subseteq X$ is the set of common zeroes of $m$ polynomials on $X$. Then every nonempty component of $Y$ has dimension at least $\dim X - m$.*

Finally, we have the following corollary about the intersection of two varieties. It follows by applying the above to the diagonal $\mathbb{P}^n \times \mathbb{P}^n$, although we omit the proof.

**Theorem 2.3.3** ([SR13, Theorem 1.24, Section 6, Chapter 1])**.** *Suppose $W, Z$ are irreducible projective varieties in $\mathbb{P}^n$ of dimensions $m_1, m_2$. Then every component of $W \cap Z$ has dimension at least $m_1 + m_2 - n$.*

*Suppose $X, Y$ are irreducible affine varieties in $\mathbb{A}^n$ of dimensions $m_1, m_2$. Then every nonempty component of $X \cap Y$ has dimension at least $m_1 + m_2 - n$.*

We say $X$ and $Y$ intersect properly if equality holds in the above.

The final thing we discuss in this chapter is the fibre dimension theorem. It states that the fibres of a surjective have dimension greater than or equal to the difference of the dimensions of the domain and codomain. This statement also holds in the linear algebraic setting: the dimension of every fibre of a surjective map is exactly equal to the dimension of the kernel of the map, which is the difference in the dimension of the domain and codomain.

**Theorem 2.3.4** ([SR13, Theorem 1.25, Section 6, Chapter 1])**.** *Let $\phi : X \to Y$ be a polynomial map. Let $n, m$ denote the dimensions of $X$ and $Y$ respectively, and assume that $\phi$ is surjective. Then $n \geqslant m$. Further,*

1. *For every $\mathbf{b} \in Y$, every component of the fibre $\phi^{-1}(\mathbf{b})$ has dimension at least $n - m$.*
2. *There exists an open set $U$ of points $\mathbf{b} \in Y$ such that $\dim \phi^{-1}(\mathbf{b}) = n - m$ for every $\mathbf{b} \in U$.*

We provide a proof sketch for this result in the special case when $Y = \mathbb{A}^n$. In our applications, we will require that appropriately sampled random points satisfy the second item above, and therefore we need some control over which points have that property, which we also do. Also note that the above statement holds if we replace surjective by dominant, in which case the first item is satisfied by every point with a nonempty fibre, and the other statements holds as is.

*Proof sketch for a special case of Theorem 2.3.4.* Since $\phi$ is a surjective map, the induced map $\phi^* : k[Y] \to k[X]$ is injective. The statement $n \geqslant m$ holds from the fact that $\operatorname{trdeg}(k(Y)) \leqslant \operatorname{trdeg}(k(X))$, which is clear from the above inclusion of $\phi^*(k[Y])$ in $k[X]$.

Now let $Y = \mathbb{A}^m$, and let $y$ be an arbitrary point in $Y$. The point $\mathbf{b}$ is defined in $Y$ by $m$ equations, namely $y_1 = b_1, \ldots, y_m = b_m$, where $y_i$ are the coordinate functions of $\mathbb{A}^m$. [2] The fibre in $X$ is therefore defined by the equations $\phi_1 = b_1, \ldots, \phi_m = b_m$, where $\phi_1, \ldots, \phi_m$ are the coordinate functions of $\phi_i$. By Corollary 2.3.2, every nonempty component of the fibre has dimension at

[2] When $Y$ is an arbitrary variety, it is not always true that a point is fixed by dimension many equations, and we have to pass to open subsets.

least $n - m$. Finally, that the fibre is nonempty follows from the fact that $\phi$ is surjective. [3]

We now prove the second item. The ring $k[Y]$ is generated by $m$ algebraically independent elements $y_1, \dots, y_m$. Under $\phi^*$, these map to $\phi_1, \dots, \phi_m$, whence these are algebraically independent elements of $k[X]$. The transcendence degree of $k(X)$ is $n$. Let $x_1, \dots, x_n$ be a transcendental basis for $k(X)$ ordered so that $\phi_1, \dots, \phi_m, x_{m+1}, \dots, x_n$ are algebraically independent. Let $A_i$ denote the annihilator of $x_i, \phi_1, \dots, \phi_m, x_{m+1}, \dots, x_n$, for $i = 1, \dots, m$.

Now let $\mathbf{b}$ be a point in $Y$, and consider its fibre $\phi^{-1}(\mathbf{b})$. Let $W$ be an irreducible component of its fibre. The ring $k[W]$ is generated by $x_1, \dots, x_n$. Suppose $\mathbf{b}$ is such that $A_i(x_i, \phi_1(\mathbf{b}), \dots, \phi_m(\mathbf{b}), x_{m+1}, \dots, x_n)$ is nonzero for every $i$. Then in $k[W]$, each $x_i$ for $i = 1, \dots, m$ depends algebraically on $x_{m+1}, \dots, x_n$. This shows that the transcendence degree of $k[W]$, and therefore the dimension of $W$ is at most $n - m$. Combined with item 1 in the theorem, it shows that $W$ has dimension exactly $n - m$. Finally note that if the $A_i$ are nonzero after specialization, then the above relationships fold for every component of the fibre, whence the fibre itself has dimension $W$.

We therefore just have to prove that the set of points $\mathbf{b}$ in $Y$ such that the polynomials $A_i(x_i, \phi_1(\mathbf{b}), \dots, \phi_m(\mathbf{b}), x_{m+1}, \dots, x_n)$ are nonzero form an open set. For this, we look at $A_i$ as a polynomial in $x_i, \dots, x_{m+1}, \dots, x_n$ with coefficients in $\phi_i$, and let $A_i'$ be the highest degree coefficients. It suffices that $A_i'$, which is now a polynomial in $k[Y]$, is nonzero for every $i$ for the above to hold. Therefore, we can pick $U$ to be the complement of the unions of the zerosets of $A_i'$, which is an open set. □

### Degree

We now define the degree of a variety. This is a far more involved notion than the dimension, and therefore we just state the definition and theorem we require. The following definition is from [Hei83]. For an irreducible affine variety $X \subseteq \mathbb{A}^n$ of dimension $r$ we define its degree to be the supremum of $|X \cap H|$, where $H$ is an affine subspace of dimension $n - r$ such that $\dim X \cap H = 0$. It holds that a general linear subspace attains this supremum. That the supremum is finite follows from the facts that it is a variety and hence has an irreducible decomposition. We use $\deg X$ to denote the degree of $X$.

Suppose $Y$ is an arbitrary variety, with irreducible decomposition $\cup Y_i$ with no $Y_i$ containing another. Then we define the degree of $Y$ to be the sum of the degrees of $Y_i$. It no longer holds that this is cardinality of the intersection of $Y$ with a general affine subspace of dimension $n - \dim Y$, unless every $Y_i$ has the same dimension as $Y$.

We note that the above definition of the degree of a variety is different from that in the algebraic geometry literature. In the latter, they extend the first definition to every variety, and as noted above, these match only if every component of the variety has the same dimension.

[3] The remark about dominant maps is clear, and the same proof works, only changing the last sentence.

The degree of a hypersurfaces matches the degree of the polynomial that defines it. We also have the following theorem that controls the degree of the intersection of two varieties. We refer to it as Bézout's Theorem.

**Theorem 2.3.5** (Bézout's Theorem, [Hei83, Theorem 1])**.** *Suppose* $X$ *and* $Y$ *are varieties. Then* $\deg X \cap Y \leqslant \deg X \times \deg Y$.

Finally, the degree of the projective closure of a variety is the same as the degree of the original variety.

# Chapter 3

# Hyperplane intersection and Noether Normalization

Given a variety of dimension $r$, intersecting it with a hyperplane chosen *randomly* reduces the dimension by 1. The set of hyperplanes that do not have this property form a subvariety in the space of all hyperplanes. In order to apply this result, we have to get bounds on the bad set of hyperplanes. We do this in the first part of this chapter. In the second part, we discuss Noether normalization, which is a fundamental result from commutative algebra and algebraic geometry. We will state some basic results about finite maps that will be useful in later chapters. We then use results from the first part to get bounds on the projections that are not Noether normalizing. All of the results presented in this chapter are folklore.

## 3.1 Hyperplane intersection

We prove the result for both projective and affine varieties. In the case of projective varieties, the intersection theorem guarantees that the intersection of a hyperplane with a variety of dimension at least 1 is nonempty. Proving then that a random hyperplane reduces the dimension by exactly one reduces only to proving that the dimension does not remain the same. When dealing with affine varieties, a new complication arises. The intersection with a hyperplane might be empty: for example consider the intersection of two parallel hyperplanes whose defining equations have different constant terms. We must therefore also bound the probability of this event happening. We will first prove the projective case, and use it to prove the affine case.

We formally state the first lemma that we will prove.

**Lemma 3.1.1.** *Let* $V \subseteq \mathbb{P}^n$ *be a projective variety of degree* $D$*. Let* $S$ *be a subset of the field* $k$ *that does not contain* $0$*. Let* $h = \sum_{i=0}^{n} c_i x_i$ *be a linear equation, with each coefficient* $c_i$ *picked independently and independently from the subset* $S$*, and let* $H$ *be* *the variety* $V(h)$*. Then with probability at least* $1 - D/|S|$ *we have that* $\dim V \cap H = \dim V - 1$.

*Proof of Lemma 3.1.1.* Let $V = \cup_{i=1}^{d}$ be the decomposition of $V$ into irreducible components. Since $\deg V = \sum_{i=1}^{d} \deg V_i$ and $\deg V_i \geqslant 1$ for each $i$, we have $d \leqslant D$. Also pick a point $p_i$ from each component $V_i$.

By the intersection theorem (Theorem 2.3.1), for a fixed irreducible component $V_j$, the intersection $V_j \cap H$ has dimension $V_j - 1$ unless $V_j \subseteq H$ in which case $\dim V_j \cap H = \dim V_j$. The event $V_j \subseteq H$ implies in particular that $p_j \in H$. The probability that $p_j \notin H$ is at least $1/|S|$. To see this, suppose $(p_j)_{j'}$ is the last nonzero coordinate of $p_j$. Then for any setting of all the $c_i$ other than $i = j'$, there is at most one value of $c_{j'}$ that makes $\sum c_i(p_j)_i = 0$. Therefore, the probability that $V_j \subseteq H$ is bounded above by $1/|S|$.

To complete the proof, we use the union bound. The condition that for every $i$ the intersection $V_i \cap H$ has dimension one less than that of $V_i$ guarantees that the intersection $V \cap H$ has dimension one less than that of $V$. By the union bound, the probability that for some $i$ we have $\dim V_i \cap H = \dim V_i$ is at most $d/|S|$, whence with probability alt east $1 - d/|S|$ we have that $\dim V \cap H = \dim V - 1$. The proof is completed by using the initial observation that $k \leqslant D$. □

In most of our applications, the degree $D$ will be at most single exponential in the input size. We can therefore sample from subsets of size $\mathcal{O}(D)$ in polynomial time, and still guarantee that the intersections behave as expected with high probability.

We now state and prove the affine case.

**Lemma 3.1.2.** *Let* $V \subseteq \mathbb{A}^n$ *be an affine variety of degree* $D$*. Let* $S$ *be a subset of the field* $k$ *that does not contain* $0$*. Let* $h = c_0 + \sum_{i=1}^{n} c_i$ *be a linear equation, with each coefficient* $c_i$ *picked uniformly and independently from the subset* $S$*, and let* $H$ *be the variety* $V(h)$*. Then with probability at least* $1 - 2D/|S|$ *we have that* $\dim V \cap H = \dim V - 1$.

As stated before, the difficulty arises in ensuring that the intersection is nonempty. In order to do this, we consider the projective closures of the varieties involved, and bound the probability that the intersection occurs in the hyperplane at infinity.

*Proof of Lemma 3.1.2.* Let $V^p$ be the projective closure of $V$, and $H^p$ be the projective closure of $H$. The variety $H^p$ is defined by the equation $\sum_{i=1}^{n} c_i x_i = 0$. We have $\deg V = \deg V^p$ and $\dim V = \dim V^p$.

By the intersection theorem (Theorem 2.3.1), the intersection $V^p \cap H^p$ is non empty (unless $\dim V = 0$). Therefore, $V \cap H = \emptyset$ implies that the intersection $V^p \cap H^p$ is contained in the hyperplane at infinity $\mathbb{P}_n^\infty$. In order for the event $\dim V \cap H = \dim V - 1$ to hold, it is therefore sufficient that the following two conditions hold:

- $\dim V^p \cap H^p = \dim V^p - 1$.

- $V^p \cap H^p \not\subseteq \mathbb{P}^n_\infty$ unless $V^p \cap H^p = \emptyset$.

All of the coefficients $c_0, \ldots, c_n$ were chosen randomly from $S$, and we are therefore in the setting of Lemma 3.1.1. By applying the lemma to $V^p$ and $H^p$, we get that the first condition holds with probability at least $1 - D/|S|$.

We now prove that if the first condition holds, then the second condition also holds with probability $1 - D/|S|$. Since no component of $V^p$ is contained in $\mathbb{P}^n_\infty$, the variety $V^p \cap \mathbb{P}^n_\infty$ has dimension $\dim V^p - 1$. By Bézout's theorem, $\deg V^p \cap \mathbb{P}^n_\infty \leqslant D$. We can therefore apply Lemma 3.1.1 to the variety $V^p \cap \mathbb{P}^\infty_n$ and $H^p$ to get that with probability at least $1-D/|S|$ the intersection $V^p \cap \mathbb{P}^n_\infty \cap H^p$ has dimension $\dim V^p \cap \mathbb{P}^n_\infty - 1 = \dim V^p - 2$. If this is the case then $V^p \cap H^p$ cannot be a subset of $\mathbb{P}^n_\infty$, since if it were, then we would have $V^p \cap H^p \cap \mathbb{P}^n_\infty = V^p \cap H^p$, and the latter has dimension $\dim V - 1$. Therefore, with probability at least $1 - D/|S|$ we have that $V^p \cap H^p \not\subseteq \mathbb{P}^n_\infty$.

A union bound on the two conditions completes the proof of the lemma. □

In our applications, we will frequently use the above lemma iteratively to reduce the dimension of a variety to 0. It is clear from the above lemmas that if a variety $V$ has dimension $r$, then intersecting $V$ with $r$ many random linear hyperplanes will achieve this with high probability. We will however sometimes require that the intersecting hyperplanes have some structure. In particular, we will require that only the first hyperplane depends on $x_1$, only the first two depend on $x_2$, and so on. In the following lemmas we prove that the intersection will still behave as expected.

**Lemma 3.1.3.** *Let $V \subseteq \mathbb{P}^n$ be a projective variety of dimension $r$ and degree $D$. Let $S$ be a subset of $k$ that does not contain $0$. Let $h$ be a linear form that depends only on $n - r + 1$ variables, and let $H$ be the hyperplane it defines. If the coefficients of each variable is picked uniformly and independently from $S$ then with probability at least $1 - D/|S|$ we have* $\dim V \cap H = \dim V - 1$.

*Let $W \subseteq \mathbb{A}^n$ is an affine variety of dimension $r \geqslant 1$ and degree $D$. Let $\ell$ be a linear equation that depends on $n - r + 1$ variables, and let $L$ be the hyperplane it defines. If the coefficients of each variable is picked uniformly and independently from $S$ then with probability at least $1 - 2D/|S|$ we have* $\dim W \cap L = \dim W - 1$. *If $r = 0$, then $\ell$ must depend on all the variables, and have a constant term.*

The following examples show that this is essentially tight, that we cannot always pick a hyperplane whose equation has fewer nonzero coefficients. Suppose $V$ is the projective variety in $n$ dimensions defined by $x_0 = x_1 = x_2 = 0$. This variety has dimension $n - 3$. Any hyperplane of the form $a_0x_0 + a_1x_1 + a_2x_2 = 0$ always contains $V$, irrespective of the $a_i$. Therefore, such a hyperplane can never property intersect $V$. We emphasize that the above is a worst case statement. If instead we picked an equation of the form $a_2x_2 + a_3x_3 + a_4x_4 = 0$ then it is possible to get proper intersection. In the affine case, we must pick $n - r + 1$ coefficients, not including the constant term. This ensures that the intersection on the hyperplane at infinity is proper, which was essential in the proof of Lemma 3.1.2. For example, consider the variety $W$ defined by $x_1 = \cdots = x_{n-2} = 0$. This has dimension 2. If we define $\ell = b_0 + \sum_{i=1}^{n-2} b_i x_i$, then the intersection of $W$ and $L$ is empty (if $b_0 \neq 0$) or not proper (if $b_0 = 0$). For the statement about $r = 0$, assume that $W = \mathbf{0}$. Then if $\ell$ does not have a constant term, the intersection will never be proper. We now prove the lemma.

*Proof of Lemma 3.1.3.* This proof is similar to the proofs of Lemma 3.1.1 and Lemma 3.1.2, and we only focus on the differences here. We first prove the projective case. Assume without loss of generality that $x_0, x_1, \ldots, x_{n-r}$ are the $n-r+1$ variables that $h$ depends on, that is, $h = \sum_{i=0}^{n-r} c_i x_i$. Let $V = \cup_{i=1}^{d} V_i$ be the irreducible decomposition. In the proof of Lemma 3.1.1, the hyperplane properly intersected every $V_i$, which ensured that the dimension of $V$ drops by 1. It is sufficient however that the hyperplane property intersect all of the components of $V$ that have dimension $r$. If this happens, even if the intersection with a lower dimension component is non proper, the intersection with $V$ will still be. We can therefore drop all of the lower dimensional components from the above union, and assume that $V_1, \ldots, V_k$ are the irreducible components of $V$ of dimension $r$. The bound $d \leqslant D$ clearly still holds.

We now want to pick a point $p_i$ from each component. It might happen however that the first $n-r+1$ coordinates of $p_i$ are all zero. If this is the case, then the previous proof fails, since $h(p_i) = 0$ irrespective of the coefficients. The proof of Lemma 3.1.1 fails since it involves picking the coefficient $c_{j'}$ corresponding to a nonzero coordinate of $p_i$. Therefore, we want to ensure that we pick $p_i$ such that not all of the first $n-r+1$ coordinates are zero. Fix a component $V_j$, and suppose that every $p_j \in V_j$ has all of the first $n-r+1$ coordinates zero. Then $V_j$ would be contained in the subspace defined by $x_0 = x_1 = \cdots = x_{n-r} = 0$. But this subspace has dimension $r-1$, whence $V_j$ would have dimension at most $r-1$. This contradicts the assumption that $V_j$ is a component of dimension $r$.

Once such $p_i$ are picked from each component, the same arguments as in the proof of Lemma 3.1.1 work, and we get the required result.

For the affine case, we have two sufficient conditions for the intersection to be proper. The first condition is that $L^p$ intersects $W^p$ properly. The second condition is that $L^p \cap \mathbb{P}^n_\infty$ intersects $W^p \cap \mathbb{P}^n_\infty$ properly. In the proof of Lemma 3.1.2, these were ensured by invoking Lemma 3.1.1 twice. In order to prove the affine case of this lemma, we follow the same proof, and replace invocations of Lemma 3.1.1 with the projective version of this lemma instead. All we have to show is that the assumptions are satisfied.

