Morphisms of Rings and Applications to Complexity

A Thesis Submitted
in Partial Fulfilment of the Requirements
for the Degree of
Doctor of Philosophy

by
Nitin Saxena

to the

DEPARTMENT OF COMPUTER SCIENCE & ENGINEERING
INDIAN INSTITUTE OF TECHNOLOGY KANPUR

June, 2006

CERTIFICATE

Certified that the work contained in the thesis entitled "*Morphisms of Rings and Applications to Complexity*", by "*Nitin Saxena*", has been carried out under my supervision and that this work has not been submitted elsewhere for a degree.

(Dr. Manindra Agrawal)
Professor,
Department of Computer Science & Engineering,
Indian Institute of Technology,
Kanpur.

June, 2006

To my parents
and
members of my family

Synopsis

One of the main goals of theoretical computer science is to understand the complexity of various problems. This work mainly focuses on problems that are of algebraic flavor but are related to problems in number theory and graph theory. This thesis builds a framework that gives new insights into the complexity of various seemingly unrelated open problems and also derandomizes some problems that were previously known to have efficient but randomized solutions.

The framework that this thesis keeps alluding to is that of the *morphisms* of finitely presented *rings*. Rings are fundamental algebraic objects with associated natural operations of addition and multiplication. A morphism is a map from a ring R_1 to a ring R_2 such that it preserves the underlying ring operations of addition and multiplication. An *automorphism* of a ring is a bijective morphism from the ring to itself. An *isomorphism* from a ring R_1 to another ring R_2 is a bijective morphism from R_1 to R_2. We begin with defining general morphism problems of rings and then move on to specific applications.

The ring morphism problems that we study are – deciding whether a ring has a nontrivial automorphism (RA), deciding whether there is an isomorphism between two given rings (RI); finding a nontrivial ring automorphism (FRA), finding a ring isomorphism (FRI); computing the number of automorphisms of a given ring (#RA), computing the number of isomorphisms between two given rings (#RI); testing whether a given map is a ring automorphism (TRA), testing whether a given map is a ring isomorphism (TRI). A study of these problems, when the rings are finite and are given in the *basis representation*, shows that none of these can be NP-hard (unless the polynomial hierarchy collapses) but they can be harder than some well-known problems – like, graph isomorphism, polynomial equivalence, integer factoring and

polynomial factoring.

Next, we show an interesting connection of the isomorphism problem of rings to the problem of polynomial equivalence. Given two polynomials $f(\overline{x})$, $g(\overline{x}) \in \mathbb{F}[x_1, \ldots, x_n]$, polynomial equivalence is the problem of checking whether there is a linear transformation $\tau \in \mathbb{F}^{n \times n}$ such that $f(\tau\overline{x}) = g(\overline{x})$. In most of the cases this problem easily reduces to the ring isomorphism problem. More interestingly, we show that the isomorphism problem for finite dimensional commutative $\mathbb{F}$-algebras (rings defined over a field $\mathbb{F}$) reduces to solving the equivalence problem for cubic forms (homogeneous polynomials of degree 3). Since we have shown that graph isomorphism reduces to commutative $\mathbb{F}$-algebra isomorphism, this means that graph isomorphism reduces to cubic forms equivalence over any field $\mathbb{F}$. This can be taken as a new way of attacking graph isomorphism or as an evidence to the structural hardness of cubic forms equivalence.

Next, we apply the properties of rings to solve a special case of the identity testing problem. Given an arithmetic circuit $\mathcal{C}(x_1, \ldots, x_n)$, the identity testing problem is to check whether $\mathcal{C} \equiv 0$ in time polynomial in the size of the circuit $\mathcal{C}$. There is an efficient randomized algorithm for identity testing since a long time but there has been very little progress on the derandomization front. The difficulty of derandomizing the identity testing problem was partly explained in 2003 by showing that such a derandomization would imply proving lower bounds. In this work we assume that $\mathcal{C}$ is a depth 3 circuit with bounded top fanin and give the first deterministic polynomial time algorithm for identity testing in this case. The algorithm can be viewed as solving a special case of the ring isomorphism problem and is based on the philosophy that polynomials over *local* rings imitate the properties of polynomials over a field.

Finally, we apply the framework of rings to attack a famous problem – primality testing. Primality testing is the problem of checking whether a given number n is prime and the algorithm should take time polynomial in the number of input bits $\log n$. Prior to this work various *randomized* algorithms were known for primality testing but the challenge was to eliminate the use of randomness. Here we consider the cyclotomic ring $R := (\mathbb{Z}/n\mathbb{Z})[x]/(x^r - 1)$ and study its Frobenius-like map

$\sigma_n : a(x) \mapsto a(x)^n$. We show that if σ_n is an automorphism of R then we get strong conditions on n. This study culminates with the AKS algorithm – the first deterministic polynomial time algorithm for primality testing.

Acknowledgements

I am greatly indebted to Manindra Agrawal for his advise, mentoring and collaboration. His clarity of thoughts, insights and simplicity in exposition will remain inspirational to me. I thank him and his family for the care and the dinners they provided me over all these years.

IIT Kanpur and the Department of Computer Science has been my home since my undergraduate days. I am grateful to the institute and my professors for creating this wonderful environment. A special thanks go to Somenath Biswas, Sumit Ganguly and Pankaj Jalote for their teaching and encouragement in all these years. I thank Infosys Technologies Limited for funding my graduate studies.

I would like to thank my colleague and friend Neeraj Kayal with whom I did this research. His enthusiasm and clear thinking were contagious. I learnt a lot from him.

Thanks to Hendrik W. Lenstra for various illuminating discussions. Some of his observations and questions guided my thesis in the right direction.

I am grateful to Bernard Chazelle and Princeton University for hosting me in 2003-04. I am also thankful to P. S. Thiagarajan and National University of Singapore for hosting me in 2004-05. Thanks to all the people in Princeton and NUS with whom I interacted and who made my visit memorable. A special thanks to Ankur Dhanik, Shien Jin Ong, Hemalnam Rathod and Comandur Seshadhri.

Thanks to all my friends and fellow post-graduate students in the department for the discussions and the food. A special thanks to my office-mates Atul Gupta and Vibhu Saujanya Sharma for all the bull-sessions.

There are many others, whose names I cannot continue listing, who have helped in my development as a person and a researcher. I express my sincere gratitude to

them all.

Finally, I would like to thank my family members for supporting me and my decisions. My grandfather, parents, Nalin and Gauri provided the shelter conditions under which this work could take place: thanks to them for this and many other things.

Contents

Chapter 1

Introduction

The primary goal of the study of computation is to ascertain *upper* and *lower bounds* for the complexity of a specific problem at hand. It was realized early on that the assumption of randomness, i.e., ability to toss coins, usually helps in upper bounding the complexity of problems. However, it is widely believed that the use of *randomness* in algorithms is dispensable. This belief is supported by the recent results showing equivalence between lower bounds and derandomizations. All these years of computer science research have had considerable successes in finding upper bounds or efficient (maybe randomized) algorithms for many computational problems but not much is known about proving lower bounds or getting general derandomizations.

This work focuses on the various *natural* problems of algebraic flavor that are not known to be in P but are not believed to be NP-hard. Some of these problems are known to have randomized polynomial time algorithms and there are others that do not even have randomized subexponential time algorithms yet. A better understanding of these problems of intermediate complexity would hopefully give us new insights into NP computation. We give upper and lower bounds for these problems as close as we can and also derandomize some problems that were previously known to have only randomized algorithms.

The mathematical object that keeps recurring in this thesis is a *ring*. Rings are algebraic structures with addition and multiplication operations defined on them

and in all the applications we give in this thesis they are commutative with unity. Study of the morphisms of these rings and the computational variants of morphism problems forms the core of the next four chapters of this thesis. It was remarked by Lenstra [Len04]: *One has the strong feeling that essentially ANY problem in mathematics can be 'hidden' in a finite local commutative ring!* This work shows that indeed many computational problems of intermediate complexity reduce to questions of rings.

1.1 The Framework

The framework in this thesis constitutes of algebraic structures called rings and computational problems defined on rings. A *ring* is a set R equipped with two binary operations $+$ and $\cdot$, called addition and multiplication, such that (a, b, c are general elements in R):

1. $(R, +)$ is an *abelian group* with identity element 0, i.e., R satisfies the following properties:
   - Associativity: $(a + b) + c = a + (b + c)$
   - Commutativity: $a + b = b + a$
   - Identity: $0 + a = a + 0 = a$
   - Inverse: $\forall a\ \exists(-a)$ such that $a + -a = -a + a = 0$
2. $(R, \cdot)$ is a *monoid* with identity element 1 also called the *unity*, i.e., R satisfies the following properties:
   - Identity: $1 \cdot a = a \cdot 1 = a$
   - Associativity: $(a \cdot b) \cdot c = a \cdot (b \cdot c)$
3. Multiplication *distributes* over addition, i.e., R satisfies the following properties:
   - $a \cdot (b + c) = (a \cdot b) + (a \cdot c)$

- $(a + b) \cdot c = (a \cdot c) + (b \cdot c)$

If the multiplication operation satisfies commutativity then R is called a *commutative ring*. If $(R \setminus \{0\}, \cdot)$ is an abelian group too then R becomes a *field*.

Example $R_0 := (\mathbb{Z}/n\mathbb{Z}, +, \cdot)$ is a ring, it is a field iff n is prime. $R_1 := R_0[x]/(x^r - 1)$ is a commutative ring but never a field for $r > 1$. The set $R_2 := \left\{A \mid A \in R_0^{2\times 2}\right\}$ is a noncommutative ring under matrix addition and multiplication in R_0. ∎

1.1.1 Ring Representations

Normally, in this work we express commutative rings in the form:

$$R = (\mathbb{Z}/n\mathbb{Z})[x_1, \ldots, x_k]/(f_1, \ldots, f_m)$$

where, $f_1, \ldots, f_m \in (\mathbb{Z}/n\mathbb{Z})[x_1, \ldots, x_k]$ are multivariate polynomials. This notation means that all the polynomials $\sum_{i=1}^{m} g_i f_i$ – where, $g_1, \ldots, g_m \in (\mathbb{Z}/n\mathbb{Z})[x_1, \ldots, x_k]$ – are zero in the ring R. This representation of rings, called the **polynomial representation**, is very convenient but in the computational problems that we define on rings we will need a more verbose way of representing rings in the input. We will consider the following two ways of presenting a ring R:

Table Representation: Here, we assume that ring R has finitely many elements, say s, and provide two $s \times s$ addition and multiplication tables, thus defining R completely.

Basis Representation: Here, ring R can be infinite but it should be finite *dimensional*, i.e. the additive group of R should be decomposable as:

$$(R, +) \;\cong\; (R_1, +) \oplus \cdots \oplus (R_n, +) \tag{1.1}$$

where $R_1, \ldots, R_n$ are special rings, namely, $\mathbb{Z}$, $\mathbb{Z}/m\mathbb{Z}$, or a field. Thus, there are 'basis' elements $b_1, \ldots, b_n \in R$ such that $(R, +) = (R_1, +)b_1 \oplus \cdots \oplus (R_n, +)b_n$ and, hence, to describe R it is sufficient to give the products $b_i \cdot b_j$ as a 'linear' combination of b_k's.

In the basis representation of a ring R if the component rings of the additive group are fields, say $R_1 = \cdots = R_n = \mathbb{F}$, then R is called an $\mathbb{F}$-*algebra*.

Example Consider the ring $R := \mathbb{Q}[x]/(x^2 - x + 1)$. Here, 1 and x can be taken as basis elements and $(R, +) = \mathbb{Q} \cdot 1 \oplus \mathbb{Q} \cdot x$. Multiplication on the basis elements is defined as: $1 \cdot 1 = 1 \cdot 1 + 0 \cdot x,\ 1 \cdot x = x \cdot 1 = 0 \cdot 1 + 1 \cdot x$ and $x \cdot x = (-1) \cdot 1 + 1 \cdot x$. Also, note that R is a 2 dimensional commutative $\mathbb{Q}$-algebra. ∎

Note that the basis representation is more compact as it can represent a ring of size s in $O(\log^4 s)$ space whereas table representation requires $\Theta(s^2 \log s)$ space. This exponential compactness of basis representation as compared to the table representation suggests that the complexity of problems of rings would be different for these two different representations.

In much of this thesis we will assume that the rings, whenever given as input to an algorithm, are in the basis representation and the groups are in terms of generators.

1.1.2 Ring Morphisms

A *homomorphism* ϕ from a ring R to S is a map that preserves addition and multiplication operations, i.e., for all $a, b \in R$:

- $\phi(a + b) = \phi(a) + \phi(b)$ in the ring S.
- $\phi(a \cdot b) = \phi(a) \cdot \phi(b)$ in the ring S.

A bijective homomorphism from ring R to S is called an *isomorphism*. A bijective homomorphism from ring R to itself is called an *automorphism*. Observe that to specify a homomorphism on a ring, given in the basis representation, it is enough to specify the images of the basis elements together with a description of the homomorphism on the component rings $R_1, \ldots, R_n$ in Equation (1.1).

Example Let $R := \mathbb{F}_p[x]/(x^2)$. Then the map $\phi : 1 \mapsto 1,\ x \mapsto 0$ is a homomorphism from R to $\mathbb{F}_p$. The map $\phi : 1 \mapsto 1,\ x \mapsto ax$ (where $a \in \mathbb{F}_p \setminus \{0\}$) is an automorphism of R. ∎

Study of automorphisms of fields has been very fruitful in understanding field extensions. It was Galois who initiated this study and subsequently showed that the roots of a general quintic polynomial cannot be expressed in terms of radicals. In this work we study computational aspects of automorphism and isomorphism problems of rings.

1.2 Our Contributions

Our contributions are twofold:

1) We study the complexity of problems related to computing ring morphisms and relate it to the complexities of some well-known problems.

2) We design efficient algorithms for solving certain special cases of morphism problems which, in turn, yield efficient algorithms for some well-known problems.

1.2.1 Complexity of Ring Morphism Problems

The computational problems of ring automorphisms that we study in this thesis are: the *ring automorphism* problem (RA) to determine whether a given ring has nontrivial automorphisms, the *finding ring automorphism* problem (FRA) to find a nontrivial automorphism of a given ring, the *counting ring automorphisms* problem (#RA) to compute the number of automorphisms of a given ring, and the *testing ring automorphism* problem (TRA) to test whether a given map is an automorphism of a given ring. Similarly, the computational problems of ring isomorphisms that we study in this thesis are: the *ring isomorphism* problem (RI) to determine whether two given rings are isomorphic or not, the *finding ring isomorphism* problem (FRI) to find an isomorphism between two given rings, the *counting ring isomorphisms*

problem (#RI) to compute the number of isomorphisms between two given rings, and the *testing ring isomorphism* problem (TRI) to test whether a given map is an isomorphism between two given rings.

This work shows that for finite rings given in the basis representation all these problems are low for the polynomial hierarchy and, hence, are unlikely to be NP-hard. We also lower bound the complexity of these problems by giving reductions from well known problems of intermediate complexity, namely, graph isomorphism, polynomial equivalence, integer factoring and polynomial factoring.

Graph Isomorphism: The problem is to determine whether two given graphs are isomorphic. This is a fundamental open problem with no efficient algorithm known yet. Schoning [Sch88] showed that this problem is unlikely to be NP-hard. Using group-theoretic ideas, an algorithm was given by Luks [Luk82] that works in polynomial time for graphs of bounded degree. This work shows that graph isomorphism reduces to #RA, RI, FRI and #RI.

Polynomial Equivalence: Given two polynomials $f,\ g$ the problem is to determine whether there is a linear transformation that when applied on the variables of f makes it equal to g. Not much is known about this problem (see [Har75, Pat96]) except that it is unlikely to be NP-hard over finite fields. We show that most of the cases of this problem reduce to #RA, RI, FRI and #RI. More interestingly, the ring isomorphism problem for finite dimensional commutative $\mathbb{F}$-algebras reduces to cubic forms equivalence. This, as a corollary, gives us that the graph isomorphism problem reduces to testing equivalence of cubic forms over *any* field.

Integer Factoring: Given a composite number n the problem is to find a nontrivial factor. There is no efficient algorithm known but the algorithms used in practice are based on the number field sieve [LL93] and elliptic curves [Len87]. The best known algorithm is conjectured to run in expected $2^{O(\log^{\frac{1}{3}} n \log\log^{\frac{2}{3}} n)}$ time. This is a longstanding open problem that is of both theoretical and practical interest. We show that integer factoring reduces to all of FRA, #RA, FRI and #RI.

Polynomial Factoring: Given a univariate polynomial over a finite field the problem is to find a nontrivial factor. There are randomized polynomial time algorithms known, for example, Berlekamp's [Berl70]. Also, a deterministic subexponential algorithm was given by Ronyai [Ron88] assuming the extended Riemann Hypothesis (ERH). We show that polynomial factoring deterministically reduces to FRA assuming ERH.

1.2.2 Efficient Algorithms for the Special Cases

Using the framework of rings we solve the problem of Identity Testing for depth 3 arithmetic circuits of bounded top fanin and the problem of Primality Testing.

Identity Testing: Given an arithmetic circuit $\mathcal{C}$ the problem is to check whether $\mathcal{C} \equiv 0$. The first randomized efficient algorithm was given by Schwartz, Zippel [Sch80, Zip79] and no deterministic polynomial time algorithm is known yet. Impagliazzo and Kabanets [IK03] showed that derandomizing identity testing would mean proving lower bounds. In this work we solve a special case of the ring isomorphism problem that consequently gives the first deterministic polynomial time algorithm for the case of depth 3 circuits ($\Sigma\Pi\Sigma$ circuits) having a bounded top fanin. We view the problem of identity testing for $\Sigma\Pi\Sigma$ circuits of bounded top fanin as a special case of the ring isomorphism problem in the polynomial representation. We utilise the nice structure of this special case to give a recursive solution invoking the properties of commutative local rings.

Primality Testing: The problem is to determine whether a given number n is prime or not. Several randomized polynomial time primality tests are there ([Mil76, Rab80, SoS77]). A deterministic subexponential time algorithm was given by Adleman, Pomerance and Rumely [APR83]. In this work we view the problem of primality testing as a special case of testing whether a given map is an automorphism of a given ring (recall the TRA problem) and eventually give the first deterministic polynomial time primality test. The ring in this case is the cyclotomic ring: $R :=$

$(\mathbb{Z}/n\mathbb{Z})[x]/(x^r - 1)$ and the map is the Frobenius map σ_n that sends any element $a(x) \in R$ to $a(x)^n$.

1.3 Organization of the Thesis

The results to be presented in this thesis first appeared in the following five papers: [AKS04, KS05, AS05, AS06, KS06]. This thesis expands on these published results and gives a self-contained treatment based on the framework of ring automorphism and isomorphism problems. For an alternative treatment of primality testing and identity testing, and the full proof of RA $\in$ P we refer the reader to the manuscript [Kay06].

Chapter 2 studies the various morphism problems of rings, inspired from the graph isomorphism problem, and gives upper and lower bounds for their complexity. Connections are shown to graph isomorphism, integer factoring and polynomial factoring. This chapter deals with *finite rings*.

Chapter 3 discusses the problem of polynomial equivalence. The emphasis is on the equivalence problem of cubic forms and its relation to the isomorphism problems of $\mathbb{F}$-algebras and graphs. It also studies the cubic forms that we construct out of $\mathbb{F}$-algebras. This chapter deals with *finite dimensional commutative rings*.

Chapter 4 solves a special case of ring isomorphism that immediately yields an identity test for $\Sigma\Pi\Sigma$ arithmetic circuits of bounded top fanin. The chapter also has some new $\Sigma\Pi\Sigma$ identities that are of high rank. This chapter deals with *local rings*.

Finally, the AKS algorithm for primality testing and the related results are discussed in Chapter 5 using the ring automorphism framework. This chapter deals with *cyclotomic rings*.

The basic notions of complexity theory and rings are given in chapter 2 and the appendix with brief proofs. A familiarity with rings would be very helpful to the reader in understanding most of the thesis.

Chapter 2

The Ring Morphism Problems

A *ring* consists of a set of elements together with addition and multiplication operations. These structures are fundamental objects of study in mathematics and particularly so in algebra and number theory. It has long been recognized that the group of automorphisms of a ring provides valuable information about the structure of the ring. Galois [Gal] initiated the study of the group of automorphisms of a field and it was later applied by Abel [Ros95] to prove the celebrated theorem that there does not exist any formula for finding the roots of a quintic (degree 5) polynomial. However, to the best of our knowledge, the computational complexity of the ring isomorphism and automorphism related problems has not been investigated so far. In this chapter, we initiate such a study and show interesting connections to some well known problems.

In this chapter we will restrict our attention to *finite* rings. We show that the ring isomorphism problems are of intermediate complexity but are hard in the sense that well-known problems of graph isomorphism and integer factoring reduce to them.

The results of this chapter mostly appear in [KS05].

2.1 Basics of Groups and Rings

A *group* is a set of elements with a suitably defined operation of multiplication while a *ring* is a set of elements with two operations of addition ($+$) and multiplication ($\cdot$)

defined. There are two useful groups living in a ring R. Firstly, $(R, +)$ is a group with respect to addition called the *additive group*. If R^* is the set of elements in R having multiplicative inverse then $(R^*, \cdot)$ is the second group called the *multiplicative group*.

2.1.1 Representing Rings

For concreteness we first fix the way we are going to present the finite rings and their homomorphisms in the input or the output.

Definition 2.1 Basis representation of rings: *A finite ring R is given by first describing its additive group in terms of n additive generators and then specifying multiplication by giving for each pair of generators, their product as an element of the additive group. More concretely, R is presented as:*

$$(R, +, .) := \langle (d_1, d_2, d_3, \cdots, d_n), ((a_{i,j,k}))_{1 \leq i,j,k \leq n} \rangle$$

where, for all $1 \leq i, j, k \leq n,\ 0 \leq a_{i,j,k} < d_k$ and $a_{i,j,k} \in \mathbb{Z}$.

This specifies a ring R generated by n elements $b_1, b_2, \cdots b_n$ with each b_i having additive order d_i and $(R, +) = (\mathbb{Z}/d_1\mathbb{Z})b_1 \oplus (\mathbb{Z}/d_2\mathbb{Z})b_2 \cdots \oplus (\mathbb{Z}/d_n\mathbb{Z})b_n$. Moreover, multiplication in R is defined by specifying the product of each pair of additive generators as an integer linear combination of the generators: for $1 \leq i, j \leq n$, $b_i \cdot b_j = \sum_{k=1}^{n} a_{i,j,k} b_k$.

Definition 2.2 Representation of maps on rings: *Suppose R_1 is a ring given in terms of its additive generators $b_1, \ldots, b_n$ and ring R_2 given in terms of $c_1, \ldots, c_n$. In this chapter maps on rings would invariably be homomorphisms on the additive group. Then to specify any map $\phi : R_1 \to R_2$, it is enough to give the images $\phi(b_1), \ldots, \phi(b_n)$. So we represent ϕ by an $n \times n$ matrix of integers A, such that for all $1 \leq i \leq n$:*

$$\phi(b_i) = \sum_{j=1}^{n} A_{i,j} c_j$$

and for all $1 \leq i, j \leq n,\ 0 \leq A_{i,j} <$ additive order of c_j.

Example Consider the ring $R := (\mathbb{Z}/3\mathbb{Z})[x]/(x^2 - x + 1)$. Here, 1 and x can be taken as basis elements and $(R, +) = (\mathbb{Z}/3\mathbb{Z}) \cdot 1 \oplus (\mathbb{Z}/3\mathbb{Z}) \cdot x$. Multiplication on the basis elements is defined as: $1 \cdot 1 = 1 \cdot 1 + 0 \cdot x$, $1 \cdot x = x \cdot 1 = 0 \cdot 1 + 1 \cdot x$ and $x \cdot x = 2 \cdot 1 + 1 \cdot x$. Note that the map ϕ sending $1 \mapsto 1$ and $x \mapsto -1$ is a homomorphism from R to itself and with respect to the basis $\{1, x\}$ it can be represented as: $A = \begin{pmatrix} 1 & 0 \\ 2 & 0 \end{pmatrix}$. ∎

2.1.2 The Problems

Now we define the ring isomorphism and related problems that we are going to explore.

- The *ring automorphism problem* is to decide whether a given ring has a nontrivial ring automorphism. If we let $Aut(R)$ denote the group of automorphisms of a ring R then the language corresponding to the ring automorphism problem is:

$$\text{RA} := \{R \mid R \text{ is a ring in basis form s.t. } \#Aut(R) > 1\}$$

- The *ring isomorphism problem* is to decide whether two given rings are isomorphic. The corresponding language we define as:

$$\text{RI} := \{(R_1, R_2) \mid \text{ rings } R_1, R_2 \text{ are given in the basis form and } R_1 \cong R_2\}$$

- FRA is the functional problem of *computing a nontrivial automorphism* of a ring R given in the basis form.

- FRI is the functional problem of *computing an isomorphism* (if one exists) between two rings given in basis form.

- #RA is defined as the functional problem of *computing the number of automorphisms* of a given ring. Its decision version can be viewed as the language:

$$\text{cRA} := \{(R, k) \mid R \text{ is a ring in basis form s.t. } \#Aut(R) \geq k\} \tag{2.1}$$

- #RI is defined as the functional problem of *computing the number of isomorphisms* between two rings given in the basis form.

- *Testing ring automorphism* is the problem of deciding whether a given map is an automorphism of a ring given in basis form. The corresponding language we define as:

$$TRA := \{(R, \phi) \mid R, \phi \text{ are given in basis form and } \phi \in Aut(R)\}$$

Remark: If the map is given as a circuit $\mathcal{C}$ computing the value of ϕ then the problem of primality testing becomes a special case of TRA where $R = (\mathbb{Z}/n\mathbb{Z})[x]/(x^r - 1)$ and $\phi : a(x) \mapsto a(x)^n$ (see chapter 5). ∎

- *Testing ring isomorphism* is the problem of deciding whether a given map is an isomorphism between two rings given in basis form. The corresponding language we define as:

$$TRI := \left\{(R_1, R_2, \phi) \mid R_1, R_2, \phi \text{ are given in basis form and } R_1 \overset{\phi}{\cong} R_2\right\}$$

2.1.3 The Preliminaries

If G, H are two groups then we use $H \leq G$ to denote that H is a subgroup of G. For a finite group G: $H \leq G$ implies that $\#H$ divides $\#G$. The converse does not hold in general but if for a prime p, $p^k | \#G$ then there always exist a subgroup of size p^k. If p^k is the highest power of p dividing $\#G$ then a subgroup of size p^k is called a *p-Sylow subgroup* of G. A p-Sylow subgroup S_p of size p^k can be broken into a *composition series*, i.e., there are groups G_i of size p^{k-i} such that:

$$S_p = G_0 \ > \ G_1 \ > \ G_2 > \ldots \ > \ G_k = \{1\}.$$

In analysing a ring R we use special subgroups of $(R, +)$ called *ideals*.

Definition 2.3 *A subset $I \subseteq R$ is an* ideal *of R if:*

- *$(I, +)$ is a subgroup of $(R, +)$, and*

- *for all $i \in I,\ r \in R$, both $i \cdot r$ and $r \cdot i$ are in I. This can also be stated as: $\forall r \in R$ both $r \cdot I,\ I \cdot r \subseteq I$.*

Ideals can be multiplied together to give new (smaller) ideals.

Definition 2.4 *Let $\mathcal{I}, \mathcal{J}$ be two ideals of a ring R. We define their product as:*

$$\mathcal{I} \cdot \mathcal{J} := \text{ ring generated by the elements } \{ij \mid i \in \mathcal{I}, j \in \mathcal{J}\}$$

Powering of ideals, $\mathcal{I}^t$ for positive integer t, is defined similarly. It is easy to see that $\mathcal{I} \cdot \mathcal{J}$ is again an ideal of R.

Algebraic structures mostly break into simpler objects. In the case of rings we get the following simpler rings. This is discussed in more detail in the appendix.

Definition 2.5 Indecomposable or Local ring: *A ring R is said to be indecomposable or local if there do not exist rings R_1, R_2 such that $R \cong R_1 \times R_2$, where $\times$ denotes the natural composition of two rings with component wise addition and multiplication.*

Commutative local rings have nice properties (see [McD74]). For instance, if R is a finite commutative local ring then for all $r \in R$ either r is invertible or r is a *nilpotent* i.e., $\exists k,\ r^k = 0$. This makes $\mathcal{M} := R \setminus R^*$ an ideal of R and it can be shown that $\mathcal{M}$ is the *unique maximal ideal* of R.

Example Let $n = p^2q$ where $p,\ q$ are distinct primes and define a natural ring $R := (\mathbb{Z}/n\mathbb{Z}, +, \cdot)$. Then observe that R decomposes as $(\mathbb{Z}/p^2\mathbb{Z}, +, \cdot) \times (\mathbb{Z}/q\mathbb{Z}, +, \cdot)$ where the two *component* rings are local. ∎

Example Consider a ring $R := \mathbb{F}[x, y]/(x^3, y^2)$. The subset yR, denoted as (y), is an ideal of R. Similarly, $xR + yR$, denoted by (x, y), is also an ideal of R. Note that the product of these two ideals is $(y) \cdot (x, y) = (xy, y^2) = (xy)$. Similarly in R,

$(x,y)^2 = (x^2, xy)$, $(x,y)^3 = (x^2y)$ and $(x,y)^4 = 0$. Moreover, it can be shown that R is a local ring with $\mathcal{M} = (x,y)$ as its unique maximal ideal. ∎

Example It is an interesting exercise to show that $R_1 := \mathbb{F}[x,y]/(x^3, y(x+y))$ is a nonzero local ring while $R_2 := \mathbb{F}(y)[x]/(x^3, y(x+y))$ is the zero ring, where, $\mathbb{F}(y)$ denotes a rational function field. ∎

We collect some of the known results about groups and rings. Their proofs can be found in algebra texts, e.g., [McD74, Lang].

There is a classification known for finite commutative groups. Basically, each such group completely decomposes into a bunch of *cyclic* groups.

Proposition 2.1 [Structure theorem for abelian groups] *If R is a finite ring then its additive group $(R,+)$ can be uniquely (up to permutations) expressed as:*

$$(R,+) \cong \bigoplus_i (\mathbb{Z}/{p_i}^{\alpha_i}\mathbb{Z})$$

where p_i's are primes (not necessarily distinct) and $\alpha_i \in \mathbb{Z}^{\geq 1}$.

Remark: This theorem can be used to check in polynomial time whether for two rings, given in basis form, the additive groups are isomorphic or not. Suppose the two additive groups are $G := (\mathbb{Z}/d_1\mathbb{Z}) \oplus \cdots \oplus (\mathbb{Z}/d_n\mathbb{Z})$ and $G' := (\mathbb{Z}/d'_1\mathbb{Z}) \oplus \cdots \oplus (\mathbb{Z}/d'_n\mathbb{Z})$. Consider the set $D = \{d_i \mid \ i \in [n]\} \cup \{d'_i \mid \ i \in [n]\}$. We take *gcd*s of all pairs of integers from the set D and expand D in each such *gcd*-operation as: if $\alpha,\ \beta \in D$ have a nontrivial *gcd* then replace them by $\frac{\alpha}{gcd(\alpha,\beta)}$, $\frac{\beta}{gcd(\alpha,\beta)}$ and $gcd(\alpha,\beta)$. We can keep repeating this process on the new expanded D till all the elements of D become mutually coprime. It is guaranteed to stop in polynomial time, for D can expand to a maximum size of $\log(\#G \cdot \#G')$ as the number of prime factors of a number N are less than $\log N$. Now factor d_i 's and d'_j 's as much as possible using the numbers from D. Say, $d_i = d_{i,1}^{e_1} \cdots d_{i,k}^{e_k}$ where $d_{i,1}, \ldots, d_{i,k} \in D$ are mutually coprime. We can refine the decomposition of G by breaking $(\mathbb{Z}_{d_i}, +)$ as:

$$(\mathbb{Z}/d_{i,1}^{e_1}\mathbb{Z}) \oplus \cdots \oplus (\mathbb{Z}/d_{i,k}^{e_k}\mathbb{Z}).$$

At the end of all this refining of d_i's and d'_j's using D, let the *finer* structural decompositions be: $G \cong (\mathbb{Z}/m_1\mathbb{Z}) \oplus \cdots \oplus (\mathbb{Z}/m_{n'}\mathbb{Z})$ and $G' \cong (\mathbb{Z}/m'_1\mathbb{Z}) \oplus \cdots \oplus (\mathbb{Z}/m'_{n'}\mathbb{Z})$. Now by invoking the structure theorem: G will be isomorphic to G' if and only if the *multi-sets* (i.e. elements with repetition) $\{m_i\}_{i\in[n']}$ and $\{m'_i\}_{i\in[n']}$ are equal. ∎

Using the structure theorem of abelian groups, we can compute $\#Aut(R,+)$ of a ring R presented in terms of additive generators having prime-power additive orders.

Proposition 2.2 *Given a ring R in terms of additive generators, all having prime-power additive orders, we can compute the number of automorphisms of the additive group of R, $\#Aut(R,+)$, in polynomial time.*

Proof: Automorphisms of the additive group $(R,+)$ are nothing but the invertible linear maps on the additive generators of R. Thus, to compute $\#Aut(R,+)$ we compute the number of invertible linear maps or the number of invertible matrices.

Let $(R,+)$ be given as $\cong \bigoplus_{i=1}^{l} \bigoplus_j (\mathbb{Z}/{p_i}^{\alpha_{i,j}}\mathbb{Z})$, where p_i's are distinct primes and $\alpha_{i,j} \in \mathbb{Z}^{\geq 1}$. For $1 \leq i \leq l$ define subrings R_i of R as:

$$R_i := \{r \in R \mid r \text{ has power-of-}p_i \text{ additive order}\}$$

Observe that

$$R \cong R_1 \times \cdots \times R_l$$

this is because if $r_i \in R_i$ and $r_j \in R_j$ $(i \neq j)$ then for some $c_i, c_j \in \mathbb{Z}^{\geq 0}$, $p_i^{c_i} r_i r_j = p_j^{c_j} r_i r_j = 0$ which implies that $r_i r_j = 0$ (since $\exists a, b \in \mathbb{Z}$ such that $a p_i^{c_i} + b p_j^{c_j} = 1$) and by a similar argument $r_1 \in R_1, \ldots, r_l \in R_l$ are *linearly independent*.

This decomposition of R gives us:

$$\#Aut(R,+) = \prod_{i=1}^{l} \#Aut(R_i,+)$$

Thus, it suffices to show how to compute $\#Aut(R,+)$ when $(R,+)$ is given as $\cong \bigoplus_{i=1}^{n} (\mathbb{Z}/p^{\alpha_i}\mathbb{Z})$ where p is a prime and $\alpha_i \in \mathbb{Z}^{\geq 1}$.

Suppose we are given R in terms of the following additive basis:

$$\begin{aligned}(R,+) &= (\mathbb{Z}/p^{\beta_1}\mathbb{Z})e_{1,1} \oplus \ldots \oplus (\mathbb{Z}/p^{\beta_1}\mathbb{Z})e_{1,n_1} \oplus \ldots \\ &\quad \ldots \oplus (\mathbb{Z}/p^{\beta_m}\mathbb{Z})e_{m,1} \oplus \ldots \oplus (\mathbb{Z}/p^{\beta_m}\mathbb{Z})e_{m,n_m}\end{aligned}$$

where, $n_1 + \ldots + n_m = n$ and $1 \leq \beta_1 < \ldots < \beta_m$.

Observe that $\phi \in Aut(R,+)$ iff the matrix A describing the map ϕ is invertible (mod p) and preserves the additive orders of $e_{i,j}$'s. Our intention is to count the number of all such matrices A. To do that let us see how A looks:

$$A = \begin{pmatrix} B_{1,1} & B_{1,2} & \ldots & B_{1,m} \\ B_{2,1} & B_{2,2} & \ldots & B_{2,m} \\ \vdots & \ldots & \ddots & \vdots \\ B_{m,1} & B_{m,2} & \ldots & B_{m,m} \end{pmatrix}_{n \times n}$$

where the block matrices $B_{i,j}$'s are integer matrices of size $n_i \times n_j$. The properties of these block matrices which make A describe an automorphism of $(R,+)$ are:

- for $1 \leq j < i \leq m$: entries in $B_{i,j}$ are from $\{0, 1, \ldots, p^{\beta_j} - 1\}$.
- for $1 \leq i \leq m$: entries in $B_{i,i}$ are from $\{0, 1, \ldots, p^{\beta_i} - 1\}$ and $B_{i,i}$ is invertible (mod p).
- for $1 \leq i < j \leq m$: entries in $B_{i,j}$ are from $\{0, 1, \ldots, p^{\beta_j} - 1\}$ and $B_{i,j} \equiv 0 \pmod{p^{\beta_j - \beta_i}}$.

It is not difficult to see that the number of matrices satisfying these conditions can be found in time polynomial in $(n_1\beta_1 + \ldots + n_m\beta_m)(\log p)$, and hence the number of A's which describe an automorphism of $(R,+)$. ∎

Remark: When a ring R is given in terms of generators having composite additive orders then computing $\#Aut(R,+)$ entails factoring integers. For example, suppose $n = pq$ where $p \neq q$ are primes and ring R is given as $(\mathbb{Z}/n\mathbb{Z}, +, \cdot)$. Then $\#Aut(R,+) = (p-1)(q-1) = \phi(n)$ and if we compute $\phi(n)$ then we can factorize n in randomized polynomial time (see [Mil76]). ∎

Unlike commutative groups, a classification of commutative rings is not known yet. But as a first step rings can be decomposed uniquely into indecomposable rings.