The defining equation of $L^p$ is the homogenization of $\ell$, and the equation of $L^p \cap \mathbb{P}^n_\infty$ is the degree 1 part of $\ell$. For the first condition we are intersecting $W^p$ which has dimension $r$, and hence we require the homogenization of $\ell$ to depend on $n+r-1$ variables. For the second condition we are intersecting with $W^p \cap \mathbb{P}^n_\infty$. The underlying space here is $\mathbb{P}^{n-1}_\infty$, which has dimension $n-1$. We therefore require the degree 1 part of $\ell$ to depend on $(n-1)-(r-1)+1 = n-r+1$ variables. Both these requirements are satisfied by the assumption. This completes the proof. □

We now give a simple corollary of the above lemma. The corollary basically states that we can repeatedly apply the above lemma $r$ times to a variety of dimension $r$ to obtain a variety of dimension $0$. The corollary also states that $r+1$ intersections results in the empty variety. This last statement can alternatively be seen as the fact that a random linear space of dimension $n-r-1$ avoids a variety of dimension $r$. While the proof of this corollary is obvious, we state it here so we can easily invoke the result later on, instead of having to invoke one of the previous lemmas multiple times inductively.

**Corollary 3.1.4.** *Let $V$ be a projective variety of dimension $r$ and degree $D$. Let $S$ be a subset of $k$ not containing $0$. Let $h_1, \dots, h_{r+1}$ be linear forms such that $h_i$ depends on $n+2-i$ variables, and each coefficient is picked uniformly and independently from $S$. Let $H_i$ be the hyperplane defined by $h_i$. Then the intersections $V \cap H_2 \cap \cdots \cap H_{r+1}$ has dimension $0$ with probability at least $1 - rD/|S|$. Further, the intersections $V \cap H_1 \cap \cdots \cap H_{r+1}$ is empty with probability at least $1-(r+1)D/|S|$.*

*Let $W$ be an affine variety of dimension $r$ and degree $D$. Let $\ell_1, \dots, \ell_{r+1}$ be linear forms such that $\ell_i$ depends on $n+2-i$ variables (except $\ell_1$, which depends on $n$ variables and has a constant term), and each coefficient is picked uniformly and independently from $S$. Let $L_i$ be the hyperplane defined by $\ell_i$. Then the intersections $W \cap L_2 \cap \cdots \cap L_{r+1}$ has dimension $0$ with probability at least $1-2rD/|S|$. Further, the intersections $W \cap L_1 \cap \cdots \cap L_{r+1}$ is empty with probability at least $1-2(r+1)D/|S|$.*

*Proof of Corollary 3.1.4.* The proof of both the projective and affine versions essentially follows from the repeated application of Lemma 3.1.3. In order to ensure the assumptions, we first intersect $V$ (resp. $W$) with $H_{r+1}$ (resp. $L_{r+1}$), then $H_r$ (resp. $L_r$) and so on. In both cases, we use Bézout's theorem after each intersection to guarantee that the variety obtained after intersecting with $H_i$ has degree at most $D$. The lower bound on the probabilities are obtained by a union bound on the failure of each of the intersections. □

Throughout this section, our model for random affine subspaces was to pick defining equations uniformly and independently, and considering their zerosets. In some places, we will have to consider a slightly different model. Suppose we have a map from $\mathbb{A}^{n-r}$ to $\mathbb{A}^n$ with linear coordinate functions. The image of this map is a linear subspace of $\mathbb{A}^n$ of dimension at most $n-r$. If the coordinate functions are picked randomly, then the image is a random subspace. We show that given a variety of dimension $r$ and degree $D$, with high probability the image of such a map will properly intersect $V$. We will require the statement only for the case of affine varieties, but we prove it both for the projective case and the affine case. As was the case before, we will use the former to prove the latter.

**Lemma 3.1.5.** *Suppose $V$ is a projective variety of degree $D$ and dimension $r$. Suppose $\psi_0, \dots, \psi_n$ are linear homogeneous polynomials in $z_0, z_1, \dots, z_m$ with each coefficient picked uniformly and independently from a subset $S$ of $k$. Let $\psi$ be the linear map from $\mathbb{P}^m$ to $\mathbb{P}^n$ with coordinate functions $\psi_i$, and let $H$ be its image. Then with probability at least $1 - n^3D/|S|$ we have that the above map is well defined, $\dim H = n-r$, and $\dim H \cap V = 0$.*

*Proof of Lemma 3.1.5.* Suppose each $\psi_i$ is of the form $\sum_{j=0}^{n-r} a_{i,j} z_j$. That $\psi$ is well defined requires this map to be injective (we cannot have anything other than $0$ mapping to $0$, since this point is not part of $\mathbb{P}^n$). This requires that the matrix $A_{ij} = a_{i,j}$ has full rank. Fixing any submatrix of size $(n-r) \times (n-r)$, its determinant is a polynomial in $a_{ij}$ of degree $n-r$. By the polynomial identity lemma, with probability at least $1-(n-r)/|S|$ this determinant is nonzero, the map is well defined and $\dim H = n - r$.

Now suppose we actually pick $\psi_0, \ldots, \psi_n$ such that each of them is a homogeneous equation in $n+1$ variables $z_0, \ldots, z_n$. Let $\Psi_m$ for $m = n-r, \ldots, n$ denote the restriction of the map $\psi : \mathbb{P}^n \to \mathbb{P}^n$ to the space defined by $z_{m+1} = \cdots = z_n = 0$. By definition the original map that we started with is $\Psi_{n-r}$, after identifying $\mathbb{P}^m$ with the subspace $z_{n-r+1} = \cdots = z_n = 0$. Let $L_m$ denote the image of $\Psi_m$. By the same argument as above, each $L_m$ has dimension $m$ with probability at least $1 - m/|S|$. We have $L_n = \mathbb{P}_n$, and therefore $L_n \cap V = V$, and $\dim L_n \cap V = n$. We will now show that $\dim L_{m-1} \cap V = \dim L_m \cap V - 1$ for $m = n-r+1, n-r+2, \ldots, n$ with high probability. This, combined with the above statement and a union bound will give us our desired result.

Let $W := L_m \cap V$. By Bézout's theorem we have $\deg W \leqslant D$. Let $W = \cup_{i=1}^{d} W_i$ be the irreducible decomposition of $W$, and let $p_i$ be a point in $W_i$. Each $p_i$ lies in $L_m$. The subspace $L_{m-1}$ is a linear subspace of $L_m$. If we can show that $p_i$ is not in $L_{m-1}$ then the hyperplane defined by $L_{m-1}$ properly intersects $W_i$, using the same arguments as in the proof of Lemma 3.1.1. The condition that $p_i$ is not in $L_{m-1}$ is equivalent to the condition that the coordinate vector of point $p_i$ depends linearly on the first $m$ columns of $A$. By considering a minor and applying the polynomial identity lemma, this happens with probability at most $(n-m+1)/|S|$. The probability that it does not happen for any $p_i$ is at least $1-(n-m)D/|S|$ by the union bound. If this happens, we have $\dim L_{m-1} \cap V = \dim L_m \cap V - 1$ as required.

We can now take a union bound over the above events for all $m$. With probability at least $1 - n^3 D/|S|$ therefore we have $\dim H \cap V = 0$. [1] □

We can now prove a similar statement for affine varieties.

**Lemma 3.1.6.** *Suppose* $W$ *is an affine variety of degree* $D$ *and dimension* $r$*. Suppose* $\psi_1, \ldots, \psi_n$ *are linear polynomials in* $z_1, \ldots, z_{n-r}$ *with each coefficient picked uniformly and independently from a subset* $S$ *of* $k$*. Let* $\psi$ *be the linear map from* $\mathbb{A}^m$ *to* $\mathbb{A}^n$ *with coordinate functions* $\psi_i$*, and let* $H$ *be its image. Then with probability at least* $1 - 2n^3 D/|S|$ *we have* $\dim H = n - r$ *and* $\dim H \cap W = 0$.

*Proof of Lemma 3.1.6.* Suppose each $\psi_i$ is of the form $a_{i0} + \sum_{j=1}^{n-r} a_{ij} z_j$. We use $\psi_i^h$ to denote the homogenization of $\psi_i$ using the variable $z_0$. Let $W^p$ be the projective closure of $W$. We now invoke the projective version of the theorem on $W^P$, using $\phi_0 = z_0$. We continue to use $\phi_0 = z_0$ even in the step when we assume that we had random polynomials in $n+1$ variables. In the projective case, our steps required that certain matrices were full rank. In the special case

[1] The $n^3$ is just a lazy estimate.

with $\phi_0 = z_0$, the first row of all of the matrices will be of the form $1, 0, \dots, 0$ follows by the first coordinate of $p_i$. In any case, this specialization does not force the matrices to have lower rank, and therefore the proofs go through. With probability at least $1 - n^3D/|S|$ therefore the image of the homogeneous version of $\phi$ intersects $W^p$ properly.

We finally have to show that there is a point in the intersection in $\mathbb{A}^n$, so that the intersections of the affine variety and linear subspace is also of dimension 0. This is similar to the proof of Lemma 3.1.2 The variety $W^p \cap \mathbb{P}n$ has dimension $r - 1$. The restriction of the image of the homogenized version of $\psi$ to the hyperplane at infinity is given by the degree 1 part of the polynomials $\psi_1, \dots, \psi_n$. This image has dimension $n - r - 1$. We want to invoke the projective version of this lemma on this image and $W^p$. In this case, the sum of the dimension of the linear space and variety is 1 less than that of the ambient space, and therefore by repeating the induction step in the proof on additional time, we obtain that the two varieties have intersection of dimension $-1$ with high probability. The bound on the probability of the bad event is $n^3D/|S|$ as before. As in the proof of Lemma 3.1.2 this is a sufficient condition for what we require.

The final result holds by taking a union bound over the above two results.

□

## 3.2 Noether Normalization

We now discuss a fundamental theorem of commutative algebra called the Noether normalization lemma. We first state and prove a version of the theorem. We then discuss its applications in algebraic geometry. We use [SR13] and [Gat13] as references for this section.

### 3.2.1 The Noether Normalization lemma

We first state the main lemma. The following statement is from [Gat13].

**Lemma 3.2.1.** *Let $R$ be a finitely generated $k$-algebra, with generators $x_1, \dots, x_n$. There is an injective map $k[z_1, \dots, z_r] \to R$ with indeterminates $z_i$ that make $R$ into a finite extension of $k[z_1, \dots, z_r]$. Further, if $k$ is infinite, then the images of $z_i$ can be chosen to be linear combinations of the generators $x_i$.*

The lemma essentially states that any finitely generated ring extension $R$ is an integral extension of a polynomial ring. We do not prove here that every finite extension is integral, a proof can be found in [AM94, Chapter 5]. We also use the fact that a finite extension of a finite extension is itself finite. To prove the above, we use an auxiliary lemma which proves that a multivariate polynomial can be made monic with an appropriate shift.

**Lemma 3.2.2.** *Let $f$ be a nonzero polynomial in $k[x_1, \dots, x_n]$ where $k$ is an infinite field. Then there exist $a_1, \dots, a_{n-1} \in k$ and $\lambda \in k$ such that $g(\mathbf{y}) := \lambda f(y_1 + a_1y_n, \dots, y_{n-1} + a_{n-1}y_n, y_n)$ is monic in $y_n$.*

*Proof of Lemma 3.2.2.* Suppose $f$ has degree $d$. The coefficient of $y_n^d$ in the polynomial $g(\mathbf{y})$ is $\lambda f_d(a_1, \ldots, a_{n-1}, 1)$, where $f_d$ is the degree $d$ part of $f$. We need to pick $\mathbf{a}$ such that $f_d(\mathbf{a}, 1)$ is nonzero. If we can do this, then we can pick $\lambda = f_d(\mathbf{a}, 1)^{-1}$ and complete the proof.

That such a $\mathbf{a}$ exists follows by induction. Write $f_d = \sum_{i=1}^{d} x_1^i g_i$. The polynomial $f_d$ is nonzero, so some $g_i$ is nonzero. The polynomial $g_i$ is homogeneous of degree $d - i$ in $n - 1$ variables, and so pick can pick $a_2, \cdots, a_{n-1}$ by induction so that $g_i(a_2, \ldots, a_n, 1)$ is nonzero. Then $f_d(x_1, a_2, \ldots, a_{n-1}, 1)$ is a univariate and has only finitely many roots, and we can pick $a_1$ to avoid any such roots. This last step requires $k$ to be infinite. □

We now use Lemma 3.2.2 to prove Lemma 3.2.1.

*Proof of Lemma 3.2.1.* We induct on $n$, the number of generators of $R$ as a $k$-algebra. When $n = 0$, the statement vacuously holds. Let $n$ be greater than 0. We consider two cases. Suppose $x_1, \ldots, x_n$ are algebraically independent. Then $k[x_1, \ldots, x_n]$ is isomorphic to $k[z_1, \ldots, z_n]$, with the map sending $z_i$ to $x_i$. The lemma is clearly true in this case.

Assume now that $\mathbf{x}$ are not algebraically independent, and that $f$ is a polynomial such that $f(\mathbf{x}) = 0$. Let $a_1, \ldots, a_{n-1}$ and $\lambda$ be as in Lemma 3.2.2. Set $y_i = x_i - a_i x_n$ for $i < n$ and set $y_n = x_n$. The $\mathbf{y}$ form a generating set for $R$, since $x_n = y_n$ and $x_i = y_i + a_i y_n$ for $i < n$. Further, $R$ is an integral extension of $k[y_1, \ldots, y_{n-1}]$, since $\lambda f(y_1 + a_1 y_n, \ldots, y_{n-1} + a_{n-1} y_n, y_n)$ is a monic equation for $y_n$. By induction, $R[y_1, \ldots, y_{n-1}]$ can be written as a finite extension of a polynomial ring. Therefore $R$ is a finite extension of a finite extension of a polynomial ring, and is hence itself is a finite extension of a polynomial ring. The statement that the map sends each $z_i$ to a linear combination of the $x_i$ follows from the way we constructed the $y_i$. □

In a later chapter, we will use the above result to prove the Nullstellensatz. We now make an important observation. The $\mathbf{a}$ chosen in the proof are such that $(\mathbf{a}, 1)$ is not a root of a homogeneous polynomial. A sufficiently random choice of $\mathbf{a}$ satisfies this property. Therefore, if we map each $z_i$ to a sufficiently random linear combination of the $\mathbf{x}$, the induced ring extension will still be integral.

In the next subsection, we discuss the notion of finite maps.

### 3.2.2 Finite maps

Consider a dense map $\phi : X \to Y$ between affine varieties. Since $\phi$ is dense, the map $\phi^* : k[Y] \to k[X]$ is an injection. We identify $k[Y]$ with its isomorphic copy in $k[X]$ via $\phi^*$. The map $\phi$ is called *finite* if $k[X]$ is an integral extension of $k[Y]$.

Finite maps have a number of useful properties. Before we list these, we state a consequence of the Noether normalization lemma. Suppose $X \subseteq \mathbb{A}^n$ is an irreducible affine variety. The coordinate ring of $X$ is generated by $x_1, \ldots, x_n$. If we apply Lemma 3.2.1 to $k[X]$, we get a map $\psi : k[z_1, \ldots, z_r] \to k[X]$ such that each $z_i$ is mapped to a linear combination of the $x_i$, and that $k[X]$ is an integral extension of the image of $k[z]$. The variety corresponding to $k[z]$ is $\mathbb{A}^r$. The map $\psi$ induces a map $\psi^* : X \to \mathbb{A}^r$. The map $\psi^*$ is a finite map by definition. Further, since $k[X]$ is an integral extension of $k[z]$, the field extension $k(X)/k(z)$ is algebraic. [2] In particular this means that that $\operatorname{trdeg}(k(X)) = \operatorname{trdeg}(k(z)) = r$. The $r$ obtained is therefore the same as the dimension of the variety. In fact, the above discussion provides some motivation for the definition of the dimension of the variety. We now state the algebraic version of the Noether normalization lemma.

**Lemma 3.2.3.** *Given an irreducible affine variety $X \subseteq \mathbb{A}^n$, there exists a finite map $\phi : X \to \mathbb{A}^r$ where $r$ is the dimension of $X$. Further, $\phi$ is the composition of the inclusion map and a linear map.*

Further, as discussed after the proof of Lemma 3.2.1, a random linear combination satisfies the above condition. We will soon make this more precise. Before that, we state some properties of finite maps.

**Lemma 3.2.4.** *Suppose $f : X \to Y$ is a finite map between affine varieties. Then $f$ is surjective, and every point $y \in Y$ has finite fibres.*

We only prove the second part of the lemma. The proof of the first part is slightly more involved (it is essentially an application of the Nullstellensatz followed by an application of Nakayama's lemma) and can be found in [SR13, Chapter 1, Section 5, Theorem 1.12].

*Proof of part of Lemma 3.2.4.* Each $x_i \in k[X]$ satisfies some monic equation $f_i$ with coefficients in $k[Y]$. For a fixed $b \in Y$, the set $f^{-1}(b)$ is defined by the equations $y_i = b_i$, where $y_i$ are the coordinate functions. On the set $f^{-1}(b)$ therefore each $x_i$ satisfies a specialization of the the equation $f_i$ with $y_i = b_i$. These equations only have finitely many roots, and therefore each $x_i$ can only take finitely many values on $f^{-1}(b)$, proving that the latter set is finite. □

Another property that we state here without proof is that finiteness is local property. This means that if $f : X \to Y$ is a map between affine varieties, and every point $y \in Y$ has an affine neighbourhood $V$ such that $f^{-1}(V)$ is affine and the restricted map $f : f^{-1}(V) \to V$ is finite, then $f$ itself is finite. This makes it natural to extend the definition of finite maps. If $f : X \to Y$ is an arbitrary map between quasiprojective varieties, we say that $f$ is finite it every point $y \in Y$ has an affine neighbourhood $V$ with affine preimage $f^{-1}(V)$ such that the map $f : f^{-1}(V) \to V$ is finite. We therefore have a notion of finite maps between projective varieties. Further, Lemma 3.2.3 extends to projective varieties too, although we not prove it here.

We now return to the statement that a random map is Noether normalizing. First we define the notion of a projection with a centre. Suppose $H$ is a linear

[2] A short proof of this fact. The field $k(X)$ is generated by all $a/1, 1/a$ for $a \in k[X]$. Suppose $a$ satisfies a monic equation $f$ with coefficients in $k[z]$. Then $a/1$ is algebraic over $k(z)$ because of $f$, and $1/a$ is algebraic because of $x^{\deg f} f(1/x)$. Finally, since all the generators are algebraic, so is the extension.

subspace of $\mathbb{P}^n$ of dimension $r$ defined by equations $h_1 = \cdots = h_{n-r} = 0$. The projection with center $H$ is the map with coordinate functions $(h_1 : \cdots : h_{n-r})$. If $X$ is a projective variety contained in $\mathbb{P}^n \setminus X$ then the projection with center $H$ is a well defined map from $X$ to $\mathbb{P}^{n-d-1}$. The following theorem gives a characterization of maps that are Noether normalizing for projective varieties. It is a direct consequence of [SR13, Chapter 1, Section 5, Theorem 1.15]. We do not prove it here.

**Theorem 3.2.5.** *Let $X \subseteq \mathbb{P}^n$ be a projective variety of dimension $d$, and let $H$ be a linear subspace of dimension $n-d-1$ which avoids $X$. Then the projection with center $H$ is a Noether normalizing map for $X$.*

The above theorem along with Lemma 3.1.3 lets us formalize the statement that a random linear map is Noether normalizing. The statement for projective varieties is immediate, while that for affine varieties requires an argument involving projective closures.

**Theorem 3.2.6.** *Let $V \subseteq \mathbb{P}^n$ be a projective variety of dimension $r$ and degree $D$. Let $h_1, \ldots, h_{r+1}$ be linear forms such that $h_i$ depends on $n+2-i$ variables, and each coefficient of $h_i$ is picked uniformly and independently from a subset $S$ of $k$ that does not contain $0$. Then with probability at least $1-(r+1)D/|S|$, the map with coordinate functions $h_i$ is Noether normalizing.*

*Let $W \subseteq \mathbb{A}^n$ be an affine variety of dimension $r$ and degree $D$. Let $\ell_1, \ldots, \ell_r$ be linear polynomials such that $\ell_i$ depends on $n+1-i$ variables, and each coefficient of $\ell_i$ is picked uniformly and independently from a subset $S$ of $k$ that does not contain $0$. The $\ell_i$ can also be linear forms (without constants). Then with probability at least $1-(r+1)D/|S|$, the map with coordinate functions $\ell_i$ is Noether normalizing.*

*Proof of Theorem 3.2.6.* By Corollary 3.1.4 with probability at least $1-(r+1)D/|S|$, the subspace defined by the equations $h_i$ avoids $V$. This also automatically ensures that the dimension of the subspace is $n-r-1$, since if it was more than this, it could not avoid $V$. The projection from this subspace is given by the map with coordinate functions $h_i$, and this map is Noether normalizing for $V$ by Theorem 3.2.5.

We now show the affine statement. We consider $W^p$, the projective closure of $W$, and find a normalizing map for $W^p$. Not every normalizing map for $W^p$ will give us a normalizing map for $W$. In order to ensure we get a normalizing map for $W$, we pick $\ell_0 = x_0$, and use this as the first coordinate for our Normalizing map for $W^p$. This ensures that the $\mathbb{P}^n_\infty$ is mapped to $\mathbb{P}^r_\infty$, and that the affine chart containing $W$ is mapped to the affine space $\mathbb{P}^r \setminus \mathbb{P}^r_\infty$. Since no component of $W^p$ is contained in $\mathbb{P}^n_\infty$, we have $\dim W^p \cap \mathbb{P}^n_\infty = r-1$. By Lemma 3.1.3 applied to $W^p \cap \mathbb{P}^n_\infty$, the projective closure of the subspace defined by $\ell_1, \ldots, \ell_r$ avoids $W^p \cap \mathbb{P}^n_\infty$. The subspace defined by $\ell_0, \ldots, \ell_r$ then avoids $W^p$, and the projection from this subspace is Noether normalizing. The map $\mathbb{A}^n \to \mathbb{A}^r$ with coordinate functions $\ell_1, \ldots, \ell_r$ is therefore Noether normalizing for $W$. □

# Chapter 4

# The Nullstellensatz

In this chapter, we discuss Hilbert's Nullstellensatz (aka the zero-locus-theorem). This is a foundational result that establishes a fundamental relationship between geometry and algebra. The Nullstellensatz (more precisely one of its many forms) states that a set of polynomials do not have a common zero if and only if the ideal they generate is the trivial ideal. In other words, the Nullstellensatz proves the existence of a natural certificate that a given set of polynomials do not have a common zero.

In the first part of the chapter, we state and prove the original non constructive formulation(s) of the Nullstellensatz. In the second part of the chapter, we discuss the ideal membership problem. In the third part of the chapter, we discuss the effective Nullstellensatz. This includes a proof of the effective Nullstellensatz, and a brief literature survey of related results.

## 4.1 The non-constructive Nullstellensatz

We use [CLO07], [Vak17], and [Gat13] as references for this section.

The Nullstellensatz has a number of different (and mostly equivalent formulations). The following two are the most common ones. Despite their names, they are equivalent.

**Theorem 4.1.1** (The Weak Nullstellensatz)**.** *Let $k$ be an algebraically closed field, and let $I$ be a nontrivial ideal of $k[x_1, \dots, x_n]$. Then $\mathbf{V}(I) \neq \emptyset$.*

**Theorem 4.1.2** (The Strong Nullstellensatz)**.** *Let $k$ be an algebraically closed field, and let $I$ be a nontrivial ideal of $k[x_1, \dots, x_n]$. Then $\mathbf{I}(\mathbf{V}(I)) = \sqrt{I}$.*

If $I = k[\mathbf{x}]$, then $\mathbf{V}(I) = \emptyset$ by definition, since the constant polynomial 1 has no roots. The weak Nullstellensatz states that the converse holds too. The strong Nullstellensatz completes the ideal-variety correspondence. A discussion about this correspondence can be found in chapter 2. We will first prove the equivalence of the above two statements. We then state a third form of the

Nullstellensatz, which is slightly more general than the above two forms. We then prove the third form, and finish by showing how it implies the above.

*Equivalence of Theorem 4.1.1 and Theorem 4.1.2.* That the strong Nullstellensatz implies the weak Nullstellensatz is straightforward: Suppose $I$ is an ideal such that $\mathbf{V}(I) = \emptyset$. By the strong Nullstellensatz, $\sqrt{I} = \mathbf{I}(\mathbf{V}(I)) = \mathbf{I}(\emptyset) = k[\mathbf{x}]$. This implies that $1 \in \sqrt{I}$, which implies that $1 \in I$.

We now prove the other direction. Suppose first that $h^e \in I$. Then $h^e = 0$ at every point in $\mathbf{V}(I)$, and therefore $h = 0$ at every point in $\mathbf{V}(I)$. This implies that $h \in \mathbf{I}(\mathbf{V}(I))$, and hence $\sqrt{I} \subseteq \mathbf{I}(\mathbf{V}(I))$. For the reverse inclusion we use the Rabinowitsch trick ([Rab30]). Let $I = \langle f_1, \dots, f_m \rangle$, and let $g \in \mathbf{I}(\mathbf{V}(I))$. Consider the ring $k[\mathbf{x}, y]$, where $y$ is a new variable. Let $J = \langle f_1, \dots, f_m, 1 - gy \rangle$, where $f_i$ are considered as elements of $k[\mathbf{x}, y]$. We show that $\mathbf{V}(J) = \emptyset$. Suppose $(c_1, \dots, c_n, c_{n+1})$ is an element of $k^{n+1}$. If $(c_1, \dots, c_n) \in \mathbf{V}(I)$, then it is also a root of $g$, and therefore cannot be a root of $1 - yg$ for any value of $c_{n+1}$, and therefore $\mathbf{c} \notin \mathbf{V}(J)$. But if $(c_1, \dots, c_n) \notin \mathbf{V}(I)$, then $\mathbf{c} \notin \mathbf{V}(J)$ since every element of $I$ is also in $J$. Therefore, $\mathbf{V}(J) = \emptyset$, and by the Weak Nullstellensatz we get that $J = k[\mathbf{x}, y]$. We can write then $1 = h_0(1-yg) + \sum_{i=1}^{m} h_i f_i$ where $h_i \in k[\mathbf{x}, y]$. In the above equation, we substitute $y = 1/g$, and clear denominators on the right hand side. The term $h_0(1 - yg)$ vanishes after the substitution, and we get $1 = f/g^e$ for some $f \in I$. This gives us $g^e \in I$ and therefore $g \in \sqrt{I}$, proving that $\mathbf{I}(\mathbf{V}(I)) \subseteq \sqrt{I}$. □

A third statement, equivalent to the above is that every maximal ideal of $k[\mathbf{x}]$ is of the form $\langle x_1 - a_1, \dots, x_n - a_n \rangle$. We use $\mathfrak{m}_{\mathbf{a}}$ to denote ideals of the above form. To show this equivalence, we use the fact that $f \in \mathfrak{m}_{\mathbf{a}}$ if and only if $f(\mathbf{a}) = 0$. [1] If $g$ is any polynomial not in $\mathfrak{m}$, then $g(\mathbf{a}) \neq 0$. But we also have $g(\mathbf{a}) \in \mathfrak{m}_{\mathbf{a}} + \langle g \rangle$, whence $\mathfrak{m}_{\mathbf{a}} + \langle g \rangle = k[\mathbf{x}]$. Conversely, if $J$ is an arbitrary ideal of $k[\mathbf{x}]$, then by the weak Nullstellensatz, there is a common root $b_1, \dots, b_n$ of every polynomial in $J$. This implies that $J \subseteq \mathfrak{m}_{\mathbf{b}}$. If $J$ itself is maximal then $J = \mathfrak{m}_{\mathbf{b}}$, completing the proof of the equivalence.

We now state the final form of the Nullstellensatz. This form is slightly more general than the above, and is sometimes called Zariski's Lemma. We will use this latter name in order to distinguish it from the above.

**Lemma 4.1.3** (Zariski's Lemma)**.** *Suppose $K$ is a field, and $A$ is a finitely generated $K$-algebra that is also a field. Then $A$ is a finite extension of $K$.*

*Equivalently, if $K$ is a field and $\mathfrak{m}$ a maximal ideal of $K[\mathbf{x}]$, then $K/\mathfrak{m}$ is a finite extension of $K$.*

The above lemma does not require $K$ to be algebraically closed. The equivalence follows from the fact that any finitely generated $K$ algebra is the quotient of the polynomial ring $K[\mathbf{x}]$, and if the algebra is also a field then the kernel is a maximal ideal.

[1] This fact follows from the division algorithm applied one variable at a time.

We will prove Zariski's Lemma using the Noether Normalization Lemma. Before we do this, we show that Zariski's Lemma implies the Nullstellensatz in the above forms.

*Proof of Theorem 4.1.1 and Theorem 4.1.2 using Lemma 4.1.3.* Suppose $I$ is an ideal of $k[X]$, and $\mathfrak{m}$ is a maximal ideal containing $I$. By Zariski's Lemma, $k[\mathbf{x}]/\mathfrak{m}$ is a finite extension of $k$, and since $k$ is algebraically closed we have $k[\mathbf{x}]/\mathfrak{m} = k$. Let $a_1, \ldots, a_n$ be the images of $x_1, \ldots, x_n$ under the quotient map. The ideal $\mathfrak{m}$ consists of $x_i - a_i$ for every $i$, since these elements are mapped to 0 under the quotient map. The ideal $\langle x_1 - a_1, \ldots, x_n - a_n \rangle$ is maximal, and is therefore equal to $\mathfrak{m}$. Therefore $I \subseteq \langle x_1 - a_1, \ldots, x_n - a_n \rangle$ whence the point $(a_1, \ldots, a_n)$ is a common root for every element in $I$. This proves Theorem 4.1.1.

The proof of Theorem 4.1.2 is also complete since we proved the equivalence of the two Nullstellensatz. The following is an alternative proof of the fact that $\mathbf{I}(\mathbf{V}(I)) \subseteq \sqrt{I}$, which is the nontrivial part of the Nullstellensatz. It uses more commutative algebra than in the preliminaries. Let $I = \langle f_1, \ldots, f_m \rangle$, and $g$ be an arbitrary polynomial. Consider the quotient ring $B := k[\mathbf{x}]/I$. The condition $g \in \sqrt{I}$ is equivalent to $B_g = 0$. Suppose $B_g \neq 0$, and $\mathfrak{n}$ is a maximal ideal in $B_g$. The field $B_g/\mathfrak{n}$ is a finitely generated $k$ algebra, and hence is equal to $k$. Let $b_1, \ldots, b_n$ be the images of $x_1, \ldots, x_n$ under the map $k[\mathbf{x}] \to B \to B_g \to B_g/\mathfrak{n} = k$. Then $x_i - b_i$ generate the kernel of the map. Each $f_i$ goes to 0 under this map, and hence $\mathbf{b}$ is a root of every element in $I$. On the other hand, $g$ is nonzero under this map, since $g$ is a unit in $B_g$, and hence $\mathbf{b}$ is not a root of $g$. We have proved $g \notin \sqrt{I} \implies g \notin \mathbf{I}(\mathbf{V}(I))$, or equivalent $\mathbf{I}(\mathbf{V}(I)) \subseteq \sqrt{I}$ as required. □

We note that the last proof is basically the Rabinowitsch trick: in both places we prove that $B_g$ is 0 when $g \in \mathbf{I}(\mathbf{V}(I))$, the Rabinowitsch trick just uses a different representation of $B_g$. We finally prove Zariski's Lemma.

*Proof of Lemma 4.1.3.* Let $A$ be a finitely generated $k$ algebra that is also a field. By the Noether normalization lemma, we know that $A$ is an integral extension of some polynomial ring $k[z_1, \ldots, z_r]$. We will now prove that if $R$ is a ring and $R'$ is some integral extension of $R$, then $R$ is a field if $R'$ is a field. Once we prove this, we obtain in our setting that $k[z_1, \ldots, z_r]$ is a field, whence we must have $r = 0$. This will complete the proof, since $A$ will then be an integral extension of $k$, and therefore a finite field extension of $k$.

Suppose $R'$ is a field, and $x \in R$ is an arbitrary element. Then $x^{-1} \in R'$, and hence $x^{-1}$ is integral over $R$. Let the minimal monic equation of $x^{-1}$ be $x^{-m} + r_{m-1}x^{-m+1} + \cdots + r_0 = 0$, with $r_i \in R$. Then by multiplying by $x^{m-1}$ and rearranging we get $x^{-1} = -r_{m-1} - xr_{m-2} - \cdots - r_0 x^{m-1}$, and therefore $x^{-1} \in R$. This completes the proof. The converse of the above statement is also true if we assume that $R$ is a domain, although we do not need it here. Given an $x \in R'$, the idea there is to basically consider the minimal polynomial of $x$, argue that it has a constant term since $R$ is a domain, and that the reciprocal polynomial is satisfied by $x^{-1}$. Finally, the reciprocal polynomial can be made monic since the coefficients (which are from $R$) are units. This proves that $x^{-1} \in R'$. $\square$

The most natural question to ask given an existential statement like the Nullstellensatz, is whether we can decide by algorithm if a set of polynomials have a common root. In other words, suppose we are given polynomials $f_1, \ldots, f_n$ over $k$, and we have to check if they have common roots. By the Nullstellensatz, it suffices to check if $1 \in \langle f_1, \ldots, f_n \rangle$, or equivalently, to check if there exists $g_1, \ldots, g_n$ such that $1 = \sum f_i g_i$. We call the $g_i$ witnesses, since they witness the fact that $1 \in \langle \mathbf{f} \rangle$.

The Nullstellensatz itself does not give us any control over $g_i$. It is feasible that $g_i$ have arbitrarily high degree, and therefore no search procedure for them is guaranteed to terminate. That this is not the case was proven early in the $20^{th}$ century, by Grete Hermann [Her26]. She proved double exponential upper bounds for witnesses for arbitrary ideal membership queries. The following is a version of the theorem statement from [MM82].

**Theorem 4.1.4.** *Let $f_1, \ldots, f_m$ be polynomials of degree at most $d$ in $k[x_1, \ldots x_n]$, and $g$ be a polynomial of degree $d'$ in the ideal generated by $f_1, \ldots, f_m$. Then there exists polynomials $h_1, \ldots, h_m$ each of degree at most $d' + (md)^{2^n}$ such that $g = \sum f_i g_i$.*

A proof can be found in the appendix of [MM82]. This bound shows that the ideal membership problem, and in particular the Nullstellensatz problem is decidable, since it reduces to solving a double exponential sized linear system. Mayr and Meyer in the same paper also proved that this double exponential bound cannot be improved in general. They constructed an ideal in $10n$ variables with $10n + 1$ generators and proved the existence of a polynomial in the ideal such that every set of witness polynomials has at least one polynomial of degree $d^{2^{n-1}}$. Here, $d$ is a parameter, and every generator of their ideal is a difference of two monomials with degree at most $d + 2$. They also proved that the ideal membership problem is EXPSPACE hard, by reducing the commutative word equivalence problem to it. Finally the above double exponential bound along with the effective linear algebra results of [Csa75] show that the ideal membership problem is also in EXPSPACE, making it EXPSPACE hard.

The above discussion does not bode well for the Nullstellensatz problem, because it might be EXPSPACE complete too. However, it was proved that single exponential degree bounds exist in the special case of the Nullstellensatz, putting it in PSPACE. We discuss these results in the next section, and also prove the degree bound. The proof of single-exponential bounds for the Nullstellensatz allowed special cases of the ideal membership problem, such as the case of unmixed and zero dimensional ideals to be solved in single-exponential time [DFGS91]. In 1996, Koiran [Koi96] gave an AM protocol (conditioned on GRH) for the Nullstellensatz problem, when the underlying field is $\mathbb{C}$ and the polynomials have integer coefficients. His method is completely different from the previous methods of using the effective Nullstellensatz to reduce the system to a linear one. The positive characteristic case is an open problem, and the best known complexity remains PSPACE.

## 4.2 The effective Nullstellensatz

The first proof of a single exponential upper bound for the Nullstellensatz was given by Brownawell [Bro87]. He proved the result when $k = \mathbb{C}$ using analytic techniques. [2] A year later, an alternate algebraic proof was given by Kollár [Kol88] that worked for every characteristic. His proof used some properties of local cohomology groups. A more elementary proof using bounds on Hilbert functions was given by Sombra [Som97]. The proof that we discuss here was given by Jelonek [Jel05], and is significantly simpler than all of the proofs above. The above is far from a complete discussion of the existing literature on the problem. There are a number of other results that improve the above bounds and/or give improvements for special cases. These include but are not limited to [KPS99, KPS$^+$01, Som99].

In the projective setting, much better bounds are known. If $F_1, \ldots, F_n$ are homogeneous polynomials of degree at most $D$ without a common root in projective space, then the projective version of the Nullstellensatz states that the ideal generated by the $F_i$ contains a power of the irrelevant ideal. It follows from classical elimination theory that this power is upper bounded by $nD$ [Mac02, Laz77]. This is a quadratic bound. In particular, if we have non-homogeneous polynomials without a common root even at the hyperplane at infinity, then the improved quadratic bounds apply instead of the single exponential bounds. In the general case however, the single exponential bounds are essentially tight. An example witnessing this is given in [Bro87], and is presented later in this section.

We first state the version of the effective Nullstellensatz from [Jel05].

**Theorem 4.2.1.** *Let $k$ be algebraically closed, and let $f_1, \ldots, f_m \in k[x_1, \ldots, x_n]$ be polynomials of degrees $d_1, \ldots, d_m$. Assume that $d_1 \geqslant d_2 \geqslant \cdots \geqslant d_m$. Also assume that $f_1, \ldots, f_m$ do not have any common roots. Then there exists $g_1, \ldots, g_m$ such that $1 = \sum_{i=1}^m f_i g_i$. Further, the $g_i$ also satisfy the property that $\deg f_i g_i \leqslant \prod_{i=1}^m d_i$.*

We note that this is not the tightest version of the theorem statement in [Jel05], in the case when $m > n$. However, in our application we only use the case of $m = n + 1$, and in this case the difference is only a factor of $d_m/2$. Dropping the last factor will increase the length of the proof by a factor of at least 2, and therefore we do not do it here.