Proposition 2.3 [Structure theorem for rings] *If R is a finite ring then it uniquely (up to permutations) decomposes into indecomposable rings $R_1, \ldots, R_s$ such that*

$$R \cong R_1 \times \cdots \times R_s$$

Proof: Refer the appendix. ∎

Remark: Decomposition of a finite commutative ring R can be found in polynomial time given oracles to integer and polynomial factorizations (discussed at length in the appendix). Observe that any commutative ring R with characteristic n can be expressed as:

$$R \;\cong\; (\mathbb{Z}/n\mathbb{Z})[x_1, \ldots, x_m]/(f_1(\overline{x}), \ldots, f_\ell(\overline{x}))$$

where $\overline{x} = (x_1, x_2, \ldots, x_m)$ and $f_1, \ldots, f_\ell$ are polynomials in $x_1, \ldots, x_m$ capturing the multiplicative relations in the ring R. The above expression hints that if we can factor n into its prime factors and polynomials into irreducible factors then we can effectively factor ring R into its indecomposable components. ∎

Example Consider the ring $R := (\mathbb{Z}/p^2q^3\mathbb{Z})[x, y]/(x^4, px, y^2 - y)$. By factoring the characteristic p^2q^3 we get:

$$R \cong (\mathbb{Z}/p^2\mathbb{Z})[x, y]/(x^4, px, y^2 - y) \times (\mathbb{Z}/q^3\mathbb{Z})[x, y]/(x^4, px, y^2 - y)$$

Further, by factoring $y^2 - y$ into *coprime* irreducibles over the respective local rings in x we get:

$$\begin{aligned} R \;\cong\; & (\mathbb{Z}/p^2\mathbb{Z})[x, y]/(x^4, px, y) \times (\mathbb{Z}/p^2\mathbb{Z})[x, y]/(x^4, px, y - 1) \\ & \times (\mathbb{Z}/q^3\mathbb{Z})[x, y]/(x^4, px, y) \times (\mathbb{Z}/q^3\mathbb{Z})[x, y]/(x^4, px, y - 1) \end{aligned}$$

∎

2.2 Basics of Complexity Theory

A decision problem in computer science is represented by a *language* $L \subseteq \{0, 1\}^*$ which is the set of all 'yes' strings. We say that L is in the *complexity class* NP if

there is a polynomial time deterministic Turing Machine M and a positive number c such that:

$$L = \left\{ x \mid \ \exists y \in \{0,1\}^{|x|^c},\ M(x,y) \text{ accepts} \right\}$$

x is the *input* and y is called as *witness*, *membership proof* or *nondeterministic guess*. L is said to be in coNP iff $\overline{L} \in$ NP.

Example Consider the problem of satisfiability of boolean formulas:

$$\text{3-SAT} := \{\phi(x_1, \ldots, x_n) \mid \phi = \wedge_{i=1}^{m} (x_{i_1} \vee x_{i_2} \vee x_{i_3}) \text{ and has a satisfying assignment}\}$$

3-SAT is in NP as given a formula ϕ and a satisfying assignment $\overline{v}$ it can be verified in polynomial time whether $\phi(\overline{v})$ is 'true'. ∎

We can also define a "randomized" version of the class NP called AM (for Arthur-Merlin protocol). We will say a language L is in AM if there is a positive number c and a polynomial time deterministic Turing Machine M such that:

$$\begin{aligned} x \in L \ &\Rightarrow \ \text{Prob}_{y \in \{0,1\}^{|x|^c}}[\exists z \in \{0,1\}^{|x|^c},\ M(x,y,z) \text{ accepts}] \geq \frac{2}{3} \\ x \notin L \ &\Rightarrow \ \text{Prob}_{y \in \{0,1\}^{|x|^c}}[\exists z \in \{0,1\}^{|x|^c},\ M(x,y,z) \text{ accepts}] \leq \frac{1}{3} \end{aligned}$$

Typically, the proof of showing an $L \in$ AM goes through by giving a protocol between the *Verifier* (named Arthur – the 'king') who can do randomized polynomial time computations and the *Prover* (named Merlin – the 'advisor' to the king) who has unlimited computational resources. Arthur is interested in determining whether the input $x \in L$ and he sends (x, y) to Merlin who responds with a witness z. Arthur does some computations on (x, y, z) following M and decides whether $x \in L$ with high confidence.

A classic example of a problem in AM is that of checking whether a set is large. We keep referring to its AM protocol in this chapter.

Proposition 2.4 *Suppose S is a set whose membership can be tested in nondeterministic polynomial time and its size is either m or $2m$. Then the decision problem of testing whether S is of size $2m$ is in AM.*

Proof: The idea of the AM protocol is that if S is large then for a random hash function h there will be an $x \in S$ such that $h(x) = 0$ with high probability.

Suppose that the elements of S are represented as binary strings of length s. Arthur first increases the 'gap' in the size of S by defining a new set $T = S^4$. Now $\#T$ is either m^4 or $16m^4$. Also, the elements of T are binary strings of length $4s$. View them as a column vector. Arthur then chooses a random 0/1 matrix A of size $\lceil \log 3m^4 \rceil \times 4s$ and sends it to Merlin. Merlin returns a column vector $t \in \{0,1\}^{4s}$ with a membership (in T) proof t'. Arthur accepts iff $t \in T$ and $A \cdot t = 0 \pmod 2$.

To analyse this AM protocol note that for a given $x \in \{0,1\}^{4s} \setminus \{0\}^{4s}$:

$$\text{Prob}_{A \in \{0,1\}^{\lceil \log 3m^4 \rceil \times 4s}} [A \cdot x = 0 \pmod 2] = \frac{1}{2^{\lceil \log 3m^4 \rceil}}$$

Thus by linearity of expectation:

$$E_{A \in \{0,1\}^{\lceil \log 3m^4 \rceil \times 4s}} [\#\{t \in T \mid A \cdot t = 0 \pmod 2\}] \;=\; \frac{\#T}{2^{\lceil \log 3m^4 \rceil}}.$$

Now Markov inequalities give us that:

$$\#T = 16m^4 \;\Rightarrow\; \text{Prob}_{A \in \{0,1\}^{\lceil \log 3m^4 \rceil \times 4s}} [\exists t \in T,\ A \cdot t = 0 \pmod 2] \geq \frac{5}{8}$$

$$\#T = m^4 \;\Rightarrow\; \text{Prob}_{A \in \{0,1\}^{\lceil \log 3m^4 \rceil \times 4s}} [\exists t \in T,\ A \cdot t = 0 \pmod 2] \leq \frac{1}{3}$$

This shows that with high probability Arthur accepts only when set S is large.

Also, note that this AM protocol uses $O(s \log m)$ random bits (for A) and $O(s + |t'|)$ nondeterministic bits (for t and t'). ∎

If a problem L is in NP ∩ coNP then intuition suggests that it should not be "hard". Similarly, if a problem L is in NP ∩ coAM (or AM ∩ coAM) then L is 'unlikely' to be NP-hard. What makes these classes interesting is that there are many problems in NP ∩ coAM that are not known to be in P. Such problems are called problems of "intermediate" complexity. To make these notions more precise we need to form a polynomial-time hierarchy.

Let us denote NP by Σ_1 and define $\Sigma_2 = \text{NP}^{\text{NP}}$, where by $\text{NP}^{\mathcal{C}}$ we mean set of languages L such that there is a polynomial time deterministic Turing Machine M using an *oracle* to $\mathcal{C}$ and a positive number c such that:

$$L = \left\{x \mid\ \exists y \in \{0,1\}^{|x|^c},\ M(x,y) \text{ accepts}\right\}$$

Similarly, $\Sigma_k := \mathrm{NP}^{\Sigma_{k-1}}$. The union of all these Σ's is called the *polynomial-time hierarchy*: $\mathrm{PH} = \cup_{k \geq 1} \Sigma_k$.

It is mostly believed that $\Sigma_1, \Sigma_2, \ldots$ are all distinct complexity classes and hence there is no k such that PH collapses to Σ_k. Coming back to the intermediate complexity classes, it is easy to see that if $\mathrm{NP} \cap \mathrm{coNP}$ has a NP-hard problem then $\mathrm{PH} = \Sigma_1$. Also, if $\mathrm{NP} \cap \mathrm{coAM}$ (or $\mathrm{AM} \cap \mathrm{coAM}$) has a NP-hard problem then it was shown in [Sch88, Klap89] that PH collapses to the second level Σ_2. The proof goes through by showing that $\mathrm{AM} \cap \mathrm{coAM}$ is *low for* Σ_2, i.e., $\Sigma_2^{\mathrm{AM} \cap \mathrm{coAM}} = \Sigma_2$ and thus, $\mathrm{NP} \subseteq \mathrm{AM} \cap \mathrm{coAM}$ implies $\Sigma_3 = \Sigma_2^{\mathrm{NP}} \subseteq \Sigma_2^{\mathrm{AM} \cap \mathrm{coAM}} = \Sigma_2$ which eventually results in collapsing PH to Σ_2.

This notion of intermediate complexity can be generalized to *functional problems*. We define FP to be the set of functional problems computable in polynomial time. Now the functional problems in $\mathrm{FP}^{\mathrm{AM} \cap \mathrm{coAM}}$ are of intermediate complexity. If a function $f \in \mathrm{FP}^{\mathrm{AM} \cap \mathrm{coAM}}$ is NP-hard (i.e. $\mathrm{NP} \subseteq \mathrm{P}^f$) then the techniques of Schoning [Sch88] essentially show that PH collapses to Σ_2, an 'unlikely' event. Further, define *functional AM* – denoted by fnAM – to contain functions $f : \{0,1\}^* \rightarrow \{0,1\}^*$ such that there is a deterministic polynomial time Turing machine M (that *outputs* a string) and a positive number c such that, for all $x,\ t \in \{0,1\}^*$:

$$f(x) = t \quad \text{iff} \quad \mathrm{Prob}_{y \in \{0,1\}^{|x|^c}}[\exists z \in \{0,1\}^{|x|^c}\ M(x,y,z) = t] \geq \frac{2}{3} \tag{2.2}$$

Remark: The above definition says that for "most" of the y 's there is a z such that $M(x,y,z)$ outputs the correct value of $f(x)$. On the other hand, for "most" of the y 's there is no z such that $M(x,y,z)$ outputs an incorrect value. ∎

Again the techniques of Schoning [Sch88] essentially show that fnAM is low for Σ_2, i.e. $\Sigma_2^{\mathrm{fnAM}} = \Sigma_2$. Thus, if a function $f \in \mathrm{fnAM}$ is NP-hard (i.e. $\mathrm{NP} \subseteq \mathrm{P}^f$) then PH collapses to Σ_2. We sketch the proof here for the sake of completeness. Define for all $k \geq 1$, $\Pi_k :=$ co-Σ_k.

Proposition 2.5 $\Sigma_2^{fnAM} = \Sigma_2$.

Proof: Let a language $L \in \Pi_2^{\mathrm{fnAM}}$. Then, by definition, there is a positive number c and a polynomial time deterministic Turing Machine A using functions from fnAM

as *oracles* such that:

$$L = \big\{x \mid \ (\forall y \in \{0,1\}^{|x|^c})(\exists z \in \{0,1\}^{|x|^c})\ [A^{\{f_1,\ldots,f_m\}}(x,y,z) \text{ accepts}] \\ \text{where, } f_1,\ldots,f_m \in \text{fnAM and } m \leq |x|^c\} \tag{2.3}$$

Suppose on input x, A queries f_i at strings $w_{i,j} \in \{0,1\}^{|x|^c}$ where i,j are upper-bounded by $|x|^c$. Now from defining-Equation (2.2) we have that there is a deterministic polynomial time Turing machine M_i (that *outputs* a string) and a positive number c_i such that:

$$f_i(w_{i,j}) = t_{i,j} \quad \text{iff} \quad \text{Prob}_{y\in\{0,1\}^{|x|^{c_i}}}[\exists z \in \{0,1\}^{|x|^{c_i}}\ M_i(w_{i,j},y,z) = t_{i,j}] \geq \frac{2}{3} \tag{2.4}$$

Now combining Equations (2.4) for various i,j (after probability amplification) and then plugging in Equation (2.3) we get that there is a deterministic polynomial time Turing machine B (that basically simulates M_i 's to compute f_i 's and then runs A to decide L) and a positive number d such that:

$$\begin{aligned}
L &= \Big\{x \mid \ (\forall y \in \{0,1\}^{|x|^c})(\exists z \in \{0,1\}^{|x|^c}) \\
&\qquad \text{Prob}_{u\in\{0,1\}^{|x|^d}}[\exists v \in \{0,1\}^{|x|^d},\ B(u,v,x,y,z) \text{ accepts}] \geq \frac{2}{3}\Big\} \\
&= \Big\{x \mid \ (\forall y \in \{0,1\}^{|x|^c})\ \text{Prob}_{u\in\{0,1\}^{|x|^{d'}}}[(\exists z \in \{0,1\}^{|x|^c}) \\
&\qquad (\exists v \in \{0,1\}^{|x|^d})\ B'(u,v,x,y,z) \text{ accepts}] \geq \frac{2}{3}\Big\} \\
&\qquad [\because \text{By Swapping lemma there is a } d' \text{ and } B' \text{ such that the above holds}] \\
&= \big\{x \mid \ (\forall y \in \{0,1\}^{|x|^c})(\forall u_1 \in \{0,1\}^{|x|^e})(\exists u_2 \in \{0,1\}^{|x|^e})(\exists z \in \{0,1\}^{|x|^c}) \\
&\qquad (\exists v \in \{0,1\}^{|x|^d})\ [B''(u_1,u_2,v,x,y,z) \text{ accepts}]\big\} \\
&\qquad [\because e \text{ and } B'' \text{ exists by Lemma A.14}] \\
&\in \Pi_2
\end{aligned}$$

Consequently, $\Pi_2^{\text{fnAM}} = \Pi_2$ and hence, $\Sigma_2^{\text{fnAM}} = \Sigma_2$. ∎

The definitions of ring isomorphism problems are inspired from graph isomorphism (GI) problems that have been open for a long time. But the graph isomorphism problems are not believed to be NP-hard. The AM protocol for graph

nonisomorphism was one of the first interactive protocols (see [GMR85]) proving that GI $\in$ NP $\cap$ coAM.

The results in this chapter mostly *reduce* one problem L to another problem L'. If there is a function $f : \{0,1\}^* \to \{0,1\}^*$ in class $\mathcal{C}$ such that $x \in L$ iff $f(x) \in L'$ then we say that L is *many-one reducible* to L' and denote it by $L \leq_m^{\mathcal{C}} L'$.

If a problem L can be solved in class $\mathcal{C}$ by using L' as an oracle then we say that L is *Turing reducible* to L' and denote it by $L \leq_T^{\mathcal{C}} L'$.

In the reductions given in this chapter $\mathcal{C}$ is either P or ZPP – the set of languages (functions) that can be decided (computed) in *expected* polynomial time.

2.3 The Complexity of Ring Isomorphism Problem

In this section we prove upper and lower bounds on the complexity of Ring Isomorphism problem. Specifically, we show that RI is in NP $\cap$ coAM and the Graph Isomorphism problem reduces to RI.

2.3.1 An Upper Bound

This work has been unable to solve the ring isomorphism problem in polynomial time or even subexponential time. But we show in this section that at least the problem is unlikely to be NP-hard. Thus, RI becomes a natural example of an intermediate problem which also has a rich algebraic flavor to it.

Theorem 2.1 *RI $\in$ NP $\cap$ coAM.*

Proof: We start with the easier part,

Claim 2.1.1 *RI $\in$ NP.*

Proof of Claim 2.1.1. Suppose we are given two rings R and R' together with a map $\phi : R \to R'$. Following the remark of Proposition 2.1, we have an algorithm that gives us a description of the rings R, R' over the same additive basis, say,

$$(\mathbb{Z}/m_1\mathbb{Z}) \oplus \ldots \oplus (\mathbb{Z}/m_n\mathbb{Z})$$

Thus, we can assume without loss of generality that the rings R, R' are provided as:

$$(R, +) = (\mathbb{Z}/m_1\mathbb{Z})b_1 \oplus \ldots \oplus (\mathbb{Z}/m_n\mathbb{Z})b_n$$

$$(R', +) = (\mathbb{Z}/m_1\mathbb{Z})b'_1 \oplus \ldots \oplus (\mathbb{Z}/m_n\mathbb{Z})b'_n$$

Now ϕ is an isomorphism from $R \to R'$ iff it satisfies the following conditions:

- ϕ *preserves addition*: check whether for all $1 \leq i \leq n,\ m_i \cdot \phi(b_i) = 0$.
- ϕ *preserves multiplication*: check whether for all $1 \leq i, j \leq n,\ \phi(b_i) \cdot \phi(b_j) = \sum_{k=1}^{n} a_{i,j,k}\phi(b_k)$, where $((a_{i,j,k}))_{i,j,k\in[n]}$ is the same matrix as given in the description of R.
- ϕ *is an invertible map from* $(R, +)$ *to* $(R', +)$: check whether $det(A) \in (\mathbb{Z}/(m_1 m_2 \ldots m_n)\mathbb{Z})^*$, where A is the $n \times n$ integer matrix describing the map $\phi : R \to R'$.

The first two conditions above imply that ϕ is a homomorphism between the two rings. The third condition ensures that ϕ is bijective. All these three conditions can be checked in polynomial time. □

The next question is whether there are short certificates to prove that two given rings are nonisomorphic i.e., is RI ∈ coNP? We are able to tweak the AM protocol for graph nonisomorphism to show that RI is in the randomized version of coNP.

Claim 2.1.2 *RI ∈ coAM.*

Proof of Claim 2.1.2. Arthur has two rings R_1, R_2 in basis forms and he wants a *proof* of their non-isomorphism from Merlin. Arthur checks whether $(R_1, +) \cong (R_2, +)$ (see the remark of Proposition 2.1), if not then Arthur already has a proof of non-isomorphism. So assume that $(R_1, +) \cong (R_2, +)$ and now Merlin can provide the descriptions of $(R_1, +), (R_2, +)$ in the form:

$$\begin{aligned} (R_1, +) &= \bigoplus_{i=1}^{n} (\mathbb{Z}/p_i^{\alpha_i}\mathbb{Z})b_i \text{ and} \\ (R_2, +) &= \bigoplus_{i=1}^{n} (\mathbb{Z}/p_i^{\alpha_i}\mathbb{Z})c_i, \text{ where } p_i\text{'s are primes and } \alpha_i \in \mathbb{Z}^{\geq 1}. \end{aligned}$$

Arthur checks the primality of p_i's and that the above is a basis representation of the rings R_1 and R_2. Let us define sets $C(R_1), C(R_2)$ that we will be using to give an AM protocol for ring non-isomorphism. They will have the nice property that their sizes can be computed *easily* and that $C(R_1) = C(R_2)$ if and only if $R_1 \cong R_2$.

$$\begin{aligned} C(R_1) \;:=\; & \{\, \left\langle ((a_{i,j,k}))_{i,j,k\in[n]}, A_\phi \right\rangle \mid \exists \pi \in Aut(R_1,+) \text{ s.t.} \\ & \text{for all } i,j \in [n],\ \pi(b_i)\cdot\pi(b_j) = \sum_{k=1}^{n} a_{i,j,k}\pi(b_k); \\ & \text{for all } i,j,k \in [n],\ 0 \le a_{i,j,k} < p_k^{\alpha_k}; \\ & A_\phi \text{ is an integer matrix describing some } \phi \in Aut(R_1) \\ & \text{with respect to the additive basis } \{\pi(b_i)\}_{i=1}^n \text{ of } R_1 \,\}. \end{aligned}$$

$C(R_2)$ is defined similarly by replacing the b_i's above by the c_i's and R_1 by R_2. (Note that in the case of graph isomorphism we consider all permutations on the vertices, here we consider all automorphisms of the additive group.)

Observe that: $\#C(R_1)$

$$\begin{aligned} &= \left(\text{number of representations } ((a_{i,j,k}))_{i,j,k\in[n]} \text{ of ring } R_1 \text{ over } \bigoplus_{i=1}^{n} \mathbb{Z}/p_i^{\alpha_i}\mathbb{Z} \right) \cdot \#Aut(R_1) \\ &= \frac{\#Aut(R_1,+)}{\#Aut(R_1)} \cdot \#Aut(R_1) \\ &= \#Aut(R_1,+) \end{aligned}$$

that can be computed in polynomial time when $(R_1, +)$ is given in terms of basis elements all having prime-power additive orders (see Proposition 2.2). Thus, Arthur can compute $s := \#C(R_1) = \#C(R_2)$.

Define $C(R_1, R_2) := C(R_1) \cup C(R_2)$. Note that:

$$\begin{aligned} R_1 \cong R_2 \;&\Rightarrow\; C(R_1) = C(R_2) \\ &\Rightarrow\; \#C(R_1,R_2) = \#C(R_1) = s. \\ R_1 \ncong R_2 \;&\Rightarrow\; C(R_1) \cap C(R_2) = \emptyset \\ &\Rightarrow\; \#C(R_1,R_2) = \#C(R_1) + \#C(R_2) = 2s. \end{aligned}$$

Thus, the size of the set $C(R_1, R_2)$ has a gap factor of 2 between the cases of $R_1 \cong R_2$ and $R_1 \ncong R_2$, which can be distinguished by the AM protocol of Proposition 2.4.

Note that this AM protocol for ring nonisomorphism requires:

$$O\left((\log^4 \#R_1) \cdot (\log s)\right) = O(\log^7 \#R_1)$$

random bits, and $O(\log^4 \#R_1)$ nondeterministic bits. □

The two claims show that RI is in NP ∩ coAM. ∎

This shows that the ring isomorphism problem cannot be NP-hard (unless polynomial hierarchy collapses to Σ_2 [Sch88]). It also follows easily from the above proof that the problems of testing ring automorphism and testing ring isomorphism can be solved in deterministic polynomial time.

Corollary 2.1 *TRA and TRI are in P.*

Proof: Clearly, it is sufficient to show that TRI is in P. Suppose rings R_1, R_2 and a map ϕ between them are given in the basis representation. It is clear from Claim 2.1.1 that there is a deterministic polynomial time algorithm to determine whether ϕ is an isomorphism from R_1 to R_2. ∎

2.3.2 A Lower Bound: Reduction from Graph Isomorphism

The proofs above were all similar in spirit to those for graph isomorphism which hints a connection to graph isomorphism. Indeed, we can lower bound the complexity of RI by graph isomorphism (GI). The reduction gives a way to construct a local commutative $\mathbb{F}$-algebra out of a given graph.

Theorem 2.2 *GI* $\leq_m^P$ *RI.*

Proof: The proof involves constructing a local commutative $\mathbb{F}$-algebra. We associate variables to each vertex (x-variable) and capture the "connectivity" of the graph by defining the edges-polynomial – $\sum_{(u,v) \text{ is an edge}} x_u x_v$ – as zero in the ring.

Let G be an undirected graph with n vertices and no self loops. Choose any field $\mathbb{F}$ of characteristic not equal to 2. Define the following commutative $\mathbb{F}$-algebra:

$$R(G) := \mathbb{F}[x_1, \ldots, x_n]/\mathcal{I}$$

where, ideal $\mathcal{I}$ has the following relations:

1. x's are nilpotents of degree 2, i.e., for all $i \in [n]$: $x_i^2 = 0$.
2. the edges-polynomial is zero, i.e., $\sum_{\substack{1 \le i < j \le n \\ (i,j) \in E(G)}} x_i x_j = 0$.
3. all cubic terms are zero, i.e., for all $i, j, k \in [n]$: $x_i x_j x_k = 0$.

Suppose (i_0, j_0) is an edge in G such that $1 \le i_0 < j_0 \le n$. Then the additive structure of the ring is:

$$(R(G), +) = \mathbb{F} \cdot 1 \oplus \bigoplus_{i \in [n]} \mathbb{F} \cdot x_i \oplus \bigoplus_{\substack{i < j \in [n] \\ (i,j) \neq (i_0, j_0)}} \mathbb{F} \cdot (x_i x_j)$$

Thus, the dimension of the ring over $\mathbb{F}$ is $\binom{n+1}{2}$. Multiplication satisfies the associative law simply because the product of any three *variables* (in any order) is zero. Also, $R(G)$ is a local commutative $\mathbb{F}$-algebra.

Observe that if $G \cong G'$ then any graph isomorphism ϕ induces a natural isomorphism between rings $R(G)$ and $R(G')$. So we only have to prove the converse:

Claim 2.2.1 *Let G and G' be two undirected graphs having no self-loops. Further, assume that graphs G and G' are not a disjoint union of a clique and a set of isolated vertices. Then, $R(G) \cong R(G')$ implies $G \cong G'$.*

Proof of Claim 2.2.1. Suppose ϕ is an isomorphism from $R(G) \to R(G')$. Let

$$\phi(x_i) = c_{i,0} + c_{i,1}x_1 + \ldots + c_{i,n}x_n + (\text{quadratic terms}). \tag{2.5}$$

where all $c_{i,j}$'s in the coefficients are in $\mathbb{F}$.

By squaring the above we get:

$$0 = \phi(x_i^2) = \phi(x_i)^2 = c_{i,0}^2 + (\text{linear and quadratic terms})$$

which means that $c_{i,0} = 0$. The next observation about ϕ is that there is at most one nonzero linear term in $\phi(x_i)$. Let $C_i = \{j \in [n] \mid c_{i,j} \neq 0\}$ be of size > 1. Then $\phi(x_i)^2 = 0$ gives:

$$\sum_{j<k \in C_i} (2c_{i,j}c_{i,k})x_j x_k \;=\; 0 \text{ in } R(G')$$

We know that in $R(G')$ the quadratic relations are $x_i^2 = 0$ and $\sum_{\substack{1\leq i<j\leq n \\ (i,j)\in E(G')}} x_i x_j = 0$. This means that the above equation holds only if there is a $\lambda \in \mathbb{F}$:

$$\sum_{\substack{1\leq j<k\leq n \\ j,k\in C_i}} (2c_{i,j}c_{i,k})x_j x_k = \lambda \cdot \sum_{\substack{1\leq i<j\leq n \\ (i,j)\in E(G')}} x_i x_j = 0$$

This equality interpreted in graph terms means that G' is a union of a clique on C_i and a set of $(n - \#C_i)$ isolated vertices (remember that $2 \neq 0$ in $\mathbb{F}$). This we ruled out in the hypothesis, thus size of $C_i \leq 1$. If $\#C_i = 0$ then for any j, $\phi(x_i x_j) = 0$ which contradicts the assumption that ϕ is an isomorphism. Thus, for all $i \in [n]$, $\#C_i = 1$. Define a map $\pi : [n] \to [n]$ such that the nonzero linear term occurring in $\phi(x_i)$ is $x_{\pi(i)}$.

Suppose π is not a permutation on $[n]$ then there are $i \neq j$ such that $\pi(i) = \pi(j)$. But then there will exist $a, b \in \mathbb{F}^*$ such that there is no nonzero linear term in $\phi(ax_i + bx_j)$. Whence, we get that $\phi(ax_i x_k + bx_j x_k) = 0$ for all $k \in [n]$ which contradicts the assumption that ϕ is an isomorphism. Hence, π is a permutation on $[n]$. Now look at the action of ϕ on the edges-polynomial:

$$\begin{aligned} 0 &= \phi\left(\sum_{\substack{1\leq i<j\leq n \\ (i,j)\in E(G)}} x_i x_j\right) \\ &= \sum_{\substack{1\leq i<j\leq n \\ (i,j)\in E(G)}} \phi(x_i)\phi(x_j) \\ &= \sum_{\substack{1\leq i<j\leq n \\ (i,j)\in E(G)}} c_{i,\pi(i)}c_{j,\pi(j)}x_{\pi(i)}x_{\pi(j)} \end{aligned}$$

Since the above is a zero relation in the ring $R(G')$, we get that the polynomial $\sum_{\substack{1\leq i<j\leq n \\ (i,j)\in E(G')}} x_i x_j$ divides the above. Hence, $(\pi(i), \pi(j)) \in E(G')$ if $(i, j) \in E(G)$.

By symmetry this shows that π is an isomorphism from $G \to G'$. $\square$

The theorem follows from the claim. ∎

Remark: The above reduction does not work for fields $\mathbb{F}$ of characteristic 2. We can modify the ring $R(G)$ slightly to make the reduction go through even when $\mathbb{F}$ is a field of characteristic 2. Define the ring $R(G)$ from a graph G, having n vertices, as:

$$R(G) := \mathbb{F}[x_1, \ldots, x_n]/\mathcal{I}$$

where, ideal $\mathcal{I}$ has the following relations:

1. x's are nilpotents of degree 3, i.e., for all $i \in [n]$: $x_i^3 = 0$.
2. the *modified* edges-polynomial is zero, i.e., $\sum_{\substack{1 \le i < j \le n \\ (i,j) \in E(G)}} (x_i^2 x_j + x_i x_j^2) = 0$.
3. all *quartic* terms are zero, i.e., for all $i, j, k, l \in [n]$: $x_i x_j x_k x_l = 0$.

A similar proof as above shows that isomorphism problem for rings like $R(G)$ solves the graph isomorphism problem too. ∎

Note that even if graph G is rigid (i.e., G has no nontrivial automorphism) the ring $R(G)$ has lots of nontrivial automorphisms, for example, $\phi : x_i \mapsto x_i + x_1 x_2$. Thus, unfortunately, this reduction does not reduce the problem of testing rigidity of graphs to testing rigidity of rings.

2.3.3 Table Representation: Is it any easier?

One can also consider a different, exponentially larger, representation for rings: when the rings are given in terms of the addition and multiplication tables of all its elements. We do not know if the ring isomorphism problem even under this representation can be solved in time polynomial in the size of the representation. However, one suspects that this version of ring isomorphism should be easier as there is a simple subexponential time algorithm: Suppose rings R_1, R_2 are of size n. Then the additive group of R_1 will have $O(\log n)$ generators and there are $n^{O(\log n)}$ ways to

map these generators into R_2. Thus, a brute-force search over all these maps yields a $n^{O(\log n)}$ time algorithm for ring isomorphism.

Here we give another theoretical evidence that the problem is easy by showing that it is "almost" in NP $\cap$ coNP.

Let us give this problem a name:

$$\mathrm{RI}_{TF} := \{(R_1, R_2) \mid R_1, R_2 \text{ are given in terms of tables, } R_1 \cong R_2\}$$

It is easy to see that $\mathrm{RI}_{TF} \in$NP. The nontrivial part is to show:

Theorem 2.3 *There exists an NP-machine that decides all but $2^{\log^{11} n}$ instances of $\overline{RI}_{TF}$ of length n and is always correct when the input rings are nonisomorphic.*

Proof: The proof is basically the one given by Arvind and Toran [AT04] applied to the case of rings.

We showed in Claim 2.1.2 that $\overline{\mathrm{RI}}_{TF} \in \mathrm{AM}(\log^7 n)$, where the parameter bounds the number of random bits used by Arthur. We interpret this result to mean that there is an advice-taking NP machine $M(\cdot,\cdot)$ for $\overline{\mathrm{RI}}_{TF}$ such that:

$$\forall \text{ input } x \in \{0,1\}^n,\ \mathrm{Prob}_{y \in \{0,1\}^{\log^7 n}}[M(x,y) \text{ is correct}] \geq \frac{2}{3}.$$

Notice that since a ring is completely defined once we specify the multiplication on the additive generators, we have that the number of binary strings of length n that define a ring, in table form, is no more than $2^{\log^4 n}$. Thus, using probability amplification we modify M to get an advice-taking NP machine M' for $\overline{\mathrm{RI}}_{TF}$ such that:

$$\mathrm{Prob}_{y \in \{0,1\}^{\log^{11} n}}[\forall x \in \{0,1\}^n,\ M'(x,y) \text{ is correct}] \geq \frac{2}{3}.$$

Since we are using only a "small" number of random bits we can apply techniques of Goldreich and Wigderson [GW02] to get an NP-machine for $\overline{\mathrm{RI}}_{TF}$ that fails for at most $2^{\log^{11} n}$ inputs of size n and is always correct when the input rings are nonisomorphic. ∎

2.4 The Complexity of Counting Ring Automorphisms

This section will explore the complexity of the problem of counting ring automorphisms. We will show that this problem is unlikely to be NP-hard but both graph isomorphism and integer factoring reduce to it.

2.4.1 An Upper Bound

We will show that given a finite ring R there is an AM protocol in which Merlin sends a number ℓ and convinces Arthur that $\#Aut(R) = \ell$. The ideas in the proof are basically from Babai and Szemeredi [BS84].

Theorem 2.4 *$\#RA \in FP^{AM \cap coAM}$.*

Proof: Let R be a finite ring given in its basis form. We will first show how Merlin can convince Arthur that $\#Aut(R) \geq k$. Recall that in Equation (2.1) we defined this problem as cRA.

Claim 2.4.1 *$cRA \in AM$.*

Proof of Claim 2.4.1. Merlin can give Sylow subgroups $S_{p_1}, \ldots, S_{p_m}$ of $Aut(R)$, in terms of generators, to Arthur such that $p_1, \ldots, p_m$ are distinct primes and the product $|S_{p_1}|.\cdots.|S_{p_m}| \geq k$. Arthur now has to verify whether for a given Sylow subgroup S_p, $|S_p| = p^t$ or not. So Merlin can further provide the composition series of S_p:

$$S_p = G_t > G_{t-1} > \ldots > G_1 > G_0 = \{1\}.$$

Suppose, by induction, that Arthur is convinced about $|G_i| = p^i$. Then to prove $|G_{i+1}| = p^{i+1}$, Merlin will provide $x_{i+1} \in G_{i+1}$ to Arthur with the claim that $x_{i+1} \notin G_i$ but $x_{i+1}^p \in G_i$. Latter can be verified easily by Arthur as Merlin can give the way to produce x_{i+1}^p from the generators of G_i. Finally, the only nontrivial thing left for Arthur to verify is whether $x_{i+1} \notin G_i$, which can be verified by a standard

AM protocol (Proposition 2.4) as there is a gap in the size of the set $X :=$ (group generated by x_{i+1} and G_i):

$$\begin{aligned} x_{i+1} \notin G_i &\Rightarrow \#X = p^{i+1} \\ x_{i+1} \in G_i &\Rightarrow \#X = p^i \end{aligned}$$

To avoid too many rounds, Merlin first provides $x_0 = 1, x_1, \ldots, x_t \in Aut(R)$ with the proof of: for all $1 \leq i \leq t,\ x_i^p \in G_{i-1} :=$ (group generated by $x_0, \ldots, x_{i-1}$) to Arthur and then provides the proof of: for all $1 \leq i \leq t,\ x_i \notin G_{i-1}$ in the second round for Arthur to verify. □

Now we give the AM protocol that convinces Arthur of $\#Aut(R) \leq k$.

Claim 2.4.2 *$cRA \in coAM$.*

Proof of Claim 2.4.2. Arthur has a finite ring R and he wants a proof of $\#Aut(R) \leq k$. As in the proof of Claim 2.1.2, we can assume that R is given in terms of generators having prime-power additive orders. For concreteness let us assume:

$$(R, +) = \bigoplus_{i=1}^{n} (\mathbb{Z}/p_i^{\alpha_i}\mathbb{Z})b_i$$

Merlin sends Arthur a number $\ell \leq k$ as a candidate value for $\#Aut(R)$ and also provides some Sylow subgroups, the product of their sizes being equal to ℓ, with the AM-proofs for their sizes (as used in Claim 2.4.1). Let

$$\begin{aligned} X \ := \ \{ \left\langle ((a_{i,j,k}))_{i,j,k\in[n]} \right\rangle \mid \exists \pi \in Aut(R,+) \text{ s.t. } \pi(b_i) \cdot \pi(b_j) = \sum_{k=1}^{n} a_{i,j,k}\pi(b_k); \\ \text{for all } 1 \leq i,j,k \leq n,\ 0 \leq a_{i,j,k} < p_k^{\alpha_k} \ \} . \end{aligned}$$

Observe that $\#X = \frac{\#Aut(R,+)}{\#Aut(R)}$ and $\#Aut(R,+)$ can be computed in polynomial time when $(R,+)$ is given in terms of generators having prime-power additive orders (see Proposition 2.2). Thus, Arthur computes $s := \#Aut(R,+)$. Arthur is already convinced that $\ell | \#Aut(R)$ and he now wants to verify $\#Aut(R) \leq \ell$. A standard AM protocol (see Proposition 2.4) now follows by utilizing the gap in the size of X

in the two cases:

$$\begin{aligned} \#Aut(R) \leq \ell \quad &\Rightarrow \quad \#X \geq \frac{s}{\ell} \ . \\ \#Aut(R) > \ell \quad &\Rightarrow \quad \#Aut(R) \geq 2\ell \quad [\because \#Aut(R) \text{ has a subgroup of size } \ell] \\ &\Rightarrow \#X \leq \frac{s}{2\ell} \ . \end{aligned}$$

□

The claims above show that $\#\mathrm{RA} \in \mathrm{FP}^{\mathrm{cRA}} \subseteq \mathrm{FP}^{\mathrm{AM} \cap \mathrm{coAM}}$. ∎

Note that the AM protocols that we give for #RA not only count the number of automorphisms but give a lot more information about the automorphism group. In fact, these AM protocols compute the full automorphism group of a ring R in terms of the generators of the Sylow subgroups of $Aut(R)$. Let us denote the functional problem of *computing the group of automorphisms* of a ring given in basis form by GroupRA.

Corollary 2.2 *Function GroupRA ∈ fnAM and hence is low for* Σ_2.

Proof: Let f be the function, corresponding to GroupRA, that maps a ring R (given in basis form) to the tuple $(\#Aut(R), Aut(R))$. Since cRA is in both AM and coAM there are deterministic polynomial time Turing Machines A and B, and positive constants c, d such that:

$$\begin{aligned} \#Aut(R) \leq k \quad \text{iff} \quad &\mathrm{Prob}_{y \in \{0,1\}^{\log^c \#R}}[(\exists z \in \{0,1\}^{\log^c \#R})\ A(R,k,y,z) \text{ accepts}] \\ &\geq \left(1 - \frac{1}{2^{\log^d \#R}}\right) \\ \#Aut(R) \geq k \quad \text{iff} \quad &\mathrm{Prob}_{y \in \{0,1\}^{\log^c \#R}}[(\exists z \in \{0,1\}^{\log^c \#R})\ B(R,k,y,z) \text{ accepts}] \\ &\geq \left(1 - \frac{1}{2^{\log^d \#R}}\right) \end{aligned} \tag{2.6}$$

The parameter d above will be chosen large enough so that all the subsequent arguments go through. To show that $f \in$ fnAM we plan to run A and B in parallel. We can modify A slightly to A' by requiring that $A(R,k,y,z)$ outputs (ℓ, G) where,

ℓ is the number and G is the group, given by the generators of the (intended) Sylow subgroups, as occurred in the proof of the Claim 2.4.2. It is easy to see that:

$$\begin{aligned} & f(R) = (m, H) \\ \Rightarrow \quad & \text{Prob}_{y \in \{0,1\}^{2\log^c \#R}}[(\exists \ell' z z' \in \{0,1\}^{3\log^c \#R}),\ \text{both } A'(R, \ell', y, z) \\ & \qquad \text{and } B(R, \ell', y, z') \text{ accept and } A'(R, \ell', y, z) = (m, H)] \geq \frac{3}{4} \end{aligned} \tag{2.7}$$

The above holds because Merlin can simply send ℓ' as equal to $\#G$ and a part of the string z and z' having the group $Aut(R)$ in terms of the generators of Sylow subgroups (see the proof of Claim 2.4.2). Then Equations (2.6) give us the probability lower bound of $\frac{3}{4}$. Also, the output of $A'(R, \ell', y, z)$ for such ℓ', z will trivially be (m, H).