Before we prove this, we give the example from [Bro87] that shows that this is essentially tight. Fix some $d$, and let $f_1 = x_1^d$ and $f_n = 1 - x_{n-1}x_n^{d-1}$. For $2 \leqslant i \leqslant n-1$, let $f_i = x_{i-1} - x_i^d$. These polynomials do not have any common root, since the only common roots of $f_1, \ldots, f_{n-1}$ has first $n-1$ coordinates 0, and no such point can be a root of $f_n$. Suppose $1 = \sum f_i g_i$ for some $g_i$. In the above equation, we substitute $x_i = t^{(d-1)d^{n-i-1}}$ for $i \leqslant n-1$, and $x_n = t_n$. Under this substitution, the polynomials $f_2, \ldots, f_n$ all vanish, and therefore we get $1 = f_1'(t)g_1'(t)$. Now $f_1'(t) = t^{(d-1)d^{n-1}}$. The only way the product is equal to 1 is if $g_1$ has degree at least $d^n - d^{n-1}$, which shows the required bound.

[2] The results extend to all characteristic 0 fields via the Lefschetz principal.

A key ingredient in Jelonek's proof of the effective Nullstellensatz is the classical result of Perron which bounds the degree of the annihilator of $n+1$ polynomials in $n$ variables. The following statement is from [Pł05], while the original proof is from [Per51, Satz 57]. This theorem will also be of fundamental importance when we discuss algebraic independence.

**Theorem 4.2.2.** *Let $f_1, \dots, f_{n+1}$ be a sequence of polynomials in $k[x_1, \dots, x_n]$, with degrees $d_1, \dots, d_{n+1}$. Then there exists a polynomial $A$ in $k[y_1, \dots, y_{n+1}]$ such that*

- *$A(f_1, \dots, f_{n+1})$ is identically $0$, and*
- *$\deg_w(A) \leqslant \prod_{i=1}^{n+1} d_i$, where $\deg_w$ is the weighted degree with $\deg_w(y_i) = d_i$.*

The above bound on the weighted degree of $A$ will be referred to as the Perron bound. A proof can be found in [Pł05], and involves only linear algebra.

Before we present the proof of Theorem 4.2.1, we make an observation. Suppose we had $n+1$ polynomials that did not have a common root, and suppose the polynomial $A$ from Theorem 4.2.2 was such that it had a nonzero constant term. Then the polynomial $A$ also gives us a Nullstellensatz witness that matches the degree bound of Theorem 4.2.1. We start with the equation $A(f_1, \dots, f_{n+1}) = 0$, move the constant term to the other side and divide by it. We then collect all monomials in which $f_1$ appears, and after factoring out $f_1$ label the other factor $g_1$. We then collect all the monomials among the remaining ones in which $f_2$ appears, factor it, and call the remaining bit $g_2$, and so on. This gives us an equation of the form $\sum f_i g_i = 1$. In the case when the $f_i$ have transcendence degree $n$, the condition that $A$ has a nonzero constant term is equivalent to the condition that the $f_i$ do not have a common approximate root, for an appropriate definition of approximate roots. Therefore we can deduce that the hard case of the effective Nullstellensatz is when the polynomials do not have a common root, but have a common approximate root. A discussion on approximate common roots, and a proof of the above equivalence can be found in [GSS18].

We now prove Theorem 4.2.1.

*Proof of Theorem 4.2.1.* If $m \leqslant n$, then we add polynomials $f_{m+1} = \cdots = f_{n+1} = 0$, and we assume that $m > n$. This does not change either the assumption of empty zeroset. Further, we set $d_{m+1}, \dots, d_{n+1} = 1$ in this case. The degree bounds also therefore go unchanged.

Let $h_1, \dots, h_m$ be polynomials such that $1 = \sum f_i h_i$. Such polynomials exist by the classical Nullstellensatz. Let $z$ be a fresh variable. Define the map $\phi : \mathbb{A}^{n+1} \to \mathbb{A}^{n+m}$ as

$$\phi(x_1, \dots, x_n, z) = (x_1, \dots, x_n, z f_1(x_1, \dots, x_n), \dots, z f_m(x_1, \dots, x_n)). \tag{4.1}$$

The space $\mathbb{A}^{n+1}$ is isomorphic to its image under $\phi$, since the following polynomial map is an inverse of $\phi$ when restricted to the image:

$$\phi'(y_1, \dots, y_{n+m}) = \left( y_1, \dots, y_n, \sum_{i=1}^{m} h_i y_{n+i} \right). \tag{4.2}$$

In particular, the image $\phi(\mathbb{A}^{n+1})$ is closed, and has dimension $n+1$ and when treated as a map from $\mathbb{A}^{n+1}$ to $\phi(\mathbb{A}^{n+1})$, the map $\phi$ is finite. Let $d$ denote the degree of the image. We will prove in later sections that $d \leqslant \prod_{i=1}^{m}(d_i+1)$, but for this proof we just require that $d$ is finite.

Let $\pi : \mathbb{A}^{n+m} \to \mathbb{A}^{n+1}$ be a linear projection of the form

$$\pi(y_1, \dots, y_{n+m}) = \left(\sum_{i=1}^{m+n} a_{1,i}y_i, \sum_{i=2}^{n+m} a_{2,i}y_i, \dots, \sum_{i=n+1}^{n+m} a_{n+1,i}y_i\right). \qquad (4.3)$$

Note that the first coordinate function is a linear combination of all the variables, the second is a linear combination of all variables except the first, and so on. Suppose each $a_{ij}$ is picked uniformly and randomly from a subset $S$ of of $k$. By Corollary 3.1.4, with probability at least $1 - 2(n+2)d/|S|$, the subspace defined by the equations $\sum_{i=j}^{n+m} a_{j,i}y_i$ is disjoint from $\phi(\mathbb{A}^{n+1})$. If this is the case then by Theorem 3.2.5, the map $\pi$ is Noether normalizing for $\phi(\mathbb{A}^{n+1})$. For the rest of the proof, we assume that $\pi$ has this property.

Both the maps $\phi$ and $\pi$ are finite. Since integral extensions of integral extensions are integral, the composition $\psi := \pi \circ \phi : \mathbb{A}^{n+1} \to \mathbb{A}^{n+1}$ is also a finite map. Explicitly, the $j^{th}$ coordinate function $\psi_j$ is $l_j(x_1, \dots, x_n) + \sum_{i=j}^{m} a_{j,i}zf_i$, where $l_j$ is a linear form.

Let $\psi_j$ also denote the polynomial corresponding to this function. Each of these polynomials can be treated as a $n$ variate polynomial over the field $k(z)$. Since $\deg f_i \geqslant \deg f_{i+1}$ for all $i$, and since each $\psi_i$ is a linear combination of $f_i, f_{i+1}, \dots, f_m$, we have $\deg \psi_i = d_i$ when treated as polynomials over $k(z)$. Since there are $n+1$ of them, by Theorem 4.2.2 there exists a $n+1$ variate polynomial $A$ with coefficients in $k(z)$ such that $A(\psi_1, \dots, \psi_{n+1}) = 0$. Further, this $A$ also satisfies weighted degree bounds, that is $\deg_w A \leqslant \prod_{i=1}^{n+1} d_i$, where the $i^{th}$ variable of $A$ has weight $d_i$.

Each coefficient of $A$ is an element of $k(z)$, and by potentially clearing denominators, we can assume that the coefficients are in $k[z]$. This does not change the degree of $A$. We can therefore construct a polynomial $B$ in $n+2$ variables $v_1, \dots, v_{n+2}$ by starting with $A$, labelling its $n+1$ variables as $v_1, \dots, v_{n+1}$ and replacing the variable $z$ which occurs as part of the coefficients with the variable $v_{n+2}$. The polynomial $B$ satisfies $B(\psi_1, \dots, \psi_{n+1}, z) = 0$. By construction, $B$ has weighted degree at most $\prod_{i=1}^{n+1} d_i$, where the first $n+1$ variables $v_1, \dots, v_{n+1}$ have weights $d_1, \dots, d_{n+1}$ respectively, and $v_{n+2}$ has weight $0$.

The composed map $\psi : \mathbb{A}^{n+1} \to \mathbb{A}^{n+1}$ is finite, and therefore surjective, and in particular also dominant. The corresponding map of coordinate rings $\psi^*$ is therefore an injection, and the coordinate ring of the domain $\mathbb{A}^{n+1}$ is an integral extension of the image of $\psi^*$. More explicitly, the ring $k[x_1, \dots, x_n, z]$ is an integral extension of the ring $k[\psi_1, \dots, \psi_{n+1}]$. Let $C$ denote the minimal polynomial of $z$ over $k[\psi_1, \dots, \psi_{n+1}]$. By definition, $C$ is a monic univariate polynomial, say in the variable $u$, with coefficients from $k[\psi_1, \dots, \psi_{n+1}]$ such that $C(z) = 0$. We write $C = \sum_{i=0}^{D} p_i(\psi_1, \dots, \psi_{n+1})u^i$, where $D$ is the degree of $C$.

Since the extension is integral, the ideal of univariates in $k[\psi_1, \dots, \psi_{n+1}][u]$ with root $z$ is a principal ideal generated by $C$. [3] We can also treat $B$ as a polynomial in $k[\psi_1, \dots, \psi_{n+1}][u]$ by specializing the first $n+1$ variables of $B$ to $\psi_1, \dots, \psi_{n+1}$, and relabelling the last variable to $u$. This univariate has $z$ as a root by construction, therefore, $C$ divides $B$ in $k[\psi_1, \dots, \psi_{n+1}][u]$.

The polynomials $\psi_1, \dots, \psi_{n+1}$ themselves are algebraically independent over the field $k$: suppose there some polynomial $\rho$ such that $\rho(\psi_1, \dots, \psi_{n+1}) = 0$. Then $\rho$ would vanish on the image of the map $\psi$, but since $\psi$ is a finite map, it is surjective and has image $\mathbb{A}^{n+1}$, and therefore $\rho$ would have to be 0. Therefore, we can naturally treat $C$ as a polynomial in the $n+2$ variables $v_1, \dots, v_{n+2}$. We do this by replacing $u$ with $v_{n+2}$, and occurrences of $\psi_i$ in the coefficients with $v_i$, obtaining $C = \sum_{i=0}^{D} p_i(v_1, \dots, v_{n+1}) v_{n+2}^i$. This conversion is well defined and unique by the independence of $\psi_i$. That $C$ divides $B$ in $k[\psi_1, \dots, \psi_{n+1}][u]$ now also means that $C$ divides $B$ as polynomials in $k[v_1, \dots, v_{n+2}]$, as can be seen by applying a similar conversion to the factor $B/C$. This implies that the weighted degree of $C$ is at most the weighted degree of $B$, which itself is at most $\prod_{i=1}^{n+1} d_i$. Here the variable $v_i$ has weight $d_i$ for $i \leqslant n+1$, and $v_{n+2}$ has weight 0. In particular this means that for every $i$, the polynomial $p_i(\psi_1, \dots, \psi_{n+1})$ is a polynomial of degree at most $\prod_{i=1}^{n+1} d_i$ when treated as a polynomial in $x_1, \dots, x_n, z$.

Finally, consider the expansion $C(z) = \sum_{i=0}^{D} p_i(\psi_1, \dots, \psi_{n+1}) z^i$ written as a polynomial in $k[x_1, \dots, x_n][z]$. By definition this is the 0 polynomial, and therefore the coefficient of every $z^i$ in this expansion is 0. Consider in particular the coefficient of $z^D$, which is $\sum_{i=0}^{D} \text{coeff}_{z^{D-i}}(p_i(\psi_1, \dots, \psi_{n+1}))$. The term corresponding to $i = D$ in the summand is 1, since $C$ is monic. Every other $\text{coeff}_{z^{D-i}}(p_i(\psi_1, \dots, \psi_{n+1}))$ is a sum of multiples of the original polynomials $f_1, \dots, f_m$, since in each $\psi_i$ the variable $z$ only occurs multiplied to some $f_j$. Therefore the coefficient of $z^D$ is of the form $1 + \sum f_i g_i$ for $g_i \in k[x_1, \dots, x_n]$, and we can rearrange (and change signs) to obtain $1 = \sum f_i g_i$. By the weighted degree bound proved in the previous paragraph, each $f_i g_i$ has degree at most $\prod_{i=1}^{n+1} d_i$, which completes the proof. □

[3] The same proof that shows that a univariate polynomial ring over a field is a PID works here: the scaling step can be performed since we assume that $C$ is monic.

# Chapter 5

# Algebraic independence

In this chapter, we take a detour from our previous discussion and discuss the algebraic independence problem. The problem is to determine, given a set of polynomials $f_1, \ldots, f_n$ whether or not they are algebraically independent in the function field $k(x_1, \ldots, x_n)$.

There is evidence (see discussion in [Mul12]) that from a computational perspective, it is advantageous to define varieties by giving a polynomial map whose closure is the given variety, as opposed to describing the generators of the ideal corresponding to the given variety. The algebraic independence problem is equivalent to computing the dimension of such an explicit variety.

In this chapter, we will focus more on the second formulation of the problem, even though it requires more technical background. It turns out that the problem is easier in characteristic 0 than in finite characteristics, even though there is considerable evidence that it cannot be too difficult in the latter case. The hope is that this more complicated formulation will also afford the use of the more sophisticated tools of algebraic geometry.

This chapter is organized as follows. In the first section we will state the problem in multiple equivalent ways. In the second section we will state some results about algebraic independence and we will use the polynomial map formulation to provide proofs of these results. All of these results are well know, and all we do is provide (in some cases) alternative proofs. A survey, including many of the original proofs of this result and the history of the problem can be found in [Sin19].

## 5.1 Problem definition

Let $k$ be the underlying field, and $p$ denote its characteristic. Let $f_1, \ldots, f_m$ be $n$ variate polynomials from $k[x_1, \ldots, x_n]$ of degrees $d_1, \ldots, d_m$ respectively. In fields that are of interest to us, the problem does not change if we replace $k$ by an algebraic extension, and therefore we assume that $k$ is replaced by its algebraic closure. When required, we will be more explicit about the field in which the coefficients of $f_i$ lie.

The polynomials are said to be *algebraically independent* if and only if for every non-zero polynomial $G \in k[y_1, \ldots, y_m]$, it holds that $G(f_1, \ldots, f_m)$ is not the identically zero polynomial. Equivalently, the polynomials are said to be *algebraically dependent* if and only if there exists some polynomial $A \in k[y_1, \ldots, y_m]$ such that $A(y_1, \ldots, y_m)$ is the identically zero polynomial. For example, if $f_1 := x_1 + x_2$ and if $f_2 := (x_1 + x_2)^2$, then $f_1$ and $f_2$ are algebraically dependent, with $A = y_2 - y_1^2$. On the other hand, if $f_1 := x_1$ and $f_2 := x_1 + x_2$, then no such polynomial $A$ exists, and thus $f_1$ and $f_2$ are algebraically independent. [1] When polynomials $f_1, \ldots, f_m$ are dependent, any polynomial $A$ that satisfies $A(f_1, \ldots, f_m) \equiv 0$ will be called an *annihilator* of $f_1, \ldots, f_m$. The set of annihilating polynomials form an ideal. Note that the dependence/independence of polynomials depends on the underlying field. Consider for example polynomials $f_1 := x_1 + x_2$ and $f_2 := x_1^2 + x_2^2$. If $\operatorname{char} k = 2$, then $f_1$ and $f_2$ are dependent, with annihilator $y_2 - y_1^2$. If $\operatorname{char} k \neq 2$, then some more variable chasing will show that $f_1$ and $f_2$ are independent.

An alternative formulation of this problem is to check, given some field extensions, whether or not they are algebraic. Consider the field extensions $k(\mathbf{x})/k(\mathbf{f})$, $k(\mathbf{f})/k$ and $k(\mathbf{x})/k$. We have

$$\operatorname{trdeg}_k k(\mathbf{x}) = \operatorname{trdeg}_k k(\mathbf{f}) + \operatorname{trdeg}_{k(\mathbf{f})} k(\mathbf{x}),$$

where $\operatorname{trdeg}_K L$ denotes the transcendence degree of the extension $L/K$. We also have $\operatorname{trdeg}_k k(\mathbf{x}) = n$ by definition, and therefore $\operatorname{trdeg}_k k(\mathbf{f}) \leqslant n$. By definition, the polynomials are algebraically independent if and only if $\operatorname{trdeg}_k k(\mathbf{f}) = m$. If $m > n$ therefore, the polynomials are always dependent. Finally, the polynomials are dependent if and only if the field extension $k(\mathbf{f})$ is algebraic over $k(f_1, \ldots, f_{i-1}, f_{i+1}, \ldots, f_m)$ for some $i$. In general, we will refer to $\operatorname{trdeg}_k k(\mathbf{f})$ as the transcendence degree of the polynomials $f_1, \ldots, f_m$.

At this point, we show that going to an extension of $k$ does not change the algebraic independence of the polynomials. We will assume that the field $k$ is perfect, which is the case when $k$ is a finite field or a field of characteristic 0. Suppose $f_1, \ldots, f_m$ are polynomials with coefficients in $k$ that are dependent when considered as polynomials in an algebraic extension $K$ of $k$. Let $B$ be the annihilator of the polynomials $f_1, \ldots, f_m$ in $K[y_1, \ldots, y_m]$. It suffices to replace $K$ with the subfield that contains $k$ and all the coefficients of $B$. We assume therefore that $K$ is a finite extension, and by the primitive element theorem, is generated by a single element $\alpha$ ([Lan02, Chapter 5, Theorem 4.6]). We can now write $B$ as $B = \sum_{i=0}^{D-1} \alpha^i B_i(y_1, \ldots, y_m)$ where each $B_i$ has coefficients in $k$, and $D$ is the degree of the extension. Each $B_i$ when evaluated at $\mathbf{x}$ results in an element of $k[\mathbf{x}]$ since the coefficients of $f_i$ are from $k$. Further, since $\alpha^i$ are $k$-linearly independent, they are also $k[x_1, \ldots, x_n]$ linearly independent, and therefore each $B_i$ must be 0 when evaluated at $\mathbf{f}$. That $B$ is nonzero implies that some $B_i$ is nonzero, and this $B_i$ is an annihilator for $\mathbf{f}$ as polynomials in $k$. We

[1] Proving this involves some simple variable chasing.

can also note that since the degree of $B_i$ is at most the degree of $B$, going to an extension does not even imply the existence of annihilators of smaller degree. Therefore, we can assume that $k$ is algebraically closed.

We now present the final formulation. Given its importance in this chapter, we reserve a subsection for it, and provide some motivation.

### 5.1.1 Polynomial maps and algebraic independence

Given the ring $k[\mathbf{f}]$, a natural object to look at is the affine variety that it corresponds to. [2] Since $k[\mathbf{f}]$ is a finitely generated algebra over $k$ with $m$ generators, it is isomorphic to $k[y_1, \dots, y_m]/\mathfrak{U}$ for some ideal $\mathfrak{U}$. The isomorphism takes each $y_i$ to $f_i$. If $\mathbf{f}$ satisfy some algebraic relation, then applying the inverse of the above isomorphism we see that $\mathbf{y}$ also must satisfy the same relation. The converse also holds. Thus the ideal $\mathfrak{U}$ is exactly the ideal of all annihilators of $\mathbf{f}$.

The ring $k[\mathbf{y}]/\mathfrak{U}$ corresponds to the affine variety defined by the equations in $\mathfrak{U}$. We call this affine variety $Y$. When we want to emphasize the dependence of $Y$ on $\mathbf{f}$, we use $Y_{\mathbf{f}}$. There is a natural map $i$ from $k[\mathbf{y}]/\mathfrak{U}$ to $k[\mathbf{x}]$ that takes each $y_i$ to $f_i$. As discussed above, this is well defined since elements in $\mathfrak{U}$ are exactly the algebraic relationships in $\mathbf{f}$. Further, this map is an injection, since $\mathfrak{U}$ consists of all algebraic relationships between $\mathbf{f}$. The map $i$ corresponds to a map $i^*$ from the affine variety corresponding to $k[\mathbf{x}]$(namely $\mathbb{A}^n$) to $Y$. The map $i^*$ has $j^{th}$ coordinate function $f_j$, and is thus exactly the polynomial map with coordinates $f_1, \dots, f_m$. We will call this map $\phi_{\mathbf{f}}$. Since $i$ is an injection, the map $\phi_{\mathbf{f}}$ is dominant. In other words, $Y$ is exactly the closure of the image of $\mathbb{A}^n$ under $\phi_{\mathbf{f}}$. [3]

The above discussion shows that given $k[\mathbf{f}]$, it is natural to consider the map $\phi_{\mathbf{f}}$ with coordinate functions $f_1, \dots, f_m$. This point is driven home by the fact that the dimension of the affine variety $Y$ is exactly the transcendence degree of $\mathbf{f}$. This follows by (one of) the definition(s) of the dimension of an affine variety. We first make the following simple observation.

**Lemma 5.1.1.** *The affine variety $Y$ is irreducible.*

*Proof of Lemma 5.1.1.* The affine variety $\mathbb{A}^n$ is irreducible, and hence so is its image under the regular map $\phi_{\mathbf{f}}$. The affine variety $Y$ is thus the closure of an irreducible set [4] , and hence is itself irreducible.

Alternatively, an affine variety is irreducible if and only if its coordinate ring is a domain. That $k[\mathbf{f}]$ is a domain follows from the fact that it is a subring of the domain $k[\mathbf{x}]$. □

We can now use the definition of the dimension.

[2] Of course one has to make sure that the ring has no nilpotent elements, a property that $k[\mathbf{f}]$ satisfies.

[3] The focus of this document is the finite characteristic case, and thus unless stated otherwise, the topology is always the Zariski topology.

[4] The image is also a quasi-projective variety, just maybe not affine.

**Definition 1** ([SR13, p. 67])**.** The dimension of an irreducible affine variety is the transcendence degree of its function field.

Lemma 5.1.1 allows us to apply the above to $Y$. In particular, polynomials $f_1, \dots, f_m$ are algebraically independent if and only if $\dim Y_f = m$. Alternatively, the polynomials are independent if and only if the image of the map $\phi_f$ is dense in $\mathbb{A}^m$.

Before we prove properties about algebraic independence, we state an upper bound on the degree of $Y$ that will be useful. We collect this property, and the statement about the dimension of $Y$ in a single lemma to make it easier to invoke later. The proof of the degree bound is non trivial, and we only provide a reference.

**Lemma 5.1.2.** *The variety $Y$ has degree equal to the transcendence degree of the polynomials $f_1, \dots, f_m$. Further, the degree of $Y$ is at most $(\max d_i)^r$, where $r$ is the transcendence degree.*

*Proof of Lemma 5.1.2.* The first statement follows from the discussion preceding the lemma. A proof of the second statement can be found in [BCS97, Theorem 8.48]. The idea is to write $\phi_f$ as the composition of the Veronese embedding followed by a linear map. Studying the Hilbert polynomial shows that the Veronese embedding has degree equal to the product of the degrees, reducing the statement to the case of linear maps, where it is easily proved. □

We will now prove some known results using the above framework.

## 5.2 Some properties of algebraic independence

### 5.2.1 Basic results

We start with some fairly easy results. The first is that if the transcendence degree of $f_1, \dots, f_m$ is $m - 1$, then the ideal of annihilators is principal. The following statement is from [Kay09].

**Theorem 5.2.1** ([Kay09, Lemma 7])**.** *If $f_1, \dots, f_m$ are algebraically dependent such that no subset of them are algebraically dependent, then the ideal of annihilators $\mathfrak{U}$ is principal.*

This is a consequence of the fact that ideals corresponding to irreducible varieties of dimension $m-1$ in $\mathbb{A}^m$ are principal. The proof is immediate given this fact.

**Lemma 5.2.2** ([SR13, Theorem 1.21])**.** *If $X \subset \mathbb{A}^k$ is an irreducible affine variety of dimension $k - 1$, then the coordinate ring $k[X]$ is isomorphic to $k[\mathbf{y}]/\mathfrak{U}_X$ with $\mathfrak{U}_X$ principal.*

The next statement is the fact that if we have $m$ polynomials that have transcendence degree $r$, then taking $r$ linear combinations of these polynomials results in a set of independent polynomials. Further, taking $r + 1$ linear combinations results in a set of polynomials of transcendence degree $r$, and therefore the previous result applies. Finally, random linear combinations have the above properties. This is a consequence of the Noether Normalization lemma.

**Theorem 5.2.3.** *Let $f_1, \dots, f_m$ be a set of $m$ polynomials in $n$ variables. Let $\operatorname{trdeg}(\mathbf{f}) = r$. Then there exist polynomials $g_1, \dots, g_{r+1}$ of the form*

$$g_i = \sum_{j=i}^{m} a_{i,j} f_j$$

*such that $\operatorname{trdeg}(g_1, \dots, g_r) = r$ and also $\operatorname{trdeg}(g_1, \dots, g_{r+1}) = r$.*

*Further, suppose each $a_{ij}$ is picked uniformly and randomly from a subset $S$ of the field $k$ that does not include zero. Then the above holds with probability at least $1 - 2(r+1)D/|S|$, where $D = (\max d_i)^r$.*

*Proof of Theorem 5.2.3.* Let $\varphi_f$ be the polynomial map with coordinate functions $f_i$, and let $Y$ be the closure of its image. We have $\dim Y = r$ and $\deg Y \leqslant D$ by Lemma 5.1.2. Let $\pi$ be the linear map $\mathbb{A}^m \to \mathbb{A}^r$ with coordinate functions $\pi_i(y_1, \dots, y_m) = \sum_{j=i}^{m} a_{ij} y_j$. By Theorem 3.2.6, the map $\pi$ is Noether normalizing for $Y$ with the mentioned probability, and in particular finite. The composed map $\pi \circ \varphi_f$ is therefore dense. This map has coordinate functions $g_1, \dots, g_r$, which proves that these polynomials have transcendence degree $r$. Let $\pi' : \mathbb{A}^m \to \mathbb{A}^{r+1}$ be the map with coordinate functions $\pi'_i(y_1, \dots, y_m) = \sum_{j=i}^{m} a_{ij} y_j$. The image of $Y$ under this map has dimension $r$ at most $r$ since $Y$ has dimension $r$. The image has dimension at least $r$ since the projection $\pi$ of $\pi'$ has dimension $r$, and therefore the image has dimension exactly $r$. Therefore the polynomials $g_1, \dots, g_{r+1}$ have transcendence degree $r$. □

The next result we discuss is variable reduction. Suppose polynomials $f_1, \dots, f_m$ have transcendence degree $r$. We can then replace each $x_i$ with a linear combination of $r$ new variables $z_1, \dots, z_r$ such that the resulting $r$ variate polynomials are also algebraically independent.

**Theorem 5.2.4** ([Kay09, Claim 11.1], [Mit12, Theorem 4.2.2])**.** *Let $f_1, \dots, f_m$ be a set of $m$ polynomials in $n$ variables, with $m < n$. Let $\operatorname{trdeg}(\mathbf{f}) = r$. Then there exist a homomorphism $\psi^* : k[x_1, \dots, x_n] \to k[z_1, \dots, z_m]$ of the form*

$$\psi^*(x_i) = a_{i,0} + \sum_{j=1}^{m} a_{i,j} z_j$$

*such that $\operatorname{trdeg}(\psi^*(\mathbf{f})) = \operatorname{trdeg}(\mathbf{f}) = r$.*

In fact we will show the stronger statement that a random such homomorphism will work.

*Proof of Theorem 5.2.4.* By potentially considering just a subset of the polynomials, we can assume without loss of generality that $m = r$. Let $\mathbf{b}$ be a point in the image $\phi_{\mathbf{f}}(\mathbb{A}^n)$ that has fibre of dimension exactly $n - m$. We use $V$ to denote $\phi_{\mathbf{f}}^{-1}(\mathbf{b})$. Let $H$ be any linear subspace of dimension $m$ such that $\dim H \cap V = 0$. Let $\psi$ be a linear map from $\mathbb{A}^m$ to $\mathbb{A}^n$ that maps $\mathbb{A}^m$ isomorphically to $H$. The corresponding map $\psi^* : k[\mathbb{A}^n] \to k[\mathbb{A}^m]$ is a linear map of the stated form. We will prove that $\operatorname{trdeg}(\psi^*(\mathbf{f})) = m$.

Let $Y' := \phi_{\mathbf{f}} \circ \psi(\mathbb{A}^m)$, and let $m' := \dim Y'$ be the dimension of the image of the composed map. By construction, the point $\mathbf{b}$ is in $Y'$. Further, $\mathbf{b}$ has a finite fibre under this map, since the fibre corresponds exactly to the set $H \cap V$ which we assumed was finite. By the dimension theorem, every fibre has dimension at least $m - m'$, and therefore $m' \geqslant m$. Further, since the map $\phi_{\mathbf{f}} \circ \psi : \mathbb{A}^m \to Y'$ is dominant, we also have $m \geqslant m'$ whence we deduce that $m = m'$. Finally, $m' = \operatorname{trdeg}(\psi^*(\mathbf{f}))$ by the definition of dimension, and therefore we have $\operatorname{trdeg}(\psi^*(\mathbf{f})) = m$ as required.

We will now prove that a random map has this property. The subspaces here are the images of random linear maps, and therefore we use Lemma 3.1.6. The fibre $V$ of $\mathbf{b}$ is defined by the equations $f_1 = b_1, \ldots, f_m = b_m$. By Bézout's theorem it has degree $D$. By Lemma 3.1.6, if we pick each $a_{ij}$ from a subset $S$ of $k$ not containing 0, the image of the map $\psi$ properly intersects $V$ with probability at least $1 - n^3 D/|S|$. If we pick $|S| = 3n^3 D$, the required statement holds with probability at least $2/3$. We can sample from this set in time polynomial in $\log nD$, which is polynomial in $n, d_i$. □

### 5.2.2 The Jacobian criterion

We now prove the Jacobian criterion. This is an efficient way of checking if a given set of polynomials is algebraically independent when the underlying field has characteristic 0 (or large enough).

Given polynomials $\mathbf{f}$, define the Jacobian matrix $\mathcal{J}(\mathbf{f})$ as $\mathcal{J}(\mathbf{f})_{ij} := \partial f_i / \partial x_j$. This is a matrix with entries from the field $k(\mathbf{x})$. The following statement of the Jacobian criterion is from [PSS16], the references therein point to the places where different cases were first proved.

**Theorem 5.2.5** ([PSS16, Lemma 5])**.** *Let $f_1, \ldots, f_m$ be polynomials of degree at most $d$ and transcendence degree $r$. If* $\operatorname{char} k = 0$ *or* $\operatorname{char} k > d^r$ *then* $\operatorname{trdeg}(\mathbf{f}) = \operatorname{rank}_{k(\mathbf{x})} \mathcal{J}(\mathbf{f})$.

To prove this, we will use the notion of tangent spaces. Given a point $w$ on a variety $W$, we can define ideal $\mathfrak{m}_w$ of $k[W]$ consisting of all polynomials vanishing on $w$. This ideal is maximal, since the quotient $k[W]/\mathfrak{m}_w$ is $k$, which is a field. Further, the $k[W]$-module $\mathfrak{m}_w/\mathfrak{m}_w^2$ is annihilated by $\mathfrak{m}_w$, and is thus a $k$-vector space. The vector space $\mathfrak{m}_W/\mathfrak{m}_w^2$ is called the cotangent space at $w$, and the dual $(\mathfrak{m}_w/\mathfrak{m}_w^2)^*$ is called the tangent space. This is denoted by $\Theta_{W,w}$. This definition of tangent spaces matches the more classical definition using differentials in the case of complex numbers, but has the advantage of being purely algebraic, and thus being well defined over the closures of finite fields.

Given a regular map $\phi : W \to Z$, we have an induced map $\phi^* : k[Z] \to k[W]$. Suppose $\phi(w) = z$ for some $w \in W, z \in Z$. Then $\phi^*(\mathfrak{m}_z) \subseteq \mathfrak{m}_w$, and

$\phi^*(\mathfrak{m}_z^2) \subseteq \mathfrak{m}_w^2$. We thus get an induced map $\phi^* : \mathfrak{m}_z/\mathfrak{m}_z^2 \to \mathfrak{m}_w/\mathfrak{m}_w^2$, which in turn induces a map between $\Theta_{W,w}$ and $\Theta_{Z,z}$. We denote this map $d_w\phi : \Theta_{W,w} \to \Theta_{Z,z}$.

If $W$ is an irreducible variety, then a point $w \in W$ is called nonsingular if $\dim W = \operatorname{rank} \Theta_{W,w}$. The set of nonsingular points in a variety form a dense open set (a proof can be found in [SR13, Section 1.4]). Further, we have the following useful lemma.

**Lemma 5.2.6** ([SR13, Theorem 2.3])**.** *The dimension of the tangent space at a nonsingular point equals the dimension of the variety.*

We now consider our setting. We have a map from $\mathbb{A}^n$ to $Y$ defined by the polynomials $\mathbf{f}$. The tangent space at every point in $\mathbb{A}^n$ is a vector space of dimension $n$ since $\mathbb{A}^n$ is irreducible and nonsingular. The induced map $d_{x_0}\phi_{\mathbf{f}}$ at point $x_0$ is given by the linear map $\mathcal{J}(\mathbf{f})(x_0)$.

Given the above, we can prove one part of Theorem 5.2.5. Suppose the Jacobian has rank $r$. If $r < m$, then we can restrict our attention to some $r$ linearly independent rows, and thus we can assume without loss of generality that $r = m$. Let $x_0$ be a point such that $\mathcal{J}(\mathbf{f})(x_0)$ is rank $r$. The set of points where this does not hold is a subvariety of $\mathbb{A}^n$ of dimension at most $n-1$. The map $d_{x_0}\phi_{\mathbf{f}}$ then has image a linear space of rank $m$, whence the codomain, that is $\Theta_{Y,\phi_{\nu\mathbf{f}}(x_0)}$ must have rank at least $m$. This shows that in a dense open subset of $Y$, the tangent space has dimension at least $m$, which shows that the dimension of $Y$ is at least $m$.

The above proof does not require the characteristic of $k$ to be large, and indeed this requirement is only required for the other direction. For this, we use the following lemma.

**Lemma 5.2.7** ([SR13, Lemma 2.4])**.** *Suppose* $\operatorname{char} k = 0$. *Then there is a nonempty open subset* $V \subset X$ *such that* $d_x F$ *is surjective for* $x \in V$.

While the above lemma assumes that the characteristic is zero, the proof works as long as the characteristic is large enough, that is bigger than $d^r$. The above lemma immediately gives us the other direction of the Jacobian criterion: since the map on the tangent spaces is surjective, it must be that for a dense open subset of $Y$ we have $\operatorname{rank} \Theta_{Y,y} \leqslant m$. This implies that $\dim Y \leqslant m$, as required.

### 5.2.3 Functional dependence

Suppose we have dependent polynomials $g_1, \ldots, g_r$. In general, the dependency of any $g_i$ on the rest will is nonlinear, and we cannot write say $g_1 = H(g_2, \ldots, g_r)$ for some polynomial $H$. [PSS16] showed that while the above is not possible, if we randomly shift the polynomials and allow power series, then we can write a power of $g_1$ as a function of $g_2, \ldots, g_r$. Formally, they proved the following theorem.

**Theorem 5.2.8** ([PSS16, Theorem 10]). *Let* $\mathbf{f}$ *be a set of polynomials of transcendence degree* $r$*. Then there exist an algebraically independent subset* $\{g_1, \dots, g_r\} \subset \mathbf{f}$ *of polynomials such that for a random* $\mathbf{a} \in k^n$ *and every* $f_j$*, there is a power series* $h_j \in k[[y_1, \dots, y_r]]$ *such that* $f_j(\mathbf{x}+\mathbf{a}) = h_j(g_1(\mathbf{x}+\mathbf{a}) - g_1(\mathbf{a}), \dots, g_r(\mathbf{x}+\mathbf{a}) - g_r(\mathbf{a}))$*.*

We now prove the result. Define the polynomial map $\phi_{\mathbf{f}}$ with coordinate functions $(f_1, \dots, f_m)$. Given polynomials $\mathbf{f}$, the shifted polynomials $\mathbf{f}(\mathbf{x}+\mathbf{a}) - \mathbf{f}(\mathbf{a})$ have the property that $0$ is mapped to $0$. Further, the variety corresponding to the shifted polynomials, which we call $Y_{\mathbf{a}}$, is such that the origin of $Y_{\mathbf{a}}$ is the image of $\mathbf{a}$ in $Y$. If $\mathbf{a}$ is picked randomly, then $\mathbf{f}(\mathbf{a})$ is a general point, and thus by shifting we have made the origin of $Y_{\mathbf{a}}$ a general point. In particular, in $Y_{\mathbf{a}}$, we can assume that the origin is a nonsingular point.

The ideal $\mathfrak{m} := \mathfrak{m}_0$ in $k[Y]$ is generated by $y_1, \dots, y_m$. A subset of these elements also then generate the vector space $\mathfrak{m}/\mathfrak{m}^2$. Since the origin is nonsingular, this vector space has dimension $r$. Let $y'_1, \dots, y'_r$ be the $y_j$ whose images generate this vector space. The $y'_i$ form a system of local parameters at the origin. [5] Since the point is nonsingular, the local ring $\mathcal{O}_0$ has an isomorphic inclusion into the power series ring generated by the $y'_i$. Each $y_j$ can thus be written as a power series in the $y'_i$ on $Y_{\mathbf{a}}$. Finally, given the power series for $y_j$ in terms of $y'_i$, we can substitute $f_i(\mathbf{x}+\mathbf{a}) - f_i(\mathbf{a})$ for $y_i$ everywhere to get a power series for $f_j(\mathbf{x}+\mathbf{a}) - f_j(\mathbf{a})$. This completes the proof.

This proof essentially uses a Newton iteration like procedure to find the power series. The result can also be proved directly using Newton iteration. We provide this proof in appendix A

### 5.2.4 coAM and AM protocols

We now prove that the algebraic independence problem over finite fields is in AM and in coAM. These results are by [GSS18], and is the result that motivates this whole enterprise of polynomial maps.