To show the converse assume that there is a number m and a group H such that:

$$\begin{aligned} & \text{Prob}_{y \in \{0,1\}^{2\log^c \#R}}[(\exists \ell' z z' \in \{0,1\}^{3\log^c \#R}),\ \text{both } A'(R, \ell', y, z) \\ & \qquad \text{and } B(R, \ell', y, z') \text{ accept and } A'(R, \ell', y, z) = (m, H)] \geq \frac{3}{4} \end{aligned} \tag{2.8}$$

Now if $(m, H) \neq (\#Aut(R), Aut(R))$ then the way A' outputs, it is clear that Merlin tried to "fool" Arthur and so by the Equations (2.6) we get that for some positive d':

$$\begin{aligned} & \text{Prob}_{y \in \{0,1\}^{2\log^c \#R}}[(\exists \ell' z z' \in \{0,1\}^{3\log^c \#R}), \text{both } A'(R, \ell', y, z) \text{ and} \\ & \quad B(R, \ell', y, z') \text{ accept } \mid A'(R, \ell', y, z) \neq (\#Aut(R), Aut(R))] \leq \frac{1}{2^{\log^{d'} \#R}} \end{aligned}$$

which together with the large probability lower bound of Equation (2.8) means that: $(m, H) = (\#Aut(R), Aut(R))$. Thus,

$$\begin{aligned} & \text{Prob}_{y \in \{0,1\}^{2\log^c \#R}}[(\exists \ell' z z' \in \{0,1\}^{3\log^c \#R}),\ \text{both } A'(R, \ell', y, z) \\ & \qquad \text{and } B(R, \ell', y, z') \text{ accept and } A'(R, \ell', y, z) = (m, H)] \geq \frac{3}{4} \\ \Rightarrow \quad & f(R) = (m, H) \end{aligned} \tag{2.9}$$

Recall Equation (2.2) for the definition of fnAM, clearly, Equations (2.7) and (2.9) tell us that: $f \in$ fnAM. ∎

2.4.2 A Lower Bound: Reduction from Graph Isomorphism and Integer Factoring

This section shows that #RA is a fairly interesting intermediate problem as two well known problems – one of graphs and another of integers – reduce to it.

In the case of graphs it is easy to show that graph isomorphism (or counting graph isomorphisms) reduces to counting graph automorphisms. The same result continues to hold for rings with a slightly more involved proof. In the case of graphs we take disjoint union of graphs to construct a new graph, here we take *direct product* of rings to construct a new ring. It turns out that the number of automorphisms of this new ring can be used to find out whether the original rings were isomorphic or not.

Lemma 2.1 $\#RI \equiv_T^P \#RA$.

Proof: Suppose we are given a ring R. Clearly, we can compute $\#Aut(R)$ by giving (R, R) as input to the oracle of #RI.

Conversely, let R_1, R_2 be the two rings given in basis form. Let us assume the following about their decomposability into *distinct* local rings $S_1, \ldots, S_k$:

$R_1 \cong S_1 \times \cdots \times S_1 \times \ldots \times S_k \times \cdots \times S_k$
where, for all $1 \leq i \leq k$, indecomposable ring S_i occurs $a_i \geq 0$ times and $\#Aut(S_i) = m_i$.

$R_2 \cong S_1 \times \cdots \times S_1 \times \ldots \times S_k \times \cdots \times S_k$
where, for all $1 \leq i \leq k$, indecomposable ring S_i occurs $b_i \geq 0$ times.

The following claim relates the (non)isomorphism of the rings to counting ring automorphisms:

Claim 2.4.3 (Kayal) $R_1 \ncong R_2 \Rightarrow \#Aut(R_1 \times R_1) \cdot \#Aut(R_2 \times R_2) > (\#Aut(R_1 \times R_2))^2$.

Proof of Claim 2.4.3. Due to the uniqueness of decomposition of a ring into indecomposable rings (see Proposition 2.3):

$$\begin{aligned}\#Aut(R_1 \times R_2) &= \#Aut(\overbrace{S_1 \times \cdots \times S_1}^{a_1+b_1}) \cdots \#Aut(\overbrace{S_k \times \cdots \times S_k}^{a_k+b_k}) \\ &= (a_1+b_1)!m_1^{a_1+b_1} \cdots (a_k+b_k)!m_k^{a_k+b_k}\end{aligned}$$

Similarly,

$$\begin{aligned}\#Aut(R_1 \times R_1) &= \#Aut(\overbrace{S_1 \times \cdots \times S_1}^{2a_1}) \cdots \#Aut(\overbrace{S_k \times \cdots \times S_k}^{2a_k}) \\ &= (2a_1)!m_1^{2a_1} \cdots (2a_k)!m_k^{2a_k}\end{aligned}$$

$$\begin{aligned}\#Aut(R_2 \times R_2) &= \#Aut(\overbrace{S_1 \times \cdots \times S_1}^{2b_1}) \cdots \#Aut(\overbrace{S_k \times \cdots \times S_k}^{2b_k}) \\ &= (2b_1)!m_1^{2b_1} \cdots (2b_k)!m_k^{2b_k}\end{aligned}$$

Notice that $\binom{2a_i+2b_i}{a_i+b_i} \geq \binom{2a_i+2b_i}{2a_i}$ which implies $(2a_i)! \cdot (2b_i)! \geq (a_i+b_i)!^2$. This clearly shows:

$$\#Aut(R_1 \times R_1) \cdot \#Aut(R_2 \times R_2) \geq (\#Aut(R_1 \times R_2))^2$$

Now since $R_1 \ncong R_2$, there exists an $i_0 \in [k]$ such that $a_{i_0} \neq b_{i_0}$ in which case $(2a_{i_0})! \cdot (2b_{i_0})! \gneq (a_{i_0}+b_{i_0})!^2$. Thus,

$$\#Aut(R_1 \times R_1) \cdot \#Aut(R_2 \times R_2) > (\#Aut(R_1 \times R_2))^2.$$

□ ∎

As a corollary of this we get:

Theorem 2.5 *Graph Isomorphism* $\leq_T^P$ *#RA.*

Proof: Immediate from Theorem 2.2 and Lemma 2.1. ∎

Another interesting problem that reduces to #RA is integer factorization (IF).

Theorem 2.6 *IF* $\leq_T^{ZPP}$ *#RA.*

Proof: Let n be the *odd* integer to be factored. Consider the ring

$$R := (\mathbb{Z}/n\mathbb{Z})[x]/(x^2)$$

We will show that $\#Aut(R) = \phi(n) := |(\mathbb{Z}/n\mathbb{Z})^*|$. The theorem is then immediate as n can be factored in expected polynomial time if we are given $\phi(n)$, see [Mil76].

Suppose $\psi \in Aut(R)$ and let $\psi(x) = ax + b$, for some $a, b \in \mathbb{Z}/n\mathbb{Z}$. Since ψ is an automorphism; a, b should satisfy the following two conditions:

$$(ax + b)^2 = 0 \text{ in } R \Rightarrow ab = b^2 = 0 \pmod{n}, \text{ and}$$

$$a \in (\mathbb{Z}/n\mathbb{Z})^*.$$

These two conditions force $b = 0$ and any $a \in (\mathbb{Z}/n\mathbb{Z})^*$ will work. Thus, $\#Aut(R) = |(\mathbb{Z}/n\mathbb{Z})^*| = \phi(n)$. ∎

2.5 The Complexity of Finding a Ring Isomorphism

We have seen by now that ring isomorphism and its counting version are both of intermediate complexity and some well known problems – integer factoring and graph isomorphism – reduce to them. Another interesting variant of RI is its search version – FRI – *finding* an isomorphism between two rings given in basis form. The first question that arises here is whether we can find a ring isomorphism given oracles to RI or #RI. This is still open but in this section we show that FRI seems to have a complexity similar to that of RI and #RI.

2.5.1 An Upper Bound

FRI is unlikely to be NP hard as we show that it reduces to the problem of computing the automorphism group of a ring – GroupRA. The idea is that if we want to find an isomorphism from a ring R to R' then we consider the ring $S = R \times R'$ and compute the generator set T of $Aut(S)$. Now if $R \cong R'$ then there will be a generator $\phi \in T$

that sends some elements of R to those of R'. We construct an isomorphism from $R \to R'$ using this automorphism ϕ of $R \times R'$.

Theorem 2.7 *$FRI \in FP^{GroupRA} \subseteq fnAM$.*

Proof: Let R, R' be the two isomorphic finite rings given in basis form. Let their decomposition into indecomposable components be:

$$R = R_1 \times \cdots \times R_s$$
$$R' = R'_1 \times \cdots \times R'_s$$

Suppose an oracle to GroupRA queried on $S := R \times R'$ gives the group $Aut(S)$ in terms of a generator set T. For concreteness, let us fix an additive basis of S: $\{b_1, \ldots, b_n, b'_1, \ldots, b'_n\}$ where $\{b_1, \ldots, b_n\}$ are the basis elements of R and $\{b'_1, \ldots, b'_n\}$ are those of R'. Furthermore, as S is a direct product of R and R' we have: for all $i, j \in [n]$, $b_i \cdot b'_j = b'_i \cdot b_j = 0$. If $R \cong R'$ then there has to be an element $\phi \in T$ that maps some basis elements of R outside R. Fix such an automorphism ϕ. For $i \in [n]$, let:

$$\phi(b_i) = \sum_{j=1}^{n} a_{i,j} b_j + \sum_{j=1}^{n} a'_{i,j} b'_j$$

where, $a_{i,j}$'s and $a'_{i,j}$'s are integers modulo the characteristic of S, say N.

Now using linear algebra (over $\mathbb{Z}/N\mathbb{Z}$) we can compute an additive basis of the following subring of R:

$$K := \{r \in R \mid \phi(r) \in R\}$$

Note that K is a (proper) subring of R simply because ϕ is a ring homomorphism. Now since ϕ is an automorphism and the decomposition of a ring into indecomposable rings is unique (see Lemma A.2 for details) we get that ϕ applied on S permutes $R_1, \ldots, R_s, R'_1, \ldots, R'_s$ up to isomorphism. This means that there are $\{i_1, \ldots, i_t\} \subsetneq [s]$ such that:

$$K = R_{i_1} \times \cdots \times R_{i_t}$$

Again by linear algebra we can compute the 'other' component ring:

$$K^{\perp} := \{r \in R \mid K \cdot r = r \cdot K = 0\}$$

which can be shown to satisfy:

$$R = K \times K^{\perp}$$

Now what is the action of ϕ on these? Observe that $\phi(K) \subseteq R$ while $\phi(K^{\perp}) \subseteq R'$. To get a decomposition of R' too, define $L := \phi(K^{\perp})$ and compute:

$$L^{\perp} := \{r \in R' \mid L \cdot r = r \cdot L = 0\}$$

which can again be shown to satisfy:

$$R' = L \times L^{\perp}$$

(as ϕ is an isomorphism from $K^{\perp} \rightarrow L$ and $R \cong R'$).

Now recursively find an isomorphism ψ from K to $L^{\perp}$ using GroupRA as oracle. ϕ and ψ together give us an isomorphism from R to R'.

Thus, FRI $\in \mathrm{FP}^{\mathrm{GroupRA}}$.

∎

2.5.2 A Lower Bound: Reduction from Integer Factoring

It turns out that solving FRI would mean solving integer factoring (IF).

Theorem 2.8 (Kayal) *IF* $\leq_T^{ZPP}$ *FRI.*

Proof: Suppose n is an odd number to be factored and it is not a prime power. Pick a random $a \in (\mathbb{Z}/n\mathbb{Z})^*$ and define the rings:

$$R_1 := (\mathbb{Z}/n\mathbb{Z})[x]/(x^2 - a^2) \quad \text{and} \quad R_2 := (\mathbb{Z}/n\mathbb{Z})[x]/(x^2 - 1).$$

Query the oracle of FRI on (R_1, R_2) to get an isomorphism $\phi : R_1 \rightarrow R_2$. Let $\phi(x) = bx + c,\ b, c \in \mathbb{Z}/n\mathbb{Z}$.

Firstly, observe that if b is a zero divisor, i.e., there is a $b' \in (\mathbb{Z}/n\mathbb{Z}) \setminus \{0\}$ with $bb' = 0$ then $\phi(b'x - b'c) = b'(bx + c) - b'c = 0$ in R_2. As ϕ is an isomorphism this means that $(b'x - b'c) = 0$ in R_1 implying that $b' = 0$ in $\mathbb{Z}/n\mathbb{Z}$ which is a contradiction. Thus, b should be in $(\mathbb{Z}/n\mathbb{Z})^*$.

Secondly, $\phi(x^2 - a^2)$ should be zero in R_2 which means that:

$$\begin{aligned} & a^2 = \phi(x)^2 = (bx + c)^2 \pmod{n, x^2 - 1} \\ \Rightarrow \quad & 2bc = 0 (\text{mod } n) \text{ and } b^2 + c^2 - a^2 = 0 (\text{mod } n) \\ \Rightarrow \quad & c = 0 (\text{mod } n) \text{ and } b^2 = a^2 (\text{mod } n) \end{aligned}$$

This means that b is a square-root of a^2 modulo n. It is easily seen that when n has two or more prime factors then every square in $(\mathbb{Z}/n\mathbb{Z})^*$ has 4 or more square-roots. Thus,

$$\text{Prob}_{a \in (\mathbb{Z}/n\mathbb{Z})^*}[b \neq \pm a (\text{mod } n) \mid b = \sqrt{a^2} (\text{mod } n)] \geq \frac{1}{2}.$$

Now once we have a $b \neq \pm a (\text{mod } n)$ such that $b^2 = a^2 (\text{mod } n)$ we can factor n by using the standard trick of computing $gcd(b - a, n)$.

Thus, we can factor n in expected polynomial time given an oracle to FRI. ∎

2.6 The Complexity of Deciding and Finding Ring Automorphism

This section studies the problem of checking whether a given ring is rigid (i.e., has no nontrivial automorphism) and if not then finding a nontrivial automorphism. We will show that RA can be decided in deterministic polynomial time but finding a nontrivial automorphism (FRA) is as hard as integer factoring.

Thus, there appears to be a difference in the complexity of decision, search and counting versions of ring automorphism problems. Also, note the contrast that we (currently) have with the complexity of the corresponding versions for graph automorphism problems, for instance, GA is not known to be in P.

2.6.1 Kayal's algorithm for RA

Theorem 2.9 (Kayal) *$RA \in P$.*

Proof: We only give the outline here. Refer to [Kay06] for details.

Let R be a finite ring given in basis form. The algorithm for checking whether R is rigid or not follows from the classification of rigid rings that we prove.

Let us first dispose off the case when R is non-commutative.

Claim 2.9.1 *If R is a non-commutative ring then it has a nontrivial automorphism.*

Proof of Claim 2.9.1. It can be shown [Len04] that if the *units* in a ring R commute with the whole of R then $R = \langle R^* \rangle$, and consequently R will be commutative. Thus, if R is a noncommutative ring then there is a *unit* $r \in R$ that does not commute with the whole of R. Then clearly the map $\phi : x \mapsto rxr^{-1}$ gives a nontrivial automorphism of R. □

When R is commutative we first consider the case of odd sized R. We can show that indecomposable *components* of a rigid commutative odd-sized ring R are isomorphic to $\mathbb{Z}/p^m\mathbb{Z}$, for some odd prime p:

Claim 2.9.2 *If R is a rigid indecomposable commutative odd-sized ring then there is a prime p and $m \in \mathbb{N}$ such that, $R \cong (\mathbb{Z}/p^m\mathbb{Z}, +, \cdot)$.*

Finally, we take up the case of even sized commutative ring. It is sufficient to consider a ring R whose size is a power of 2. We can show that R is rigid only if the indecomposable rings that appear in the decomposition of R are isomorphic to either $\mathbb{Z}/2^m\mathbb{Z}$ or $(\mathbb{Z}/2\mathbb{Z})[x]/(x^2)$.

Claim 2.9.3 *If R is a rigid indecomposable commutative power-of-2 sized ring then R is either $(\mathbb{Z}/2^m\mathbb{Z}, +, \cdot)$ or $(\mathbb{Z}/2\mathbb{Z})[x]/(x^2)$.*

There is a polynomial time algorithm (see [Kay06]) that checks whether all the indecomposable components of a given finite ring R fall in one of the above categories or not, hence proving RA ∈ P. ∎

2.6.2 FRA is randomly equivalent to Integer Factoring

We just saw that deciding whether a ring has a nontrivial automorphism is in P, here we give evidence that the search version of this problem is apparently harder. We show that FRA is as hard as integer factoring (IF).

Theorem 2.10 (Kayal) *$IF \equiv_T^{ZPP} FRA$.*

Proof: We only give the outline here. Refer to [Kay06] for details.

Let us first sketch how we can find a nontrivial ring automorphism if we can do integer factoring. Suppose the given ring R is non-commutative then we know from the proof of Claim 2.9.1: there is a *unit* of R that does not commute with the whole of R and thus defines a nontrivial automorphism. So we compute the multiplicative generators of R^* in *randomized* polynomial time and surely one of the generators r will not commute with the whole of ring R. Thus, $\phi : x \mapsto rxr^{-1}$ is a nontrivial automorphism of R.

Now assume the given ring R is commutative. It can be decomposed into local rings, as remarked in Proposition 2.3, in expected polynomial time using randomized methods for polynomial factorization and an oracle of integer factorization. Once we have local rings it is easy to construct nontrivial automorphisms. For, suppose that the maximal ideal of R is $\mathcal{M}$ and $t \in \mathbb{N}$ is such that $\mathcal{M}^{t-1} \neq 0$ but $\mathcal{M}^t = 0$ also let $1, b_1, \ldots, b_n$ be an additive basis with $b_1, \ldots, b_n$ as nilpotents. Then in most cases any element $\alpha \in \mathcal{M}^{t-1}$ defines a nontrivial automorphism:

$$\phi : \begin{cases} b_1 & \mapsto b_1 + \alpha \\ b_2 & \mapsto b_2 \\ & \vdots \\ b_n & \mapsto b_n \end{cases}$$

Conversely, suppose we can find nontrivial automorphisms of rings and n is a given number. Let us assume for simplicity that input n is a product of two distinct primes p, q. Randomly choose a monic cubic polynomial $f(x) \in (\mathbb{Z}/n\mathbb{Z})[x]$. Define $R := (\mathbb{Z}/pq\mathbb{Z})[x]/(f(x))$ and suppose we can find a nontrivial automorphism ϕ of R. It follows from the distribution of irreducible polynomials over finite fields (see [LN86]) that with probability $\sim \frac{1}{9}$: $f(\text{mod } q)$ is irreducible *and* $f(\text{mod } p)$ has exactly two irreducible factors f_1, f_2, say f_1 is linear. Thus ring R decomposes as:

$$R \cong (\mathbb{Z}/p\mathbb{Z}) \times (\mathbb{Z}/p\mathbb{Z})[x]/(f_2(x)) \times (\mathbb{Z}/q\mathbb{Z})[x]/(f(x)).$$

Note that we can compute R^ϕ, the set of elements of R fixed by ϕ, using linear algebra (if at any point we cannot invert an element (mod n), we get a factor of n). As ϕ is a nontrivial automorphism of R we have that ϕ is identity on at most one of the component rings $(\mathbb{Z}/p\mathbb{Z})[x]/(f_2(x))$ or $(\mathbb{Z}/q\mathbb{Z})[x]/(f(x))$. Thus, we have three cases:

1) If ϕ fixes $(\mathbb{Z}/p\mathbb{Z})[x]/(f_2(x))$:
   Then $R^\phi \cong (\mathbb{Z}/p\mathbb{Z}) \times (\mathbb{Z}/p\mathbb{Z})[x]/(f_2(x)) \times (\mathbb{Z}/q\mathbb{Z})$. Thus, $|R^\phi| = p^3q$.

2) If ϕ fixes $(\mathbb{Z}/q\mathbb{Z})[x]/(f(x))$:
   Then $R^\phi \cong (\mathbb{Z}/p\mathbb{Z}) \times (\mathbb{Z}/p\mathbb{Z}) \times (\mathbb{Z}/q\mathbb{Z})[x]/(f(x))$. Thus, $|R^\phi| = p^2q^3$.

3) If ϕ moves both $(\mathbb{Z}/p\mathbb{Z})[x]/(f_2(x))$ and $(\mathbb{Z}/q\mathbb{Z})[x]/(f(x))$:
   Then $R^\phi \cong \mathbb{Z}_p \times \mathbb{Z}_p \times \mathbb{Z}_q$. Thus, $|R^\phi| = p^2q$.

Since, the size of R^ϕ is in no case of the form n, n^2 or n^3, the process of finding R^ϕ by doing linear algebra (mod n) is going to yield a factor of n. In particular, this means that if the matrix describing ϕ over the natural additive basis $\{1, x, x^2\}$ is:

$$A := \begin{pmatrix} 1 & 0 & 0 \\ a_0 & a_1 & a_2 \\ b_0 & b_1 & b_2 \end{pmatrix}$$

then the determinant of one of the submatrices of $(A - I)$ will have a nontrivial gcd with n.

This idea can be extended to the case of composite n having more prime factors (see [Kay06]).

Thus, the two problems: finding nontrivial automorphisms of commutative rings and integer factoring have the same complexity (with respect to randomized polynomial time reductions). ∎

2.6.3 Reduction from Polynomial Factoring to FRA

Polynomial factorization over finite fields is still not known to have a deterministic polynomial time algorithm. The randomized algorithms known for this problem (see [LN86, vzGG99]) invariably use automorphisms of rings as a tool (see [AS05]).

Here we give a specific relation of polynomial factorization to FRA assuming the extended Riemann hypothesis (ERH). ERH needs to be invoked as it gives us a deterministic polynomial time algorithm to find k-th roots in a finite field [vzGG99]. The reduction we give here uses the main idea of Evdokimov's algorithm [Evd94].

Theorem 2.11 *Assuming the ERH, Polynomial Factoring* $\leq_m^P$ *FRA.*

Proof: Suppose we want to factor a polynomial $f(x)$ over the finite field $\mathbb{F}_q$. We could assume wlog that $f(x)$ is square free and splits completely over $\mathbb{F}_q$. Let us define a ring $R := \mathbb{F}_q[x]/(f(x))$ and let d be the degree of $f(x)$. Suppose an oracle of FRA gives a nontrivial automorphism ϕ of the ring R. We will show how to find a factor of $f(x)$ assuming ERH.

We can first easily compute the subring R^ϕ of elements in R which are fixed by ϕ. If $x, \phi(x), \phi^2(x), \ldots, \phi^d(x)$ are all distinct modulo $f(x)$ then we have $(d+1)$ roots of degree-d-polynomial $f(x)$ which implies that $\exists i \neq j$ *s.t.* $gcd(\phi^i(x) - \phi^j(x), f(x))$ factors $f(x)$. So we can assume that for some $2 \leq k \leq d$, $\phi^k(x) = x$.

Wlog we can assume that there is a k-th root of unity $\zeta_k \in \mathbb{F}_q$, for otherwise, we can invoke ERH and construct an appropriate extension of $\mathbb{F}_q$ that has a k-th root of unity [vzGG99]. Consider the element:

$$\beta := \sum_{i=0}^{k-1} \zeta_k^i \phi^i(x) \in R$$

which satisfies $\phi(\beta) = \zeta_k^{-1}\beta$. Thus, $\beta^k \in R^\phi$ but $\beta \notin R^\phi$. Also, note that β^k has a k-th root y in the ring R^ϕ, for, β^k has a k-th root in $R \cong \times_{i=1}^{d} \mathbb{F}_q$ and R^ϕ is just a subring $\times_{i=1}^{d'} \mathbb{F}_q$ of R. Also, we can compute $y \in R^\phi$ as we are assuming ERH (the k-th root finding algorithm either gives a k-th root of β^k in R^ϕ *or* factors $f(x)$). But then we have $(k+1)$ k-th roots of β^k which are all distinct modulo $f(x)$, namely: $\beta, \zeta_k\beta, \ldots, \zeta_k^{k-1}\beta, y$; thus, there are two roots among these whose difference is a zero divisor of the ring R and hence will give a nontrivial gcd with $f(x)$. ∎

2.7 Discussion

This chapter studied the automorphism and isomorphism problems of rings. The problems were all inspired from those of graphs. The rings considered in this chapter were assumed to be finite which was used in showing that these problems are of intermediate complexity and unlikely to be NP-hard. On the other hand all the lower bound results of the chapter do not need this finiteness assumption, for example, graph isomorphism reduces to the isomorphism problem of $\mathbb{F}$-algebras for *any* field $\mathbb{F}$. This chapter showed that the automorphism problems of finite rings are related to the classical problems – like, graph isomorphism, integer factoring and polynomial factoring – and the most general automorphism problem is computing the group of automorphisms of a finite ring.

The complexity of all the morphism problems, except RA and testing automorphism/isomorphism problems, that we considered in this chapter remain open. A solution to any one of them will be very interesting as it would solve some of the classical problems as well! To understand these problems more we would like to ask the following questions:

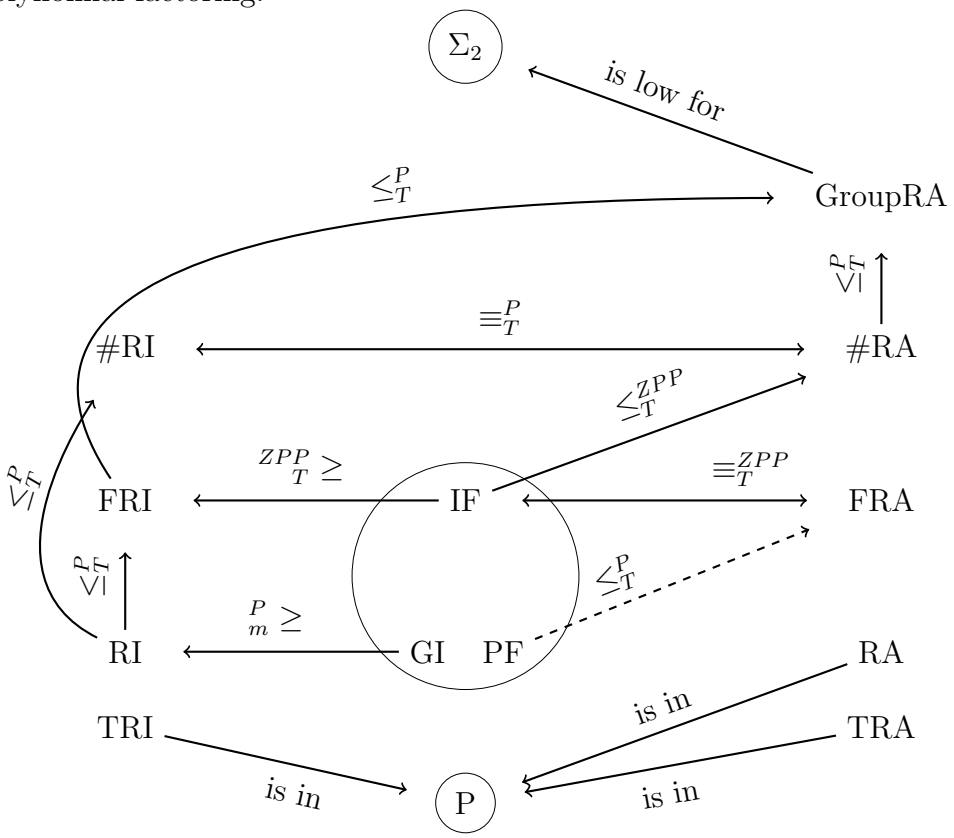
- We have seen two well-known problems of intermediate complexity reduce to #RA. Can one reduce some other such problem, e.g., finding discrete logarithm?

- The ring problems differ from the graph ones in their (in)ability to efficiently "fix" part of the automorphisms. This property allows one to prove the equivalence between computing automorphism groups, counting automorphisms, finding isomorphisms, and testing isomorphisms in the case of graphs. For rings, we cannot prove such equivalence. Does there exist some way of doing such "fixing" for rings which will allow us to prove similar equivalences?

- As #RA is an algebraic problem is there a polynomial time quantum algorithm for it, i.e., is #RA $\in$ BQP ?

- Consider the ring isomorphism problem over rationals: $\mathrm{RI}_{\mathbb{Q}}$. It is not even clear if this problem is decidable.

The following figure shows the various relations we proved in this chapter. The arrows are labelled by the type of reduction or relation and the dotted arrow signifies a conditional result (assuming ERH). The well-known problems are in the central circle and labelled as: IF for integer factoring, GI for graph isomorphism and PF for polynomial factoring.



Chapter 3

Polynomial Equivalence

Suppose we are given two polynomials $f(x_1, \ldots, x_n)$ and $g(x_1, \ldots, x_n)$ of total degree d with coefficients in a field $\mathbb{F}$. We say that f is *equivalent* to g, denoted by $f \sim g$, if there is an invertible linear transformation τ sending each x_i to a linear combination of $x_1, \ldots, x_n$ such that:

$$f\left(\tau(x_1), \ldots, \tau(x_n)\right) = g(x_1, \ldots, x_n).$$

The polynomials f, g are assumed to be provided in the input in *expanded* form:

$$\sum_{0 \le i_1 + \ldots + i_n \le d} a_{i_1, \ldots, i_n} x_1^{i_1} \cdots x_n^{i_n}$$

Example Suppose $f(x, y) = x^2 + y^2$ and $g(x, y) = 2x^2 + 2y^2$ are polynomials over $\mathbb{Q}$. Then the map $\tau : \begin{cases} x \mapsto x + y \\ y \mapsto x - y \end{cases}$ applied on f gives g, i.e., $\tau \circ f(x, y) = g(x, y)$. Thus, $f \sim g$ over rationals. ∎

Example Consider $f(x) = x^2$ and $g(x) = 2x^2$. Then f and g are not equivalent over $\mathbb{Q}$ but they are equivalent over $\mathbb{R}$ as $\tau : x \mapsto \sqrt{2}x$ is an equivalence. ∎

The computational problem of *polynomial equivalence* is to check whether two input polynomials $f, g \in \mathbb{F}[\overline{x}]$ are equivalent, in time polynomial in the size of the input. We treat d as a constant while n varies. We show in this chapter that this easily defined problem is apparently harder than commutative $\mathbb{F}$-algebra

isomorphism ($\mathbb{F}$-algebras given in the basis form) and hence as a corollary the graph isomorphism problem too reduces to polynomial equivalence. Also, in the other direction most cases of polynomial equivalence reduce to the commutative $\mathbb{F}$-algebra isomorphism problem.

Previous research on polynomial equivalence has primarily focussed on a restricted case – when f, g are homogeneous polynomials called *forms*. The most celebrated case is perhaps when f, g are *quadratic forms* – homogeneous polynomials of degree 2. The classification of quadratic forms is known due to the works of Minkowski [Minkow], Hasse [Has21] and Witt [Witt]. The classification theorem of quadratic forms is effective in the sense that it gives algorithms for deciding and finding quadratic forms equivalence over "interesting" fields like finite fields, $\mathbb{Q}, \mathbb{R}$ and $\mathbb{C}$.

In this work we focus on polynomial equivalence for homogeneous polynomials of degree 3 – *cubic forms*. This case of polynomial equivalence seems to be significantly harder than quadratic forms equivalence as we show that a fairly general case of ring isomorphism – commutative $\mathbb{F}$-algebra isomorphism – reduces to cubic forms equivalence. This reduction implies that graph isomorphism reduces to cubic forms equivalence too. Moreover, we also give evidence that the problem of equivalence for higher degree forms reduces to that of cubic forms. Thus, cubic forms seem to be the most important restricted case of polynomial equivalence. Cubic forms equivalence has been well studied in mathematics (for instance see [Har75, HP88, MH74, Rup03]). Over the last ten years, it has been found to be useful in computer science as well: [Pat96, CGP98] propose a cryptosystem based on the hardness of the cubic forms equivalence over finite fields.

The results of this chapter mostly appear in [AS05, AS06].

3.1 The Complexity of Polynomial Equivalence

For a given field $\mathbb{F}$ and degree d let us define the language for the problem of polynomial equivalence over $\mathbb{F}$ as:

$$\text{polyEquiv}_{d,\mathbb{F}} := \{(f, g) \mid f, g \text{ are polynomials of total degree } d \text{ over } \mathbb{F} \text{ and } f \sim g\}$$

3.1.1 Upper Bounds

The complexity of polynomial equivalence depends upon the base field. In this section we give upper bounds on polynomial equivalence for various "interesting" fields.

Theorem 3.1 *For any fixed $d \in \mathbb{Z}^{>0}$, the problem of polynomial equivalence satisfies:*

1) For a finite field $\mathbb{F}$, $polyEquiv_{d,\mathbb{F}} \in NP \cap coAM$.

2) When $\mathbb{F} = \mathbb{R}$, $polyEquiv_{d,\mathbb{F}} \in EEXP$.

3) For an algebraically closed field $\mathbb{F}$ (eg. $\mathbb{C}$), $polyEquiv_{d,\mathbb{F}} \in PSPACE$.

Proof: [of 1)] Let $\mathbb{F}$ be a finite field of size q. Given a linear transformation τ on the variables $x_1, \ldots, x_n$, it is easy to check whether $f(\tau x_1, \ldots, \tau x_n) = g(x_1, \ldots, x_n)$ simply by substituting for τ in f and doing the computations in time $poly(n^d, \log q)$. Thus, polynomial equivalence over $\mathbb{F}$ is in NP.

Let us now see an AM protocol for $\overline{\text{polyEquiv}}_{\mathbb{F}}$. Suppose $f, g \in \mathbb{F}[x_1, \ldots, x_n]$ are two given polynomials. We call an invertible linear transformation $\phi \in (\mathbb{F}^{n\times n})^*$ an *automorphism of f* if $f(\phi\overline{x}) = f(\overline{x})$. Let us define a set $C(f)$ as:

$$C(f) := \left\{(f(\tau\overline{x}), \phi) \mid \tau, \phi \in \left(\mathbb{F}^{n\times n}\right)^* \text{ and } \phi \text{ is an automorphism of } f(\tau\overline{x})\right\}$$

If s is the number of invertible $n \times n$ matrices over $\mathbb{F}$ then observe that the size of the set $C(f)$ is:

$$\begin{aligned} \#C(f) &= (\text{number of polynomials } \sim f(\overline{x})) \cdot \#Aut(f) \\ &= \frac{s}{\#Aut(f)} \cdot \#Aut(f) \\ &= s \end{aligned}$$

Similarly, we have the set $C(g)$ and we define $C(f, g) = C(f) \cup C(g)$. It is a simple exercise to show that given $\mathbb{F}_q$ and n we can compute the number s of $n \times n$ invertible matrices over $\mathbb{F}_q$ in polynomial time. Now let us see how $C(f, g)$ behaves:

$$f \not\sim g \Rightarrow C(f) \cap C(g) = \emptyset \Rightarrow \#C(f, g) = 2s.$$
$$f \sim g \Rightarrow C(f) = C(g) \Rightarrow \#C(f, g) = s.$$

Thus, the set $C(f,g)$ is larger by a factor of 2 when $f \not\sim g$ and we can give an AM protocol for $\overline{polyEquiv}_{d,\mathbb{F}}$ as in Proposition 2.4. ∎

Proof: [of 2)] When $\mathbb{F} = \mathbb{R}$, we consider the equivalence as a matrix A over $\mathbb{R}$ in n^2 unknowns $((a_{i,j}))$ and then solve the system of equations that we get from:

$$f(A\overline{x}) = g(\overline{x})$$

This system of equations can be solved in EEXP due to the result of Tarski on the decidability of first-order equations over reals [DH88]. ∎

Proof: [of 3)] When $\mathbb{F}$ is an algebraically closed field, we consider the equivalence as a matrix A over $\mathbb{F}$ in n^2 unknowns $((a_{i,j}))$ and then solve the system of equations that we get from:

$$f(A\overline{x}) = g(\overline{x})$$

This system of equations can be solved over $\mathbb{F}$ in PSPACE by using Hilbert's Nullstellensatz [Bro87]. ∎

Remark: When $\mathbb{F} = \mathbb{Q}$, it is not yet known if the problem is decidable. ∎

3.1.2 Reduction to $\mathbb{F}$-algebra Isomorphism (in some cases)

At the first glance, the problem of polynomial equivalence does not appear to be related to the problems of ring isomorphism. But in this section we exhibit a connection of polynomial equivalence to the ring isomorphism problem. We show that the problem of polynomial equivalence restricted to homogeneous polynomials reduces to the ring isomorphism problem for most cases.

Theorem 3.2 *Suppose $\mathbb{F}$ is a field having d^{th} roots, i.e. $\forall \alpha \in \mathbb{F}\ \alpha^{\frac{1}{d}} \in \mathbb{F}$. Then equivalence of homogeneous polynomials of degree d over $\mathbb{F}$ is many-one polynomial time reducible to $\mathbb{F}$-algebra isomorphism.*

Proof: Suppose f, g are homogeneous polynomials of degree d in n variables over $\mathbb{F}$. Then construct a commutative $\mathbb{F}$-algebra R_f from f as:

$$R_f := \mathbb{F}[x_1, \ldots, x_n] / (f, \mathcal{I}_{d+1})$$

where, the ideal $\mathcal{I}_{d+1}$ is generated by all the monomials of degree $d+1$. We claim that the rings R_f and R_g are isomorphic iff $f \sim g$.

Suppose ψ is an equivalence that sends f to g. Then ψ easily extends to an isomorphism from R_f to R_g.