We first provide an intuitive idea for why the result holds, in the special case of $m = n$. In order to check for independence, Arthur can pick a random point $\mathbf{b}$ in $\mathbb{A}^m$, and ask Merlin for a point $\mathbf{a}$ in $\mathbb{A}^n$ such that $\phi_{\mathbf{f}}(\mathbf{a}) = \mathbf{b}$. If the polynomials are independent, then the image of the map $\phi_{\mathbf{f}}$ is dense, and with high probability such a $\mathbf{a}$ exists, and if the polynomials are dependent, then $\mathbf{b}$ is not in the image, and no such $\mathbf{a}$ exists. In order to check for dependence, Arthur can pick a point in $\mathbf{a} \in \mathbb{A}^n$, and ask Merlin for a list of points that also map to $\phi_{\mathbf{f}}(\mathbf{a})$. If the polynomials are dependent, then the fibres have dimension at least 1, and therefore Merlin can produce an arbitrary number of points with this property. If the polynomials are independent, then we can prove a bound on the sizes of the fibres, and therefore Merlin can only produce a list of bounded size.

Both the above protocols are incorrect, since the interactions are not polynomial sized. The actual protocols will use the Sipser-Goldwasser protocol.

[5] The definition of systems of local parameters can be found in [SR13], in Chapter 2.

**Lemma 5.2.9.** *Let $S \subseteq \{0,1\}^m$ be a set, such that membership in $S$ can be verified in AM. Let $K$ be a number between $0$ and $2^m$. If $S$ is promised to have size either less than $K$ or greater than $2K$, then there is an AM protocol to verify that the size of $S$ is at least $2K$.*

The setting of the above is very similar to our second protocol. The idea is that instead of listing 2K members of S, which would have size exponential, Arthur sends a random hash function to Merlin, and asks Merlin to send a string which is not only a member of S, but also hashes to a particular randomly chosen string. If S is large then this holds with high probability. The details regarding the size of the hash function etc can be found in [AB09]. Standard statements of the above require that membership in S can be NP, but the same proof works when testing is done in AM, by adding two more rounds to the protocol and using the result that constant round AM protocols are equivalent to 2 round protocols.

If $m > n$ the polynomials are always dependent. If $m < n$, then by Theorem 5.2.4 we can replace the input variables by random linear combinations of $n$ variables. We assume therefore that $m = n$. Let $f_1, \dots, f_n$ be the input polynomials. We continue to use the notation from section 5.1.1, that is $\phi_{\mathbf{f}}$ denotes the polynomial map defined by the input polynomials, and $Y$ is the closure of the image of the map. We also use $D$ to denote $\prod_{i=1}^{n} d_i$, where $d_i = \deg f_i$. Since we are aiming for a protocol that runs in polynomial time, it will be important to address the issues of representing elements of the field, and performing operations on them. To this end, we assume that the inputs $f_i$ have coefficients in the field $\mathbb{F}_q$, and that operations in this field take unit time. We will require operating in a field extension of degree $e$, where $e$ will be fixed later. It suffices to ensure that $e$ is polynomial in the inputs, in which case operations in $\mathbb{F}_{q^e}$ take polynomial time. We will still use $k$ to denote the algebraic closure $\overline{\mathbb{F}_q}$. We will first show some facts about sets of independent and dependent polynomials, then use them to state and prove the formal protocols.

**Independent polynomials:** Suppose polynomials $f_1, \dots, f_n$ are algebraically independent. For $i = 1, \dots, n$ let $A_i$ be the annihilator of the polynomials $f_1, \dots, f_n, x_i$. Each $A_i$ is a polynomial in $n+1$ variables $z_1, \dots, z_{n+1}$. By the Perron bound (Theorem 4.2.2), each $A_i$ has weighted degree $D$.

Let $\mathbf{a}$ be a random point in $\mathbb{A}^n$. Write $A_i$ as a polynomial in the variable $z_{n+1}$ and consider the coefficient of the highest degree term. This is a polynomial in the variables $z_1, \dots, z_n$, call this polynomial $A'_i$. It follows from the degree bound on $A_i$ that the weighted degree of $A'_i$ is at most $D$. Therefore the polynomial $B_i := A'_i(f_1, \dots, f_n)$ has degree at most $D$. Note that this polynomial is nonzero since $\mathbf{f}$ are independent.

Suppose $\mathbf{a}$ is such that $B_i(\mathbf{a})$ is nonzero for all $i$. Then the polynomials $A_i$ are nonzero after specializing the first $n$ variables to $f_i(\mathbf{a})$. Since the equations $A_i(\mathbf{f}(\mathbf{a}), x_i)$ holds on the fibre of $\phi_{\mathbf{f}}(\mathbf{a})$, the different $i^{th}$ coordinates on the fibre is bounded by the degree of this polynomial, which is $D$. The fibre itself therefore has size at most $D^n$.

By the polynomial identity lemma, if $\mathbf{a}$ is picked randomly from $\mathbb{F}_{q^e}$, and if $q^e > D$, then $B_i(a)$ is nonzero with probability $1 - D/q^e$ for a fixed $i$. By a union bound, every $B_i$ is nonzero with probability $1 - nD/q^e$, and in this case the fibre of $\phi_f(a)$ has size at most $D^n$.

**Dependent polynomials:** Let $f_1, \dots, f_n$ be a set of dependent polynomials. Let $A$ be an annihilator of $f_1, \dots, f_n$, which by Theorem 4.2.2 has degree at most $D$. We use $\phi'_f$ to denote the restriction of $\phi_f$ to $\mathbb{F}^n_{q^e}$. Since the coefficients of $\mathbf{f}$ all lie in $\mathbb{F}_{q^e}$, the image of $\phi'_f$ also lies in $\mathbb{F}^n_{q^e}$. Further, every point in the image satisfies the polynomial $A$. Therefore, by the polynomial identity lemma the image has size at most $Dq^{e(n-1)}$. We now crudely bound the number of points $a$ in the domain $\mathbb{F}^n_{q^e}$ such that $\phi'_f(a)$ has fibre of size at most $2D^n$. Suppose $T$ is this set of points. The images of the elements of $T$, namely $\phi'_f(T)$, has size at most $Dq^{e(n-1)}$, since this is a upper bound on the entire image space itself. Each point in $\phi'_f(T)$ can have at most $2D^n$ elements in the fibre by the definition of $T$, whence the size of $T$ is at most $2D^{n+1}q^{e(n-1)}$.

We can now prove the main statements. Informally, AM protocol for dependence is as follows: Arthur picks a random point $a$ in $\mathbb{A}^n$. He then asks Merlin to prove to him that there are at least 2D points in $\mathbb{F}_{q^e}$ all lie in the fibre of $\phi_f(a)$.

**Theorem 5.2.10** ([GSS18]). *Given polynomials $f_1, \dots, f_n$ with coefficients in $\mathbb{F}_q$, there is an AM protocol to check if they are algebraically dependent.*

*Proof of Theorem 5.2.10.* We will use the notation from the above discussion. Set $e$ such that $q^e > 6nD^{n+1}$. This requires $e$ to be polynomial in $\log q$, $n$, $d_i$, which is polynomial in the input size.

Arthur picks a random point $a$ in $\mathbb{F}^n_{q^e}$. If the polynomials are dependent, with probability at least 2/3 the fibre of the point $\phi_f(a)$ has at least 2D elements in $\mathbb{F}^n_{q^e}$. If the polynomials are independent, with probability at least 5/6, the fibre of the point $\phi_f(a)$ has less than D elements. The Goldwasser Sipser protocol Lemma 5.2.9 can therefore be used to provide an interactive AM proof for dependence. Note that membership in the fibre can be checked in P itself. □

Informally, AM protocol for independence is as follows: Arthur asks Merlin to prove that there are more than $2Dq^{e(n-1)}$ points in the image of $\phi'_f$. If Merlin can prove this, Arthur accepts that the polynomials are independent.

**Theorem 5.2.11** ([GSS18]). *Given polynomials $f_1, \dots, f_n$ with coefficients in $\mathbb{F}_q$, there is an AM protocol to check if they are algebraically independent.*

*Proof of Theorem 5.2.11.* As before, set $e$ such that $q^e > 6nD^{n+1}$. If the polynomials are dependent, then there are only $Dq^{e(n-1)}$ points in the image of the map $\phi'_f$.

Suppose the polynomials are independent. Then the number of points in $a \in \mathbb{F}^n_{q^e}$ that are such the fibres of $\phi_f(a)$ have size greater than $D^n$ is at most $Dq^{e(n-1)}$. By a counting argument, the number of points in the image of the map is at least $(q^{en} - Dq^{e(n-1)}/D^n$, assuming the worst case where every other point has fibre of size $D^n$. By our choice of $e$, this is greater than $2Dq^{e(n-1)}$.

The Goldwasser Sipser protocol Lemma 5.2.9 can therefore be used to provide an interactive AM proof for independence. Membership in the set (that is, the images) can be checked in NP, with a point in the preimage acting as the certificate. □

# Chapter 6

# Efficient algorithms for polynomials with low transcendence degree

In this chapter, we give algorithms for the Nullstellensatz and transcendence degree computation that depend on the transcendence degree of the polynomials. When the input polynomials are independent, the complexities of our algorithms match known algorithms, but when the transcendence degree is constant (or logarithmic, with constant degree polynomials), our algorithms perform significantly better.

The workhorse of all our algorithms will be an algorithm by Lakshman and Lazard [LL91] that can check if a variety is zero dimensional, given generators for the ideal. This algorithm itself is fairly nontrivial, and we do not state or prove the correctness of the algorithm, we just use it as a blackbox.

Certain radical membership methods were developed by Gupta [Gup14] in his work on deterministic polynomial identity testing algorithms for restricted depth-four circuits. The focus there however was on a *deterministic algorithm* for the above problem. Further, he restricts his attention to systems where the underlying field is $\mathbb{C}$.

The results of this chapter are from [GS20]. We first state the three main results of this chapter, and provide rough proof sketches. We then prove each of these results.

## 6.1 Main results

Our algorithms will be Monte Carlo algorithms. We assume that our base field $\mathrm{k}$ is algebraically closed, but our algorithms only use operations in the field in which the coefficients of the inputs lie, which we denote by $\mathrm{k}_i$. For example, $\mathrm{k}_i$ might be $\mathbb{F}_p$, and $\mathrm{k}$ would then be $\overline{\mathbb{F}_p}$. By time complexity we mean operations in $k_i$, where operations include arithmetic operations, finding roots, and computing GCD of polynomials. Our results are valid for any field where the above procedures are efficient, for example finite fields.

We have three main results. We relate the complexity of radical membership, and the degree bounds in effective Nullstellensatz, to the transcendence degree of the input set of polynomials. We do this by showing that given a system of polynomials, we can reduce both the number of variables and the number of polynomials to one more than the transcendence degree, while preserving the existence and non-existence of common roots.

Before we state our results, we provide a motivating example. Suppose $f_1, \dots, f_m$ are polynomials in $n$ variables. Suppose further that $h_1, \dots, h_n$ are polynomials in $r$ variables. Then the polynomials $f_1(\mathbf{h}), \dots, f_m(\mathbf{h})$ have transcendence degree $r$. If this $r$ is small, then our algorithms will be faster.

### 6.1.1 Radical membership

Our first result is an improvement in the complexity of radical membership.

**Theorem 6.1.1** (Radical membership)**.** *Suppose* $f_1, \dots, f_m$ *and* $g$ *are polynomials, in variables* $x_1, \dots, x_n$*, of degrees* $d_1, \dots, d_m$ *and* $d_g$ *respectively, given as blackboxes. Suppose that* $\text{trdeg}(f_1, \dots, f_m) \leqslant r$*. Define* $d := \max(\max_i d_i, d_g)$.

*Then testing if* $g$ *belongs to the radical of the ideal generated by* $f_1, \dots, f_m$ *can be done in time polynomial in* $n, m$ *and* $d^r$*, with randomness.*

**Remarks:**

**(1)** The transcendence degree $r$ can be much smaller than $n$, and this improves the complexity significantly to $d^r$ from the prior $d^n$ [LL91]. On the other hand, the usual reduction from SAT to HN results in a set of polynomials with transcendence degree $n$, due to the presence of polynomials $x_i^2 - x_i$ (that enforce the binary 0/1 values). It is therefore unlikely that this complexity can be improved.

**(2)** We also show that the transcendence degree itself can be computed in time $d^r$, independent of the characteristic (Theorem 6.1.3). In the above statement therefore, we can always pick $r = \text{trdeg}(\mathbf{f})$, and we can assume that $r$ is not part of the input.

**(3)** The transcendence degree is upper bounded by the number of polynomials, and therefore we generalize the case of few polynomials. It is surprising if one contrasts this case with that of *ideal* membership— where the instance with three polynomials (i.e. tr.deg=3) is as *hard* as the general instance making it EXPSPACE-complete. For completeness, we present a reduction of the general ideal membership problem to the case of membership with ideals generated by three elements here. This transformation is from [Sap19]. Suppose $g \in \langle f_1, \dots, f_m \rangle$ is an instance of ideal membership. This is equivalent to $z_1^m z_2^m g \in \left\langle z_1^{m+1}, z_2^{m+1}, \sum_i f_i z_1^i z_2^{m-i} \right\rangle$. Here, $z_1, z_2$ are new variables.

### 6.1.2 Nullstellensatz certificates

Next, we show that taking constant-free random linear combinations preserves the zeroset of the polynomials, if the number of linear combinations is at least one more than the transcendence degree. This allows us to get bounds on the Nullstellensatz certificates that depend on the transcendence degree.

**Theorem 6.1.2** (Effective Nullstellensatz)**.** *Suppose* $f_1, \dots, f_m$ *are polynomials in* $x_1, \dots, x_n$*, of degrees* $d_1 \geqslant \cdots \geqslant d_m$ *respectively, with an empty zeroset. Suppose further that* $\operatorname{trdeg}(\mathbf{f}) = r$*.*

*Then, there exist polynomials* $h_i$ *such that* $\deg f_i h_i \leqslant \prod_{i=1}^{r+1} d_i$ *that satisfy* $\sum f_i h_i = 1$.

**Remark:** The prior best degree-bound for the case of small transcendence degree is $\prod_{i=1}^{m} d_i$ [Jel05]. Our bound is significantly better when the transcendence degree $r$ is smaller than the number of polynomials $m$.

### 6.1.3 Computing transcendence degree

Finally, we show that the transcendence degree of a given system of polynomials can be computed in time polynomial in $d^r$ (and $m, n$), where $d$ is the maximum degree of the input polynomials, and $r$ is their transcendence degree. The algorithm is output-sensitive in the sense that the time-complexity depends on the output number $r$.

**Theorem 6.1.3** (Computing transcendence degree)**.** *Given as input polynomials* $f_1, \dots, f_m$*, in variables* $x_1, \dots, x_n$*, of degrees at most* $d$*, we can compute the transcendence degree* $r$ *of the polynomials in time polynomial in* $d^r, n, m$*.*

**Remark:** In the case when the characteristic of the field is greater than $d^r$, there is a much more efficient (namely, randomized polynomial time) algorithm using the Jacobian criterion discussed in the previous chapter [BMS13]. The algorithm presented here is useful when the characteristic is 'small'; whereas the previous best known time-complexity was $> d^{r^2}$ if one directly implements the PSPACE algorithm.

### 6.1.4 Proof ideas

We prove brief proof ideas for each of the above three theorems before providing the complete proofs.

*Proof idea for Theorem 6.1.1:* We first use the Rabinowitsch trick to reduce to HN: the case $g = 1$. Next, we perform a random linear variable reduction. We show that replacing each $x_i$ with a linear combination of $r$ new variables $z_j$ preserves the existence of roots. This is done by using the fact that a general linear hyperplane intersects a variety properly (Corollary 3.1.4). Once we are able to reduce the variables, we can interpolate to get dense representation of our polynomials, and invoke existing results about testing nonemptiness of varieties (Theorem 6.2.1).

*Proof idea for Theorem 6.1.2:* For the second theorem, we show that random linear combinations of the input polynomials, as long as we take at least $r+1$ many of them, preserve the zeroset. For this, we study the image of the polynomial map defined by the polynomials. We will use Corollary 3.1.4 and Theorem 3.2.6 for this. In order to get the degree bounds, we must allow these hyperplanes to depend on fewer variables, and allow their equations to be constant free. Once this is proved, we can use an existing bound on the Nullstellensatz certificates for the new polynomials to obtain a bound for the original polynomials.

*Proof idea for Theorem 6.1.3:* The image of the polynomial map defined by the polynomials is such that the general fibre has codimension equal to the transcendence degree. We first show that a random point, with coordinates from a subset which is not 'too large', satisfies this property. In order to efficiently compute the dimension of this fibre, we take intersections with hyperplanes; and apply Corollary 3.1.4 and Theorem 6.2.1.

## 6.2 Proofs of main results

We will need the following algorithm for checking if a variety has dimension 0. The statement assumes that the polynomials are given in the monomial (also called *dense*) representation. We only state the part of the theorem that we require. We note that the below theorem itself invokes results from [Laz81], section 8 of which proves that the operations occur in a field extension of degree at most $d^n$ of the field $k_i$.

**Theorem 6.2.1.** [LL91, Part of Thm.1] *Let $f_1, \dots, f_m$ be polynomials of degree at most $d$ in $n$ variables. There exists a randomized algorithm that checks if the dimension of the zeroset of $f_1, \dots, f_m$ is 0 or not, in time polynomial in $d^n, m$. The error-probability is $2^{-d^n}$.*

In the special case when $r$ is a constant, we can alternatively use the dimension computation result of [GHS93]. The complexity is slightly worse, but the proof is a bit simpler.

We will also continue to use notation from section 5.1.1, which we recall here. Given polynomials $f_1, \dots, f_m$, we use $\phi_f$ to denote the map $\mathbb{A}^n \to \mathbb{A}^m$ with coordinate functions $f_i$. We use $Y$ to denote the closure of the image of this map.

### 6.2.1 Proof of radical membership

*Proof of Theorem 6.1.1.* We first assume $g = 1$, which is the Nullstellensatz problem HN. Define $D := \prod_{i=1}^m d_i$, and $V := V(\langle \mathbf{f} \rangle)$. The set of common zeroes of these polynomials is the fibre of the point $\mathbf{0}$ under the map $\phi_f$. The problem HN is thus equivalent to testing if a particular fibre of a polynomial map is nonempty. By the fibre dimension theorem (Theorem 2.3.4), the codimension of the zeroset—if it is nonempty—is bounded above by the dimension of the image of the map, which by Lemma 5.1.2 is $r$. The zeroset $V$ is therefore either empty, or has dimension at least $n-r$. Assume that $V$ is nonempty. By repeated applications of Bézout's theorem (Theorem 2.3.5), $\deg V \leqslant D$. Let $S$ be a subset of the underlying field $k_i$ (or an extension) of size at least $6(n-r)D$ that does not contain 0. We can sample from $S$ in time polynomial in $d, n, m$, since $S$ has size exponential in these parameters. Further, if we were required to go to an extension to form $S$, the degree of the extension would be polynomial in $d, n, m$. Pick $n-r$ random linear polynomials $\ell_1, \dots, \ell_{n-r}$ with coefficients from $S$, and call their zero sets $H_1, \dots, H_{n-r}$ respectively. By Corollary 3.1.4, we get $\dim V \cap H_1 \cap H_{n-r} \geqslant 0$ with probability at least 2/3.

Therefore, when the polynomials $\mathbf{f}$ have nonempty zeroset and are restricted to the $r$ dimensional affine subspace $\cap H_i$, the new zeroset has dimension at least 0, and in particular is nonempty. If the zeroset of the polynomials was empty to begin with, then the restriction to the linear subspace also results in an empty zeroset.