Conversely, suppose ϕ is an isomorphism from $R_f \to R_g$. Then $\phi(f)$ has to map to 0 in R_g thus, there is a $c \in \mathbb{F}$ such that:

$$\phi(f) = cg(\overline{x}) + (\text{terms of degree } d+1 \text{ or more}) \tag{3.1}$$

Since, $x_i^{d+1} = 0$ in R_f, $\phi(x_i)$ cannot have a constant term otherwise $\phi(x_i)^{d+1} \neq 0$. Let us denote the linear part of $\phi(x_i)$ by $\psi(x_i)$. Hence, for all $i \in [n]$:

$$\phi(x_i) = \psi(x_i) + (\text{quadratic and higher degree terms})$$

Since, f is homogeneous of degree d, the degree d terms of $\phi(f)$ are exactly those in $\psi(f)$. Thus:

$$\phi(f) = \psi(f) + (\text{terms of degree } d+1 \text{ or more}) \tag{3.2}$$

The Equations (3.1) and (3.2) imply that $\psi(f) = cg$. Now since $\mathbb{F}$ has d^{th} roots and g is homogeneous of degree d we further get:

$$f(\psi(x_1), \ldots, \psi(x_n)) = g(c^{\frac{1}{d}}x_1, \ldots, c^{\frac{1}{d}}x_n)$$

Thus, $f \sim g$.

Hence, $R_f \cong R_g$ iff $f \sim g$. ∎

Remark: If one slightly generalizes the definition of polynomial equivalence as $f \sim g$ iff there is a $\tau \in \mathbb{F}^{n \times n}$ and a $c \in \mathbb{F}$ such that $f(\tau(\overline{x})) = c \cdot g(\overline{x})$ then this theorem works for all fields $\mathbb{F}$. ∎

3.1.3 A Lower Bound: Reduction from $\mathbb{F}$-algebra Isomorphism

Here, we will show that a fairly general case of the ring isomorphism problem – commutative $\mathbb{F}$-algebra isomorphism – reduces to the equivalence problem of polynomials having total degree 3 (called cubic polynomials).

An isomorphism of $\mathbb{F}$-algebras has to preserve all the multiplicative relations, which are $\sim n^2$ if there are n basis elements. On the other hand an equivalence of polynomials has to satisfy only one equation. It is interesting that there is a way to combine the various multiplicative relations of a commutative $\mathbb{F}$-algebra into one polynomial such that its equivalence gives an $\mathbb{F}$-algebra isomorphism.

Theorem 3.3 *Commutative $\mathbb{F}$-algebra Isomorphism $\leq_m^P$ cubic polynomial equivalence.*

Proof: Let R be a commutative $\mathbb{F}$-algebra with additive basis $b_1, \ldots, b_n$ over $\mathbb{F}$. Furthermore, multiplication in R is defined as: for all $1 \leq i \leq j \leq n$,

$$b_i \cdot b_j = \sum_{k=1}^{n} a_{i,j,k} b_k, \quad \text{where, } a_{i,j,k} \in \mathbb{F}$$

Let us define a polynomial that *captures* the multiplicative relations defining ring R:

$$f_R(\overline{z}, \overline{b}) := \sum_{1 \leq i \leq j \leq n} z_{i,j} \left(b_i b_j - \sum_{1 \leq k \leq n} a_{i,j,k} b_k \right) \tag{3.3}$$

Note that here $\overline{z} = (z_{1,1}, \ldots, z_{n,n})$ and $\overline{b} = (b_1, \ldots, b_n)$ are formal variables and f_R is a polynomial in $\mathbb{F}[\overline{z}, \overline{b}]$. Similarly, for another commutative $\mathbb{F}$-algebra R' the polynomial would be:

$$f_{R'}(\overline{z}, \overline{b}) := \sum_{1 \leq i \leq j \leq n} z_{i,j} \left(b_i b_j - \sum_{1 \leq k \leq n} a'_{i,j,k} b_k \right)$$

An isomorphism from R to R' easily gives an equivalence from f_R to $f_{R'}$:

Claim 3.3.1 *If $R \cong R'$ then $f_R \sim f_{R'}$.*

Proof of Claim 3.3.1. Let ϕ be an isomorphism from R to R'. Note that ϕ sends each b_i to a linear combination of b's and for all $i \leq j \in [n]$: $\phi(b_i)\phi(b_j) - \sum_{1 \leq k \leq n} a_{i,j,k} \phi(b_k) = 0$ in R'. This implies that there exist constants $c_{i,j,k,\ell} \in \mathbb{F}$ such that:

$$\phi(b_i)\phi(b_j) - \sum_{1 \leq s \leq n} a_{i,j,s} \phi(b_s) = \sum_{1 \leq k \leq \ell \leq n} c_{i,j,k,\ell} \left(b_k b_\ell - \sum_{1 \leq s \leq n} a'_{k,\ell,s} b_s \right)$$

This immediately suggests that the linear transformation τ that sends:

$$\text{for all } 1 \leq i \leq n, \qquad b_i \mapsto \phi(b_i)$$

$$\text{for all } 1 \leq k \leq \ell \leq n, \qquad \left(\sum_{1\leq i\leq j\leq n} c_{i,j,k,\ell} z_{i,j}\right) \mapsto z_{k,\ell}$$

makes f_R equal to $f_{R'}$. The linear transformation τ is an invertible map because $\tau\mid_{\overline{b}} = \phi$ is invertible and $\tau\mid_{\overline{z}}$ has a range space of full dimension implying that $\tau\mid_{\overline{z}}$ is invertible too. □

The converse, i.e., getting an $\mathbb{F}$-algebra isomorphism from a polynomial equivalence, is more involved to show.

Claim 3.3.2 *If $f_R \sim f_{R'}$ then $R \cong R'$.*

Proof of Claim 3.3.2. Let ϕ be a linear transformation such that

$$\sum_{1\leq i\leq j\leq n} \phi(z_{i,j}) \left(\phi(b_i)\phi(b_j) - \sum_{1\leq k\leq n} a_{i,j,k}\phi(b_k)\right) = \sum_{1\leq i\leq j\leq n} z_{i,j}\left(b_i b_j - \sum_{1\leq k\leq n} a'_{i,j,k} b_k\right) \tag{3.4}$$

By comparing the cubic terms on both sides we get:

$$\sum_{1\leq i\leq j\leq n} \phi(z_{i,j})\phi(b_i)\phi(b_j) = \sum_{1\leq i\leq j\leq n} z_{i,j} b_i b_j \tag{3.5}$$

We aim to show that $\phi(b_i)$ has no z's, *i.e.*, $\phi(b_i)$ is a linear combination of only b's. We will be relying on the following property of the RHS of Equation (3.5): if τ is an invertible linear transformation on the z's then for all $1 \leq i \leq j \leq n$, the coefficient of $z_{i,j}$ in $\sum_{1\leq i\leq j\leq n} \tau(z_{i,j}) b_i b_j$ is nonzero.

Suppose $\phi(b_{i_0})$ has z's, i.e.,

$$\phi(b_{i_0}) = \sum_j c_{i_0,j} b_j + \sum_{j,k} c_{i_0,j,k} z_{j,k}$$

We can apply an invertible linear transformation τ on z's in Equation (3.5) such that τ maps $\sum_{j,k} c_{i_0,j,k} z_{j,k}$ to $z_{1,1}$. Then apply an evaluation map val that substitutes $z_{1,1}$ by $\left(-\sum_j c_{i_0,j} b_j\right)$. Now $val \circ \tau \circ \phi(b_{i_0}) = 0$ and thus, Equation (3.5) becomes:

$$\sum_{\substack{1\leq j\leq k\leq n \\ j,k\neq i_0}} val \circ \tau \circ \phi(z_{j,k} b_j b_k) \;=\; \sum_{\substack{1\leq j\leq k\leq n \\ (j,k)\neq(1,1)}} z_{j,k}(\text{quadratic } b\text{'s}) + (\text{cubic } b\text{'s}) \tag{3.6}$$

Notice that the LHS of Equation (3.5) had $\binom{n+1}{2}$ summands while the LHS of Equation (3.6) has at most $\left\{\binom{n+1}{2} - n\right\}$ summands. These summands on the LHS of Equation (3.6) are of two kinds: those that have a nonzero occurrence of a z-variable and those that are cubic in b's. So we repeat this process of applying invertible linear transformations on z's and fixing z's in Equation (3.6) so that for all $1 \leq j \leq k \leq n,\ j,k \neq i_0,\ val \circ \tau \circ \phi(z_{j,k}b_jb_k)$ either maps to zero or to a cubic in b's. Thus, after $\left\{1 + \binom{n+1}{2} - n\right\}$ z-fixings the LHS of Equation (3.5) is a cubic in b's while the RHS still has $\binom{n+1}{2} - \left\{1 + \binom{n+1}{2} - n\right\} = (n-1)$ unfixed z's, which is a contradiction.

Since $\phi(b_i)$'s have no z's and there are no cubic b's in the RHS of Equation (3.4) we can ignore the b's in $\phi(z_{j,k})$'s. Thus, now $\phi(z_{j,k})$'s are linear combinations of z's and $\phi(b_i)$'s are linear combinations of b's. Again looking at Equation (3.4), this means that $\left(\phi(b_i)\phi(b_j) - \sum_{1\leq s\leq n} a_{i,j,s}\phi(b_s)\right)$ is a linear combination of $(b_kb_\ell - \sum_{1\leq s\leq n} a'_{k,\ell,s}b_s)$ for $1 \leq k \leq \ell \leq n$; implying that $\left(\phi(b_i)\phi(b_j) - \sum_{1\leq s\leq n} a_{i,j,s}\phi(b_s)\right) = 0$ in ring R'. This combined with the fact that $\phi|_{\bar{b}}$ is an invertible linear transformation on $\bar{b}$ means that ϕ induces an isomorphism from ring R to R'. □

The above two claims complete the proof. ∎

3.2 Another Lower Bound: $\mathbb{F}$-algebra Isomorphism reduces to Cubic Forms Equivalence

We had seen in Theorem 3.3 how to construct non homogeneous cubic polynomials that capture the multiplicative relations of a given $\mathbb{F}$-algebra. Now what happens if we homogenize those cubic polynomials, does an equivalence between such cubic forms give us isomorphism between the original $\mathbb{F}$-algebras?

In this section we first give a reduction from commutative $\mathbb{F}$-algebra isomorphism to local commutative $\mathbb{F}$-algebra isomorphism. Then from these local commutative $\mathbb{F}$-algebras we construct cubic forms (obtained by homogenizing Equation (3.3)) and prove that an equivalence between these cubic forms induces an isomorphism between the local commutative $\mathbb{F}$-algebras. Thus, cubic forms equivalence problem

is at least as hard as the isomorphism problem of commutative $\mathbb{F}$-algebras. Consequently, for any field $\mathbb{F}$, cubic forms equivalence problem is at least as hard as the graph isomorphism problem.

3.2.1 Commutative $\mathbb{F}$-algebras reduce to local $\mathbb{F}$-algebras

An $\mathbb{F}$-algebra is *local* if it cannot be broken into simpler $\mathbb{F}$-algebras, *i.e.*, if it cannot be written as a direct product of algebras. Given a commutative $\mathbb{F}$-algebra this direct product decomposition can be done by factoring polynomials over the field $\mathbb{F}$. Any non-unit r in a finite dimensional local commutative $\mathbb{F}$-algebra is *nilpotent*, i.e., there is an m such that $r^m = 0$. For more details on local rings refer the appendix or the text: [McD74].

In this section we give a many-to-one reduction from commutative $\mathbb{F}$-algebra isomorphism to local commutative $\mathbb{F}$-algebra isomorphism. Moreover, the local commutative $\mathbb{F}$-algebras that we construct have basis elements most of whose products vanish. We exploit the properties of this local $\mathbb{F}$-algebra to give a reduction from commutative $\mathbb{F}$-algebra to cubic forms in the next subsection.

Theorem 3.4 *$\mathbb{F}$-algebra isomorphism $\leq_m^P$ Local $\mathbb{F}$-algebra isomorphism.*

Proof: Given two $\mathbb{F}$-algebras R and S, Theorem 3.3 constructs two cubic polynomials p and q respectively such that p, q are equivalent iff R, S are isomorphic. These polynomials live in $\mathbb{F}[z_{1,1}, \ldots, z_{n,n}, b_1, \ldots, b_n]$ and look like:

$$p(\overline{z}, \overline{b}) := \sum_{1 \leq i \leq j \leq n} z_{i,j} \left(b_i b_j - \sum_k a_{i,j,k} b_k \right)$$

$$q(\overline{z}, \overline{b}) := \sum_{1 \leq i \leq j \leq n} z_{i,j} \left(b_i b_j - \sum_k a'_{i,j,k} b_k \right)$$

Let

$$p_3(\overline{z}, \overline{b}) := \sum_{1 \leq i \leq j \leq n} z_{i,j} b_i b_j \quad \text{and} \quad p_2(\overline{z}, \overline{b}) := - \sum_{1 \leq i \leq j \leq n} \left(z_{i,j} \sum_k a_{i,j,k} b_k \right) \tag{3.7}$$

Similarly define $q_3(\overline{z}, \overline{b})$ and $q_2(\overline{z}, \overline{b})$ from q. Thus, $p = p_3 + p_2$ and $q = q_3 + q_2$, where p_3, q_3 are homogeneous of degree 3 and p_2, q_2 are homogeneous of degree 2.

Using p, q we construct the following commutative $\mathbb{F}$-algebras:

$$\begin{aligned} R' &:= \mathbb{F}[\overline{z}, \overline{b}, u]/\left\langle p_3, up_2, u^2, \mathcal{I}\right\rangle \\ S' &:= \mathbb{F}[\overline{z}, \overline{b}, u]/\left\langle q_3, uq_2, u^2, \mathcal{I}\right\rangle \end{aligned} \tag{3.8}$$

where, $\mathcal{I}$ is the ideal generated by all possible products of 4 variables (with repetition) from the set:

$$\{z_{1,1}, \ldots, z_{1,n}, \ldots, z_{n,1}, \ldots, z_{n,n}, b_1, \ldots, b_n, u\}$$

Note that all the variables in R', S' are nilpotent and hence the two rings are *local* commutative $\mathbb{F}$-algebras (see the appendix). The following claim tells us that it is enough to consider the isomorphism problem for these local structures. Recall that $R \cong S$ iff p, q are equivalent polynomials.

Claim 3.4.1 *$p(\overline{z}, \overline{b}), q(\overline{z}, \overline{b})$ are equivalent polynomials* iff $R' \cong S'$.

Proof of Claim 3.4.1. If p, q are equivalent then the same equivalence, extended by sending $u \mapsto u$, gives an isomorphism from R' to S'.

Conversely, say ϕ is an isomorphism from R' to S'. Our intention is to show that the *linear part* of ϕ, i.e., ignoring the quadratic or higher degree terms in $\phi(v)$, where variable $v \in \{z_{1,1}, \ldots, z_{n,n}, b_1, \ldots, b_n, u\}$, induces an equivalence from p to q. Note that since $\overline{z}, \overline{b}, u$ are nilpotents in R', therefore $\forall i \leq j \in [n], k \in [n]$, $\phi(z_{i,j}), \phi(b_k), \phi(u)$ can have no constant term.

Let us see where ϕ sends u. Since, $\phi(u)^2 = 0$ in S', while for all i, j: $z_{i,j}^2$ and b_i^2 are nonzero in S', thus, we deduce that the linear part of $\phi(u)$ can have no $\overline{z}$, $\overline{b}$'s. Further, as ϕ is an isomorphism $\phi(u)$ should have at least one linear term. Thus,

$$\phi(u) \;=\; c \cdot u + (\text{terms of degree 2 or more}), \text{ where } c \in \mathbb{F}^*. \tag{3.9}$$

Now by the definition of ϕ there are $c_1, c_2 \in \mathbb{F}$ such that:

$$\phi(p_3) \;=\; c_1 \cdot q_3 + c_2 \cdot uq_2 + (\text{linear terms in } \overline{z}, \overline{b}, u) \cdot u^2 + (\text{terms of degree 4 or more})$$

By substituting $u = 0$ we get,

$$\phi(p_3)\mid_{u=0} \;=\; c_1 q_3 + (\text{terms of degree 4 or more}) \tag{3.10}$$

Also, there are $d_1, d_2 \in \mathbb{F}$ such that:

$$\phi(up_2) \;=\; d_1 \cdot q_3 + d_2 \cdot uq_2 + (\text{linear terms in } \overline{z}, \overline{b}, u) \cdot u^2 + (\text{terms of degree 4 or more})$$

Using Equation (3.9) we deduce that $d_1 = 0$. Thus,

$$\phi(up_2) \;=\; d_2 \cdot uq_2 + (\text{linear terms in } \overline{z}, \overline{b}, u) \cdot u^2 + (\text{terms of degree 4 or more})$$

As $c \neq 0$ in Equation (3.9), we deduce that there is a $d_2' \in \mathbb{F}$:

$$u\phi(p_2) \;=\; d_2' \cdot uq_2 + (\text{linear terms in } \overline{z}, \overline{b}, u) \cdot u^2 + (\text{terms of degree 4 or more})$$

Factoring out u and substituting $u = 0$ gives us:

$$\phi(p_2) \mid_{u=0} = \; d_2' \cdot q_2 + (\text{terms of degree 3 or more}) \tag{3.11}$$

Let ψ be the linear part of $\phi \mid_{u=0}$, that is:

$$\begin{aligned} \text{for all } i \leq j,\ \psi(z_{i,j}) \;&:=\; \text{linear terms of } \phi(z_{i,j}) \text{ other than } u, \text{ and} \\ \text{for all } i,\ \psi(b_i) \;&:=\; \text{linear terms of } \phi(b_i) \text{ other than } u \end{aligned}$$

By comparing degree 3 and degree 2 terms on both sides of Equations (3.10) and (3.11) respectively, we get:

$$\psi(p_3) \;=\; c_1 q_3 \tag{3.12}$$

$$\psi(p_2) \;=\; d_2' q_2 \tag{3.13}$$

Note that since ϕ is an isomorphism, ψ has to be an invertible map and thus, $\psi(p_3), \psi(p_2) \neq 0$. As a result c_1 and d_2' are both non-zero. Consider the map $\psi' := (\frac{d_2'}{c_1}) \circ \psi$. The above two equations give us: $\psi'(p_3 + p_2) = \frac{d_2'^3}{c_1^2} \cdot (q_3 + q_2)$. Denote $\frac{d_2'^3}{c_1^2}$ by c. Thus,

$$\psi'(p(\overline{z}, \overline{b})) = c \cdot q(\overline{z}, \overline{b})$$

Now we can get rid of the extra factor of c by defining a map ψ'':

$$\begin{aligned} \forall i, j,\ \psi''(z_{i,j}) \;&:=\; \frac{1}{c}\psi'(z_{i,j}) \\ \forall i,\ \psi''(b_i) \;&:=\; \psi'(b_i) \end{aligned}$$

It follows that $\psi''(p) = \frac{1}{c}\psi'(p) = q$ and thus, $p(\overline{z}, \overline{b}),\ q(\overline{z}, \overline{b})$ are equivalent under the map ψ''. □

Thus, $R \cong S$ iff $R' \cong S'$ and hence it is sufficient to study $\mathbb{F}$-algebra isomorphism over local commutative $\mathbb{F}$-algebras of the form occurring in Equation (3.8). ∎

3.2.2 Local commutative $\mathbb{F}$-algebras reduce to Cubic Forms

Here, we show that local commutative $\mathbb{F}$-algebra isomorphism reduces to cubic forms equivalence. This result when combined with the last subsection shows that cubic forms equivalence is at least as hard as the commutative algebra isomorphism and graph isomorphism (from Theorem 2.2).

We construct cubic forms from the rings of Equation (3.8) and then heavily use the properties of the underlying local commutative $\mathbb{F}$-algebra to study the equivalences of these cubic forms. The reduction that we exhibit in the following theorem holds for *any* field $\mathbb{F}$.

Theorem 3.5 *Commutative $\mathbb{F}$-algebra isomorphism $\leq_m^P$ $\mathbb{F}$-cubic forms equivalence.*

Proof: Given commutative $\mathbb{F}$-algebras $R,\ S$ we will construct cubic forms $\phi_R,\ \phi_S$ such that the cubic forms are equivalent iff the algebras are isomorphic. The construction involves first getting the local $\mathbb{F}$-algebras $R',\ S'$ (as in Theorem 3.4) and then the cubic forms out of these local commutative algebras.

Let $b_1, \ldots, b_n$ be the additive basis of R over $\mathbb{F}$. Let the multiplication in the algebra be defined as:

$$\text{for all } i, j \in [n]:\ b_i \cdot b_j = \sum_{k=1}^{n} a_{i,j,k} b_k,\ \text{where } a_{i,j,k} \in \mathbb{F}$$

Consider the following local ring R' constructed from R:

$$R' \ := \ \mathbb{F}[\overline{z}, \overline{b}, u] / \left\langle p_3, up_2, u^2, \mathcal{I} \right\rangle \tag{3.14}$$

where, $p_3(\overline{z}, \overline{b}) := \sum_{1 \leq i \leq j \leq n} z_{i,j} b_i b_j$ and $p_2(\overline{z}, \overline{b}) := \sum_{1 \leq i \leq j \leq n} z_{i,j} \left(\sum_{k=1}^{n} a_{i,j,k} b_k\right)$. $\mathcal{I}$ is the set of all possible products of 4 variables (with repetition) from $\{z_{1,1}, \ldots, z_{n,n}, b_1, \ldots, b_n, u\}$.

Similarly, construct S' from S and we know from Theorem 3.4 that $R \cong S$ iff $R' \cong S'$. Now we move on to constructing cubic forms from these local commutative algebras R' and S'.

A natural set of generators of the ring R' is: $\{1\} \cup \{z_{i,j}\}_{1 \le i \le j \le n} \cup \{b_i\}_{1 \le i \le n} \cup \{u\}$. For simplicity let us call them $1, x_1, \ldots, x_g, u$ respectively, where $g := \binom{n+1}{2} + n$. A natural additive basis of R' over $\mathbb{F}$ is:

$$\begin{aligned} &\{1\} \cup \{x_i\}_{1 \le i \le g} \cup \{u\} \cup \{x_i x_j\}_{1 \le i \le j \le g} \cup \{u x_i\}_{1 \le i \le g} \cup \{x_i x_j x_k\}_{1 \le i \le j \le k \le g} \\ &\quad \cup \{u x_i x_j\}_{1 \le i \le j \le g} \text{ minus one term each from } p_3 \text{ and } u p_2 \end{aligned} \tag{3.15}$$

For simplicity denote the elements of this additive basis by $1, c_1, \ldots, c_d$ respectively, where,

$$d := g + 1 + \binom{g+1}{2} + g + \binom{g+2}{3} + \binom{g+1}{2} - 2 = 2g + 2\binom{g+1}{2} + \binom{g+2}{3} - 1$$

Finally, we construct a cubic form ϕ_R using R' as follows:

$$\phi_R(\overline{y}, \overline{c}, v) := \sum_{1 \le i \le j \le d} y_{i,j} c_i c_j - v \sum_{1 \le i \le j \le d} y_{i,j} \left(\sum_{k=1}^{d} \tilde{a}_{i,j,k} c_k \right) \tag{3.16}$$

where $\forall i, j,\ c_i \cdot c_j = \sum_{k=1}^{d} \tilde{a}_{i,j,k} c_k$ in R', for some $\tilde{a}_{i,j,k} \in \mathbb{F}$.

Observe that the v terms in this cubic form are "few" because most of the $\tilde{a}$ are zero. This property is useful in analysing the equivalence of such forms. Let us first bound the number of v terms in ϕ_R.

Claim 3.5.1 *The number of nonzero v terms in RHS of Equation (3.16) is less than $(3d - 6)$.*

Proof of Claim 3.5.1. The number of nonzero v terms in RHS of Equation (3.16) is:

$$\le \ \#\{(k, \ell) \mid 1 \le k \le \ell \le d,\ c_k c_\ell \neq 0 \text{ in } R'\} + 3\,[\#(\text{terms in } p_3) + \#(\text{terms in } p_2)]$$

The first expression above accounts for all the relations in R' of the form $c_k c_\ell = c_m$. The second expression takes care of the relations that arise from $p_3 = 0$ and $u p_2 = 0$.

The factor of 3 above occurs because a term $x_i x_j x_k$ in p_3, up_2 can create v terms in at most 3 ways: from $(x_i) \cdot (x_j x_k)$ or $(x_j) \cdot (x_i x_k)$ or $(x_k) \cdot (x_i x_j)$.

$$\begin{aligned}
\leq\ & \#\left\{(k,\ell) \mid k \leq \ell,\ c_k, c_\ell \in \{x_i\}_{1\leq i\leq g}\right\} + \#\left\{(k,\ell) \mid\ c_k \in \{x_i\}_{1\leq i\leq g}, c_\ell = u\right\} \\
& +\#\left\{(k,\ell) \mid\ c_k \in \{x_i\}_{1\leq i\leq g}, c_\ell \in \{x_i x_j\}_{1\leq i\leq j\leq g}\right\} \\
& +\#\left\{(k,\ell) \mid\ c_k \in \{x_i\}_{1\leq i\leq g}, c_\ell \in \{u x_i\}_{1\leq i\leq g}\right\} \\
& +\#\left\{(k,\ell) \mid\ c_k = u, c_\ell \in \{x_i x_j\}_{1\leq i\leq j\leq g}\right\} + 3\left[\#(\text{terms in } p_3) + \#(\text{terms in } p_2)\right] \\
\leq\ & \left[\binom{g+1}{2} + g + g\cdot\binom{g+1}{2} + g^2 + \binom{g+1}{2}\right] + 3\left[\binom{n+1}{2} + \binom{n+1}{2}\cdot n\right]
\end{aligned}$$

Note that the dominant term in the above expression is $\frac{g^3}{2}$ while in that of d it is $\frac{g^3}{6}$. Thus, the above expression should be around $3d$. Exact computation gives the following bound:

$$< \ (3d-6)$$

□

Construct a cubic form ϕ_S from ring S in a way similar to that of Equation (3.16).

$$\phi_S(\overline{y}, \overline{c}, v) := \sum_{1\leq i\leq j\leq d} y_{i,j} c_i c_j - v \sum_{1\leq i\leq j\leq d} y_{i,j} \left(\sum_{k=1}^{d} \tilde{e}_{i,j,k} c_k\right) \tag{3.17}$$

where $\forall i, j,\ c_i \cdot c_j = \sum_{k=1}^{d} \tilde{e}_{i,j,k} c_k$ in S' for some $\tilde{e}_{i,j,k} \in \mathbb{F}$.

The following claim is what we intend to prove now.

Claim 3.5.2 *$\phi_R(\overline{y}, \overline{c}, v)$ is equivalent to $\phi_S(\overline{y}, \overline{c}, v)$ iff $R' \cong S'$ iff $R \cong S$.*

Proof of Claim 3.5.2. The part of this claim that needs to be proved is $\phi_R \sim \phi_S \Rightarrow R' \cong S'$. Suppose ψ is an equivalence from $\phi_R(\overline{y}, \overline{c}, v)$ to $\phi_S(\overline{y}, \overline{c}, v)$. We will show how to extract from ψ an isomorphism from R' to S'.

We have the following starting equation to analyze:

$$\sum_{1\leq i\leq j\leq d} \psi(y_{i,j})\psi(c_i)\psi(c_j) - \psi(v) \sum_{1\leq i\leq j\leq d} \psi(y_{i,j}) \left(\sum_{k=1}^{d} \tilde{a}_{i,j,k} \psi(c_k)\right)$$

$$= \sum_{1\le i\le j\le d} y_{i,j}c_ic_j - v\sum_{1\le i\le j\le d} y_{i,j}\left(\sum_{k=1}^{d}\tilde{e}_{i,j,k}c_k\right) \tag{3.18}$$

The main property of this huge equation that we would like to show is: $\psi(c_i)$ *consists of only* $\bar{c}$ *terms.* Thus, $\psi(c_i)$ has enough information to extract a ring isomorphism from R' to S'. In the rest of the proof we will "rule out" the unpleasant cases of $\psi(c_i)$ having $\overline{y}, v$ terms and $\psi(v)$ having $\overline{y}$ terms.

Let for every $i \in [d]$, $\psi(c_i) = \sum_j \alpha_{i,j}c_j + \sum_{j,k}\beta_{i,j,k}y_{j,k} + \gamma_i v$ where α, β, γ's $\in \mathbb{F}$. For obvious reasons we will call the expression $\sum_{j,k}\beta_{i,j,k}y_{j,k}$ as the $\overline{y}$ *part of* $\psi(c_i)$. $\overline{y}$ parts of $\psi(v)$ and $\psi(y_{i,j})$ are defined similarly. We will show that the rank of the $\overline{y}$ part of $\psi(c_1), \ldots, \psi(c_d), \psi(v)$ is less than 3.

Assume that for some i, j, k the $\overline{y}$ parts of $\psi(c_i), \psi(c_j), \psi(c_k)$ are linearly independent over $\mathbb{F}$. By a *term* on LHS of Equation (3.18) we mean expressions of the form $\psi(y_{\ell,s})\psi(c_\ell)\psi(c_s)$ or $\psi(v)\psi(y_{\ell,s})\psi(c_t)$, where $\ell, s, t \in [d]$. Let T_0 be the set of all terms on LHS of Equation (3.18). There are at least $d + (d-1) + (d-2) = (3d-3)$ terms on LHS of Equation (3.18) that have an occurrence of $\psi(c_i), \psi(c_j)$ or $\psi(c_k)$, denote this set of terms by T_1 and the set of the remaining terms by T_2. Let us build a maximal set Y of linearly independent $\overline{y}$ parts and a set T of corresponding terms as follows:
Start with keeping $\overline{y}$ parts of $\psi(c_i), \psi(c_j), \psi(c_k)$ in Y and setting $T = T_1$. Successively add a new $\overline{y}$ part to Y that is linearly independent from the elements already in Y and that occurs in a term $t \in T_0 \setminus T$, also, add t to T. When Y has grown to its maximal size, it is easy to see that:

$$\begin{aligned}
\#Y &\le 3 + \#T_2 \quad [\because \text{ initially, } \#Y = 3 \text{ and there are } \#T_2 \text{ terms outside } T] \\
&= 3 + \left[\binom{d+1}{2} + \#(\text{terms having } \psi(v)) - \#T_1\right] \\
&< 3 + \left[\binom{d+1}{2} + (3d-6) - (3d-3)\right] \quad [\text{by Claim 3.5.1 and } \because \#T_1 \ge (3d-3)] \\
&= \binom{d+1}{2} \\
&= \#\left\{y_{i,j}\right\}_{1\le i\le j\le d}
\end{aligned}$$

Now apply an invertible linear transformation τ on the $\overline{y}$ variables in Equation (3.18) such that all the $\overline{y}$ parts in Y are mapped to distinct *single* $\overline{y}$ variables, let $\tau(Y)$ denote the set of these variables. By substituting suitable linear forms, having only $\overline{c}, v$'s, to variables in $\tau(Y)$ we can make all the terms in $\tau(T)$ zero and the rest of the terms, *i.e.* $\tau(T_0 \setminus T)$, will then have no occurrence of $\overline{y}$ variables (as Y is the *maximal* set of linearly independent $\overline{y}$ parts). Thus, LHS of Equation (3.18), after applying τ and the substitutions, is completely in terms of $\overline{c}, v$ while RHS still has at least one free $\overline{y}$ variable (as we fixed only $\#\tau(Y) < \#\left\{y_{i,j}\right\}_{1\leq i\leq j\leq d}$ $\overline{y}$ variables and as τ is an invertible linear transformation). This contradiction shows that the $\overline{y}$ part of $\psi(c_i), \psi(c_j), \psi(c_k)$ cannot be linearly independent, for any i, j, k. Using a similar argument it can be shown that the $\overline{y}$ part of $\psi(c_i), \psi(c_j), \psi(v)$ cannot be linearly independent, for any i, j. Thus, the rank of the $\overline{y}$ part of $\psi(c_1), \ldots, \psi(c_d), \psi(v)$ is ≤ 2. For concreteness let us assume that the rank is *exactly* 2, the proof we give below will easily go through even when the rank is 1.

Again let Y be a maximal set of linearly independent $\overline{y}$ parts occurring in $\{\psi(y_{i,j})\}_{1\leq i\leq j\leq d}$ with the extra condition that $\overline{y}$ parts in Y are also linearly independent from those occurring in $\psi(c_1), \ldots, \psi(c_d), \psi(v)$. As we have assumed the rank of the $\overline{y}$ part of $\psi(c_1), \ldots, \psi(c_d), \psi(v)$ to be 2 we get $\#Y = \binom{d+1}{2} - 2$. Let $(i_1, j_1), (i_2, j_2)$ be the two tuples such that the $\overline{y}$ parts of $\psi(y_{i_1,j_1}), \psi(y_{i_2,j_2})$ do not appear in Y. To make things easier to handle let us apply an invertible linear transformation τ_1 on the variables in Equation (3.18) such that:

- the $\overline{y}$ parts of $\tau_1 \circ \psi(c_1), \ldots, \tau_1 \circ \psi(c_d), \tau_1 \circ \psi(v)$ are all linear combinations of only y_{i_1,j_1} and y_{i_2,j_2}.
- for all (i, j) other than (i_1, j_1) and (i_2, j_2), the $\overline{y}$ part of $\tau_1 \circ \psi(y_{i,j})$ is equal to $y_{i,j}$.
- τ_1 is identity on $\overline{c}, v$.

For clarity let $\psi' := \tau_1 \circ \psi$. Rest of our arguments will be based on comparing the coefficients of $y_{i,j}$, for $(i, j) \neq (i_1, j_1), (i_2, j_2)$, on both sides of the equation:

$$\sum_{1\leq i\leq j\leq d} \psi'(y_{i,j}) \left(\psi'(c_i c_j) - \psi'(v) \sum_{k=1}^{d} \tilde{a}_{i,j,k} \psi'(c_k) \right)$$

$$= \sum_{1 \le i \le j \le d} y_{i,j}(\text{quadratic terms in } \overline{c}, v) \tag{3.19}$$

For any c_i, choose distinct basis elements c_j, c_k and c_ℓ satisfying $c_i c_j = c_i c_k = c_i c_\ell = 0$ in R' (note that there is an ample supply of such j, k, ℓ), such that by comparing coefficients of $y_{i,j}, y_{i,k}, y_{i,\ell}$ (assumed to be other than y_{i_1,j_1}, y_{i_2,j_2}) on both sides of Equation (3.19) we get:

$$\begin{aligned}
\psi'(c_i c_j) + (e_{i,j,1}E_1 + e_{i,j,2}E_2) &= (\text{quadratic terms in } \overline{c}, v) \\
\psi'(c_i c_k) + (e_{i,k,1}E_1 + e_{i,k,2}E_2) &= (\text{quadratic terms in } \overline{c}, v) \\
\psi'(c_i c_\ell) + (e_{i,\ell,1}E_1 + e_{i,\ell,2}E_2) &= (\text{quadratic terms in } \overline{c}, v)
\end{aligned} \tag{3.20}$$

where, $e_{i,j,1}, e_{i,j,2}, e_{i,k,1}, e_{i,k,2}, e_{i,\ell,1}, e_{i,\ell,2} \in \mathbb{F}$ and

$$E_1 = \psi'(c_{i_1} c_{j_1}) - \psi'(v) \sum_{k=1}^{d} \tilde{a}_{i_1,j_1,k} \psi'(c_k)$$

$$E_2 = \psi'(c_{i_2} c_{j_2}) - \psi'(v) \sum_{k=1}^{d} \tilde{a}_{i_2,j_2,k} \psi'(c_k)$$

Now there exist $\lambda_1, \lambda_2, \lambda_3 \in \mathbb{F}$ (not all zero) such that Equations (3.20) can be combined to get rid of E_1, E_2 and get:

$$\psi'(c_i)\left(\lambda_1 \psi'(c_j) + \lambda_2 \psi'(c_k) + \lambda_3 \psi'(c_\ell)\right) = (\text{quadratic terms in } \overline{c}, v)$$

This equation combined with the observation that both $\psi'(c_i)$ and $(\lambda_1 \psi'(c_j) + \lambda_2 \psi'(c_k) + \lambda_3 \psi'(c_\ell))$ are non-zero (as ψ' is invertible) implies that:

$$\forall i, \quad \psi'(c_i) = (\text{linear terms in } \overline{c}, v) \tag{3.21}$$

This means that the $\overline{y}$-variables are only in $\psi'(y_{i,j})$'s and possibly $\psi'(v)$. Again apply an invertible linear transformation τ_2 on the $\overline{y}$-variables in Equation (3.19) such that $\tau_2 \circ \psi'(v)$ has only y_{i_0,j_0} in the $\overline{y}$ part and the $\overline{y}$ part of $\tau_2 \circ \psi'(y_{i,j})$ is equal to $y_{i,j}$ for all (i,j) except possibly (i_0, j_0). For clarity let $\psi'' := \tau_2 \circ \psi'$. Our equation now is:

$$\sum_{1 \le i \le j \le d} \psi''(y_{i,j}) \left(\psi''(c_i c_j) - \psi''(v) \sum_{k=1}^{d} \tilde{a}_{i,j,k} \psi''(c_k) \right)$$

$$= \sum_{1 \le i \le j \le d} y_{i,j}(\text{quadratic terms in } \overline{c}, v) \tag{3.22}$$

By comparing coefficients of $y_{i,j}$ (other that y_{i_0,j_0}) on both sides of the above equation we get:

$$\left(\psi''(c_i c_j) - \psi''(v)\sum_{k=1}^{d} \tilde{a}_{i,j,k}\psi''(c_k)\right) + e \cdot \left(\psi''(c_{i_0} c_{j_0}) - \psi''(v)\sum_{k=1}^{d} \tilde{a}_{i_0,j_0,k}\psi''(c_k)\right)$$
$$= (\text{quadratic terms in } \overline{c}, v), \quad \text{for some } e \in \mathbb{F}.$$

Pick i, j such that $\sum_{k=1}^{d} \tilde{a}_{i,j,k} c_k \neq 0$ in R'. Now if $\psi''(v)$ has a nonzero y_{i_0,j_0} term then by comparing coefficients of y_{i_0,j_0} on both sides of the above equation we deduce:

$$\sum_{k=1}^{d} \tilde{a}_{i,j,k}\psi''(c_k) + e \cdot \sum_{k=1}^{d} \tilde{a}_{i_0,j_0,k}\psi''(c_k) = 0 \tag{3.23}$$

But again we can pick i, j suitably so that $\left(\sum_{k=1}^{d} \tilde{a}_{i,j,k} c_k\right) \notin \left\{0,\ -e \cdot \sum_{k=1}^{d} \tilde{a}_{i_0,j_0,k} c_k\right\}$ and hence avoiding Equation (3.23) to hold. Thus, proving that $\psi''(v)$ has no y_{i_0,j_0} term. So we now have:

$$\psi''(v) = (\text{linear terms in } \overline{c}, v)$$

and

$$\forall i, \quad \psi''(c_i) = (\text{linear terms in } \overline{c}, v) \tag{3.24}$$

Since, $\overline{y}$-variables are present only in $\psi''(y_{i,j})$'s, comparing coefficients of $y_{i,j}$'s on both sides of Equation (3.22) gives:

$$\forall i, j, \ \ \psi''(c_i c_j) - \psi''(v)\sum_{k=1}^{d} \tilde{a}_{i,j,k}\psi''(c_k) = (\text{quadratic terms in } \overline{c}) - v(\text{linear terms in } \overline{c}) \tag{3.25}$$

Using this equation we will prove now that $\psi''(c_i)$ has only $\overline{c}$-variables.