This restriction can be performed by a variable reduction, as follows. Treating $\mathbb{A}^n$ as a vector space of dimension $n$ over $k$, let $H_0$ be the linear subspace corresponding to the affine subspace $H := \cap H_i$. The space $H_0$ has dimension $r$, and hence has basis $a_1, \dots, a_r$. Further, let vector $b$ be such that $H = H_0 + b$. Define linear forms $c_1, \dots, c_n$ in new variables $z_1, \dots, z_r$ as $c_i := \sum_{j=1}^{r} a_{ji} z_j + b_i$, where $a_{ji}$ is the $i^{th}$ component of $a_j$. Define $f_i' := f_i(c_1, \dots, c_n)$. Then by construction, the zeroset of $f_1', \dots, f_m'$ is equal to $V \cap (\cap H_i)$. Further, $\deg f_i' = \deg f_i$, and these polynomials are in $r$ variables. Also, the construction of these $f_i'$ can be done in a blackbox manner, given blackboxes for $f_i$. This construction takes time polynomial in $m, r, n$.

We now repeatedly invoke Theorem 6.2.1 to check if $f_i'$s have a common root. First we must convert them to a sparse representation. The polynomial $f_i'$ has at most $\binom{r+d_i}{r}$ many monomials, and therefore we can find every coefficient in time polynomial in $\binom{r+d_i}{r}$ by simply solving a linear system. Applying Theorem 6.2.1, we can test whether the dimension of the zeroset of $f_1', \dots, f_m'$ is 0 or not. However, we want to check if the dimension is at least 0. For this, we randomly sample $r$ more hyperplanes $H_1', \dots, H_r'$ as in the previous part of the proof, this time in the new variables $z_1, \dots, z_r$. Let $V'$ be the zeroset of $f_1', \dots, f_m'$. We first use Theorem 6.2.1 to check if $V'$ has dimension 0. If not, then we check if $V' \cap H_1'$ has dimension 0. If not, then we check $V' \cap H_1' \cap H_2'$, and so on. We return success if any one of the above iterations returns success (implying that the corresponding variety has dimension 0). By Lemma 3.1.2 with high probability each intersection reduces the dimension by 1. If $V'$ originally had dimension $r'$, then after intersecting with $r'$ hyperplanes, the algorithm of Theorem 6.2.1 returns success. If $V'$ was empty, then the algorithm does not return success in any of the above iterations. This allows us to decide if $V'$ has dimension at least 0. Finally, using the fact that the dimension of the zeroset of $f_1', \dots, f_m'$ is at least 0 if and only if $\dim V \geqslant 0$, we get the required algorithm for HN.

We now estimate the time taken. Computing the dense representation takes time polynomial in $d^r$ and $m$. Each of the at most $r$ applications of Theorem 6.2.1 also take the same amount of time. The sampling steps take time polynomial in $\log nD$ (in turn polynomial in $d, m$) and only requires an extension of degree polynomial in $n$ and $\log d$. The total time taken is therefore polynomial in $m, d^r$.

Now assume that $g$ is an arbitrary polynomial. We reduce the problem to the case of $g = 1$ using Rabinowitsch trick [Rab30], as in the proof of the equivalence of Theorem 4.1.1 and Theorem 4.1.2. The polynomial $g$ belongs to the radical of the ideal $\langle \mathbf{f} \rangle$ if and only if the polynomials $\mathbf{f}, 1 - yg$ have no common root (here $y$ is a new variable). Further, if $\mathbf{f}$ have transcendence degree $r$, then the set $\mathbf{f}, 1 - yg$ has transcendence degree $r + 1$. We therefore reduce the radical membership problem to HN problem, with a constant increase in the transcendence degree, number of polynomials and the number of variables. By the result in the previous paragraph, we can solve this in time polynomial in $n, m$ and $d^r$. □

### 6.2.2 Proof of improved Nullstellensatz certificates

We first prove that by taking random linear combinations of the input polynomials, we can reduce the number of polynomials to be one more than the transcendence degree while preserving the existence of roots. This reduction gives degree bounds for the Nullstellensatz certificates. Note that this reduction does not help in the above radical membership procedure, since we will only be saving a factor in $m$ if we reduce the number of polynomials. This theorem can be seen as an extension of [Hei83, Lemma 3].

**Theorem 6.2.2** (Generator reduction)**.** *Let $f_1, \ldots, f_m$ be polynomials in $x_1, \ldots, x_n$ of degrees atmost $d$, and of transcendence degree $r$. Let $g_1, \ldots, g_{r+1}$ be polynomials defined as $g_i := \sum_{j=i}^{m} c_{ij} f_j$, where each $c_{ij}$ is randomly picked from a finite subset $S$ of $k$. Then with probability at least $1 - d^{(r+1)m}/|S|$, we have $V(\langle \mathbf{f} \rangle) = V(\langle \mathbf{g} \rangle)$.*

*Proof of Theorem 6.2.2.* We use $y_1, \ldots, y_m$ to denote the coordinate functions of $\mathbb{A}^m$, the space in which $Y$ lies. By Lemma 5.1.2, $Y$ has dimension $r$ and degree at most $D := (\max_i d_i)^r$. Let $\ell_1, \ldots, \ell_{r+1}$ be the linear polynomials $\ell_i := \sum_{j=i}^{m} c_{ij} y_j$. Further, let $L$ be the map from $\mathbb{A}^m$ to $\mathbb{A}^r$ with coordinate functions $\ell_1, \ldots, \ell_r$, and let $M$ be the map from $\mathbb{A}^m \to \mathbb{A}^{r+1}$ with coordinate functions $\ell_1, \ldots, \ell_{r+1}$.

By Theorem 3.2.6, with probability at least $1 - (r + 1)D/|S|$, the map $L$ is Noether normalizing for $Y$. Suppose this is the case. By Lemma 3.2.4, $L$ when restricted to $Y$ is surjective onto $\mathbb{A}^r$ and every point has finite fibres. Let $Q$ be the fibre of $\mathbf{0}$ under $L$ when restricted to $Y$. We want to bound the size of $Q$. The image $\mathbb{A}^r$ is normal, and hence $|Q|$ is bounded by the degree of the map [SR13, Theorem 2.28]. Here, by the degree of the map we mean the degree of $k(Y)$ over the pullback $L^*(k(\mathbb{A}^r))$. Note that $k(Y) = k(f_1, \ldots, f_m)$. Applying the same isomorphism to $L^*(k(\mathbb{A}^r)) = k(\ell_1, \ldots, \ell_r)$ we get $k(\ell_1(\mathbf{f}), \ldots, \ell_r(\mathbf{f}))$. We therefore need to compute the degree of the field extension $k(f_1, \ldots, f_m)/k(\ell_1(\mathbf{f}), \ldots, \ell_r(\mathbf{f}))$, which is algebraic since the extension

$k[f_1, \dots, f_m]/k[\ell_1(\mathbf{f}), \dots, \ell_r(\mathbf{f})]$ is integral. By Perron's bound, for each $i$ there exists an annihilator of $f_i, \ell_1(\mathbf{f}), \dots, \ell_r(\mathbf{f})$ of degree at most $d^{r+1}$. The degree of the extension, and hence $|Q|$, is bounded by $d^{(m+1)r}$.

No point of $Q$ other than $\mathbf{0}$ has all of the last $m - r$ coordinates as zero. This follows from the fact that $L^{-1}(\mathbf{0})$ is a linear space of dimension $m - r$, and its intersection with $y_{r+1} = y_{r+2} = \cdots = y_m = 0$ has dimension 0. Consider now the linear form $\ell_{r+1}$. For every $\mathbf{0} \neq q \in Q$, the probability that $\ell_{r+1}(q) = 0$ is at most $1/|S|$. Therefore, with probability at least $1 - d^{(r+1)m}/|S|$, the polynomial $\ell_{r+1}$ is nonzero on every nonzero point of $Q$.

Consider the polynomials $g_1, \dots, g_{r+1}$, and let $\phi_{\mathbf{g}}$ be the polynomial map $\mathbb{A}^n \to \mathbb{A}^{r+1}$ with coordinate functions $g_i$. By the choice of $\ell_i$ in the previous paragraph, the map $\phi_{\mathbf{g}}$ is exactly the composition of the map $\phi_{\mathbf{f}} : \mathbb{A}^n \to \mathbb{A}^m$ with $M : \mathbb{A}^m \to \mathbb{A}^{r+1}$. Let $Q$ be as defined earlier, the fibre of $\mathbf{0}$ under $L$. By construction, the set $M^{-1}(\mathbf{0})$ is a subset of $Q$. But since the polynomial $\ell_{r+1}$ is nonzero on every nonzero point of $Q$, the set $M^{-1}(\mathbf{0})$ consists only of $\mathbf{0}$. Therefore, $\phi_{\mathbf{f}}^{-1}(M^{-1}(\mathbf{0})) = \phi_{\mathbf{f}}^{-1}(\mathbf{0})$. Since $\phi_{\mathbf{g}} = M \circ \phi_{\mathbf{f}}$ we get $\phi_{\mathbf{g}}^{-1}(\mathbf{0}) = \phi_{\mathbf{f}}^{-1}(\mathbf{0})$; which is the same as $V(\langle \mathbf{f} \rangle) = V(\langle \mathbf{g} \rangle)$. The probability bound follows from a union bound. □

That we pick the linear combinations so that the first involves all polynomials, the second involves all except $f_1$, the third involves all except $f_1, f_2$ and so on is crucial for the improvement in the degree bounds for the Nullstellensatz certificates. We now prove the second main result of the chapter.

*Proof of Theorem 6.1.2.* Using Theorem 6.2.2, there exists polynomials $g_1, \dots, g_{r+1}$ of degrees $d_1, \dots, d_{r+1}$ that are linear combinations of $f_1, \dots, f_m$ that do not have a common root. By Theorem 4.2.1, there exist $h'_1, \dots, h'_{r+1}$ with $\deg g_i h'_i \leqslant \prod_{i=1}^{r+1} d_i$ such that $\sum g_i h'_i = 1$. In this equation, substituting back $f_1, \dots, f_m$ for each $g_i$ we get the equation $\sum f_i h_i = 1$ with the required degree bound. □

### 6.2.3 Algorithm for computing transcendence degree

We now give an algorithm to compute the transcendence degree. For this, we use the effective version of the fibre dimension theorem.

**Lemma 6.2.3.** *Let $h_1, \dots, h_m$ be polynomials of degree at most $d$ in $n$ variables, and let $W$ be the Zariski closure of the image of the map $\mathbf{h}$ with coordinates $h_i$. Let $S \subset k$ be of size $6nd^n$. If $a_1, \dots, a_n$ are randomly picked from $S$, then with probability at least $5/6$, the fibre of $(h_1(\mathbf{a}), \cdots, h_m(\mathbf{a}))$ has codimension exactly $\dim W$.*

*Proof.* First assume that the $h_i$ are algebraically independent. Then $W = \mathbb{A}^m$. Let the input variables be labelled such that $x_1, \dots, x_{n-m}, h_1, \dots, h_m$ are algebraically independent, and let $A_j(z_0, z_1, \dots, z_{n-m}, w_1, \dots, w_m)$ be the (minimal) annihilator of $x_j$ over this set of variables, that is $A_j(x_j, x_1, \dots, x_{n-m}, h_1, \dots, h_m) = 0$. By the proof of the fibre dimension theorem (Theorem 2.3.4), a sufficient condition for point $a_1, \dots, a_n$ to be such that $\mathbf{h}(\mathbf{a})$ has fibre of dimension exactly $n - m$ is that $A_j(x_j, x_1, \dots, x_{n-m}, h_1(\mathbf{a}), \dots, h_m(\mathbf{a}))$ is a nonzero polynomial. The polynomial $A_j$, when treated as polynomials in variables $z_0, \ldots, z_{n-m}$ with coefficients in $k[w_1, \ldots, w_m]$ are such that the leading monomial has coefficient a polynomial in $w_1, \ldots, w_m$ of weighted-degree at most $\prod_{i=1}^{m} d_i$ (by Theorem 4.2.2). By the polynomial identity lemma if we pick each $a_i$ randomly from a set of size $6 \prod_{i=1}^{m} d_i$ then, with probability at least 5/6, none of the polynomials $A_j(x_j, x_1, \ldots, x_{n-m}, h_1(\mathbf{a}), \ldots, h_m(\mathbf{a}))$ is zero. In this case, the codimension of the fibre of $\mathbf{h}(\mathbf{a})$ is exactly $m$.

In the general case, the $h_i$ may be algebraically dependent, and $W$ is a subvariety of $\mathbb{A}^m$. Suppose $\dim W = \operatorname{trdeg}((\mathbf{h})) =: s$. Then we take $s$ many random linear combinations $g_i$ of the $h_i$, as in the proof of Theorem 6.1.1. The map defined by the $g_i$ is dense in $\mathbb{A}^s$ and therefore the $g_i$ ($i \in [s]$) are algebraically independent. By the previous paragraph, point $\mathbf{a}$ picked coordinatewise from $S$ is such that the fibre of $\mathbf{g}(\mathbf{a})$ has codimension $s$. The fibre of $\mathbf{h}(\mathbf{a})$ is a subset of the fibre of $\mathbf{g}(\mathbf{a})$, and therefore it has codimension at least $s$. Finally, by the fibre dimension theorem (Theorem 2.3.4) the fibre has codimension at most $s$, whence the fibre of $\mathbf{h}(\mathbf{a})$ has codimension $s$. □

We can now use this to compute the transcendence degree.

*Proof of Theorem 6.1.3.* For each $i$ from 1 to $n$, we do the following steps. We iterate till $i$ reaches the transcendence degree $r$ of the $m$ polynomials. In the $i$-th iteration, we intersect $\mathbb{A}^n$ with $n-i$ random hyperplanes $\ell_1, \ldots, \ell_{n-i}$, as in the proof of Theorem 6.1.1. Here, the coefficients are picked from a set $S$ of size at least $n \cdot 18 \prod_{i=1}^{m} d_i$. We therefore reduce the problem to $i$ variables.

Randomly pick point $\mathbf{a}$ where each coordinate (of the $n$ many) is picked from $S$. By Lemma 6.2.3, with error-probability $\leqslant 1/6n$, the point $\mathbf{f}(\mathbf{a})$ has intersected fibre of dimension $(n-r)-(n-i) = (i-r)$. We need to check this algorithmically, which is done by interpolating the polynomials $\mathbf{f}$ after hyperplane intersections, and then using Theorem 6.2.1. If the intersected fibre dimension is zero, we have certified that the transcendence degree is $r$; so we halt and return $i$ as output. If not, we move to the next $i \mapsto i+1$. The interpolation step above is performed by solving a linear system which has size polynomial in $d^i$ which is the count of the monomials of degree at most $d$ in $i$ variables.

For $i < r$, with probability $\leqslant 1/6n$, the fibre of $\mathbf{f}(\mathbf{a})$ has an empty intersection with $\ell_1, \ldots, \ell_{n-i}$ and hence gets verified by Theorem 6.2.1. By a union bound therefore, with error-probability $\leqslant 1/6$, the above algorithm gives the correct answer. For each $i$, the time complexity of the above steps is polynomial in $d^i, m$, which is the time taken for the interpolation step and to verify zero-dimension of the fibre. Therefore the algorithm as a whole takes time polynomial in $d^r, n, m$ as claimed. □

# Chapter 7

# Conclusion

In this thesis, we first provided exposition of some results in commutative algebra and algebraic geometry, namely Noether normalization, hyperplane intersection and the Nullstellensatz. We proved effective versions of most of the results we discussed.

We proved degree bounds in existence statements such as the Nullstellensatz. We also explicitly controlled bad choices when picking random hyperplanes to intersect a variety, and random linear maps to Noether normalize.

We then discussed the algebraic independence problem, and framed it as a problem in computational algebraic geometry. We used this view to give alternative proofs of a number of known results. Finally, we used all of the above to give improved algorithms for radical membership and transcendence degree computation, and improved bounds for Nullstellensatz certificates in the special case of polynomials with low transcendence degree.

There are some natural directions in which the above can be extended, which we list.

- We can try to further improve the dependency on the transcendence degree in some of the above algorithms. For example, our algorithms are polynomial time when either the transcendence degree of the polynomials is constant, or when the transcendence degree is logarithmic and the degrees of the polynomials are constant. We can look for algorithms that are polynomial time when the transcendence degree is logarithmic irrespective of the degree. There is some evidence that the above results cannot be greatly improved, since the Nullstellensatz problem is NP hard.
- Another natural problem is proving that the Nullstellensatz is in the polynomial hierarchy in the finite characteristic case. This result holds (assuming the GRH) in $\mathbb{Z}$, but the methods use do not extend to fields of finite characteristic.
- Finally, another open problem is to come up with a randomized polynomial time algorithm for computing the transcendence degree. This problem is unlikely to be NP hard, since it is in $\mathsf{AM} \cap \mathsf{coAM}$, and given that

the characteristic 0 case is in BPP, it seems likely that there exists a BPP algorithm for the finite field case.

# Bibliography

[AB09] Sanjeev Arora and Boaz Barak. *Computational complexity: a modern approach*. Cambridge University Press, 2009. 38

[AM94] M.F. Atiyah and I.G. MacDonald. *Introduction To Commutative Algebra*. Addison-Wesley series in mathematics. Avalon Publishing, 1994. 18

[BCS97] Peter Bürgisser, Michael Clausen, and Mohammad Amin Shokrollahi. *Algebraic complexity theory*, volume 315 of *Grundlehren der mathematischen Wissenschaften*. Springer, 1997. 33, 55

[BMS13] M. Beecken, J. Mittmann, and N. Saxena. Algebraic independence and blackbox identity testing. *Information and Computation*, 222:2 – 19, 2013. (Also, 38th International Colloquium on Automata, Languages and Programming, ICALP 2011). 43

[Bro87] W. Dale Brownawell. Bounds for the degrees in the Nullstellensatz. *Annals of Mathematics*, 126(3):577–591, 1987. 26

[CLO07] David A. Cox, John Little, and Donal O'Shea. *Ideals, Varieties, and Algorithms: An Introduction to Computational Algebraic Geometry and Commutative Algebra, 3/e (Undergraduate Texts in Mathematics)*. Springer-Verlag, Berlin, Heidelberg, 2007. 4, 22

[Csa75] L. Csanky. Fast parallel matrix inversion algorithms. *16th Annual Symposium on Foundations of Computer Science (SFCS 1975)*, pages 11–12, 1975. 25

[DFGS91] Alicia Dickenstein, Noaï Fitchas, Marc Giusti, and Carmen Sessa. The membership problem for unmixed polynomial ideals is solvable in single exponential time. *Discrete Applied Mathematics*, 33(1-3):73–94, 1991. 25

[DL78] Richard A. Demillo and Richard J. Lipton. A probabilistic remark on algebraic program testing. *Information Processing Letters*, 7(4):193 – 195, 1978. 4

[DSS18] Pranjal Dutta, Nitin Saxena, and Amit Sinhababu. Discovering the roots: Uniform closure results for algebraic classes under factoring.