Consider a c_i such that $c_i^2 = 0$ in R', then from Equation (3.25):

$$\psi''(c_i)^2 = (\text{quadratic terms in } \overline{c}) - v(\text{linear terms in } \overline{c}) \tag{3.26}$$

Now if $\psi''(c_i)$ has a nonzero v term then there will be a v^2 term above on LHS which is absurd. Thus, $\psi''(c_i)$ has only $\overline{c}$-variables when $c_i^2 = 0$ in R'. When $c_i^2 \neq 0$

then $c_i^2 = \sum_{k=1}^{d} \tilde{a}_{i,i,k} c_k$ in R' where the c_k's with nonzero $\tilde{a}_{i,i,k}$ satisfy $c_k^2 = 0$. This happens because the way $\overline{c}$'s are defined in Equation (3.15) the expression of c_i^2 will have only quadratic or cubic terms in $\overline{x}$ and the square of these terms would clearly be zero in R'. Thus, again if $\psi''(c_i)$ has a v term then there will be an uncancelled v^2 term on LHS of the equation:

$$\psi''(c_i)^2 - \psi''(v) \sum_{k=1}^{d} \tilde{a}_{i,i,k} \psi''(c_k) = (\text{quadratic terms in } \overline{c}) - v(\text{linear terms in } \overline{c})$$

Thus, we know at this point that $\psi''(v)$ has only $\overline{c}, v$ terms and $\psi''(c_i)$ has only $\overline{c}$ terms. Since, τ_1, τ_2 act only on $\overline{y}$'s we have what we intended to prove in the beginning (recall Equation (3.18)):

$$\psi(v) = (\text{linear terms in } \overline{c}, v)$$

and

$$\forall i, \quad \psi(c_i) = (\text{linear terms in } \overline{c}) \tag{3.27}$$

We have now almost extracted a ring isomorphism from the cubic form equivalence ψ, just few technicalities are left which we resolve next.

Apply an invertible linear transformation τ_3 on the $\overline{y}$-variables in Equation (3.18) such that the $\overline{y}$ part of $\tau_3 \circ \psi(y_{i,j})$ is equal to $y_{i,j}$ for all $i \leq j \in [d]$. Of course, we assume that τ_3 is identity on the $\overline{c}, v$ variables. So, on comparing coefficients of $y_{i,j}$ on both sides of the Equation (3.18) after applying τ_3 we get:

$$\forall i, j, \ \ \tau_3 \circ \psi(c_i c_j) - \tau_3 \circ \psi(v) \sum_{k=1}^{d} \tilde{a}_{i,j,k} \tau_3 \circ \psi(c_k) = \sum_{i \leq j} \lambda_{i,j} \left(c_i c_j - v \sum_{k=1}^{d} \tilde{e}_{i,j,k} c_k \right) \tag{3.28}$$

for some $\lambda_{i,j} \in \mathbb{F}$.

Substitute $v = 1$ in the expression for $\tau_3 \circ \psi(v) = \gamma_{v,v} v + \sum_i \alpha_{v,i} c_i$ and denote the result by m. Observe that $\gamma_{v,v} \neq 0$ and $\forall i, \ c_i$ is a nilpotent element in S' and hence m is a *unit* in the ring S'. On substituting $v = 1$ in Equation (3.28) we get:

$$\forall i, j, \quad \tau_3 \circ \psi(c_i) \cdot \tau_3 \circ \psi(c_j) - m \cdot \sum_{k=1}^{d} \tilde{a}_{i,j,k} \tau_3 \circ \psi(c_k) = 0 \ \text{ in } S'$$

If we define $\Psi := \frac{\tau_3 \circ \psi}{m}$ then we get:

$$\forall i, j, \quad \Psi(c_i)\Psi(c_j) - \sum_{k=1}^{d} \tilde{a}_{i,j,k}\Psi(c_k) = 0 \ \text{ in } S' \tag{3.29}$$

Now observe that if for some λ_i's $\in \mathbb{F}$, $\Psi(\sum_{i=1}^{d} \lambda_i c_i) = 0$ in S' then $\tau_3 \circ \psi(\sum_{i=1}^{d} \lambda_i c_i) = 0$ in S'. Since $\tau_3 \circ \psi$ is an invertible linear map from R' to equi-dimensional S' this means that $\sum_{i=1}^{d} \lambda_i c_i = 0$ in R'. Therefore, Ψ is a *bijection* from R' to S'. Together with Equation (3.29) this tells us that Ψ is an isomorphism from R' to S'. $\square$

This completes the reduction from commutative $\mathbb{F}$-algebra isomorphism to cubic form equivalence. ∎

3.3 Equivalence of Forms: Known results

The last two sections indicate that the problem of cubic forms equivalence is quite an interesting special case of polynomial equivalence. Not much is known about the structure of cubic forms. On the other hand, structure of quadratic forms is well understood. We collect in this section the main ideas that have been around to understand forms equivalence. The notions of regularity and decomposability of cubic forms given here will be used to study our cubic forms (that appeared in Equation (3.16)) in the next section.

3.3.1 Quadratic Forms Equivalence

In this subsection we sketch the classification theorem known for quadratic forms. As a byproduct we also get algorithms for solving quadratic forms equivalence over finite fields, $\mathbb{Q}, \mathbb{R}$ and $\mathbb{C}$. The detailed proofs can be found in [Serre], we present the main ideas here simply for their beauty.

Here we will assume that char $\mathbb{F} \neq 2$. Let $f \in \mathbb{F}[x_1, \ldots, x_n]$ be a quadratic form and let V be the vector space $\mathbb{F}^n$. Observe that the map $\Theta : V \times V \to \mathbb{F}$ defined as $\Theta(u, v) = \frac{f(u+v)-f(u)-f(v)}{2}$ is symmetric and *bilinear*, i.e., $\Theta(u, v) = \Theta(v, u)$ and $\Theta(u + u', v) = \Theta(u, v) + \Theta(u', v)$. Also, f is recoverable from Θ as $f(u) =$

$\Theta(u, u)$. Thus, there is a $1-1$ correspondence from quadratic forms to symmetric bilinear maps on the underlying vector space and this connection is quite fruitful in classifying quadratic forms.

The Algorithm

Suppose we are given two nonzero quadratic forms $f, g \in \mathbb{F}[x_1, \ldots, x_n]$. We will show how to check $f \sim g$ over $\mathbb{F}$.

Step 0:(Base case) If $f = a_i x_i^2$ and $g = b_j x_j^2$ then $f \sim g$ iff $\frac{a_i}{b_j}$ is a square in $\mathbb{F}$.

Step 1:(Diagonalization) Let us express f as a matrix product:

$$\begin{aligned} f(x_1, \ldots, x_n) &= \sum_{i=1}^{n} a_{i,i} x_i^2 + \sum_{1 \le i < j \le n} 2a_{i,j} x_i x_j \\ &= (x_1 \ \ldots \ x_n)\, A\, (x_1 \ \ldots \ x_n)^{\mathrm{T}} \end{aligned}$$

where, A is a symmetric matrix with $a_{i,j}$ as the $(i,j)^{\text{th}}$ and $(j,i)^{\text{th}}$ entries. Since A is a symmetric matrix over a field we can apply Gaussian elimination to get an invertible matrix C such that CAC^{T} is diagonal, say $\mathrm{diag}[b_1 \ \ldots \ b_n]$. Then we have,

$$\begin{aligned} f\left((x_1 \ \ldots \ x_n)C\right) &= (x_1 \ \ldots \ x_n) CAC^{\mathrm{T}} (x_1 \ \ldots \ x_n)^{\mathrm{T}} \\ &= \sum_{i=1}^{n} b_i x_i^2 \end{aligned}$$

Thus, from now on we can assume that the input quadratic forms f, g are given as sums of squares. Note that in this step we needed $\mathrm{char}\ \mathbb{F} \neq 2$.

Step 2:(Root-finding) Let $f = \sum_{i=1}^{n} a_i x_i^2$ and $g = \sum_{i=1}^{n} b_i x_i^2$, where a_i, b_i's are nonzero in $\mathbb{F}$. Find a root $(\alpha_1, \ldots, \alpha_n) \in \mathbb{F}^n$ of the diagonal quadratic equation:

$$\sum_{i=1}^{n} a_i x_i^2 = b_n \tag{3.30}$$

Step 3:(Witt's decomposition) Let Θ be the symmetric bilinear map corresponding to f. Using simple linear algebra compute the subspace:

$$U := \left\{ u \in \mathbb{F}^n \mid \Theta\left((\alpha_1 \cdots \alpha_n)^T, u\right) = 0 \right\}$$

Now Witt's theorem states that subspace U and the "orthogonal" vector $(\alpha_1 \cdots \alpha_n)^T$ span the full space V:

$$V = \mathbb{F} \begin{pmatrix} \alpha_1 \\ \vdots \\ \alpha_n \end{pmatrix} \oplus U$$

This means that any $v \in V$ can be written as $\lambda(\alpha_1 \cdots \alpha_n)^T + u$, where $\lambda \in \mathbb{F}$ and $u \in U$. Thus,

$$\begin{aligned} f(v) &= \Theta(v, v) \\ &= \Theta\left(\lambda(\alpha_1 \cdots \alpha_n)^T + u, \lambda(\alpha_1 \cdots \alpha_n)^T + u\right) \\ &= \lambda^2 \Theta\left((\alpha_1 \cdots \alpha_n)^T, (\alpha_1 \cdots \alpha_n)^T\right) + \Theta(u, u) \\ &= \lambda^2 f\left((\alpha_1 \cdots \alpha_n)^T\right) + f(u) \\ &= \lambda^2 b_n + f(u) \end{aligned}$$

This simply means that $f \sim b_n x_n^2 + f_1(x_1, \ldots, x_{n-1})$ for some quadratic form $f_1 \in \mathbb{F}[x_1, \ldots, x_{n-1}]$.

Step 4:(Witt's cancellation) So, we now have $f(x_1, \ldots, x_n) \sim b_n x_n^2 + f_1(x_1, \ldots, x_{n-1})$ and $g(x_1, \ldots, x_n) = b_n x_n^2 + \sum_{i=1}^{n-1} b_i x_i^2$. Witt's cancellation lemma says that:

$$b_n x_n^2 + f_1(x_1, \ldots, x_{n-1}) \sim b_n x_n^2 + \sum_{i=1}^{n-1} b_i x_i^2$$

iff

$$f_1(x_1, \ldots, x_{n-1}) \sim \sum_{i=1}^{n-1} b_i x_i^2$$

So, now we can recursively do steps 0-3 on these smaller quadratic forms of rank $n-1$.

Observe that steps 0, 1 and 3 are 'easy' to do, so the only part that needs explanation is step 2 – solving diagonal quadratic equations.

Solving diagonal quadratic equations

Here we are interested in solving Equation (3.30) in step 2. We will show how to find roots when $\mathbb{F}$ is a finite field, $\mathbb{C}, \mathbb{R}$ and $\mathbb{Q}$.

Suppose $\mathbb{F}$ is a finite field, say $\mathbb{F}_q$. If $n = 1$ we need to solve $a_1 x_1^2 = b_n$ which is just finding square-roots. If $n \geq 2$ a classic theorem of Weil (see [Bac96]) states that for a random choice of $x_1, \ldots, x_{n-1} \in \mathbb{F}_q$ there exists an $x_n \in \mathbb{F}_q$ satisfying the Equation (3.30). Thus, in all the cases we can find roots of the Equation (3.30) over $\mathbb{F}_q$ in randomized polynomial time.

Suppose $\mathbb{F}$ is $\mathbb{R}$ or $\mathbb{C}$ then it is easily seen that roots of the Equation (3.30) can be found in deterministic polynomial time.

Suppose $\mathbb{F} = \mathbb{Q}$. If $n = 1$ then solving $a_1 x_1^2 = b_n$ is just finding square-roots over rationals. The first nontrivial case is $n = 2$ when we need to solve $a_1 x_1^2 + a_2 x_2^2 = b_n$. We can first pre-process the equation by clearing the denominators of a_1, a_2, b_n and then taking the square parts of the integer coefficients 'in' x_1, x_2 to get an equation: $ax^2 + by^2 = z^2$ where a, b are *square-free* integers and we want *coprime* $x, y, z \in \mathbb{Z}$. We now demonstrate an algorithm, due to Legendre, to solve this equation. We just need to define the *norm* of elements in the number field $\mathbb{Q}(\sqrt{a})$. Elements of $\mathbb{Q}(\sqrt{a})$ are of the form $(\alpha + \beta\sqrt{a})$ for some $\alpha, \beta \in \mathbb{Q}$ and we define the norm function $N : \mathbb{Q}(\sqrt{a}) \to \mathbb{Q}$ as: $N(\alpha + \beta\sqrt{a}) = \alpha^2 - a\beta^2$. Observe that it is a multiplicative function.

Wlog assume $|a| < |b|$. If $ax^2 + by^2 = z^2$ has a solution then for any prime $p|b$, p cannot divide x (otherwise $p|z \Rightarrow p^2|by^2 \Rightarrow p|y \Rightarrow x, y, z$ are not coprime). Thus, a is a square mod p. As a is a square mod p for every prime $p|b$ we get that a is a square mod b. Thus, there is a $t \in \mathbb{Z}$ such that $|t| \leq \frac{|b|}{2}$ and $a = t^2 (\text{mod } b)$. Let $b' \in \mathbb{Z}$ be such that:

$$t^2 = a + bb' \quad \text{over } \mathbb{Z} \tag{3.31}$$

Now we claim that $ax^2 + by^2 = z^2$ has a solution iff $ax^2 + b'y^2 = z^2$ has a solution. This happens because (say) if $ax^2 + by^2 = z^2$ has a solution then:

$$b = N\left(\frac{z + x\sqrt{a}}{y}\right)$$

Also, from Equation (3.31):

$$\begin{aligned} bb' &= N(t + \sqrt{a}) \\ \Rightarrow b' &= N\left(\frac{yt + y\sqrt{a}}{z + x\sqrt{a}}\right) \end{aligned}$$

Which on rationalizing the denominator effectively gives an integral solution of $ax^2 + b'y^2 = z^2$. Conversely, if $ax^2 + b'y^2 = z^2$ has a solution then $ax^2 + by^2 = z^2$ can be shown to have solutions in the exact same way as above.

Now notice that the equation $ax^2 + b'y^2 = z^2$ is a "smaller" equation, for:

$$\begin{aligned} |a| + |b'| &= |a| + \left|\frac{t^2 - a}{b}\right| \\ &\leq |a| + \left|\frac{t^2}{b}\right| + \left|\frac{a}{b}\right| \\ &< |a| + \frac{|b|}{4} + 1 \\ &< |a| + |b| \end{aligned}$$

Thus, the above procedure can be repeatedly applied till we reach the equation $\pm x^2 \pm y^2 = z^2$ or $\pm x^2 = z^2$ which are trivial to solve over integers.

The interesting thing to note in the above algorithm is that it constructively shows that the equation $ax^2 + by^2 + cz^2 = 0$ has a solution over $\mathbb{Q}$ iff it has a solution over $\mathbb{R}$ and mod p for all primes p. This property is famously known as the *local-global principle*.

Rational root-finding for diagonal quadratic equations when $n > 2$ uses the above algorithm and the 'tool' of local-global principle.

This completes the sketch of algorithms for quadratic forms equivalence and we collect the results in the following theorem.

Theorem 3.6 (Hasse, Witt et al) *1. Over finite fields, quadratic forms equivalence can be decided in P and found in ZPP.*

2. Over $\mathbb{R}$ and $\mathbb{C}$, quadratic forms equivalence can be decided and found in P.

3. Over $\mathbb{Q}$, quadratic forms equivalence can be done in EXP.

3.3.2 Cubic Forms Equivalence

Unlike quadratic forms the theory of cubic forms is still in its infancy. We collect here some known notions useful in "pre-processing" a given cubic form (see Harrison [Har75]).

Let $f(x_1, \ldots, x_n)$ be a cubic form over $\mathbb{F}$. In this section we will assume that characteristic of $\mathbb{F}$ is not 2 or 3. Let $V = \mathbb{F}^n$. We say that a map $\Theta : V \times V \times V \to \mathbb{F}$ is *symmetric* if for any permutation π on $\{1, 2, 3\}$ and any $v_1, v_2, v_3 \in V$, $\Theta(v_1, v_2, v_3) = \Theta(v_{\pi(1)}, v_{\pi(2)}, v_{\pi(3)})$. Θ is said to be 3-*linear* if it is linear in all the 3 arguments, where linear in the first argument means that: for all $u, u', v, w \in V$, $\Theta(u + u', v, w) = \Theta(u, v, w) + \Theta(u', v, w)$. Now the claim is that we can define a symmetric 3-linear map on V from any given cubic form $f(x_1, \ldots, x_n) = \sum_{1 \le i \le j \le k \le n} a_{i,j,k} x_i x_j x_k$. Let $\overline{x}_1 = \begin{pmatrix} x_{1,1} \\ \vdots \\ x_{n,1} \end{pmatrix}, \overline{x}_2, \overline{x}_3$ be vectors in $V = \mathbb{F}^n$. Define a map Θ from the cubic form f as:

$$\begin{aligned} \Theta\left(\overline{x}_1, \overline{x}_2, \overline{x}_3\right) &= \Theta\left(\begin{pmatrix} x_{1,1} \\ \vdots \\ x_{n,1} \end{pmatrix}, \begin{pmatrix} x_{1,2} \\ \vdots \\ x_{n,2} \end{pmatrix}, \begin{pmatrix} x_{1,3} \\ \vdots \\ x_{n,3} \end{pmatrix}\right) \\ &= \frac{1}{6} \sum_{\alpha} D_\alpha(f) \cdot x_{\alpha(1),1} x_{\alpha(2),2} x_{\alpha(3),3} \end{aligned}$$

where α ranges over all maps from $\{1, 2, 3\} \to \{1, 2, \ldots, n\}$ and the coefficient $D_\alpha(f)$ is given as:

$$D_\alpha(f) := \frac{\partial^3 f(x_1, \ldots, x_n)}{\partial x_{\alpha(1)} \partial x_{\alpha(2)} \partial x_{\alpha(3)}}$$

It is easily seen that this map Θ is symmetric 3-linear and moreover:

$$\Theta\left(\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}, \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}, \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}\right) = f(x_1, \ldots, x_n)$$

Thus, we have a $1-1$ correspondence between the cubic forms and the symmetric 3-linear maps on the underlying vector space (compare this with a similar observation

for quadratic forms in section 4.2).

Example Let $f(x, y) = x^3 + x^2y$ be a cubic form. Then the corresponding symmetric 3-linear map Θ on $V = \mathbb{F}^2$ is defined as:

$$\Theta\left(\begin{pmatrix}x_1\\y_1\end{pmatrix}, \begin{pmatrix}x_2\\y_2\end{pmatrix}, \begin{pmatrix}x_3\\y_3\end{pmatrix}\right) = x_1x_2x_3 + \frac{1}{3}x_1x_2y_3 + \frac{1}{3}x_1x_3y_2 + \frac{1}{3}x_2x_3y_1$$

and verify that:

$$\Theta\left(\begin{pmatrix}x\\y\end{pmatrix}, \begin{pmatrix}x\\y\end{pmatrix}, \begin{pmatrix}x\\y\end{pmatrix}\right) = f(x, y)$$

∎

Regularity

The first thing we would like to ensure about a given cubic form f is that there should not be "extra" variables in f, i.e., there is no invertible linear transformation τ such that $f(\tau x_1, \ldots, \tau x_n)$ has less than n variables. Such a cubic form is called *regular*.

Example The cubic form $f(x) = x^3$ is regular while $f(x, y) = (x + y)^3$ is not regular as the invertible map:

$$\tau : \begin{cases} x + y & \mapsto x \\ y & \mapsto y \end{cases}$$

reduces the number of variables of f. ∎

By *regularizing* a given cubic form f we mean finding an invertible linear transformation that applied on f makes it regular.

Proposition 3.1 (Harrison) *A given cubic form can be regularized in deterministic polynomial time.*

Proof: Suppose $f \in \mathbb{F}[x_1, \ldots, x_n]$ is a given cubic form and $\Theta(\cdot,\cdot,\cdot)$ is its corresponding symmetric 3-linear map on $V = \mathbb{F}^n$. Suppose $f(x_1, \ldots, x_n)$ is not regular and its regularized form is $f^{reg}(x_1, \ldots, x_m)$ in smaller number of variables $1 \leq m < n$. Further, let Θ^{reg} be the symmetric 3-linear map corresponding to f^{reg} and A be the invertible matrix in $\mathbb{F}^{n\times n}$ such that for all $\overline{x}_1, \overline{x}_2, \overline{x}_3 \in V$:

$$\Theta(A\overline{x}_1, A\overline{x}_2, A\overline{x}_3) = \Theta^{reg}\left(\begin{pmatrix} x_{1,1} \\ \vdots \\ x_{m,1} \end{pmatrix}, \begin{pmatrix} x_{1,2} \\ \vdots \\ x_{m,2} \end{pmatrix}, \begin{pmatrix} x_{1,3} \\ \vdots \\ x_{m,3} \end{pmatrix}\right)$$

Now observe that the RHS above is independent of the last coordinates, i.e. $x_{n,1}, x_{n,2}, x_{n,3}$.

Thus, if we fix $\overline{x}_1$ to be $\begin{pmatrix} 0 \\ \vdots \\ 0 \\ 1 \end{pmatrix}$ then for all $\overline{x}_2, \overline{x}_3 \in V$:

$$\Theta\left(A\begin{pmatrix} 0 \\ \vdots \\ 0 \\ 1 \end{pmatrix}, A\overline{x}_2, A\overline{x}_3\right) = \Theta^{reg}\left(0, \begin{pmatrix} x_{1,2} \\ \vdots \\ x_{m,2} \end{pmatrix}, \begin{pmatrix} x_{1,3} \\ \vdots \\ x_{m,3} \end{pmatrix}\right) = 0$$

As A is invertible $v := A\begin{pmatrix} 0 \\ \vdots \\ 0 \\ 1 \end{pmatrix} \neq 0$ and we have $\Theta(v,\cdot,\cdot) = 0$.

More interestingly, we will now see that the converse holds too, i.e., if there is a nonzero $v \in V$ such that $\Theta(v,\cdot,\cdot) = 0$ then f is not regular. Consider the following equation in the variables $x_{1,1}, x_{2,1}, \ldots, x_{n,1}$:

$$\text{for all } \overline{x}_2, \overline{x}_3 \in V,\ \Theta\left(\overline{x}_1, \overline{x}_2, \overline{x}_3\right) = 0 \tag{3.32}$$

If we compare the coefficient of $x_{i,2}x_{j,3}$ on both sides of the equation we get a linear equation and hence as i, j vary over all of $\{1, 2, \ldots, n\}$ we get a system of

homogeneous linear equations, say:

$$M \begin{pmatrix} x_{1,1} \\ \vdots \\ x_{n,1} \end{pmatrix} = 0$$

Now, if there is a nonzero $v \in V$ such that $\Theta(v, \cdot, \cdot) = 0$ then it means that $Mv = 0$ and hence, $\text{rank}(M) < n$. Now, by applying Gaussian elimination on M we get invertible matrices C, D such that the last $(n - \text{rank}(M))$ columns of $DMC =: M'$ are zero. Thus, the elements of the column vector $M(C\overline{x}_1) = (D^{-1}M')\overline{x}_1$ are independent of $x_{\text{rank}(M)+1,1}, \ldots, x_{n,1}$. In other words, $\Theta\left(C\overline{x}_1, \overline{x}_2, \overline{x}_3\right)$ is independent of the last $(n - \text{rank}(M))$ coordinates of $\overline{x}_1$. Now since Θ is symmetric 3-linear and C is an invertible linear transformation, the system of equations in the variables $\overline{x}_2$ that we get from the following equality:

$$\text{for all } \overline{x}_1, \overline{x}_3 \in V,\ \Theta\left(C\overline{x}_1, \overline{x}_2, \overline{x}_3\right) = 0$$

is equivalent to the system: $M\overline{x}_2 = 0$. Thus, as before, $M(C\overline{x}_2)$ is independent of the last $(n - \text{rank}(M))$ coordinates of $\overline{x}_2$ implying that $\Theta\left(C\overline{x}_1, C\overline{x}_2, \overline{x}_3\right)$ is independent of the last $(n - \text{rank}(M))$ coordinates of $\overline{x}_1$ and that of $\overline{x}_2$. Repeating this same argument again, we deduce: $\Theta\left(C\overline{x}_1, C\overline{x}_2, C\overline{x}_3\right)$ is independent of the last $(n - \text{rank}(M))$ coordinates of $\overline{x}_1, \overline{x}_2, \overline{x}_3$.

Thus, $f\left(C\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}\right) = \Theta\left(C\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}, C\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}, C\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}\right)$ is independent of $x_{\text{rank}(M)+1}, \ldots, x_n$ and regular over the variables $x_1, \ldots, x_{\text{rank}(M)}$.

Note that all the steps in the above discussion require simple linear algebra and hence can be executed in deterministic polynomial time. ∎

∎ *Decomposability*

Cubic forms do not satisfy the nice property of diagonalization unlike quadratic forms, for example: $x^3 + x^2y$ cannot be written as a sum of cubes. But there is a notion of decomposability of cubic forms into simpler cubic forms. We call a cubic

form $f(x_1, \ldots, x_n)$ *decomposable* if there is an invertible linear transformation τ, an $i \in [n]$ and cubic forms g, h such that:

$$f(\tau x_1, \ldots, \tau x_n) = g(x_1, \ldots, x_i) + h(x_{i+1}, \ldots, x_n)$$

This is also denoted by: $f \sim g \oplus h$.

Example The cubic form $f_1(x, y) = x^3 + y^3$ is decomposable while the cubic form $f_2(x, y) = x^3 + xy^2$ is indecomposable. ∎

It is interesting that given a cubic form f the decomposition of f can be found algorithmically. To show this we need the notion of centre of a cubic form that captures the symmetries of the underlying 3-linear map.

Definition 3.1 *Let f be a cubic form and Θ be the corresponding symmetric 3-linear map on the space V. The* center, *Cent(f), of the cubic form f is defined as:*

$$\left\{M \in \mathbb{F}^{n \times n} \mid \textit{for all } v_1, v_2, v_3 \in V,\ \Theta(Mv_1, v_2, v_3) = \Theta(v_1, Mv_2, v_3)\right\}$$

Example Let $f(x)$ be the cubic form x^3 then $\mathrm{Cent}(f) = \mathbb{F}$. If $f(x, y) = x^3 + y^3$ then $\mathrm{Cent}(f) \cong \mathrm{Cent}(x^3) \times \mathrm{Cent}(y^3) \cong \mathbb{F} \times \mathbb{F}$. ∎

The following properties of the center were first proved by Harrison [Har75]:

Lemma 3.1 *Suppose $f(x_1, \ldots, x_n)$ is a regular cubic form and Θ is the corresponding symmetric 3-linear map on $V = \mathbb{F}^n$.*

(1) Cent(f) is a commutative $\mathbb{F}$-algebra.

(2) f is indecomposable if and only if Cent(f) is indecomposable.

Proof: [(1)] Suppose $M_1, M_2 \in \mathrm{Cent}(f)$ then $M_1 + M_2$ is also in the centre and it is routine to show that $(\mathrm{Cent}(f), +)$ is an abelian group.

To see that $M_1 \cdot M_2 \in \mathrm{Cent}(f)$ observe that for any $u, v, w \in V$:

$$\begin{aligned}
\Theta(M_1 \cdot M_2 u, v, w) &= \Theta(M_2 u, v, M_1 w) \quad [\because M_1 \in \mathrm{Cent}(f)] \\
&= \Theta(u, M_2 v, M_1 w) \quad [\because M_2 \in \mathrm{Cent}(f)] \\
&= \Theta(u, M_1 \cdot M_2 v, w) \quad [\because M_1 \in \mathrm{Cent}(f)]
\end{aligned}$$

Thus, by definition $M_1 \cdot M_2$ is in $\mathrm{Cent}(f)$. Multiplication in $\mathrm{Cent}(f)$ is associative simply because it is matrix multiplication. To see commutativity observe that:

$$\begin{aligned}\Theta(M_1 \cdot M_2 u, v, w) &= \Theta(M_2 u, v, M_1 w) \quad [\because M_1 \in \mathrm{Cent}(f)] \\ &= \Theta(u, M_2 v, M_1 w) \quad [\because M_2 \in \mathrm{Cent}(f)] \\ &= \Theta(M_1 u, M_2 v, w) \quad [\because M_1 \in \mathrm{Cent}(f)] \\ &= \Theta(M_2 \cdot M_1 u, v, w) \quad [\because M_2 \in \mathrm{Cent}(f)]\end{aligned}$$

Thus, $\Theta\left((M_1 \cdot M_2 - M_2 \cdot M_1)u, \cdot, \cdot\right) = 0$. As f is regular this means that $(M_1 \cdot M_2 - M_2 \cdot M_1)u = 0$ (refer the proof of the Proposition 3.1). Since, this happens for all $u \in V$ we have that $(M_1 \cdot M_2 - M_2 \cdot M_1) = 0$ implying that $M_1 \cdot M_2 = M_2 \cdot M_1$.

Also, $\mathbb{F}$ is clearly contained in $\mathrm{Cent}(f)$. Thus, $\mathrm{Cent}(f)$ is a commutative $\mathbb{F}$-algebra. ∎

Proof: [(2)] Here, we need a property of local commutative rings proved in the appendix: a finite dimensional commutative algebra R is decomposable iff there is a nontrivial idempotent element, i.e., there is a $r \in R \setminus \{0, 1\},\ r^2 = r$.

If the cubic form f decomposes as $f_1 \oplus f_2$ then it is easy to show that $\mathrm{Cent}(f)$ decomposes as $\mathrm{Cent}(f_1) \times \mathrm{Cent}(f_2)$.

Conversely, suppose $\mathrm{Cent}(f)$ is decomposable. Then there is a matrix $M \in \mathrm{Cent}(f)$ such that $M^2 = M$ but $M \neq 0, I$. Now we want to decompose f using M.

Firstly, observe that if there is a $v \in MV \cap (I - M)V$ then $Mv = (I - M)v = 0$ and by adding the two we get $v = 0$. Next, observe that for any $u, v, w \in V$:

$$\begin{aligned}\Theta(Mu, (I - M)v, w) &= \Theta(u, M(I - M)v, w) \quad [\because M \in \mathrm{Cent}(f)] \\ &= 0 \quad [\because M^2 = M]\end{aligned}$$

Thus, for any $v_1 \in MV, v_2 \in (I - M)V,\ \Theta(v_1, v_2, \cdot) = 0$ or in other words: $MV, (I - M)V$ are *orthogonal* subspaces of V with respect to Θ. This means that for any $v \in V$ if we express v as $v = v_1 + v_2$, where $v_1 \in MV, v_2 \in (I - M)V$, then:

$$\begin{aligned}f(v) &= \Theta(v, v, v) \\ &= \Theta(v_1 + v_2, v_1 + v_2, v_1 + v_2) \\ &= \Theta(v_1, v_1, v_1) + \Theta(v_2, v_2, v_2) \quad [\because \Theta \text{ is linear and } v_1, v_2 \text{ are orthogonal}]\end{aligned}$$

If f_1 is the cubic form corresponding to Θ acting on MV and f_2 is the cubic form corresponding to Θ acting on $(I - M)V$ then the above equation says that: $f \sim f_1 \oplus f_2$. ∎

Note that given a cubic form f we can compute the center in terms of a basis over $\mathbb{F}$ as it just requires linear algebra computations. Thus, the above lemma gives a method of decomposing the cubic form if we can decompose its centre.

Proposition 3.2 (Harrison) *Cubic form decomposition can be done in polynomial time given an oracle of polynomial factoring over* $\mathbb{F}$.

Proof: Suppose f is a cubic form. Assume wlog that f is regular as otherwise we can regularize f by applying Proposition 3.1. Now compute its centre, $\text{Cent}(f)$, in deterministic polynomial time. As $\text{Cent}(f)$ is a commutative $\mathbb{F}$-algebra, recall the remark of Proposition 2.3, we can find the decomposition of the centre, using polynomial factoring over $\mathbb{F}$, into local commutative rings. In particular, if $\text{Cent}(f)$ is decomposable we can compute a nontrivial decomposition:

$$\text{Cent}(f) = R_1 \times R_2$$

from where we get a nontrivial idempotent, for example, the element of $\text{Cent}(f)$ corresponding to $(0, 1)$ (where 0 is the zero of R_1 and 1 is the unity of R_2). Now, the proof of Lemma 3.1 outlines a way of decomposing f using this nontrivial idempotent of $\text{Cent}(f)$. ∎

3.4 Our Cubic Forms

The cubic forms that we worked with in this chapter were of a special form. They owe their origin to local commutative $\mathbb{F}$-algebras. Suppose R is such an $\mathbb{F}$-algebra and $\mathcal{M}$ is its unique maximal ideal (refer to Definition 2.5). Let $b_1, \ldots, b_n$ be a basis of $\mathcal{M}$ over $\mathbb{F}$ and the multiplication in R is defined as:

$$\text{for all } 1 \leq i \leq j \leq n,\ b_i \cdot b_j = \sum_{1 \leq k \leq n} a_{i,j,k} b_k, \quad \text{where, } a_{i,j,k}\text{'s are in } \mathbb{F} \tag{3.33}$$

Now if we combine these multiplicative relations by considering b_i's as formal variables, homogenizing variable u and 'fresh' formal variables $z_{j,k}$'s then we get the following cubic form f from $\mathcal{M}$:

$$f(u, \overline{b}, \overline{z}) = \sum_{1 \leq i \leq j \leq n} z_{i,j} \left(b_i b_j - u \sum_{1 \leq k \leq n} a_{i,j,k} b_k \right)$$

These are more involved versions of *hyperbolic* cubic forms: $\sum_{1 \leq i \leq j \leq n} z_{i,j} b_i b_j$ (see [Keet93]). If R_1, R_2 are two $\mathbb{F}$-algebras with maximal ideals $\mathcal{M}_1, \mathcal{M}_2$ and the corresponding cubic forms f_1, f_2 then the proof of Claim 3.3.1 essentially says that an isomorphism from R_1 to R_2 gives an equivalence from f_1 to f_2.

In this section we show that these cubic forms are regular and indecomposable over any field $\mathbb{F}$ of char $\neq 2, 3$.

Theorem 3.7 *Let $\mathbb{F}$ be a field with char $\neq 2, 3$. Let $\mathcal{M}$ be a maximal ideal of a local commutative $\mathbb{F}$-algebra R such that $\mathcal{M}^2 \neq 0$. The multiplicative relations of $\mathcal{M}$ are given by Equation (3.33) and additionally $b_{n-1}^2 = 0,\ b_n \mathcal{M} = 0$. Define a cubic form f as:*

$$f(u, \overline{b}, \overline{z}) = \sum_{1 \leq i \leq j \leq n} z_{i,j} \left(b_i b_j - u \sum_{1 \leq k \leq n} a_{i,j,k} b_k \right)$$

Then,

(1) f is regular.

(2) f is indecomposable.

Proof: [(1)] As $\mathcal{M}^2 \neq 0$ note that f above is not u-free. Let Θ be the symmetric 3-linear map corresponding to f. Define the vector space $V := \mathbb{F}^m$, where $m := 1 + n + \binom{n+1}{2}$. Let us fix the notation for specifying the coordinates of a vector v_i in V as:

$$(u_i, b_{1,i}, \ldots, b_{n,i}, z_{1,1,i}, \ldots, z_{1,n,i}, z_{2,2,i}, \ldots, z_{2,n,i}, \ldots, z_{n,1,i}, \ldots, z_{n,n,i})^{\mathrm{T}}$$

or more compactly as:

$$\begin{pmatrix} u_i \\ \overline{b}_i \\ \overline{z}_i \end{pmatrix}$$

If f is not regular then there is a nonzero $v \in V$ such that $\Theta(v,\cdot,\cdot) = 0$. So consider the following equation in the variables $u_1, \overline{b}_1, \overline{z}_1$:

$$\text{for all } \begin{pmatrix} u_2 \\ \overline{b}_2 \\ \overline{z}_2 \end{pmatrix}, \begin{pmatrix} u_3 \\ \overline{b}_3 \\ \overline{z}_3 \end{pmatrix} \in V,\ \Theta\left(\begin{pmatrix} u_1 \\ \overline{b}_1 \\ \overline{z}_1 \end{pmatrix}, \begin{pmatrix} u_2 \\ \overline{b}_2 \\ \overline{z}_2 \end{pmatrix}, \begin{pmatrix} u_3 \\ \overline{b}_3 \\ \overline{z}_3 \end{pmatrix}\right) = 0 \tag{3.34}$$

Therefore, by considering the coefficient of $z_{i,i,3}$ in the above equation we get:

$$\frac{b_{i,1}b_{i,2}}{3} - \frac{u_1}{6}\sum_{1\le k\le n} a_{i,i,k}b_{k,2} - \frac{u_2}{6}\sum_{1\le k\le n} a_{i,i,k}b_{k,1} \ = \ 0 \tag{3.35}$$

and by considering the coefficient of $z_{i,j,3}$, for $1 \le i < j \le n$, we get:

$$\frac{b_{i,1}b_{j,2}}{6} + \frac{b_{j,1}b_{i,2}}{6} - \frac{u_1}{6}\sum_{1\le k\le n} a_{i,j,k}b_{k,2} - \frac{u_2}{6}\sum_{1\le k\le n} a_{i,j,k}b_{k,1} \ = \ 0 \tag{3.36}$$

If $u_1 = 0$ then the coefficient of $b_{i,2}$ in Equation (3.35) gives: $b_{i,1} = 0$. As i varies over $[1 \ldots n]$ we get: $\begin{pmatrix} u_1 \\ \overline{b}_1 \end{pmatrix} = 0$.

If $u_1 \neq 0$ then considering the coefficient of $b_{k,2}$ in Equation (3.35) we get: $a_{i,i,k} = 0$ for all $k \in [n] \setminus \{i\}$. Thus, in the ideal $\mathcal{M}$: $b_i^2 = a_{i,i,i}b_i$ or $b_i(b_i - a_{i,i,i}) = 0$. This implies that $a_{i,i,i} = 0$ for otherwise $(b_i - a_{i,i,i})$ is invertible (as b_i is in the unique maximal ideal $\mathcal{M}$) forcing $b_i = 0$. Thus, in the ideal $\mathcal{M}$: $b_i^2 = 0$ for all $i \in [n]$. Similarly, considering the coefficient of $b_{k,2}$ in Equation (3.36) we get: $a_{i,j,k} = 0$ for all $k \in [n] \setminus \{i,j\}$. Thus, in the ideal $\mathcal{M}$: $b_ib_j = a_{i,j,i}b_i + a_{i,j,j}b_j$. Multiplying this equation by b_i and using $b_i^2 = 0$ we get: $a_{i,j,j}b_ib_j = 0$ and symmetrically, $a_{i,j,i}b_ib_j = 0$. So if $b_ib_j \neq 0$ then $a_{i,j,i} = a_{i,j,j} = 0$ and hence $b_ib_j = 0$. Thus, in the ideal $\mathcal{M}$: $b_ib_j = 0$ for all $1 \le i \le j \le n$. But this contradicts the hypothesis that $\mathcal{M}^2 \neq 0$.

Thus, a solution of Equation (3.34) must satisfy: $\begin{pmatrix} u_1 \\ \overline{b}_1 \end{pmatrix} = 0$. Using this we can

now expand Equation (3.34) as:

$$\text{for all } \begin{pmatrix} u_2 \\ \overline{b}_2 \\ \overline{z}_2 \end{pmatrix}, \begin{pmatrix} u_3 \\ \overline{b}_3 \\ \overline{z}_3 \end{pmatrix} \in V, \sum_{1 \le i \le j \le n} z_{i,j,1} \left(\frac{b_{i,2}b_{j,3}}{6} + \frac{b_{j,2}b_{i,3}}{6} - \frac{u_2}{6} \sum_{1 \le k \le n} a_{i,j,k}b_{k,3} \right.$$
$$\left. - \frac{u_3}{6} \sum_{1 \le k \le n} a_{i,j,k}b_{k,2} \right)$$
$$= 0$$

The above equation clearly means that: $z_{i,j,1} = 0$ for all $1 \le i \le j \le n$. Thus, $\begin{pmatrix} u_1 \\ \overline{b}_1 \\ \overline{z}_1 \end{pmatrix} = 0$ and hence f is regular. ∎

Proof: [(2)] We compute the center of f and then show that it is an indecomposable $\mathbb{F}$-algebra which means, by Lemma 3.1, that f is indecomposable.

Let Θ be the symmetric 3-linear map corresponding to f. Define the vector space $V := \mathbb{F}^m$, where $m := 1 + n + \binom{n+1}{2}$. Let us fix the notation of specifying the coordinates of a vector v_i in V as:

$$(u_i, b_{1,i}, \ldots, b_{n,i}, z_{1,1,i}, \ldots, z_{1,n,i}, z_{2,2,i}, \ldots, z_{2,n,i}, \ldots, z_{n,1,i}, \ldots, z_{n,n,i})^{\mathrm{T}}$$

or more compactly as:

$$\begin{pmatrix} u_i \\ \overline{b}_i \\ \overline{z}_i \end{pmatrix}$$

Recall that $\mathrm{Cent}(f)$ consists of matrices $M \in \mathbb{F}^{m \times m}$ such that:

$$\forall \begin{pmatrix} u_1 \\ \overline{b}_1 \\ \overline{z}_1 \end{pmatrix}, \begin{pmatrix} u_2 \\ \overline{b}_2 \\ \overline{z}_2 \end{pmatrix}, \begin{pmatrix} u_3 \\ \overline{b}_3 \\ \overline{z}_3 \end{pmatrix} \in V, \qquad \Theta \left(M \begin{pmatrix} u_1 \\ \overline{b}_1 \\ \overline{z}_1 \end{pmatrix}, \begin{pmatrix} u_2 \\ \overline{b}_2 \\ \overline{z}_2 \end{pmatrix}, \begin{pmatrix} u_3 \\ \overline{b}_3 \\ \overline{z}_3 \end{pmatrix} \right)$$
$$= \Theta \left(\begin{pmatrix} u_1 \\ \overline{b}_1 \\ \overline{z}_1 \end{pmatrix}, M \begin{pmatrix} u_2 \\ \overline{b}_2 \\ \overline{z}_2 \end{pmatrix}, \begin{pmatrix} u_3 \\ \overline{b}_3 \\ \overline{z}_3 \end{pmatrix} \right) \tag{3.37}$$

Consider the matrix M in block form as: $\begin{pmatrix} M_{11} & M_{12} \\ M_{21} & M_{22} \end{pmatrix}$ such that M_{11} is $(n+1) \times (n+1)$ and M_{22} is $\binom{n+1}{2} \times \binom{n+1}{2}$. We prove properties of these block matrices in the subsequent claims.

Claim 3.7.1 $M_{12} = 0$.

Proof of Claim 3.7.1. Substitute $\begin{pmatrix} u_1 \\ \overline{b}_1 \end{pmatrix} = \begin{pmatrix} u_3 \\ \overline{b}_3 \end{pmatrix} = 0$ in Equation (3.37) to get:

$$\forall \begin{pmatrix} 0 \\ \overline{z}_1 \end{pmatrix}, \begin{pmatrix} u_2 \\ \overline{b}_2 \\ \overline{z}_2 \end{pmatrix}, \begin{pmatrix} 0 \\ \overline{z}_3 \end{pmatrix} \in V, \quad \Theta \left(\begin{pmatrix} M_{12}\overline{z}_1 \\ M_{22}\overline{z}_1 \end{pmatrix}, \begin{pmatrix} u_2 \\ \overline{b}_2 \\ \overline{z}_2 \end{pmatrix}, \begin{pmatrix} 0 \\ \overline{z}_3 \end{pmatrix} \right) = 0$$

If $M_{12} \neq 0$ then we can assign $\overline{z}_1 = v_1 \in \mathbb{F}^{\binom{n+1}{2} \times \binom{n+1}{2}}$ such that $M_{12}v_1 \neq 0$ and:

$$\forall \begin{pmatrix} u_2 \\ \overline{b}_2 \\ \overline{z}_2 \end{pmatrix}, \begin{pmatrix} 0 \\ \overline{z}_3 \end{pmatrix} \in V, \quad \Theta \left(\begin{pmatrix} M_{12}v_1 \\ M_{22}v_1 \end{pmatrix}, \begin{pmatrix} u_2 \\ \overline{b}_2 \\ \overline{z}_2 \end{pmatrix}, \begin{pmatrix} 0 \\ \overline{z}_3 \end{pmatrix} \right) = 0 \tag{3.38}$$

Notice that we can now run the proof of the regularity of f, as equations similar to Equation (3.35) and Equation (3.36) can be obtained by comparing the coefficients of $z_{i,i,3}, z_{i,j,3}$ in the Equation (3.38), to deduce $M_{12}v_1 = 0$. This contradiction shows that $M_{12} = 0$. □

Thus, an $M \in \text{Cent}(f)$ looks like: $M = \begin{pmatrix} M_{11} & 0 \\ M_{21} & M_{22} \end{pmatrix}$. Let τ be a linear transformation on V induced by M, i.e.,

$$M \begin{pmatrix} u_i \\ \overline{b}_i \\ \overline{z}_i \end{pmatrix} = \begin{pmatrix} \tau(u_i) \\ \tau(b_{1,i}) \\ \vdots \\ \tau(b_{n,i}) \\ \tau(z_{1,1,i}) \\ \vdots \\ \tau(z_{n,n,i}) \end{pmatrix}$$

Claim 3.7.2 *There is an $\alpha \in \mathbb{F}$ such that $M_{11} = \alpha \cdot I$.*

Proof of Claim 3.7.2. To understand M more let us substitute: $\overline{z}_1 = \overline{z}_2 = 0, \begin{pmatrix} u_3 \\ \overline{b}_3 \end{pmatrix} = 0$ in the Equation (3.37):

$$\begin{aligned} \forall \begin{pmatrix} u_1 \\ \overline{b}_1 \\ 0 \end{pmatrix}, \begin{pmatrix} u_2 \\ \overline{b}_2 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ \overline{z}_3 \end{pmatrix} \in V, \qquad & \Theta \left(\begin{pmatrix} M_{11} \begin{pmatrix} u_1 \\ \overline{b}_1 \end{pmatrix} \\ M_{21} \begin{pmatrix} u_1 \\ \overline{b}_1 \end{pmatrix} \end{pmatrix}, \begin{pmatrix} u_2 \\ \overline{b}_2 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ \overline{z}_3 \end{pmatrix} \right) \\ = \ & \Theta \left(\begin{pmatrix} u_1 \\ \overline{b}_1 \\ 0 \end{pmatrix}, \begin{pmatrix} M_{11} \begin{pmatrix} u_2 \\ \overline{b}_2 \end{pmatrix} \\ M_{21} \begin{pmatrix} u_2 \\ \overline{b}_2 \end{pmatrix} \end{pmatrix}, \begin{pmatrix} 0 \\ \overline{z}_3 \end{pmatrix} \right) \end{aligned} \tag{3.39}$$

In the above equation comparing the coefficient of $z_{i,j,3}$, for $1 \leq i \leq j \leq n$, gives:

$$\begin{aligned} & \frac{\tau(b_{i,1})b_{j,2}}{6} + \frac{\tau(b_{j,1})b_{i,2}}{6} - \frac{\tau(u_1)}{6} \sum_{1 \leq k \leq n} a_{i,j,k} b_{k,2} - \frac{u_2}{6} \sum_{1 \leq k \leq n} a_{i,j,k} \tau(b_{k,1}) \\ = \quad & \frac{b_{i,1}\tau(b_{j,2})}{6} + \frac{b_{j,1}\tau(b_{i,2})}{6} - \frac{u_1}{6} \sum_{1 \leq k \leq n} a_{i,j,k} \tau(b_{k,2}) - \frac{\tau(u_2)}{6} \sum_{1 \leq k \leq n} a_{i,j,k} b_{k,1} \end{aligned} \tag{3.40}$$

We have $b_n \mathcal{M} = 0$ in R thus, $b_n^2 = 0$ in R and so $a_{n,n,k} = 0$ for all $k \in [n]$. Thus, the Equation (3.40) for $(i, j) = (n, n)$ is simply:

$$\frac{\tau(b_{n,1})b_{n,2}}{3} = \frac{b_{n,1}\tau(b_{n,2})}{3}$$

Since, the above equation holds for all $\begin{pmatrix} u_1 \\ \overline{b}_1 \\ 0 \end{pmatrix}, \begin{pmatrix} u_2 \\ \overline{b}_2 \\ 0 \end{pmatrix} \in V$ we deduce that there is an $\alpha \in \mathbb{F}$ such that $\tau(b_{n,1}) = \alpha \cdot b_{n,1}$.

Note that $b_i b_n = 0$ in R, for any $i \in [n]$, so $a_{i,n,k} = 0$ for all $k \in [n]$. Thus,

Equation (3.40) for $(i,j) = (i,n)$, where $1 \leq i < n$, becomes:

$$\begin{aligned}
&\frac{\tau(b_{i,1})b_{n,2}}{6} + \frac{\tau(b_{n,1})b_{i,2}}{6} = \frac{b_{i,1}\tau(b_{n,2})}{6} + \frac{b_{n,1}\tau(b_{i,2})}{6} \\
\Rightarrow \quad &\frac{\tau(b_{i,1})b_{n,2}}{6} + \frac{\alpha b_{n,1}b_{i,2}}{6} = \frac{\alpha b_{i,1}b_{n,2}}{6} + \frac{b_{n,1}\tau(b_{i,2})}{6} \qquad [\because \tau(b_{n,1}) = \alpha \cdot b_{n,1}] \\
\Rightarrow \quad &(\tau(b_{i,1}) - \alpha b_{i,1})\, b_{n,2} = b_{n,1}\,(\tau(b_{i,2}) - \alpha b_{i,2})
\end{aligned}$$

Since, the above equation holds for all $\begin{pmatrix} u_1 \\ \overline{b}_1 \\ 0 \end{pmatrix}, \begin{pmatrix} u_2 \\ \overline{b}_2 \\ 0 \end{pmatrix} \in V$ we deduce that there is a $\beta \in \mathbb{F}$ such that

$$\tau(b_{i,1}) - \alpha b_{i,1} = \beta \cdot b_{n,1} \quad \text{for all } i \in [n-1] \tag{3.41}$$

Since, $b_{n-1}^2 = 0$ in R, we have $a_{n-1,n-1,k} = 0$ for all $k \in [n]$ and thus, Equation (3.40) for $(i,j) = (n-1, n-1)$ becomes:

$$\frac{\tau(b_{n-1,1})b_{n-1,2}}{3} = \frac{b_{n-1,1}\tau(b_{n-1,2})}{3}$$

Since, the above equation holds for all $\begin{pmatrix} u_1 \\ \overline{b}_1 \\ 0 \end{pmatrix}, \begin{pmatrix} u_2 \\ \overline{b}_2 \\ 0 \end{pmatrix} \in V$ we deduce that there is a $\gamma \in \mathbb{F}$ such that $\tau(b_{n-1,1}) = \gamma \cdot b_{n-1,1}$. This together with Equation (3.41) gives:

$$\begin{aligned}
&\tau(b_{n-1,1}) = \gamma \cdot b_{n-1,1} = \alpha \cdot b_{n-1,1} + \beta \cdot b_{n,1} \\
\Rightarrow \quad &\gamma = \alpha \text{ and } \beta = 0
\end{aligned}$$

Finally, this together with Equation (3.41) gives us a nice form for τ:

$$\tau(b_{i,1}) = \alpha \cdot b_{i,1} \quad \text{for all } i \in [n] \tag{3.42}$$

Now choose $i \leq j \in [n]$ such that $b_i b_j \neq 0$ in R so that there is a $k \in [n]$ such that $a_{i,j,k} \neq 0$. Plugging Equation (3.42) in Equation (3.40) we get:

$$\begin{aligned}
&\frac{\tau(u_1)}{6} \sum_{1 \leq k \leq n} a_{i,j,k} b_{k,2} + \frac{\alpha u_2}{6} \sum_{1 \leq k \leq n} a_{i,j,k} b_{k,1} = \frac{\alpha u_1}{6} \sum_{1 \leq k \leq n} a_{i,j,k} b_{k,2} + \frac{\tau(u_2)}{6} \sum_{1 \leq k \leq n} a_{i,j,k} b_{k,1} \\
\Rightarrow \quad &(\tau(u_1) - \alpha u_1) \sum_{1 \leq k \leq n} a_{i,j,k} b_{k,2} = (\tau(u_2) - \alpha u_2) \sum_{1 \leq k \leq n} a_{i,j,k} b_{k,1}
\end{aligned}$$

If $b_i b_j \neq 0$ in R then there is a $k \in [n]$ such that $a_{i,j,k} \neq 0$ and as the above equation holds for all $\begin{pmatrix} u_1 \\ \overline{b}_1 \\ 0 \end{pmatrix}, \begin{pmatrix} u_2 \\ \overline{b}_2 \\ 0 \end{pmatrix} \in V$ we deduce that there is a $\gamma \in \mathbb{F}$ such that:

$$\tau(u_1) - \alpha u_1 = \gamma \cdot \sum_{1 \leq k \leq n} a_{i,j,k} b_{k,1} \quad \text{where } r := \sum_{1 \leq k \leq n} a_{i,j,k} b_{k,1} \neq 0$$

If $\gamma \neq 0$ then since the LHS of the above equation is independent of i, j we will have that for all $i \leq j \in [n]$ either $b_i b_j = 0$ or r. Thus, $r^2 = c \cdot r$ for some $c \in \mathbb{F}$. As r is a nonzero element of the maximal ideal $\mathcal{M}$ this implies that $r = 0$. This contradiction means that $\gamma = 0$ and hence:

$$\tau(u_1) = \alpha u_1$$

This together with Equation (3.42) gives:

$$\begin{aligned} M_{11} \begin{pmatrix} u_1 \\ \overline{b}_1 \end{pmatrix} &= \begin{pmatrix} \tau(u_1) \\ \tau(b_{1,1}) \\ \vdots \\ \tau(b_{n,1}) \end{pmatrix} = \begin{pmatrix} \alpha u_1 \\ \alpha b_{1,1} \\ \vdots \\ \alpha b_{n,1} \end{pmatrix} \qquad (3.43) \\ \Rightarrow \quad M_{11} &= \alpha \cdot I \end{aligned}$$

□

Claim 3.7.3 *$M_{22} = \alpha \cdot I$, where α is the same as in the last claim.*

Proof of Claim 3.7.3. Let us start by substituting: $\begin{pmatrix} u_1 \\ \overline{b}_1 \end{pmatrix} = 0, \overline{z}_2 = \overline{z}_3 = 0$ in the

Equation (3.37):

$$\Theta\left(\begin{pmatrix}0\\M_{22}\overline{z}_1\end{pmatrix},\begin{pmatrix}u_2\\\overline{b}_2\\0\end{pmatrix},\begin{pmatrix}u_3\\\overline{b}_3\\0\end{pmatrix}\right)=\Theta\left(\begin{pmatrix}0\\\overline{z}_1\end{pmatrix},\begin{pmatrix}M_{11}\begin{pmatrix}u_2\\\overline{b}_2\end{pmatrix}\\M_{21}\begin{pmatrix}u_2\\\overline{b}_2\end{pmatrix}\end{pmatrix},\begin{pmatrix}u_3\\\overline{b}_3\\0\end{pmatrix}\right)$$

$$\Rightarrow\quad\Theta\left(\begin{pmatrix}0\\M_{22}\overline{z}_1\end{pmatrix},\begin{pmatrix}u_2\\\overline{b}_2\\0\end{pmatrix},\begin{pmatrix}u_3\\\overline{b}_3\\0\end{pmatrix}\right)=\Theta\left(\begin{pmatrix}0\\\overline{z}_1\end{pmatrix},\begin{pmatrix}\alpha\begin{pmatrix}u_2\\\overline{b}_2\end{pmatrix}\\M_{21}\begin{pmatrix}u_2\\\overline{b}_2\end{pmatrix}\end{pmatrix},\begin{pmatrix}u_3\\\overline{b}_3\\0\end{pmatrix}\right)\tag{3.44}$$

$$\Rightarrow\quad\Theta\left(\begin{pmatrix}0\\M_{22}\overline{z}_1\end{pmatrix},\begin{pmatrix}u_2\\\overline{b}_2\\0\end{pmatrix},\begin{pmatrix}u_3\\\overline{b}_3\\0\end{pmatrix}\right)=\Theta\left(\begin{pmatrix}0\\\overline{z}_1\end{pmatrix},\begin{pmatrix}\alpha\begin{pmatrix}u_2\\\overline{b}_2\end{pmatrix}\\0\end{pmatrix},\begin{pmatrix}u_3\\\overline{b}_3\\0\end{pmatrix}\right)$$

$$\Rightarrow\quad\Theta\left(\begin{pmatrix}0\\(M_{22}-\alpha I)\overline{z}_1\end{pmatrix},\begin{pmatrix}u_2\\\overline{b}_2\\0\end{pmatrix},\begin{pmatrix}u_3\\\overline{b}_3\\0\end{pmatrix}\right)=0$$

As the above equation holds for all $\begin{pmatrix}0\\\overline{z}_1\end{pmatrix},\begin{pmatrix}u_2\\\overline{b}_2\\0\end{pmatrix},\begin{pmatrix}u_3\\\overline{b}_3\\0\end{pmatrix}\in V$ we deduce:

$$M_{22}=\alpha I$$

□

Thus, any element M in the center of f looks like:

$$\begin{pmatrix}0&0\\M_{12}&0\end{pmatrix}+\alpha I\quad\text{where, }\alpha\in\mathbb{F}$$

Now if M is idempotent then:

$$\begin{aligned}&M^2=M\\\Rightarrow\quad&M(M-I)=0\end{aligned}$$

But one of the matrices M or $(M - I)$ will always be invertible and hence $M = 0$ or $M = I$. Thus, Cent(f) is an indecomposable $\mathbb{F}$-algebra and, hence, f is indecomposable by Lemma 3.1. ∎

3.5 Discussion

This chapter studied the complexity of the problem of polynomial equivalence. Over finite fields this problem is of intermediate complexity and, hence, unlikely to be NP-hard. Over infinite fields we know very little about this general problem! The special case of quadratic forms is completely understood due to the works of Minkowski [Minkow], Hasse [Has21] and Witt [Witt]. Inspired from quadratic forms we considered "slightly" more general case of cubic forms and proved some interesting results. We gave a reduction from commutative $\mathbb{F}$-algebra isomorphism to $\mathbb{F}$-cubic forms equivalence for any field $\mathbb{F}$. Two of its consequences are: Graph isomorphism reduces to the problem of cubic forms equivalence over any field $\mathbb{F}$, and equivalence of higher degree d-forms reduces to cubic forms equivalence over fields $\mathbb{F}$ having d-th roots. Clearly, cubic forms equivalence seems to be the most important special case of the problem of polynomial equivalence.

We hope that the rich structure of cubic forms will eventually give us more insights about the isomorphism problems of commutative $\mathbb{F}$-algebras and graphs. As a first step to understanding cubic forms, we believe that the decidability of cubic forms equivalence over $\mathbb{Q}$ should be shown.

In the case of quadratic forms over $\mathbb{Q}$ the problem of equivalence reduced to questions of finding $\mathbb{Q}$-roots of a quadratic form. In particular, if two quadratic forms are equivalent over $\mathbb{R}$ and represent the same set of points over $\mathbb{Q}$ then they are equivalent over $\mathbb{Q}$. Here, we show that such a result does not hold for cubic forms, thus, giving evidence that $\mathbb{Q}$-root finding of a cubic form may not be related to the problem of equivalence of cubic forms. Let us define two rings:

$$R := \mathbb{Q}[x]/(x^2 - 1) \quad \text{and} \quad S := \mathbb{Q}[x]/(x^2 - 2)$$

Notice that the $\mathbb{Q}$-algebras R, S are isomorphic over $\mathbb{R}$ but nonisomorphic over $\mathbb{Q}$. Thus, using the construction given in Theorem 3.5 we get two cubic forms $\phi_R(\overline{y}, \overline{c}, v), \phi_S(\overline{y}, \overline{c}, v)$ that are equivalent over $\mathbb{R}$ but nonequivalent over $\mathbb{Q}$. But what are the rational points that these cubic forms represent? If we choose an i such that the coefficient of $y_{i,i}$ in ϕ_R is c_i^2 then:

$$\phi_R(0, \ldots, y_{i,i}, \ldots, 0, \overline{c}, v) = y_{i,i}c_i^2$$

Clearly, there exists such an i (recall the way we constructed ϕ_R) and, hence, ϕ_R represents all points in $\mathbb{Q}$. Similarly, ϕ_S represents all points in $\mathbb{Q}$. This gives us two cubic forms that are equivalent over $\mathbb{R}$, represent the same set of points over $\mathbb{Q}$ but are yet nonequivalent over $\mathbb{Q}$.

Finally, we pose some questions whose answers might unfold more structure of cubic forms:

- What are the invariants of cubic forms (under equivalence)?
- If cubic forms f, g are equivalent over $\mathbb{R}$ and are equivalent modulo p^k, for all primes p (except finitely many primes) and $k \in \mathbb{Z}^{\geq 1}$, then are they equivalent over $\mathbb{Q}$?
- Can we reduce $\mathbb{F}$-cubic forms equivalence problem to that of $\mathbb{F}$-algebra isomorphism, over *all* fields $\mathbb{F}$?

Chapter 4

Identity Testing

Given a polynomial $f(x_1, \ldots, x_n)$ over a field $\mathbb{F}$, we want to test whether it is the zero polynomial or not. For example, over $\mathbb{F}_2$, $x^2 - x$ is a nonzero polynomial while $(x + y)^2 - x^2 - y^2$ is a zero polynomial. It is a trivial problem if f is given in the expanded form, i.e., each of its coefficients are explicitly given. But suppose f is given in a more compact form, say, as an arithmetic circuit $\mathcal{C}$ having addition and multiplication gates, variables $x_1, \ldots, x_n$ and constants from the field $\mathbb{F}$. Then the problem of checking whether $\mathcal{C}(x_1, \ldots, x_n) = 0$ in time polynomial in the size($\mathcal{C}$) becomes more interesting and is called *identity testing*. Several randomized algorithms for the problem are known. Schwartz and Zippel [Sch80, Zip79] gave the first such algorithm, it evaluates f at a random point $\overline{a} \in \mathbb{F}^n$ and accepts iff $f(\overline{a}) = 0$. There are more involved randomized algorithms that require lesser number of random bits [CK97, LV98, AB99, KS01].

The study of this simply-defined algebraic problem has led to many exciting results in complexity theory. The results like – PSPACE has interactive protocols [LFKN92, Sha92], NP has probabilistically checkable proofs [AS97, AS98, ALM+98], equivalence testing of read-once branching programs [BCW80], multiset equality testing [BK95], perfect matching is in RNC [Lov79, MVV87], primality testing is in P [AKS04] – all have identity testing at their heart. Recently, identity testing gained even more significance when its connection to proving lower bounds was shown. Impagliazzo and Kabanets [IK03] showed that finding a deterministic polynomial time

algorithm for identity testing is essentially equivalent to proving super polynomial circuit lower bounds for NEXP.

Thus, derandomization of identity testing is most sought-after. The derandomization results currently known are all for restricted classes of circuits $\mathcal{C}$. When $\mathcal{C}$ is a noncommutative formula identity testing can be done in deterministic polynomial time [RS04]. For $\mathcal{C}$ of depth 3 with a bounded fanin top-gate, Dvir and Shpilka [DS05] gave a deterministic quasi polynomial identity test. They achieved this by giving a structural result about zero circuits of depth 3 with a bounded fanin top-gate.

In this chapter we too focus on the special case of $\mathcal{C}$ being a depth 3, bounded top fanin circuit. We give the first deterministic polynomial time algorithm using the machinery of local commutative rings. We view the identity testing problem for $\mathcal{C}$ as an isomorphism problem of rings given in the polynomial representation and then solve this special case.

The results of this chapter mostly appear in [KS06].

4.1 $\Sigma\Pi\Sigma$ Circuits

Proving lower bounds for general arithmetic circuits is one of the central problems of complexity theory. Due to the difficulty of the problem research has focussed on restricted models like monotone circuits and bounded depth circuits. For monotone arithmetic circuits, exponential lower bounds on the size [ShS77, JS80] and linear lower bounds on the depth [ShS80, TT94] have been shown. However, only weak lower bounds are known for bounded depth arithmetic circuits [Pud94, RS01]. Thus, a more restricted model was considered – the model of depth 3 arithmetic circuits. A depth 3 circuit computes a sum of products of linear functions *or* a product of sums of terms. Exponential lower bounds on the size of depth 3 arithmetic circuits has been shown over finite fields [GK98]. For general depth 3 circuits over infinite fields only the quadratic lower bound of [SW99] is known.

No efficient algorithm for identity testing of depth 3 circuits is known. Note that if the top gate of a depth 3 circuit $\mathcal{C}$ is a multiplication gate then $\mathcal{C} = 0$ iff one of the

inputs to the top gate is zero, which in turn is easy to check. Thus, the hard case is when the top gate is an addition gate and the next two layers are of multiplication and addition gates respectively. Such a circuit is called a ΣΠΣ circuit. It is a sum of products of linear functions and looks like:

$$\mathcal{C}(\overline{x}) = \sum_{i=1}^{k} \prod_{j=1}^{d_i} L_{i,j}(\overline{x}) \tag{4.1}$$

where, $L_{i,j}$'s are (wlog) homogeneous linear functions called *linear forms*. The identities of "small" ΣΠΣ circuits seem very natural, for example, the identities taught in high-school algebra are mostly identities of this kind.

Example The zero circuit $\mathcal{C}(x_1, \ldots, x_n) := (x_1 + \cdots + x_n)^2 - \sum_{1 \leq i,j \leq n} x_i x_j$ is clearly a $O(n^2)$-sized ΣΠΣ circuit involving nontrivial linear forms. ∎

Ben-Or [SW99] showed that polynomial-sized ΣΠΣ circuits can compute some very nontrivial functions, for example, they can compute all symmetric polynomials (of degree $n^{O(1)}$) over $x_1, \ldots, x_n$. This gives a related identity for ΣΠΣ circuits over infinite fields.

Example [Ben-Or] There are constants (not all zero) $a_0, \ldots, a_n \in \mathbb{Q}$ such that the $O(n^2)$-sized ΣΠΣ circuit:

$$\mathcal{C}(x_1, \ldots, x_n) := \sum_{i=0}^{n} a_i (x_1 + i) \cdots (x_n + i)$$

is a zero circuit. ∎

Here, we are interested in studying the identity testing problem for a restricted case of ΣΠΣ circuits – when the top fanin is bounded. This case was posed as a challenge by Klivans and Spielman [KS01] and a *quasi polynomial time* algorithm was given by Dvir and Shpilka [DS05].

4.2 Previous Approaches

Let $\mathcal{C}$ be a ΣΠΣ circuit, as in Equation (4.1), computing the zero polynomial. We will call $\mathcal{C}$ to be *minimal* if no proper subset of the multiplication gates of $\mathcal{C}$ sums

to zero. We say that $\mathcal{C}$ is *simple* if there is no linear function that appears in all the multiplication gates (up to a multiplicative constant). *Rank* of $\mathcal{C}$ is the rank of the linear forms appearing in $\mathcal{C}$.

Example The circuit $\mathcal{C}_1(x_1, x_2) := x_1^2 - x_2^2 - x_1^2 + x_2^2$ is not minimal as a sub-circuit is zero: $x_1^2 - x_1^2 = 0$. The circuit $\mathcal{C}_2(x_1, x_2) := x_1^3 - x_2^2 x_1 - (x_1 - x_2)(x_1 + x_2)x_1$ is minimal but not simple as x_1 is common to all multiplication gates. The circuit $\mathcal{C}_3(x_1, x_2) := x_1^2 - x_2^2 - (x_1 - x_2)(x_1 + x_2)$ is both minimal and simple.

All these circuits $\mathcal{C}_1, \mathcal{C}_2$ and $\mathcal{C}_3$ are of rank 2. ∎

The quasi polynomial time algorithm of Dvir and Shpilka [DS05] is based on the result that the rank of a minimal and simple $\Sigma\Pi\Sigma$ circuit with bounded top fanin and computing zero is "small". Formally, the result says:

Theorem 4.1 (Thm 1.4 of [DS05]). *Let $k \geq 3, d \geq 2$ and let $\mathcal{C}$ be a simple and minimal $\Sigma\Pi\Sigma$ zero circuit of degree d with k multiplication gates and n inputs, then rank$(\mathcal{C}) \leq 2^{O(k^2)} \log(d)^{k-2}$.*

Effectively, this means that if we have such a circuit $\mathcal{C}$ and k is a constant then we can check whether it is zero or not by completely expanding-out $\mathcal{C}$ and checking whether each of the $O(d^{\text{rank}(\mathcal{C})})$ coefficients is zero. Clearly, this takes time $O(d^{\text{rank}(\mathcal{C})}) = 2^{O(\log(d)^{k-1})}$ as number of variables in $\mathcal{C}$ can be made equal to rank$(\mathcal{C})$ by applying a linear transformation. This gave hope of finding a polynomial time algorithm if we can improve the upper bound on the rank$(\mathcal{C})$ to a constant (i.e., independent of d). In fact, Dvir and Shpilka [DS05] conjectured that rank$(\mathcal{C}) = O(k)$. Here, we give identities that contradict this conjecture. Thus, methods of Dvir and Shpilka [DS05] are unlikely to give an efficient algorithm and we give new techniques in the subsequent sections that work.

For $k = 3$, [DS05] shows that a minimal, simple $\Sigma\Pi\Sigma$ zero circuit should have rank $O(\log d)$. We show below that this bound is tight.

Lemma 4.1 *Define*

$$\begin{aligned}\mathcal{C}(x_1,\ldots,x_m,y) := & \prod_{\substack{b_1,\ldots,b_m\in\mathbb{F}_2\\ b_1+\cdots+b_m\equiv 0(mod\ 2)}} (y+b_1x_1+\cdots+b_mx_m)\\ & + \prod_{\substack{b_1,\ldots,b_m\in\mathbb{F}_2\\ b_1+\cdots+b_m\equiv 1(mod\ 2)}} (b_1x_1+\cdots+b_mx_m)\\ & + \prod_{\substack{b_1,\ldots,b_m\in\mathbb{F}_2\\ b_1+\cdots+b_m\equiv 1(mod\ 2)}} (y+b_1x_1+\cdots+b_mx_m)\end{aligned}$$

Then, over $\mathbb{F}_2$, $\mathcal{C}$ is a simple and minimal $\Sigma\Pi\Sigma$ zero circuit of degree $d = 2^{m-1}$ with $k = 3$ multiplication gates and rank$(\mathcal{C}) = \log(d) + 2$.

Proof: For brevity denote the output of the three multiplication gates by T_1, T_2, T_3 in order.

Let $a_1,\ldots,a_m \in \mathbb{F}$ be such that $(a_1+\cdots+a_m) = 1 \pmod 2$. Let us compute $\mathcal{C}$ modulo $(a_1x_1+\cdots+a_mx_m)$. Since $(a_1x_1+\cdots+a_mx_m)$ occurs as a factor of T_2 we deduce $T_2 = 0 \pmod{a_1x_1+\cdots+a_mx_m}$. Further,

$$\begin{aligned}T_1 &= \prod_{\substack{b_1,\ldots,b_m\in\mathbb{F}_2\\ b_1+\cdots+b_m\equiv 0(\mathrm{mod}\ 2)}} (y+b_1x_1+\cdots+b_mx_m)\\ &\equiv \prod_{\substack{b_1,\ldots,b_m\in\mathbb{F}_2\\ b_1+\cdots+b_m\equiv 0(\mathrm{mod}\ 2)}} (y+(a_1+b_1)x_1+\cdots+(a_m+b_m)x_m) \pmod{a_1x_1+\cdots+a_mx_m}\\ &\equiv \prod_{\substack{b_1,\ldots,b_m\in\mathbb{F}_2\\ b_1+\cdots+b_m\equiv 1(\mathrm{mod}\ 2)}} (y+b_1x_1+\cdots+b_mx_m) \pmod{a_1x_1+\cdots+a_mx_m}\\ &\equiv T_3 \pmod{a_1x_1+\cdots+a_mx_m}\end{aligned}$$

Thus, we deduce: $T_1+T_2+T_3 \equiv 0 \pmod{a_1x_1+\cdots+a_mx_m}$ for any $a_1,\ldots,a_m \in \mathbb{F}$, $(a_1+\cdots+a_m) = 1 \pmod 2$. Also, notice that $T_1 = 0 \pmod y$ (consider the linear factor of T_1 obtained by setting: $b_1 = \cdots = b_m = 0$) and $T_2 = T_3 \pmod y$ implying that $T_1+T_2+T_3 = 0 \pmod y$. Thus, we get that:

$$\left(y \cdot \prod_{\substack{b_1,\ldots,b_m\in\mathbb{F}_2\\ b_1+\cdots+b_m\equiv 1(\mathrm{mod}\ 2)}} (b_1x_1+\cdots+b_mx_m)\right) \text{ divides } \mathcal{C}(x_1,\ldots,x_m,y)$$

But the divisor above has degree higher than that of $\mathcal{C}$ implying that $\mathcal{C} = 0$ (see Claim 4.1.3).

Moreover, it is easy to see that $\mathcal{C}$ is a minimal, simple $\Sigma\Pi\Sigma$ circuit of degree 2^{m-1}. ∎

The above identity is over a very special field – $\mathbb{F}_2$. Are there minimal, simple $\Sigma\Pi\Sigma$ identities of bounded k but unbounded rank over any field $\mathbb{F}$? We are not sure about fields of characteristic 0 but over fields of prime characteristic the following lemma answers in the affirmative.

Lemma 4.2 *Let p be an odd prime. Define:*

$$\mathcal{C}(x_1, \ldots, x_m, y) := \sum_{i=0}^{p-1} \prod_{\substack{b_1,\ldots,b_m \in \mathbb{F}_p \\ b_1+\cdots+b_m \equiv i (mod\ p)}} (y + b_1x_1 + \cdots + b_mx_m)$$

Then, over $\mathbb{F}_p$, $\mathcal{C}$ is a simple and minimal $\Sigma\Pi\Sigma$ zero circuit of degree $d = p^{m-1}$ with $k = p$ multiplication gates and rank$(\mathcal{C}) = \log_p(d) + 2$.