In *Proceedings of the 50th Annual ACM SIGACT Symposium on Theory of Computing*, STOC 2018, pages 1152–1165, New York, NY, USA, 2018. ACM. 55

[Eis13] David Eisenbud. *Commutative Algebra: with a view toward algebraic geometry*, volume 150. Springer Science & Business Media, 2013. 4, 5

[Gat13] Andreas Gathmann. Commutative algebra. *Class Notes TU, Kaiserslautern*, 14, 2013. 18, 22

[GHS93] Marc Giusti, Joos Heintz, and Juan Sabia. On the efficiency of effective Nullstellensätze. *Computational Complexity*, 3:56–95, 1993. 44

[GS20] Abhibhav Garg and Nitin Saxena. Algorithms for blackbox radical membership, nullstellensatz and tr.deg. 2020. Accepted at ISSAC 2020. 1, 41

[GSS18] Zeyu Guo, Nitin Saxena, and Amit Sinhababu. Algebraic dependencies and pspace algorithms in approximative complexity. In *Proceedings of the 33rd Computational Complexity Conference*, CCC '18, 2018. 27, 37, 39

[Gup14] Ankit Gupta. Algebraic geometric techniques for depth-4 PIT & Sylvester-Gallai conjectures for varieties. *Electronic Colloquium on Computational Complexity (ECCC)*, 21:130, 2014. 41

[Hei83] Joos Heintz. Definability and fast quantifier elimination in algebraically closed fields. *Theoretical Computer Science*, 24(3):239 – 277, 1983. 10, 11, 46

[Her26] Grete Hermann. Die frage der endlich vielen schritte in der theorie der polynomideale. *Mathematische Annalen*, 95(1):736–788, Dec 1926. 25

[Jel05] Zbigniew Jelonek. On the effective Nullstellensatz. *Inventiones mathematicae*, 162(1):1–17, Oct 2005. 26, 43

[Kay09] N. Kayal. The complexity of the annihilating polynomial. In *24th Annual IEEE Conference on Computational Complexity*, pages 184–193, July 2009. 33, 34

[Koi96] Pascal Koiran. Hilbert's Nullstellensatz is in the polynomial hierarchy. *J. Complexity*, 12(4):273–286, 1996. 25

[Kol88] János Kollár. Sharp effective Nullstellensatz. *Journal of the American Mathematical Society*, 1(4):963–975, 1988. 26

[KPS99] Teresa Krick, Luis Miguel Pardo, and Martín Sombra. Arithmetic Nullstellensätze. *ACM SIGSAM Bulletin*, 33(3):17, 1999. 26

[KPS$^+$01] Teresa Krick, Luis Miguel Pardo, Martín Sombra, et al. Sharp estimates for the arithmetic Nullstellensatz. *Duke Mathematical Journal*, 109(3):521–598, 2001. 26

[Kru50] Wolfgang Krull. Jacobsonsches radikal und Hilbertscher Nullstellensatz. In *Proceedings of the International Congress of Mathematicians, Cambridge, Mass*, volume 2, pages 56–64, 1950. 1

[Lan02] Serge Lang. *Graduate Texts in Mathematics: Algebra*. Springer, 2002. 4, 31

[Laz77] Daniel Lazard. Algèbre linéaire sur $k[x\_1, \dots, x\_n]$ et élimination. *Bulletin de la Société Mathématique de France*, 105:165–190, 1977. 26

[Laz81] Daniel Lazard. Resolution des systemes d'equations algebriques. *Theoretical Computer Science*, 15(1):77 – 110, 1981. 44

[LL91] Y. N. Lakshman and Daniel Lazard. *On the Complexity of Zero-dimensional Algebraic Systems*, pages 217–225. Birkhäuser Boston, 1991. 41, 42, 44

[Mac02] Francis Sowerby Macaulay. Some formulae in elimination. *Proceedings of the London Mathematical Society*, 1(1):3–27, 1902. 26

[Mit12] Johannes Mittmann. *Independence in Algebraic Complexity Theory*. PhD thesis, Hausdorff Center for Mathematics, Bonn, 2012. 34

[MM82] Ernst W Mayr and Albert R Meyer. The complexity of the word problems for commutative semigroups and polynomial ideals. *Advances in Mathematics*, 46(3):305 – 329, 1982. 25

[Mul12] K. D. Mulmuley. Geometric complexity theory v: Equivalence between blackbox derandomization of polynomial identity testing and derandomization of noether's normalization lemma. In *2012 IEEE 53rd Annual Symposium on Foundations of Computer Science*, pages 629–638, 2012. 30

[Ore22] Øystein Ore. Über höhere kongruenzen. *Norsk Mat. Forenings Skrifter*, 1(7):15, 1922. 4

[Per51] Oskar Perron. I, Die Grundlagen. In *Algebra*, 1951. 27

[PSS16] Anurag Pandey, Nitin Saxena, and Amit Sinhababu. Algebraic independence over positive characteristic: New criterion and applications to locally low algebraic rank circuits. In *41st International Symposium on Mathematical Foundations of Computer Science, MFCS 2016, August 22-26, 2016 - Kraków, Poland*, pages 74:1–74:15, 2016. (Comput.Compl., 27(4), 617–670, 2018). 35, 36, 37

[Pł05] Arkadiusz Płoski. Algebraic dependence of polynomials after O.Perron and some applications. *Computational Commutative and Non-Commutative Algebraic Geometry*, pages 167–173, 2005. 27

[Rab30] JL Rabinowitsch. Zum Hilbertschen Nullstellensatz. *Mathematische Annalen*, 102(1):520–520, 1930. 23, 46

[Sap19] Ramprasad Saptharishi. Private Communication, 2019. 42

[Sch80] J. T. Schwartz. Fast probabilistic algorithms for verification of polynomial identities. *J. ACM*, 27(4):701–717, October 1980. 4

[Sin19] Amit Kumar Sinhababu. *Power series in complexity: Algebraic Dependence, Factor Conjecture and Hitting Set for Closure of VP*. PhD thesis, Indian Institute of Technology Kanpur, 2019. 30

[Som97] Martín Sombra. Bounds for the Hilbert function of polynomial ideals and for the degrees in the Nullstellensatz. *Journal of Pure and Applied Algebra*, 117-118:565 – 599, 1997. 26

[Som99] Martín Sombra. A sparse effective Nullstellensatz. *Advances in Applied Mathematics*, 22(2):271 – 295, 1999. 26

[SR13] I.R. Shafarevich and M. Reid. *Basic Algebraic Geometry 1: Varieties in Projective Space*. SpringerLink : Bücher. Springer Berlin Heidelberg, 2013. 4, 8, 9, 18, 20, 21, 33, 36, 37, 46

[Vak17] Ravi Vakil. The rising sea: foundations of algebraic geometry. 2017. 22

[Zip79] Richard Zippel. Probabilistic algorithms for sparse polynomials. In *Proceedings of the International Symposiumon on Symbolic and Algebraic Computation*, EUROSAM '79, page 216–226, Berlin, Heidelberg, 1979. Springer-Verlag. 4

# Appendix A

# Functional dependence and Newton iteration

In this appendix, we look at a proof of part of the functional dependence criterion via Newton Iteration. This vastly simplifies the proof. It also gives an idea of why the random shift is necessary, and the role of the inseparability.

Assume that we are given polynomials $f_1, \ldots, f_n \in k[\mathbf{x}]$. Assume that the $\mathbf{f}$ have transcendence degree $n$. Let $g$ be any polynomial in $\mathbb{F}[\mathbf{x}]$. We know that the transcendence degree of $\{\mathbf{f}, g\}$ will also be $n$, and thus $g$ depends algebraically on the $\mathbf{f}$. Let $A$ be an annihilator of $\{\mathbf{f}, g\}$. By definition, $A(\mathbf{f}, g) = 0$. We will also look at $A$ as a polynomial in one variable, say $y$, which is the variable in which we plug in $g$. We assume that $A$ is an annihilator with minimum degree in $y$. We will now show that after a random shift, we can write $g$ as a power series in $\mathbf{f}$.

We use the following formulation of Newton iteration from [DSS18], which is a slight modification of Theorem 2.31 from [BCS97]. For completeness, we provide a proof of this lemma in the last section.

**Lemma A.0.1** (Newton Iteration)**.** *Let $F(\mathbf{x}, y) \in k[[\mathbf{x}]][y]$ be a polynomial in $y$ with coefficients power series in $k[[\mathbf{x}]]$. Suppose $\mu$ is such that $F(\mathbf{0}, \mu) = 0$ and also $F'(\mathbf{0}, \mu) \neq 0$, where $F'$ is the derivative with respect to the last variable. There is then a unique element $Y \in k[[\mathbf{x}]]$ with constant term $\mu$ such that $F(\mathbf{x}, Y) = 0$. We also have*

$$y_{t+1} = y_t - \frac{F(\mathbf{x}, y_t)}{F'(\mathbf{x}, y_t)},$$

*such that* $Y \equiv y_t \pmod{\langle \mathbf{x} \rangle^{2^t}}$.

In the above lemma, it is essential that $F$ is a polynomial in $y$. This allows us to evaluate it at power series with non-zero constant term. In general, if $F$ were a power series in $y$, we would require $\mu = 0$.

In order to get $g$ as a power series in $\mathbf{f}$, we will try and use NI to get a power series in $\mathbf{x}$, and then try and argue that what we get is actually a power series in $\mathbf{f}$. First note that if $g$ depended inseparably on $\mathbf{f}$, then $A'$ would be an identically zero polynomial. In this case, we will not be able to satisfy the conditions of the lemma. Thus we assume that $g$ depends separably on $\mathbf{f}$. In general, this can be obtained by replacing $g$ by a power $g^{p^i}$. Therefore we can assume now that $g$ depends separably on $\mathbf{f}$.

A possibly bigger issue is the fact that if the $f_i$ have non-zero constant terms, then power series in $f_i$ are not valid elements in $\mathbb{F}[[\mathbf{x}]]$. To fix this, we apply the shift operator, to remove the constant term: Define $\mathcal{H}f_i := f_i(\mathbf{x}+\mathbf{z})-f_i(\mathbf{z})$, and similarly for $\mathcal{H}g = g(\mathbf{x}+\mathbf{z})-g(\mathbf{z})$. For now we treat the $\mathbf{z}$ as part of the base field, that is, we switch from working with $k$ to working with $k(\mathbf{z})$. Eventually we show that we can replace $\mathbf{z}$ by an arbitrary element from $k^n$, and the proof will continue to hold. We have $A(\mathcal{H}\mathbf{f}+\mathbf{f}(\mathbf{z}), \mathcal{H}g+g(\mathbf{z})) = A(\mathbf{f}(\mathbf{x}+\mathbf{z}), g(\mathbf{x}+\mathbf{z})) = 0$. We define $B(\mathbf{x}, y) = A(\mathcal{H}\mathbf{f}+f(\mathbf{z}), y+g(\mathbf{z})) = A(\mathbf{f}(\mathbf{x}+\mathbf{z}), y+g(\mathbf{z}))$. The polynomial $B$ has root $y = \mathcal{H}g$. Now note that $B(\mathbf{0}, 0) = A(\mathbf{f}(\mathbf{z}), g(\mathbf{z})) = 0$, since $A$ is an annihilator[1] . Further, consider $B'(\mathbf{0}, 0)$. We have

$$B = \sum_{i=0}^{d} c_i (y+g(\mathbf{z}))^i,$$

where the $c_i$ are polynomials in $\mathbf{x}$ and $\mathbf{z}$, and $d$ is the degree with respect to $y$. Differentiating, we get

$$B' = \sum_{i=0}^{d} ic_i (y+g(\mathbf{z}))^{i-1}.$$

When evaluated at $(\mathbf{0}, 0)$, each of the $c_i$ is a polynomial in $f_i(\mathbf{z})$. Therefore, $B'(\mathbf{0}, 0)$ is a polynomial in $f_i(\mathbf{z})$ and $g(\mathbf{z})$, of degree $d-1$. As a polynomial in $\mathbf{z}$, this is non-zero: if it were not, we would have an annihilator for $\mathbf{f}, g$ of degree $d-1$ in $y$, contradicting the assumption that $A$ is the annihilator with minimum $y$ degree. In general, when we replace $\mathbf{z}$ with a vector of random elements from $\mathbb{F}$, we can still say that $B'(\mathbf{0}, 0) \neq 0$ (for most choices), by using the polynomial identity lemma.

We have now satisfied the conditions of the lemma. The lemma then gives us a root $Y \in k[[\mathbf{x}]]$ such that $B(\mathbf{x}, Y) = 0$. Further, this is the unique root with constant term 0. But we know that $\mathcal{H}g$ is also a root of $B(\mathbf{x}, y)$ with constant term 0. Thus it must be that $Y = g$. All that is left to show is that we can actually get $Y$ as a power series in $\mathcal{H}\mathbf{f}$, since the lemma only promises us a power series in $\mathbf{x}$. For this, we look at the series $y_t$ whose limit is $Y$. We will inductively show that $y_t$ can be written as a power series in $\mathcal{H}\mathbf{f}$ for all $t$.

The base case is $t = 0$. We have $y_0 = 0$, and thus vacuously $y_0$ is a power series in $\mathcal{H}\mathbf{f}$. Assume inductively that $y_t$ is a power series in $\mathcal{H}\mathbf{f}$. First consider $B(\mathbf{x}, y_t) = A(\mathcal{H}\mathbf{f}+f(\mathbf{z}), y_t+g(\mathbf{z}))$. The first argument, $\mathcal{H}\mathbf{f}+f(\mathbf{z})$ is vacuously

[1] The choice of setting $\mu = 0$ is motivated by the fact that we know that the root $\mathcal{H}g$ has no constant term. We also know that this is not a repeated root, due to minimality and separability assumption. The calculation of $B(\mathbf{0}, 0)$ and $B'(\mathbf{0}, 0)$ thus also act as a sort of sanity check.

a power series in $\mathcal{H}\mathbf{f}$, and by the inductive hypothesis, so is the second argument $y_t + g(\boldsymbol{z})$. Thus $B(\mathbf{x}, y_t)$ is also a power series in $\mathcal{H}\mathbf{f}$. Now consider $(B'(\mathbf{x}, y_t))^{-1}$. The term $B'(\mathbf{x}, y_t)$ is a power series in $\mathcal{H}\mathbf{f}$ by an argument similar to the one above. In this form $B'(\mathbf{x}, y_t)$ must have a nonzero constant term, since the constant term will be exactly $B'(\mathbf{0}, 0)$, which is non-zero by assumption. Thus we have $B'(\mathbf{x}, y_t) = c_0 + D(\mathcal{H}\mathbf{f})$, where $c_0 \neq 0$, and $D$ is a power series with no constant term. But then we have

$$\begin{aligned}\frac{1}{B'(\mathbf{x}, y_t)} &= \frac{1}{c_0 + D(\mathcal{H}\mathbf{f})} \\ &= \frac{1}{c_0}\frac{1}{1 - D_1(\mathcal{H}\mathbf{f})} && \text{(Setting } D_1 = -D/c_0) \\ &= \frac{1}{c_0}\left(1 + D_1(\mathcal{H}\mathbf{f}) + D_2(\mathcal{H}\mathbf{f})^2 + \ldots\right)\end{aligned}$$

This converges since each $D_1(\mathcal{H}\mathbf{f})^i$ has $x$-adic valuation at least $i$. It is also a power series in $\mathcal{H}\mathbf{f}$. The product $B(\mathbf{x}, y_t)(B'(\mathbf{x}, y_t))^{-1}$ is thus also a power series in $\mathcal{H}\mathbf{f}$, and so is $y_{t+1}$. Note that $c_0$ is a non-zero element in $\mathbb{F}(\boldsymbol{z})$, and by Schwartz Zippel, it continues to remain non-zero after we replace $\boldsymbol{z}$ by random field elements. It is crucial that the term $c_0$ is independent of $t$, since otherwise the random choice of $\boldsymbol{z}$ would have had to be such that a countable number of equations are non-zero. This completes the proof.

We now prove the version of Newton iteration used.

*Proof of Lemma A.0.1.* In order to see the existence of $Y$, we plug in a power series with unknown coefficients, equate with zero, and compare coefficients on both sides. This gives us a system of linear equations, with unknowns corresponding to monomials, and equations also corresponding to monomials. In particular, let $Y = \sum c_{\boldsymbol{e}}\mathbf{x}^{\boldsymbol{e}}$ where the sum runs over all $\mathbb{N}$ valued vectors of length $n$. We will first show that $c_{\mathbf{0}} = \mu$ satisfies the equation corresponding to the constant term. Then we will use the $y_t$ described in the statement of the lemma, to get coefficients $c_{\boldsymbol{e}}$ in the following way: we will look at some $y_t$, and use the coefficients of monomials up to degree $2^t$ as the values for the corresponding variables in our system. We will show that these satisfy the equations corresponding to the monomials of degree at most $2^t$. Note that these equations do not have any other variables. This is equivalent to showing that $F(\mathbf{x}, y_t) \equiv 0 \pmod{\langle \mathbf{x} \rangle^{2^t}}$. When showing that the $y_t$ satisfy these equations, we will additionally show that the values for the variables that we already had from $y_{t-1}$, namely those for the coefficients of degree at most $2^{t-1}$, are the same as those in $y_{t-1}$. More succinctly, we will show that $y_t \equiv y_{t-1} \pmod{\langle \mathbf{x} \rangle^{2^{t-1}}}$. As hinted, the proof will proceed by induction on $t$.

First we show the base case, namely $t = 0$. Consider the equation $F(\mathbf{x}, Y) = 0$. The constant term in this expression is $F(\mathbf{0}, c_{\mathbf{0}})$. By assumption, since $F(\mathbf{0}, \mu) = 0$, we can set $c_{\mathbf{0}} = \mu$. This also ensures we satisfy the requirement of our $Y$ having constant term $\mu$. In the notation of the question, we also get $y_0 = \mu$. The statement about equality of coefficients holds vacuously.

Assume now that the statement holds for $t$. First note that $F(x, y_t) \equiv F(x, y_0) \pmod{\langle x \rangle}$, since $y_t \equiv y_0 \pmod{\langle x \rangle}$ by the induction hypothesis. This implies that $F'(x, y_t)$ has constant term $F'(0, \mu)$, which is non-zero by assumption. This implies that $F'(x, y_t)$ is invertible in the power series ring, and that the expression for $y_{t+1}$ is well defined. Further, by induction, we have that $F(x, y_t) \equiv 0 \pmod{\langle x \rangle^{2^t}}$. This implies that $y_{t+1} - y_t \equiv 0 \pmod{\langle x \rangle^{2^t}}$, proving the consistency requirement. Now we compute $P(x, y_{t+1})$. For this, we will use the Taylor expansion. We have

$$\begin{aligned} F(x, y_{t+1}) &= F\left(x, y_t - \frac{F(x, y_t)}{F'(x, y_t)}\right) \\ &= F(x, y_t) + \frac{F'(x, y_t)}{1!}\left(-\frac{F(x, y_t)}{F'(x, y_t)}\right) + \frac{F''(x, y_t)}{2!}\left(-\frac{F(x, y_t)}{F'(x, y_t)}\right)^2 + \dots \end{aligned}$$

On the right hand side, the first two summands cancel. All other summands, and hence the entire right hand side, are $0 \pmod{\langle x \rangle^{2^{t+1}}}$. This shows that $y_{t+1}$ has the required property.

Finally we must show that $Y$ is unique. This follows from the fact that $\mu$ is not a repeated root of $F(0, y)$. □