Proof: Fix an $i_0 \in \mathbb{F}_p$ and let $a_1, \ldots, a_m \in \mathbb{F}_p$ such that $(a_1 + \cdots + a_m) = i_0 \pmod p$. Now we compute $\mathcal{C}$ modulo $(y + a_1x_1 + \cdots + a_mx_m)$:

$$\begin{aligned}
\mathcal{C} &= \sum_{i=0}^{p-1} \prod_{\substack{b_1,\ldots,b_m \in \mathbb{F}_p \\ b_1+\cdots+b_m \equiv i (\mathrm{mod}\ p)}} (y + b_1x_1 + \cdots + b_mx_m) \\
&\equiv \sum_{\substack{i=0 \\ i \neq i_0}}^{p-1} \prod_{\substack{b_1,\ldots,b_m \in \mathbb{F}_p \\ b_1+\cdots+b_m \equiv i (\mathrm{mod}\ p)}} (y + b_1x_1 + \cdots + b_mx_m) \pmod{y + a_1x_1 + \cdots + a_mx_m} \\
&\equiv \sum_{\substack{i=0 \\ i \neq i_0}}^{p-1} \prod_{\substack{b_1,\ldots,b_m \in \mathbb{F}_p \\ b_1+\cdots+b_m \equiv i (\mathrm{mod}\ p)}} ((b_1 - a_1)x_1 + \cdots + (b_m - a_m)x_m) \pmod{y + a_1x_1 + \cdots + a_mx_m} \\
&\equiv \sum_{i=1}^{p-1} \prod_{\substack{b_1,\ldots,b_m \in \mathbb{F}_p \\ b_1+\cdots+b_m \equiv i (\mathrm{mod}\ p)}} (b_1x_1 + \cdots + b_mx_m) \pmod{y + a_1x_1 + \cdots + a_mx_m}
\end{aligned}$$

$$\equiv \sum_{i=1}^{\frac{p-1}{2}} \left(\prod_{\substack{b_1,\ldots,b_m \in \mathbb{F}_p \\ b_1+\cdots+b_m \equiv i \pmod p}} (b_1x_1 + \cdots + b_mx_m) \quad + \right.$$

$$\left. \prod_{\substack{b_1,\ldots,b_m \in \mathbb{F}_p \\ b_1+\cdots+b_m \equiv -i \pmod p}} (b_1x_1 + \cdots + b_mx_m) \right) \pmod{y + a_1x_1 + \cdots + a_mx_m}$$

$$\equiv \sum_{i=1}^{\frac{p-1}{2}} \left(\prod_{\substack{b_1,\ldots,b_m \in \mathbb{F}_p \\ b_1+\cdots+b_m \equiv i \pmod p}} (b_1x_1 + \cdots + b_mx_m) \quad + \right.$$

$$\left. (-1)^{p^{m-1}} \cdot \prod_{\substack{b_1,\ldots,b_m \in \mathbb{F}_p \\ b_1+\cdots+b_m \equiv i \pmod p}} (b_1x_1 + \cdots + b_mx_m) \right) \pmod{y + a_1x_1 + \cdots + a_mx_m}$$

$$\equiv 0 \pmod{y + a_1x_1 + \cdots + a_mx_m}$$

Thus, we deduce that for any $a_1, \ldots, a_m \in \mathbb{F}_p$:

$$\mathcal{C}(x_1, \ldots, x_m, y) \equiv 0 \pmod{y + a_1x_1 + \cdots + a_mx_m}$$

$$\Rightarrow \left(\prod_{a_1,\ldots,a_m \in \mathbb{F}_p} (y + a_1x_1 + \cdots + a_mx_m) \right) \text{ divides } \mathcal{C}(x_1, \ldots, x_m, y)$$

But the divisor above has a degree higher than that of $\mathcal{C}$ implying that $\mathcal{C} = 0$ (see Claim 4.1.3).

Moreover, it is easy to see that $\mathcal{C}$ is a minimal, simple $\Sigma\Pi\Sigma$ circuit of degree p^{m-1}. ∎

4.3 An Algorithm for bounded-$\Sigma\Pi\Sigma$

This section describes the first deterministic polynomial time identity test for $\Sigma\Pi\Sigma$ circuits of bounded top fanin. Assume that in the input we are given a circuit $\mathcal{C}$ computing a polynomial in $\mathbb{F}[x_1, x_2, \cdots, x_n]$. Let,

$$\mathcal{C} = T_1 + T_2 + \cdots + T_k$$

where, k is treated as a constant and each T_i, wlog, is a product of d linear forms:

$$T_i = L_{i,1}L_{i,2}\cdots L_{i,d}.$$

Each linear form $L_{i,j}$ looks like:

$$L_{i,j} = a_{i,j,1}x_1 + a_{i,j,2}x_2 + \cdots + a_{i,j,n}x_n, \quad a_{i,j,1}, \cdots, a_{i,j,n} \in \mathbb{F}$$

Our main idea of checking whether $\mathcal{C} = 0$ is Chinese remaindering, i.e., we pick suitable polynomials $f_1, \ldots, f_m \in \mathbb{F}[x_1, \ldots, x_n]$ and check whether $\mathcal{C} = 0$ modulo each of these f_i's. This idea is easy to demonstrate for the cases of $k = 2$ and $k = 3$.

The case $k = 2$:
In this case we need to verify if $T_1 = -T_2$. Since the ring $\mathbb{F}[x_1, \cdots, x_n]$ is a unique factorization domain and linear forms are irreducible elements in $\mathbb{F}[x_1, \cdots, x_n]$, therefore, $T_1, -T_2$ are equal if and only if there is a one-one correspondence between the linear forms on the LHS and the linear forms on the RHS and the coefficient of any one monomial occurring on the LHS equals the coefficient of that monomial on the RHS. All this can easily be checked in deterministic polynomial time. This solves the case $k = 2$.

The case $k = 3$:
By discarding the linear forms common to all the terms we can assume that T_1, T_2 and T_3 are coprime. Let,

$$L \subseteq \{L_{i,j} \mid 1 \leq i \leq 3, 1 \leq j \leq d\}$$

be the set of all distinct (up to constant multiples) linear forms occurring in the terms T_1, T_2 and T_3. We accept if and only if:

$$\forall \ell \in L, \quad \mathcal{C} = 0 \pmod{\ell}$$

Note that the ring $\mathbb{F}[x_1, \cdots, x_n]/(\ell)$ is isomorphic to the polynomial ring $\mathbb{F}[x_1, \cdots, x_{n-1}]$ and hence is also a unique factorization domain. Moreover, assuming (wlog) that ℓ occurs in T_1 we have:

$$\mathcal{C} = T_2 + T_3 \pmod{\ell}$$

Thus verification of $\mathcal{C} = 0 \pmod{\ell}$ boils down to the case $k = 2$. Now let us see what happens if $\mathcal{C} = 0$ modulo every $\ell \in L$:

$$\begin{aligned} & \forall \ell \in L, \quad \mathcal{C} = 0 \pmod{\ell} \\ \Rightarrow \quad & \mathcal{C} = 0 \pmod{\prod_{\ell \in L} \ell} \end{aligned}$$

Now if $\#L > d$ then clearly, $\mathcal{C} = 0$. If $\#L \leq d$ then $\mathcal{C} = T_1 + T_2 + T_3 \neq 0$ by the ABC theorem for polynomials [Sto81, Mas84]. This gives us a deterministic polynomial time algorithm for $k = 3$.

Unfortunately, the ABC theorem for polynomials does not extend in the desired way to sums of more than 3 terms (see [Pal93]). In order to get an algorithm for larger values of k we need to generalize the above approach and go modulo products of linear forms.

4.3.1 A special case of Ring Isomorphism

The problem of checking whether a polynomial $f(z_1, \ldots, z_n)$ is the zero polynomial over $\mathbb{F}$ can be viewed as a ring isomorphism problem since:

Claim 4.1.1 *$f(z_1, \ldots, z_n) = 0$ iff $\mathbb{F}[z_1, \ldots, z_n]/(f) \cong \mathbb{F}[z_1, \ldots, z_n]$*

Proof of Claim 4.1.1. The forward direction is easy to see.

Conversely, suppose τ is an isomorphism from $\mathbb{F}[z_1, \ldots, z_n]/(f)$ to $\mathbb{F}[z_1, \ldots, z_n]$. Thus, $\tau(f)$ is being mapped to the zero of $\mathbb{F}[z_1, \ldots, z_n]$ implying:

$$0 = \tau(f) = f(\tau(z_1), \ldots, \tau(z_n))$$

Now since f is a polynomial and $\tau(z_1), \ldots, \tau(z_n)$ are all algebraically independent transcendentals over $\mathbb{F}$ we deduce that $f = 0$. □

We aim to "solve" this version of ring isomorphism problem when f is of a special form, i.e., f is a sum of bounded-many products of linear forms or in other words f is a $\Sigma\Pi\Sigma$ circuit of bounded top fanin. But first we need some definitions and a lemma which is basically Chinese remaindering over local commutative rings.

In this section we will consider local commutative rings R of the form:

$$R = \mathbb{F}[x_1, \ldots, x_k]/(x_1^{e_1}, \ldots, x_k^{e_k}, h_1(x_1, \ldots, x_k), \ldots, h_\ell(x_1, \ldots, x_k)) \tag{4.2}$$

This ring has a unique maximal ideal $\mathcal{M}$ of nilpotents such that: $R/\mathcal{M} \cong \mathbb{F}$ (refer to Lemma A.4 in the appendix). Every element of R is of the form $(a + \alpha)$, where $a \in \mathbb{F}$ and $\alpha \in \mathcal{M}$. Moreover, there is a natural onto ring homomorphism $\phi : R \to \mathbb{F}$ such that $\phi : (a + \alpha) \mapsto a$ and thus having $\mathcal{M}$ as its kernel.

Example Let $R := \mathbb{F}[x, y]/(x^2, y(y + x))$. The elements of R look like: $a + bx + cy + dxy$. Note that in the ring R: $y^3 = -xy^2 = -x(-xy) = x^2y = 0$. Thus, both x, y are nilpotents and hence R is a local ring with $\mathcal{M} = (x, y)$ as its maximal ideal (see Lemma A.4 in the appendix).

The map ϕ, that sends $\mathcal{M}$ to zero and fixes $\mathbb{F}$, is a ring homomorphism from R to $\mathbb{F}$. Consider a polynomial $f(z) := 2z^2 + xyz + 1 \in R[z]$ then ϕ can be defined to act on f as:

$$\phi(f)(z) = \phi(2)z^2 + \phi(xy)z + \phi(1) = 2z^2 + 1$$

∎

Lemma 4.3 *Let R be a local commutative ring (as mentioned in Equation (4.2)) and $f(z_1, \ldots, z_n)$ be a polynomial living in $R[z_1, \ldots, z_n]$ of total degree d. Let $\phi : R \to \mathbb{F}$ be the natural onto ring homomorphism of R with kernel $\mathcal{M}$. Let $f_1, \ldots, f_m \in R[z_1, \ldots, z_n]$ be polynomials such that:*

- *$\phi(f_1), \ldots, \phi(f_m)$ are mutually coprime polynomials over $\mathbb{F}$.*
- *total degree of $(\phi(f_1) \cdots \phi(f_m)) > d$.*

Then,

$$R[z_1, \ldots, z_n]/(f) \cong R[z_1, \ldots, z_n] \tag{4.3}$$
$$\textit{iff}$$
$$R[z_1, \ldots, z_n]/(f, f_i) \cong R[z_1, \ldots, z_n]/(f_i), \quad \textit{for all } i \in [m]$$

Proof: Clearly, if $R[z_1, \ldots, z_n]/(f) \cong R[z_1, \ldots, z_n]$ then for all $i \in [m]$, $R[z_1, \ldots, z_n]/(f, f_i)$ is isomorphic to $R[z_1, \ldots, z_n]/(f_i)$. So the more interesting part is the converse.

Suppose for all $i \in [m]$, $R[z_1, \ldots, z_n]/(f, f_i) \cong R[z_1, \ldots, z_n]/(f_i)$. Note that if we denote the second ring by R'_i then the first ring can be viewed as: $R'_i/(f)$ where, (f) is being considered as an ideal of R'_i (or equivalently $(f) = fR'_i$). Now notice that $R'_i/(f) \cong R'_i$ iff f is zero in $R'_i = R[z_1, \ldots, z_n]/(f_i)$ which in turn happens iff $f_i \mid f$ over R. Thus, for all $i \in [m]$:

$$R[z_1, \ldots, z_n]/(f, f_i) \cong R[z_1, \ldots, z_n]/(f_i) \iff f_i \text{ divides } f \text{ over } R$$

Now what can we say about f if for all $i \in [m]$, $f_i \mid f$ over R? We answer this question by the following two claims. The first one says that $(f_1 \cdots f_m) \mid f$ over R.

Claim 4.1.2 (Kayal) *$p, g, h \in R[z_1, z_2, \cdots, z_n]$ be multivariate polynomials such that $\phi(g)$ and $\phi(h)$ are coprime. Moreover,*

$$p \equiv 0 \pmod{g}$$
$$p \equiv 0 \pmod{h}$$

Then $p \equiv 0 \pmod{g \cdot h}$.

Proof of Claim 4.1.2. We reproduce the following proof from [Kay06].

Recall that the unique maximal ideal of R is $\mathcal{M}$, $\phi : R \to \mathbb{F}$ is the natural onto ring homomorphism of R with kernel $\mathcal{M}$ and let t be the least integer such that $\mathcal{M}^t = 0$ in R. Let the (total) degrees of $\phi(g)$ and $\phi(h)$ be d_g and d_h respectively. Then by applying a suitable invertible linear transformation on the variables $z_1, \cdots, z_n$, if needed, we can assume without loss of generality that the coefficients of $z_n^{d_g}$ in g and that of $z_n^{d_h}$ in h are both units of R (see Lemma A.8). Consequently, in the product $g \cdot h$ the coefficient of $z_n^{d_g+d_h}$ is also a unit of R.

Now think of g and h as polynomials in one variable z_n with coefficients coming from the ring of fractions – $R(z_1, z_2, \ldots, z_{n-1})$ – of $R[z_1, \ldots, z_{n-1}]$. Now since $\phi(g)$ and $\phi(h)$ are coprime over $\mathbb{F}$, they are also coprime as univariate polynomials

in z_n over the function field $\mathbb{F}(z_1, \ldots, z_{n-1})$. Consequently, there exists $a, b \in \mathbb{F}(z_1, \ldots, z_{n-1})$ such that:

$$a\phi(g) + b\phi(h) = 1 \text{ over } \mathbb{F}(z_1, \ldots, z_{n-1})$$

That is, $ag + bh = 1$ in $(R/\mathcal{M})(z_1, \ldots, z_{n-1})$ (since $R/\mathcal{M} \cong \mathbb{F}$). By the well known Hensel's lifting lemma (see Lemma A.9) we get that there exist $a^*, b^* \in R(z_1, \ldots, z_{n-1})$ such that:

$$a^*g + b^*h = 1 \text{ over } (R/\mathcal{M}^t)(z_1, \ldots, z_{n-1}) \text{ which is } R(z_1, \ldots, z_{n-1}).$$

Now by the initial hypothesis:

$$\begin{aligned}
& p \equiv 0 && \pmod{g} \\
\Rightarrow\quad & p = qg && \text{for some } q \text{ in } R[z_1, \ldots, z_{n-1}][z_n] \\
\textit{also},\quad & p \equiv 0 && \pmod{h} \\
\Rightarrow\quad & qg \equiv 0 && \pmod{h} \\
\Rightarrow\quad & a^*qg \equiv 0 && \pmod{h} \text{ in } R(z_1, \ldots, z_{n-1})[z_n] \\
\Rightarrow\quad & q \equiv 0 && \pmod{h} \text{ in } R(z_1, \ldots, z_{n-1})[z_n] \\
\therefore\quad & p = ghq' && \text{for some } q' \text{ in } R(z_1, \ldots, z_{n-1})[z_n]
\end{aligned}$$

Since, the leading coefficient of z_n in gh is in R^* and p is in $R[z_1, \ldots, z_{n-1}][z_n]$, therefore by Gauss' lemma (see Lemma A.10) we get that in fact $a' \in R[z_1, \ldots, z_{n-1}][z_n]$ and so:

$$p \equiv 0 \pmod{gh} \text{ in } R[z_1, \ldots, z_n]$$

□

Since, by the hypothesis, $\phi(f_1), \ldots, \phi(f_m)$ are mutually coprime polynomials over $\mathbb{F}$, we repeatedly apply the above claim and deduce that:

$$\text{The polynomial } (f_1 \cdots f_m) \text{ divides } f \text{ over } R.$$

Notice that the total degree of $(f_1 \cdots f_m)$ is larger than that of f. The next claim shows that this means f is the zero polynomial over R.

Claim 4.1.3 *Suppose that $p, g \in R[z_1, \ldots, z_n]$ and p has total degree d_p. Moreover, g has total degree $d_g > d_p$ and contains at least one monomial of degree d_g whose coefficient is a unit in R. Then, $p \equiv 0 \ (mod\ g) \Rightarrow\ p = 0$ in $R[z_1, \ldots, z_n]$.*

Proof of Claim 4.1.3. Since $p \equiv 0 \pmod{g}$ over R we have:

$$p = qg \text{ for some } q \in R[z_1, \ldots, z_n]$$

By applying a suitable invertible linear transformation on the variables $z_1, \ldots, z_n$, if needed, we can assume that the coefficient of $z_n^{d_g}$ in g is a unit of R (see Lemma A.8). Now view p, g, q as univariate polynomials in z_n over the ring $R[z_1, \ldots, z_{n-1}]$ and let the degree of q with respect to z_n be $d_q > 0$. Then the coefficient of $z_n^{d_q+d_g}$ on the RHS is nonzero whereas all the terms on the LHS have degree at most $d_p < d_q + d_g$, a contradiction. This means that $d_q = 0$ and hence, $p = 0$ over R. □

By the hypothesis we have that the total degree of $(\phi(f_1) \cdots \phi(f_m)) > d$. Thus, $(f_1 \cdots f_m)$ has a monomial of degree larger than d whose coefficient is a unit of R. Thus, the above claim together with $(f_1 \cdots f_m) \mid f$ implies that $f \equiv 0$ over R, implying that:

$$R[z_1, \ldots, z_n]/(f) \cong R[z_1, \ldots, z_n]$$

This completes the proof of our lemma. ∎

4.3.2 Description of the Algorithm

In this section we sketch an algorithm for solving the special case of ring isomorphism problem (as occurred in the Equation (4.3)) when f is a $\Sigma\Pi\Sigma$ circuit of bounded top fanin. This section is dedicated to proving the following main theorem:

Theorem 4.2 *Let R be a local commutative ring over $\mathbb{F}$ (as mentioned in Equation (4.2)) with a unique maximal ideal $\mathcal{M}$ of nilpotents. Suppose $f \in R[z_1, \ldots, z_n]$ is the given polynomial. f is a sum of product of linear functions, i.e.,*

$$f = T_1 + T_2 + \cdots + T_k$$

where, each T_i is a product of $d_i \geq 1$ linear functions:

$$T_i = L_{i,1} \cdots L_{i,d_i}$$

where, each linear function $L_{i,j}$ looks like:

$$L_{i,j} = a_{i,j,0} + a_{i,j,1}z_1 + \cdots + a_{i,j,n}z_n, \quad a_{i,j,0} \in \mathcal{M} \text{ while } a_{i,j,1}, \cdots, a_{i,j,n} \in \mathbb{F}$$

Define $d := max_{1 \leq i \leq k}\{d_i\}$. Then the ring isomorphism problem:

$$R[z_1, \ldots, z_n]/(f) \stackrel{?}{\cong} R[z_1, \ldots, z_n]$$

can be solved in time $poly(d^k, n)$ assuming that the ring operations of R take constant time.

Proof: Recall that the unique maximal ideal of R is $\mathcal{M}$, $\phi : R \to \mathbb{F}$ is the natural onto ring homomorphism of R with kernel $\mathcal{M}$ and let t be the least integer such that $\mathcal{M}^t = 0$ in R.

If $k = 1$ then $f = T_1 = L_{1,1} \cdots L_{1,d_1}$ which is just a product of linear functions. Apply ϕ on these linear functions. Now if for all $j \in [d]$, $\phi(L_{1,j}) \neq 0$ then clearly $f \neq 0$ over R and, thus, $R[z_1, \ldots, z_n]/(f)$ is not isomorphic to $R[z_1, \ldots, z_n]$ by Claim 4.1.1. So assume that there is at least one $j \in [d_1]$ such that $\phi(L_{1,j}) = 0$ which means that the linear function $L_{1,j}$ has no z-term and it is simply equal to $a_{1,j,0}$. Collect all such linear functions as:

$$\ell_1 := \{j \in [d_1] \mid \phi(L_{1,j}) = 0\}$$

Thus,

$$f = \left(\prod_{j \in \ell_1} a_{1,j,0}\right) \cdot \left(\prod_{j \in [d_1] \setminus \ell_1} L_{1,j}\right)$$

and it is easy to see that $f = 0$ iff $\left(\prod_{j \in \ell_1} a_{1,j,0}\right) = 0$ in R. Thus, we can solve:

$$R[z_1, \ldots, z_n]/(T_1) \stackrel{?}{\cong} R[z_1, \ldots, z_n]$$

in time: $poly(d, n)$ assuming that the ring operations of R take constant time.

For larger k, we give a recursive algorithm using Lemma 4.3. As in the lemma we need to collect suitable polynomials $f_1, \ldots, f_m$ such that $\phi(f_1), \ldots, \phi(f_m)$ are coprime and the total degree of $\phi(f_1) \cdots \phi(f_m)$ is greater than the total degree of f.

Form the *largest* set $S := \{s_1, \ldots, s_m\}$ of linear forms in $R[z_1, \ldots, z_n]$ such that the elements of S satisfy:

- for each $i \in [m]$ there is a $j \in [k]$ such that $(s_i + r)$ is a linear factor of T_j for some $r \in \mathcal{M}$.
- for every $i \neq j \in [m]$, s_i, s_j are coprime over $\mathbb{F}$.

If S is empty then it means that for all $i \in [k], j \in [d_i]$, $L_{i,j} = a_{i,j,0}$ and hence we can easily compute f using just ring operations of R. Thus, assume that S is nonempty. For each $i \in [m]$, let $e_i \in [d]$ be the largest number such that $(s_i + r_{i,1}), \ldots, (s_i + r_{i,e_i})$, for some $r_{i,1}, \ldots, r_{i,e_i} \in \mathcal{M}$, are linear factors (with repetition) of some T_j, say T_{π_i}. The way we have defined e_i's we have that for any $j \in [k]$, the number of linear functions $L_{j,*}$ whose ϕ-image is s_i is at most e_i. Thus, we get the following bound:

$$(e_1 + \ldots + e_m) \geq d$$

If we define: for all $i \in [m]$, $f_i := (s_i + r_{i,1}) \cdots (s_i + r_{i,e_i})$ then one of the conditions of Lemma 4.3 is satisfied as $\phi(f_1), \ldots, \phi(f_m)$ are coprime. But what about the second condition: is the total degree of $\phi(f_1) \cdots \phi(f_m)$ larger than the total degree of f ? This is satisfied too when $(e_1 + \cdots + e_m) > d$. So we just need to handle the case: $(e_1 + \cdots + e_m) = d$.

If $(e_1 + \cdots + e_m) = d$ then form the set $U = \{T_j \mid \text{ total degree of } T_j \text{ is } d\}$. U is nonempty. Wlog let $U = \{T_1, \ldots, T_{k'}\}$. Then, for $i \in [k']$, T_i looks like:

$$T_i = \lambda_i \cdot \left(\prod_{j=1}^{e_1} (s_1 + \alpha_{i,1,j}) \right) \cdots \left(\prod_{j=1}^{e_m} (s_m + \alpha_{i,m,j}) \right)$$

where, for all $i_1 \in [k']$, $i_2 \in [m]$, $i_3 \in [e_{i_2}]$, α_{i_1,i_2,i_3}, $\lambda_{i_1} \in \mathcal{M}$. Note that the coefficient of any degree d monomial (in the variables $z_1, \ldots, z_n$) in $T_1 + \ldots + T_{k'}$ is a multiple (in $\mathbb{F}$) of:

$$\sum_{i \in [k']} \lambda_i$$

By the definition of d, this means that the coefficient of any degree d monomial in f is a multiple (in $\mathbb{F}$) of: $\sum_{i\in[k']}\lambda_i$. This can be computed using operations in R. If it is nonzero then $f \neq 0$ over R and, thus, $R[z_1,\ldots,z_n]/(f)$ is not isomorphic to $R[z_1,\ldots,z_n]$ by Claim 4.1.1. So assume that $\sum_{i\in[k']}\lambda_i = 0$ and hence the total degree of f is smaller than $d = (e_1+\cdots+e_m) =$ total degree of $(\phi(f_1)\cdots\phi(f_m))$.

Thus, $f_1,\ldots,f_m$ satisfy both the conditions of the Lemma 4.3. Before invoking the lemma we try to optimise and choose the largest $m' \in [m]$ such that:

$$d \leq \text{ total degree of } \phi(f_1)\cdots\phi(f_{m'}) \leq 2d \tag{4.4}$$

Now we apply the Lemma 4.3 on the polynomials $f_1,\ldots,f_{m'}$ and reduce the problem:

$$R[z_1,\ldots,z_n]/(f) \stackrel{?}{\cong} R[z_1,\ldots,z_n]$$

to m' smaller problems:

$$R[z_1,\ldots,z_n]/(f,f_i) \stackrel{?}{\cong} R[z_1,\ldots,z_n]/(f_i) \quad \text{for all } i \in [m'] \tag{4.5}$$

Why are the above problems smaller? First of all observe that f_i divides T_{π_i} and, hence,

$$f \equiv \sum_{j\in[k]\setminus\{\pi_i\}} T_j \pmod{f_i}$$

Next, recall: $f_i = (s_i+r_{i,1})\cdots(s_i+r_{i,e_i})$, where, s_i is a linear form in $R[z_1,\ldots,z_n]$ and $r_{i,1},\ldots,r_{i,e_i} \in \mathcal{M}$. Now if we apply a transformation τ_i on the variables $z_1,\ldots,z_n$ such that:

- τ_i maps each z_j to some linear combination $\sum_{\ell=1}^{n} a_{j,\ell}z_\ell$, where, $a_{j,\ell} \in \mathbb{F}$.
- τ_i is invertible.
- $\tau_i(s_i) = z_1$.

Then f_i transforms to a more amenable: $\tau_i(f_i) = (z_1 + r_{i,1}) \cdots (z_1 + r_{i,e_i})$, and:

$$R[z_1, \ldots, z_n]/(f, f_i) \cong R[z_1, \ldots, z_n]/(f_i)$$

iff

$$R[z_1, \ldots, z_n]/\left(\tau_i(f), (z_1 + r_{i,1}) \cdots (z_1 + r_{i,e_i})\right) \cong R[z_1, \ldots, z_n]/\left((z_1 + r_{i,1}) \cdots (z_1 + r_{i,e_i})\right)$$

iff

$$R_i[z_2, \ldots, z_n]/\left(\sum_{j \in [k] \setminus \{\pi_i\}} \tau_i(T_j)\right) \cong R_i[z_2, \ldots, z_n] \tag{4.6}$$

where,

$$R_i := R[z_1]/\left((z_1 + r_{i,1}) \cdots (z_1 + r_{i,e_i})\right) \tag{4.7}$$

This new ring R_i is also a local ring (see Lemma A.5 in the appendix) and since τ_i is a $\mathbb{F}$-linear transformation, $\sum_{j \in [k] \setminus \{\pi_i\}} \tau_i(T_j)$ is again a $\Sigma\Pi\Sigma$ circuit but with top fanin equal to $(k-1)$. Thus, Equation (4.6) is a smaller instance of the starting problem and can be recursively solved. However, there is a point to be taken care of: k reduces in the recursion but the ring R increases by one 'dimension' in every recursive call. So R can increase to at most k 'dimensions' till the recursion reaches the base step $k = 1$ and computations in that large a ring can be done in time $poly(d^k)$ (by using the special form of R_i, see Lemma A.5 in the appendix).

Let $time(k)$ be the time taken to solve the given ring isomorphism problem when f is of top fanin k. We get an easy recursive equation for $time(\cdot)$:

$$\begin{aligned} time(k) &\leq m' \cdot time(k-1) + (\text{computation-time in the intermediate base rings}) \\ &\leq m' \cdot time(k-1) + poly(d^k, n) \\ &\leq 2d \cdot time(k-1) + poly(d^k, n) \qquad [\text{by Equation (4.4)}] \\ \Rightarrow time(k) &= poly(d^k, n). \end{aligned}$$

This completes the proof of the theorem. ∎

Corollary 4.1 *Identity testing for $\Sigma\Pi\Sigma$ circuits $\mathcal{C} \in \mathbb{F}[x_1, \ldots, x_n]$, having top fanin equal to k, can be done in time: $poly(d^k, n)$ assuming that the field operations of $\mathbb{F}$ take constant time.*

Proof: This follows directly, if we put $R = \mathbb{F}$ in the statement of the above theorem and apply Claim 4.1.1. ∎

4.4 Discussion

This chapter considered the problem of identity testing for $\Sigma\Pi\Sigma$ arithmetic circuits $\mathcal{C}$. Suppose $\mathcal{C}(x_1, \ldots, x_n)$ has at most k inputs to the top addition gate and at most d inputs to the multiplication gate. Then we gave an identity test for such a circuit that works in time $poly(d^k, n)$. The machinery we used was that of local rings and a special case of their isomorphism problem. This chapter also gave examples of bounded top fanin $\Sigma\Pi\Sigma$ circuit identities, over any fixed field of prime characteristic, that have "high" rank. Are there identities of this kind over fields of characteristic 0, say $\mathbb{Q}$?

The problem of identity testing for general $\Sigma\Pi\Sigma$ arithmetic circuits remains open. It would be interesting to see if this method can be generalized for $\Sigma\Pi\Sigma\Pi$ circuits where the fanin of the topmost addition gate is bounded.

Chapter 5

Primality Testing

Primality testing – given a number test if it is prime – is one of the fundamental problems concerning numbers. Starting from ancient Chinese and Greek, many have worked on the problem of finding an efficient algorithm for primality testing. In recent times this problem has become more important from a practical perspective because of its applications in cryptography. For example, the widely used RSA public-key cryptosystem does computations modulo n, where, $n = pq$ for suitably chosen primes p and q.

An unconditional, deterministic, polynomial-time algorithm for primality testing was given for the first time in 2002 by Agrawal, Kayal and Saxena [AKS02]. In the months following the discovery new variants appeared (Lenstra 2002, Pomerance 2002, Berrizbeitia [Berr03], Cheng [Chen03], Bernstein [Bern], Lenstra and Pomerance [LP03], [AKS04]). All these algorithms are sometimes called AKS-type algorithms (see a nice survey by Granville [Gran]). The basic idea of the primality test is to give a characterization of prime numbers via cyclotomic rings $R := (\mathbb{Z}/n\mathbb{Z})[x]/(x^r - 1)$. We study the Frobenius-type map $\sigma_n : a(x) \mapsto a(x)^n$ and ask the question: when is σ_n an automorphism of R? It turns out that for a suitable r, $\sigma_n \in Aut(R)$ iff n is prime and more importantly, it is sufficient to test σ_n on a 'few' elements of R for automorphism.

The results of this chapter mostly appear in [AKS02, AS05].

5.1 Previous Work

The Sieve of Eratosthenes (ca. 240 BC) is the most ancient algorithm that works correctly for all primes, however, its time complexity (= $\Omega(n)$ where n is the input number) is exponential in the size of input. In the 17^{th} Century, Fermat proved what is referred as *Fermat's Little Theorem* stating that for any prime number p, and any number a not divisible by p, $a^{p-1} = 1 \pmod{p}$. Although the converse of this theorem does not hold (and in fact fails spectacularly for *Carmichael numbers*), this result has been the starting point for several efficient primality testing algorithms. In 1976, Miller [Mil76] used this property to obtain a deterministic polynomial-time algorithm for primality testing assuming *Extended Riemann Hypothesis (ERH)*. His test was modified by Rabin [Rab80] to yield an unconditional but randomized polynomial-time algorithm.

If we take the "square-root" of Fermat's congruence then we get: $a^{\frac{p-1}{2}} = \pm 1 \pmod{p}$. It turns out that the sign here is positive iff a is a square modulo p. This fact is usually stated in terms of Legendre symbol $\left(\frac{a}{p}\right)$ as:

$$a^{\frac{p-1}{2}} = \left(\frac{a}{p}\right) \pmod{p}$$

There is a generalization of Legendre symbol, over composite numbers n, called Jacobi symbol :

$$\left(\frac{a}{n}\right) := \prod_{i=1}^{k} \left(\frac{a}{p_i}\right) \quad \text{where, } n \text{ factors into primes as } n = \prod_{i=1}^{k} p_i.$$

It is an interesting fact that given a, n we do not know how to factor n but still we can compute $\left(\frac{a}{n}\right)$ by using Gauss' Reciprocity Law and Euclidean gcd-type algorithm (see [BS96]). Thus, the congruence: $a^{\frac{n-1}{2}} = \left(\frac{a}{n}\right) \pmod{n}$ is a candidate for a primality test and in fact was first used by Solovay and Strassen [SoS77] to design a randomized polynomial-time algorithm. Their algorithm can also be derandomized under ERH.

In 1983, Adleman, Pomerance, and Rumely [APR83] achieved a major breakthrough by giving a deterministic algorithm for primality that runs in $(\log n)^{O(\log\log\log n)}$ time (all the previous deterministic algorithms required exponential time). The

algorithm is based on an analytic number theory estimate stating that there is always an integer $m < (\log n)^{\log\log\log n}$ for which:

$$\prod_{\substack{\text{prime } q \\ (q-1)|m}} q \geq \sqrt{n}$$

In 1986, Goldwasser and Kilian [GK86] proposed a randomized algorithm based on Elliptic curves running in expected polynomial-time on almost all inputs (*all* inputs under a widely believed hypothesis) that produces a certificate for primality (until then, all randomized algorithms produced certificates for compositeness only). A similar algorithm was developed by Atkin [Atk86]. Adleman and Huang [AH92] modified Goldwasser-Kilian algorithm to obtain a randomized polynomial-time algorithm that always produced a certificate for primality.

5.2 The Beginning

Suppose p is a prime number and consider the ring $R_0 := \mathbb{Z}/p\mathbb{Z}$. Note that by Fermat's little theorem the map $\sigma_p : a \mapsto a^p$ is an automorphism of R_0. This exponentiation-map σ_p is called the *Frobenius map*. Is the Frobenius map σ_n an automorphism of the ring $\mathbb{Z}/n\mathbb{Z}$ for composite n? Note at this point that $Aut(\mathbb{Z}/n\mathbb{Z}) = \{id\}$ simply because 1 is the additive generator of the ring $\mathbb{Z}/n\mathbb{Z}$ and any automorphism fixes it.

Lemma 5.1 (Carmichael) *σ_n is an automorphism of the ring $(\mathbb{Z}/n\mathbb{Z})$ iff n is square-free and for every prime $p \mid n$, $(p-1) \mid (n-1)$.*

Proof: Suppose $\sigma_n \in Aut(\mathbb{Z}/n\mathbb{Z})$ and $n = p^s \cdot t$ where, p is some prime and $gcd(p,t) = 1$. We have the following ring decomposition:

$$\mathbb{Z}/n\mathbb{Z} \cong (\mathbb{Z}/p^s\mathbb{Z}) \times (\mathbb{Z}/t\mathbb{Z})$$

Thus, $\sigma_n \in Aut(\mathbb{Z}/n\mathbb{Z})$ implies that $\sigma_n \in Aut(\mathbb{Z}/p^s\mathbb{Z})$.

First we show that $s = 1$. Suppose $s \geq 2$. If $p = 2$ then $\sigma_n(-1) = 1 \pmod{p^s}$ while $-1 \neq 1 \pmod{p^s}$ which contradicts $\sigma_n \in Aut(\mathbb{Z}/p^s\mathbb{Z})$. On the other hand,

if $p \neq 2$ then $(\mathbb{Z}/p^s\mathbb{Z})^*$ is a cyclic group of size $p^{s-1}(p-1)$ (see Lemma A.6 in the appendix) which has a nontrivial gcd with n and hence σ_n cannot be injective on the cyclic group: $(\mathbb{Z}/p^s\mathbb{Z})^*$, again contradicting $\sigma_n \in Aut(\mathbb{Z}/p^s\mathbb{Z})$. These contradictions force $s = 1$ implying that n is square-free and:

$$\mathbb{Z}/n\mathbb{Z} \cong \times_{\text{prime } p|n} (\mathbb{Z}/p\mathbb{Z}) \tag{5.1}$$

Now $\sigma_n \in Aut(\mathbb{Z}/n\mathbb{Z})$ iff for all prime $p \mid n$, $\sigma_n \in Aut(\mathbb{Z}/p\mathbb{Z})$. But the ring $\mathbb{Z}/p\mathbb{Z}$ has only trivial automorphism implying that for a generator g of the group $(\mathbb{Z}/p\mathbb{Z})^*$: $\sigma_n(g) = g \pmod p \;\Rightarrow\; g^{n-1} = 1 \pmod p \;\Rightarrow\; (p-1) \mid (n-1)$.

Conversely, suppose n is square-free and for every prime $p \mid n$, $(p-1) \mid (n-1)$. Thus, for all prime $p \mid n$ and for all $a \in \mathbb{Z}/p\mathbb{Z}$, $a^n = a \cdot a^{n-1} = a \pmod p$. This means that σ_n is identity on $\mathbb{Z}/p\mathbb{Z}$ for all $p \mid n$ which means by Equation (5.1) that σ_n is the trivial automorphism of the ring $\mathbb{Z}/n\mathbb{Z}$.

∎

Composite numbers n as in Lemma 5.1 are called Carmichael numbers [Car10] and they are infinitely many [AGP94]. Thus, Frobenius map σ_n being an automorphism of the ring $\mathbb{Z}/n\mathbb{Z}$ imposes some conditions on n but they are not strong enough to characterize primes. What if we consider "larger" rings over $\mathbb{Z}/n\mathbb{Z}$? It was suggested at the end of the paper [AB99] that cyclotomic rings over $\mathbb{Z}/n\mathbb{Z}$ might be useful.

Let n be the given odd number to be tested for primality. For any number r coprime to n define a ring $R_{n,r} := (\mathbb{Z}/n\mathbb{Z})[x]/(x^r - 1)$ called a *cyclotomic ring*. When is the Frobenius map $\sigma_n : a(x) \mapsto a(x)^n$ an automorphism of $R_{n,r}$? We attempt to answer this question both theoretically and algorithmically in the next sections. We show that interesting things happen when r is chosen suitably and eventually we get an efficient characterization of primes.

5.3 Cyclotomic Rings Characterize Primes

The central question that we explore in this section is: when is $\sigma_n \in Aut(R_{n,r})$? We show that there is a suitable $r \sim \log^3 n$ such that $\sigma_n \in Aut(R_{n,r})$ iff n is prime. Thus,

we can get a primality test if we can efficiently test whether $\sigma_n \in Aut(R_{n,r})$ for that r. In the first subsection we show how to do this in randomized polynomial time while in the second subsection we show an Extended Riemann Hypothesis (ERH) connection.

The size of $(\mathbb{Z}/r\mathbb{Z})^*$ is classically denoted by $\phi(r)$ and ϕ is called the *Euler's totient function*. It is easily seen that r is prime iff $\phi(r) = (r-1)$. Note that $(\mathbb{Z}/r\mathbb{Z})^*$ is a finite group and, hence, for any element $a \in (\mathbb{Z}/r\mathbb{Z})^*$, $a^{\phi(r)} = 1 \pmod{r}$. We will use the notation $\mathrm{o}_r(a)$ to denote the least nonzero positive integer m such that: $a^m = 1 \pmod{r}$. $\mathrm{o}_r(a)$ is called the *order* of a modulo r. It is a simple exercise to show that $\mathrm{o}_r(a) \mid \phi(r)$. For a natural number m we will use $P(m)$ to denote the largest prime factor of m.

Theorem 5.1 *Let n be an odd number. Let r be a prime (coprime to n) such that $P(o_r(n)) > \log n$. Define the cyclotomic ring $R_{n,r} := (\mathbb{Z}/n\mathbb{Z})[x]/(x^r - 1)$ and the Frobenius map $\sigma_n : a(x) \mapsto a(x)^n$. Then,*

$$n \text{ is prime}$$
$$\text{iff}$$
$$\sigma_n \in Aut(R_{n,r})$$

Proof: If n is a prime then for any polynomial $f(x)$, $f(x)^n = f(x^n) \pmod{n}$ (proof is simply by the multinomial expansion of $f(x)^n$) and hence $\sigma_n \in Aut(R_{n,r})$ for any r. So it is the converse that we intend to show next.

Note that since r and n are coprime we have that $(x-1)$ and $\frac{x^r-1}{x-1}$ are coprime polynomials modulo n. Thus,

$$R_{n,r} = (\mathbb{Z}/n\mathbb{Z})[x]/(x^r - 1) \cong (\mathbb{Z}/n\mathbb{Z})[x]/(x-1) \times (\mathbb{Z}/n\mathbb{Z})[x]/\left(\frac{x^r-1}{x-1}\right)$$

Now $\sigma_n \in Aut(R_{n,r})$ means that $\sigma_n \in Aut((\mathbb{Z}/n\mathbb{Z})[x]/(x-1)) = Aut(\mathbb{Z}/n\mathbb{Z})$ which by Lemma 5.1 means that n is a Carmichael number and hence is square-free. Thus, $R_{n,r}$ can be decomposed into fields as:

$$\begin{aligned}(\mathbb{Z}/n\mathbb{Z})[x]/(x^r-1) &\cong \times_{\text{prime } p|n}\ (\mathbb{Z}/p\mathbb{Z})[x]/(x^r-1) \\ &\cong \times_{\text{prime } p|n}\ (\mathbb{Z}/p\mathbb{Z})[x]/(x-1)\ \times_{\text{prime } p|n}\ (\mathbb{Z}/p\mathbb{Z})[x]/(x^{r-1}+\cdots+1) \\ &\cong \times_{\text{prime } p|n}\ \mathbb{F}_p\ \times_{\text{prime } p|n} \times_{i=1}^{\frac{r-1}{\mathrm{o}_r(p)}}\ \mathbb{F}_{p^{\mathrm{o}_r(p)}} \qquad (5.2)\end{aligned}$$

The last congruence follows from the fact that the polynomial $(x^{r-1} + x^{r-2} + \cdots + 1)$ factors over the field $\mathbb{F}_p$ into $\frac{r-1}{\mathrm{o}_r(p)}$ irreducible factors each of degree $\mathrm{o}_r(p)$ (see Lemma A.7). The above decomposition tells us that $\sigma_n \in Aut(R_{n,r})$ means that $\sigma_n \in Aut(\mathbb{F}_{p^{\mathrm{o}_r(p)}})$. But $Aut(\mathbb{F}_{p^{\mathrm{o}_r(p)}})$ is generated by σ_p (see Lemma A.7). Thus, there is an $0 \leq m < \mathrm{o}_r(p)$ such that $\sigma_n \equiv (\sigma_p)^m$ over $\mathbb{F}_{p^{\mathrm{o}_r(p)}}$. Since $\mathbb{F}^*_{p^{\mathrm{o}_r(p)}}$ is cyclic there is a generator element $g \in \mathbb{F}^*_{p^{\mathrm{o}_r(p)}}$ which has to satisfy $\sigma_n(g) = (\sigma_p)^m(g)$ implying that $g^{n-p^m} = 1$ which in turn means that $n = p^m \pmod{p^{\mathrm{o}_r(p)} - 1}$. Thus,

$$\forall \text{prime } p \mid n,\ \exists m, \quad n = p^m \pmod{p^{\mathrm{o}_r(p)} - 1}$$

If for some prime $p \mid n,\ p^{\mathrm{o}_r(p)}$ is larger than n then the above equation gives us that $n = p^m$ which means that $n = p$ as n is square-free. Now observe that:

$$\mathrm{o}_r(n) \mid \prod_{\text{prime } p \mid n} \mathrm{o}_r(p)$$

Thus, the prime $P(\mathrm{o}_r(n))$ divides $\mathrm{o}_r(p_0)$ for some prime $p_0 \mid n$ and hence $p_0^{\mathrm{o}_r(p_0)} > p_0^{P(\mathrm{o}_r(n))}$ which is larger than n. Thus, we conclude that n is a prime.

But why do we expect the existence of an r satisfying $P(\mathrm{o}_r(n)) > \log n$? The reason is an estimate of prime numbers proved by Tchebycheff. Here, we use a stronger estimate due to Fouvry [Fou85] as it gives a better result for our purposes.

Claim 5.1.1 *For a sufficiently large n there always exist an $r = O^{\sim}(\log^3 n)$ such that $P(o_r(n)) > \log n$.*

Proof of Claim 5.1.1. The two analytic number-theoretic estimates that are useful for us here are the Tchebycheff's and Fouvry's estimate. Tchebycheff [Apo97] showed that for all $x \geq 2$:

$$\frac{1}{5} \cdot \frac{x}{\log x} \leq |\{q \mid q \text{ is prime}, q \leq x\}| \leq 5 \cdot \frac{x}{\log x} \tag{5.3}$$

Fouvry [Fou85, BH96] showed a much stronger result about primes – there exist constants $c > 0$ and n_0 such that, for all $x \geq n_0$:

$$|\{q \mid q \text{ is prime}, q \leq x \text{ and } P(q-1) > q^{\frac{2}{3}}\}| \geq c \cdot \frac{x}{\log x} \tag{5.4}$$

which roughly means that the density of primes q, such that $(q-1)$ has a large prime factor, is $\Theta\left(\frac{x}{\log x}\right)$.

Now consider a possible sample space for r –

$$S := \{r \mid \text{ prime } r,\ \log^3 n(\log\log n)^3 \leq r \leq d\log^3 n(\log\log n)^3,\ P(r-1) > r^{\frac{2}{3}}\}$$

where, constant $d > 0$ will be fixed later. Note that it follows from the above estimates that $|S| \geq d' \log^3 n(\log\log n)^2$ for some constant $d' > 1$ (fix d suitably). For how many r's in S is $P(\mathrm{o}_r(n)) > r^{\frac{2}{3}}$? Note that if for some $r \in S$, $P(\mathrm{o}_r(n)) \leq r^{\frac{2}{3}}$ then $P(\mathrm{o}_r(n)) < r^{\frac{1}{3}}$ (since $P(r-1) > r^{\frac{2}{3}}$ and $\mathrm{o}_r(n) \mid (r-1)$). Thus, all the r's in S with $P(\mathrm{o}_r(n)) \leq r^{\frac{2}{3}}$ divide the product:

$$\Pi = (n-1)\cdot(n^2-1)\cdots(n^{r^{\frac{1}{3}}}-1) < n^{r^{\frac{2}{3}}}$$

Thus, such r's are at most $\log\Pi = r^{\frac{2}{3}}\log n$ in number. Note that $r^{\frac{2}{3}}\log n < d^{\frac{2}{3}}\log^3 n(\log\log n)^2 < |S|$ (fix d such that $d' > d^{\frac{2}{3}}$). Thus, there is a prime $r = O^{\sim}(\log^3 n)$ in S such that $P(\mathrm{o}_r(n)) > r^{\frac{2}{3}} > \log^2 n(\log\log n)^2$ which is better than what we desired! □ ∎

This theorem shows that cyclotomic rings do give a nice algebraic characterization of prime numbers. Our next desire is to use this to find an efficient primality test. The clue lies in studying the action of σ_n on the elements of $R_{n,r}$.

5.3.1 A Randomized Algorithm

Here we will present a simple randomized algorithm to check whether $\sigma_n \in Aut(R_{n,r})$ in time $poly(r, \log n)$. The surprising thing about it will be that by checking just *two* congruences of the form $a(x)^n = a(x^n) \pmod{n,\ x^r-1}$ we can gain confidence about whether σ_n satisfies them for *all* $a(x) \in R_{n,r}$.

Theorem 5.2 *Given coprime positive integers n, r. There is a randomized algorithm to check whether $\sigma_n \in Aut(R_{n,r})$ in time $poly(r, \log n)$.*

Proof: Recall from Lemma 5.1 that if $\sigma_n \in Aut(R_{n,r})$ then n has to be square-free. This necessary condition can be checked easily by doing Fermat's little test for

all $1 \leq a \leq 4\log^2 n$ (see Lemma A.11). Thus, let us assume from now on that the given n is square-free.

Recall Equation (5.2) and for clarity let the decomposition of ring $R_{n,r}$ be:

$$R_{n,r} \cong \mathbb{F}_{p_1^{d_1}} \times \cdots \times \mathbb{F}_{p_k^{d_k}}$$

Since size of $R_{n,r}$ is n^r and each $p_i \geq 2$ we have that $k, d_1, \ldots, d_k \leq r \log n$. Firstly, we can assume that all p_i are larger than $8r \log n$ for otherwise we can factor-out these "small" p_i from n and then $\sigma_n \in Aut(\mathbb{F}_{p_i^{d_i}})$ iff n is a power of p_i modulo $(p_i^{d_i} - 1)$ which can be easily checked in time $poly(r, \log n)$.

Consider the following set G of elements in $R_{n,r}$:

$$G := \left\{ a(x) \in R_{n,r}^* \mid \sigma_n(a(x)) = a(\sigma_n(x)) \right\}$$

Clearly, $1 \in G$. Also, it is easy to see that if $a(x),\ b(x) \in G$ then $a(x) \cdot b(x) \in G$. Thus, the set G is a subgroup of $R_{n,r}^*$. Can $G = R_{n,r}^*$? Note that if $G = R_{n,r}^* = \times_{i=1}^{k} (\mathbb{F}_{p_i^{d_i}})^*$ then σ_n is an automorphism of each $\mathbb{F}_{p_i^{d_i}}$ and hence $\sigma_n \in Aut(R_{n,r})$. On the other hand if $G \neq R_{n,r}^*$ then G is a proper subgroup and hence $\#G \leq \frac{1}{2}\#R_{n,r}^*$.

Now our randomized algorithm is simple:

1. Randomly and independently choose $a(x), b(x) \in R_{n,r}$.
2. Check whether: $\sigma_n(a(x)) = a(\sigma_n(x))$ and $\sigma_n(b(x)) = b(\sigma_n(x))$ in $R_{n,r}$.
3. Output YES iff both the above tests pass.

We will show that the probability of error of the above algorithm is $\leq \frac{1}{2}$. If $\sigma_n \in Aut(R_{n,r})$ then clearly the above algorithm returns YES. So assume that $\sigma_n \notin Aut(R_{n,r})$ and thus, by the above discussion: $\#G \leq \frac{1}{2}\#R_{n,r}^*$.

Firstly, note that $a(x)$ is not a unit of $R_{n,r}$ iff its image in at least one of the fields $\mathbb{F}_{p_i^{d_i}}$ is zero. Thus,

$$\begin{aligned} \text{Prob}_{a(x),b(x) \in R_{n,r}}[a(x) \text{ or } b(x) \text{ is not in } R_{n,r}^*] &\leq 2 \cdot \sum_{i=1}^{k} \frac{1}{p_i^{d_i}} \\ &\leq 2 \cdot (r \log n) \cdot \frac{1}{8r \log n} \\ &\leq \frac{1}{4} \end{aligned} \tag{5.5}$$

Also, $\text{Prob}_{a(x)\in R^*_{n,r}}[a(x) \in G] \leq \frac{1}{2}$. Thus,

$$\text{Prob}_{a(x),b(x)\in R^*_{n,r}}[\sigma_n(a(x)) = a(\sigma_n(x)) \text{ and } \sigma_n(b(x)) = b(\sigma_n(x)) \text{ in } R_{n,r}] \leq \frac{1}{4} \quad (5.6)$$

The above two probabilities (Equations (5.5) and (5.6)) together upper bound the probability of our algorithm saying YES when $\sigma_n \notin Aut(R_{n,r})$ by $\frac{1}{2}$.

It is routine to verify that all the steps of the above algorithm can be implemented in time $poly(r, \log n)$. ∎

5.3.2 Results assuming ERH

The Extended Riemann Hypothesis (ERH) is a longstanding open conjecture in complex analysis. Our interest in ERH arises from its following connection to number theory (see [BS96]): If ERH is true then the set of primes $p \leq 2\log^2 n$ multiplicatively generate the group $(\mathbb{Z}/n\mathbb{Z})^*$. Thus, under ERH it is easy to check whether $\sigma_n \in Aut(\mathbb{Z}/n\mathbb{Z})$ as checking: $a^n = a \pmod n$ for all $1 \leq a \leq 2\log^2 n$ suffices. But for larger r it is not clear how ERH helps in checking $\sigma_n \in Aut(R_{n,r})$ as it is not known yet whether $((\mathbb{Z}/n\mathbb{Z})[x]/(x^r - 1))^*$ has "small" generators (some results in that direction are given in [Shou92]).

We study in this section the properties that r and n would have to satisfy if $\sigma_n(a(x)) = a(\sigma_n(x))$ for some "special" $a(x) \in R_{n,r}$ and then invoke ERH to get a new primality test.

Lemma 5.2 *Let n be an odd number and r be an odd prime not dividing n. Then,*

$$\sigma_n(1-x) = (1-\sigma_n(x)) \text{ in the ring } R_{n,r} \Rightarrow r^{\frac{n-1}{2}} = \left(\frac{r}{n}\right) \pmod n$$

Proof: Let $B := 16^{-1} \pmod r$. If $\sigma_n(1-x) = (1-\sigma_n(x))$ in $R_{n,r}$ then:

$$(1-x)^n = (1-x^n) \pmod{n, Q_r(x)}, \quad \text{where } Q_r(x) := \frac{x^r-1}{x-1}$$

By applying the ring automorphism $x \mapsto x^i$ (where i, r are coprime) to the above equation we get:

$$
\begin{aligned}
& (1-x^i)^n = (1-x^{in}) \pmod{n, Q_r(x)} \\
\Rightarrow \quad & \left(x^B \prod_{i=1}^{\frac{r-1}{2}} (1-x^i) \right)^n = \left(x^{Bn} \prod_{i=1}^{\frac{r-1}{2}} (1-x^{in}) \right) \pmod{n, Q_r(x)} \\
\Rightarrow \quad & \left(x^B \prod_{i=1}^{\frac{r-1}{2}} (1-x^i) \right)^n = \left(\frac{n}{r}\right) \cdot \left(x^B \prod_{i=1}^{\frac{r-1}{2}} (1-x^i) \right) \pmod{n, Q_r(x)} \text{ [by Lemma } A.12] \\
\Rightarrow \quad & \left(x^B \prod_{i=1}^{\frac{r-1}{2}} (1-x^i) \right)^{n-1} = \left(\frac{n}{r}\right) \pmod{n, Q_r(x)} \\
\Rightarrow \quad & \left((-1)^{\frac{r-1}{2}} \cdot r \right)^{\frac{n-1}{2}} = \left(\frac{n}{r}\right) \pmod{n, Q_r(x)} \quad \text{[by Lemma } A.12] \\
\Rightarrow \quad & r^{\frac{n-1}{2}} = (-1)^{\frac{r-1}{2}\frac{n-1}{2}} \left(\frac{n}{r}\right) \pmod{n, Q_r(x)} \\
\Rightarrow \quad & r^{\frac{n-1}{2}} = \left(\frac{r}{n}\right) \pmod{n, Q_r(x)} \quad \text{[by Quadratic Reciprocity Lemma } A.13] \\
\Rightarrow \quad & r^{\frac{n-1}{2}} = \left(\frac{r}{n}\right) \pmod{n}
\end{aligned}
$$

∎

Remark: Thus, for a given n checking whether $\sigma_n(1-x) = (1-\sigma_n(x))$ in the ring $R_{n,r}$ is an algebraic version of Solovay-Strassen's primality test [SoS77] and hence can be derandomized under ERH to give a 'new' cyclotomic primality test. ∎

5.4 A Deterministic and Efficient Characterization of Primes

Theorem 5.1 showed us that the condition $\sigma_n \in Aut(R_{n,r})$ forces n to be prime if $P(o_r(n))$ is large enough. Also, in the previous section we saw that checking $\sigma_n(a(x)) = a(\sigma_n(x))$ for a couple of $a(x) \in R_{n,r}$ gives us information whether $\sigma_n \in$

$Aut(R_{n,r})$. We now try to combine these two ideas by making $P(o_r(n))$ larger and testing $\sigma_n(x+a) = (\sigma_n(x) + a)$ for various "small" a's. It turns out, as we prove below, that this gives us an unconditional, deterministic, polynomial-time primality test.

Theorem 5.3 *Let n be a positive integer. Fix an integer r of magnitude $O^{\sim}(\log^6 n)$ such that $r \geq (16 \log^2 n)$ and $P(o_r(n)) > \lceil \sqrt{r} \rceil \cdot \lceil \log n \rceil$. Suppose r, n are coprime and all prime factors of n are larger than $\lceil \sqrt{r} \rceil \cdot \lceil \log n \rceil$. Define the ring $R_{n,r} := (\mathbb{Z}/n\mathbb{Z})[x]/(x^r - 1)$. Then, the following are equivalent:*

(i) n is prime.

(ii) $\sigma_n \in Aut(R_{n,r})$.

(iii) $\sigma_n(x+a) = (\sigma_n(x) + a)$ in $R_{n,r}$, for all $1 \leq a \leq \lceil \sqrt{r} \rceil \cdot \lceil \log n \rceil$.

Moreover, the condition (iii) above gives a deterministic primality test that takes time: $O^{\sim}(\log^{12} n)$.

Proof: Let $\ell := \lceil \sqrt{r} \rceil \cdot \lceil \log n \rceil$. It is easy to see that (i) implies (ii) and (ii) implies (iii). So what we intend to show now is that (iii) implies (i).

Suppose $\sigma_n(x+a) = (\sigma_n(x) + a)$ in $R_{n,r}$, for all $1 \leq a \leq \ell$. Then firstly observe that:

$$\begin{aligned} & \sigma_n(x+a) = (\sigma_n(x) + a) \pmod{n, x-1} \quad \text{for all } 1 \leq a \leq 4\log^2 n \\ \Rightarrow \quad & (a+1)^n = (a+1) \pmod{n} \quad \text{for all } 1 \leq a \leq 4\log^2 n \end{aligned}$$

But then by Lemma A.11 the above tests tell us that n is square-free. As $\mathrm{o}_r(n) \mid \prod_{\text{prime } p \mid n} \mathrm{o}_r(p)$ we get that the prime $P(\mathrm{o}_r(n))$ divides $\mathrm{o}_r(p)$ for some prime $p \mid n$. Thus, there is a prime $p \mid n$ such that $P(\mathrm{o}_r(p)) > \ell$. We will now work modulo this prime p. Note that $(x^r - 1)$ modulo p has an irreducible factor $h(x)$ of degree $> \ell$ (see Lemma A.7).

Let $n = mp$ where $gcd(m, p) = 1$. Now we have, for all $1 \leq a \leq \ell$:

$$\begin{aligned}
& (x+a)^{mp} = (x^{mp} + a) \pmod{p, x^r - 1} \\
\Rightarrow \quad & (x^p + a)^m = (x^{mp} + a) \pmod{p, x^r - 1} \quad [\because (x+a)^p = (x^p + a) \pmod{p, x^r - 1}] \\
\Rightarrow \quad & (x+a)^m = (x^m + a) \pmod{p, x^r - 1} \quad [\text{send } x \mapsto x^{p^{-1} (\text{mod } r)} \text{ in the above eqn.}]
\end{aligned}$$

Next we observe that if positive integers m_1, m_2 satisfy $(x+a)^{m_1} = (x^{m_1} + a)$ and $(x+a)^{m_2} = (x^{m_2} + a)$ in $R_{n,r}$ then :

$$\begin{aligned}
(x+a)^{m_1 m_2} &= \{(x+a)^{m_1}\}^{m_2} \pmod{n, x^r - 1} \\
&= (x^{m_1} + a)^{m_2} \pmod{n, x^r - 1} \\
&= (x^{m_1 m_2} + a) \pmod{n, x^r - 1} \\
&\quad [\text{by sending } x \mapsto x^{m_1} \text{ in } (x+a)^{m_2} = (x^{m_2} + a) \pmod{n, x^r - 1}]
\end{aligned}$$

Since $(x+a)^m = (x^m + a) \pmod{p, x^r - 1}$ and $(x+a)^p = (x^p + a) \pmod{p, x^r - 1}$, thus, we obtain from the above observations that for any positive integers i, j and for all $1 \leq a \leq \ell$:

$$\sigma_{m^i p^j}(x+a) = (\sigma_{m^i p^j}(x) + a) \pmod{p, x^r - 1} \tag{5.7}$$

Consider the set $I := \{m^i p^j \mid 0 \leq i, j < \lceil \sqrt{r} \rceil\}$. Since m, p, r are mutually coprime, we have $\#I \geq r$ and hence, I has two distinct elements with equal residue modulo r. Let $m^{i_1} p^{j_1}, m^{i_2} p^{j_2} \in I$ be two such elements.

Consider another set $J := \{(x+1)^{e_1} \cdots (x+\ell)^{e_\ell} \mid e_1, \ldots, e_\ell \in \{0, 1\}\}$ of elements in $R_{n,r}$. Note that all these elements remain distinct even in the subring $\mathbb{F}_p[x]/(h(x))$ of $R_{n,r}$, simply because all polynomials in J are of degree $\leq \ell$ while $h(x)$ is of degree $> \ell$ and because by the hypothesis we have $p > \ell$.

Thus, a generator $g(x)$ of the cyclic subgroup of $(\mathbb{F}_p[x]/(h(x)))^*$ generated by J has order $\mathrm{o}_{(p,h(x))}(g(x)) \geq \#J \geq 2^\ell$.

Now by Equation (5.7) we have that:

$$\begin{aligned}
g(x)^{m^{i_1} p^{j_1}} &= g(x^{m^{i_1} p^{j_1}}) \pmod{p, h(x)} \\
&= g(x^{m^{i_2} p^{j_2}}) \pmod{p, h(x)} \quad [\because x^{m^{i_1} p^{j_1}} = x^{m^{i_2} p^{j_2}} \pmod{h(x)}] \\
&= g(x)^{m^{i_2} p^{j_2}} \pmod{p, h(x)}
\end{aligned}$$

The above means that $g(x)^{m^{i_1}p^{j_1}-m^{i_2}p^{j_2}} = 1 \pmod{p, h(x)}$. Thus,

$$m^{i_1}p^{j_1} \equiv m^{i_2}p^{j_2} \pmod{\mathrm{o}_{(p,h(x))}(g(x))} \tag{5.8}$$

But now observe that:

$$m^{i_1}p^{j_1},\ m^{i_2}p^{j_2} < m^{\lceil\sqrt{r}\rceil}p^{\lceil\sqrt{r}\rceil} = n^{\lceil\sqrt{r}\rceil}$$

while $\mathrm{o}_{(p,h(x))}(g(x)) \geq 2^{\ell} \geq n^{\lceil\sqrt{r}\rceil}$. This means that $m^{i_1}p^{j_1} = m^{i_2}p^{j_2}$. As $gcd(m,p) = 1$ this is only possible when either $m = 1$ or $(i_1, j_1) = (i_2, j_2)$. As the latter contradicts the choice of $(i_1, j_1), (i_2, j_2)$ the only possibility left is $m = 1$ which means $n = p$, a prime.

Let us now show that there is an r of magnitude $O^{\sim}(\log^6 n)$ such that $P(o_r(n)) > \lceil\sqrt{r}\rceil \cdot \lceil\log n\rceil$. Consider a possible sample space for r –

$$S := \{r \mid \text{ prime } r,\ \log^6 n(\log\log n) \leq r \leq d\log^6 n(\log\log n),\ P(r-1) > r^{\frac{2}{3}}\}$$

where, constant $d > 0$ will be fixed later. Note that it follows from the estimates of Equations (5.3) and (5.4) that $|S| \geq d'\log^6 n$ for some constant $d' > 1$ (fix d suitably). For how many r's in S is $P(\mathrm{o}_r(n)) > r^{\frac{2}{3}}$? Note that if for some $r \in S,\ P(\mathrm{o}_r(n)) \leq r^{\frac{2}{3}}$ then $P(\mathrm{o}_r(n)) < r^{\frac{1}{3}}$ (since $P(r-1) > r^{\frac{2}{3}}$ and $\mathrm{o}_r(n) \mid (r-1)$). Thus, all the r's in S with $P(\mathrm{o}_r(n)) \leq r^{\frac{2}{3}}$ divide the product:

$$\Pi = (n-1)\cdot(n^2-1)\cdots(n^{r^{\frac{1}{3}}}-1) < n^{r^{\frac{2}{3}}}$$

Thus, such r's are at most $\log\Pi = r^{\frac{2}{3}}\log n$ in number. Note that $r^{\frac{2}{3}}\log n < d^{\frac{2}{3}}\log^5 n(\log\log n)^{\frac{2}{3}} < |S|$. Thus, there is a prime $r = O^{\sim}(\log^6 n)$ in S such that: $P(\mathrm{o}_r(n)) > r^{\frac{2}{3}} > \lceil\sqrt{r}\rceil \cdot \lceil\log n\rceil$ (as $r \geq \log^6 n(\log\log n)$).

To estimate the time taken by the algorithm just observe that the most expensive step is to compute: $(x+a)^n \pmod{n, x^r - 1}$. This can be done in time $\log n \cdot O^{\sim}(r\log n)$ by Fast Fourier multiplication techniques (see [vzGG99]). Thus, the total time complexity is:

$$\sqrt{r}\log n \cdot O^{\sim}(r\log^2 n) = O^{\sim}(r^{\frac{3}{2}}\log^3 n) = O^{\sim}(\log^{12} n)$$

∎

5.5 Discussion

This chapter studied the automorphism group of the cyclotomic ring: $R_{n,r} := (\mathbb{Z}/n\mathbb{Z})[x]/(x^r - 1)$. The aspect of $Aut(R_{n,r})$ that we are especially interested in, is whether the n-th powering map $\sigma_n \in Aut(R_{n,r})$. We showed that when r is suitably chosen then σ_n is an automorphism of the ring $R_{n,r}$ iff n is prime. So the next question was how to check $\sigma_n \in Aut(R_{n,r})$ efficiently. On further studying the action of σ_n on the elements of $R_{n,r}$ and invoking an analytic number theoretic estimate it turned out that checking $\sigma_n(x + a) = (\sigma_n(x) + a)$ in $R_{n,r}$ for a suitable r and for a "few" a 's is sufficient to decide whether n is a prime. Thus, giving us a deterministic polynomial time primality test.

The complexity of the primality test that we give is $O^{\sim}(\log^{12} n)$. Lenstra and [AKS04] improved the algebraic arguments in the proof of the Theorem 5.3 to give a faster primality test that takes time $O^{\sim}(\log^{7.5} n)$. Note that there are two groups that vaguely appear in the proof of the Theorem 5.3: first group $G_1 := (m, p) \leq (\mathbb{Z}/r\mathbb{Z})^*$ that contains I and the second group $G_2 := (x + 1, \ldots, x + \ell) \leq (\mathbb{Z}/p\mathbb{Z})[x]/(h(x))$ that contains J. Now observe that $\#G_1 > \mathrm{o}_r(n) =: t$ and it can also be shown that any two polynomials generated by $(x + 1), \ldots, (x + \ell)$ of degree $< t$ are distinct modulo $(p, h(x))$, thus, $\#G_2 > 2^t$. Thus, if we fix $t > \log^2 n$ then in Equation (5.8) we have $\mathrm{o}_{(p,h(x))}(g(x)) > 2^t$ while the numbers $m^i p^j < n^{\lceil\sqrt{t}\rceil} < 2^t$ that again forces n to be a prime! But now the requirement on r is less strong: $\mathrm{o}_r(n) > \log^2 n$ and by the Claim 5.1.1 we can find such an r of magnitude $O^{\sim}(\log^3 n)$. This gives a primality test of complexity $O^{\sim}(\log^{7.5} n)$.

A faster but more complicated primality test based on ours was given by Lenstra and Pomerance [LP03]. It takes time $O^{\sim}(\log^6 n)$ which is the best known till now.

It might be possible to get a faster cyclotomic primality test if we can show that: checking $(x + a)^n = (x^n + a) \pmod{n, x^r - 1}$ for a *constant* many a 's and a suitable r forces n to be prime. We mention the following conjecture – given in [BP01] and verified for $r \leq 100$ and $n \leq 10^{10}$ in [KS02]:

Conjecture 5.1 *If $r > \log n$ is a prime number that does not divide n and if*

$$(X - 1)^n = X^n - 1 \ (mod\ X^r - 1, n), \tag{5.9}$$

then either n *is prime or* $n^2 = 1 \ (mod\ r)$.

If this conjecture is true, we can modify the algorithm slightly to first search for an r which does not divide $n^2 - 1$. Such an r can assuredly be found in the range $[\log n, 30(\log n)(\log \log n)]$ by Tchebycheff's estimate (see [Apo97]). Thereafter, we can test whether the congruence Equation (5.9) holds or not. Verifying the congruence takes time $O^{\sim}(r \log^2 n)$. This gives a time complexity of $O^{\sim}(\log^3 n)$.

Lenstra and Pomerance [LP03b] have given a heuristic argument that the above conjecture might fail when $r = 5$.

In this chapter we also gave a randomized polynomial time test to check whether $\sigma_n \in Aut(R_{n,r})$ for *any* given coprime n and r. Is there a deterministic polynomial time test to check this? For $r = 1$, such a test would give a way to test Carmichael numbers!

Chapter 6

Conclusion and Open Problems

This work studied various morphism problems of rings and also gave efficient solutions to some specific cases, solving well-known problems of identity testing for $\Sigma\Pi\Sigma$ circuits of bounded top fanin and primality testing. We summarize below our main results and mention the questions that remain to be answered.

6.1 Ring Morphism Problems

We defined computational variants of automorphism and isomorphism problems of rings and studied their complexity in Chapter 2. The ring automorphism problems are: testing a map for ring automorphism (TRA), deciding whether there is a nontrivial ring automorphism (RA), finding a nontrivial ring automorphism (FRA) and counting ring automorphisms (#RA). The ring isomorphism problems are: testing a map for ring isomorphism (TRI), deciding whether two given rings are isomorphic (RI), finding a ring isomorphism (FRI) and counting ring isomorphisms (#RI). The complexity of these problems, of course, depends on the way rings or maps are provided in the input. We showed that if the rings are finite and presented in basis representation in the input then all of these problems are low for Σ_2 and, hence, unlikely to be NP-hard. In this case TRA, TRI and RA are in P while we lower bound the complexity of the other problems by well-known problems, namely, graph isomorphism, integer factoring and polynomial factoring. Also, all these ring

morphism problems reduce to the problem of computing the automorphism group of a ring (given in basis form in the input) which itself is low for Σ_2.

Are there more well-known problems that reduce to ring morphism problems? For example, can we reduce the problem of computing discrete logarithm to ring morphism problems?

Our reduction of graph isomorphism to RI and #RA gives us a natural algebraic formulation for the problem of isomorphism of graphs which is open even for quantum computers. Is there a quantum algorithm for #RA, i.e., is #RA $\in$ BQP ?

We have shown that RI is unlikely to be NP-hard when the rings are finite and presented in the basis representation. We believe that to further understand the complexity of ring isomorphism it might be useful to consider RI for finite dimensional $\mathbb{Q}$-algebras. The first question that arises here: is RI for finite dimensional $\mathbb{Q}$-algebras a decidable problem ?

6.2 Cubic Forms Equivalence

We studied special cases of the polynomial equivalence problem in Chapter 3. We focussed on the equivalence of homogeneous polynomials, also known as *forms*. We connect the complexity of the problem of equivalence of degree r forms to that of ring isomorphism by showing that if a field $\mathbb{F}$ has r-th roots then r-forms equivalence over $\mathbb{F}$ reduces to $\mathbb{F}$-algebra isomorphism. More interestingly, we prove a converse: for any field $\mathbb{F}$, finite dimensional commutative $\mathbb{F}$-algebra isomorphism reduces to $\mathbb{F}$-cubic forms equivalence. Thus, cubic forms equivalence seems to be the "hardest" case of forms equivalence and subsumes the isomorphism problem of algebras. Moreover, new insights into cubic forms might help us in tackling the graph isomorphism problem as graph isomorphism reduces to commutative $\mathbb{F}$-algebra isomorphism, thus, reduces to $\mathbb{F}$-cubic forms equivalence over any field $\mathbb{F}$.

We study the cubic forms obtained from $\mathbb{F}$-algebras (thus, from graphs too) and show that they satisfy the known notions of indecomposability and regularity (or non degeneracy). We conjecture that cubic forms equivalence over $\mathbb{Q}$ is decidable and such an algorithm might give us new insights into the structure of cubic forms.

The first question that we ask towards this end: If cubic forms f, g are equivalent over $\mathbb{R}$ and are equivalent modulo p^k, for all primes p (except finitely many primes) and $k \in \mathbb{Z}^{\geq 1}$, then are they equivalent over $\mathbb{Q}$?

For any field $\mathbb{F}$, can we reduce r-forms equivalence, over $\mathbb{F}$, to commutative $\mathbb{F}$-algebra isomorphism? Currently, we know such a reduction only for fields $\mathbb{F}$ having r-th roots.

6.3 Identity Testing

We studied a special case of the identity testing problem in Chapter 4. We gave the first deterministic, polynomial-time identity test for $\Sigma\Pi\Sigma$ arithmetic circuits of bounded top fanin. Suppose the given circuit $\mathcal{C}$, over a field $\mathbb{F}$, has top fanin k, total degree d and n variables. Then the problem of identity testing is equivalent to testing whether:

$$\mathbb{F}[\overline{x}]/(\mathcal{C}(\overline{x})) \cong \mathbb{F}[\overline{x}]$$

Using the nice structure of the circuit $\mathcal{C}$, we reduce this ring isomorphism question to at most d recursive questions of the form:

$$R_i[\overline{x}]/(\mathcal{C}_i(\overline{x})) \cong R_i[\overline{x}]$$

where, $\mathcal{C}_i$ is of smaller fanin and R_i is a local ring of dimension at most d times that of the older one. This easily gives us a complexity of $poly(d^k, n)$.

The obvious question is: how can we generalize this algebraic solution to unbounded fanin k? In our algorithm the application of linear transformations on $\mathcal{C}$ was very useful and we hope that it will be instrumental in derandomizing identity testing for 'larger' k too.

Dvir and Shpilka [DS05] in their study of the structure of $\Sigma\Pi\Sigma$ identities conjectured that: if a minimal, simple, $\Sigma\Pi\Sigma$ circuit of top fanin k is zero then its rank should be $O(k)$. We refuted this conjecture for fields of prime characteristic by giving minimal, simple $\Sigma\Pi\Sigma$ identities having large rank. However, we believe that the conjecture of Dvir-Shpilka might hold over fields of characteristic 0.

6.4 Primality Testing

We studied the classical problem of primality testing in Chapter 5. We gave the first deterministic, polynomial-time primality test. Given a number n we relate its primality to the testing of the Frobenius map $\sigma_n : a(x) \mapsto a(x)^n$ for automorphism of the cyclotomic ring:

$$R_{n,r} := (\mathbb{Z}/n\mathbb{Z})[x]/(x^r - 1)$$

It turns out that for a "suitably" chosen $r \sim poly(\log n)$ there is an $l \sim poly(\log n)$ such that: for all $1 \leq a \leq l$, $\sigma_n(x + a) = \sigma_n(x) + a$ in $R_{n,r}$ iff $\sigma_n \in \mathrm{Aut}(R_{n,r})$ iff n is a prime.

Currently, there are many variants known based on the above idea. But none of them are within the realm of practical usage. We make a conjecture below that has the potential of yielding a "practical" primality test. The following conjecture was given in [BP01] and verified for $r \leq 100$ and $n \leq 10^{10}$ in [KS02]:

Conjecture 6.1 *If $r > \log n$ is a prime number that does not divide n and if $(X-1)^n = X^n - 1 \ (mod\ X^r - 1, n)$ then either n is prime or $n^2 = 1 \ (mod\ r)$.*

In Chapter 5, Theorem 5.2 gave a randomized polynomial time test to check whether $\sigma_n \in Aut(R_{n,r})$ for *any* given coprime n and r. Is there a deterministic polynomial time test to check this? For $r = 1$, such a test would give an efficient and deterministic way to test Carmichael numbers.

Appendix A

Appendix: Useful Facts

We first collect some results related to decomposition of rings into simpler rings. A ring R is said to be *decomposable* if there are subrings R_1, R_2 such that:

- $R_1 \cdot R_2 = R_2 \cdot R_1 = 0$, i.e., for all $r_1 \in R_1, r_2 \in R_2$, $r_1 \cdot r_2 = r_2 \cdot r_1 = 0$.
- $R_1 \cap R_2 = \{0\}$.
- $R = R_1 + R_2$, i.e., for every $r \in R$ there are $r_1 \in R_1, r_2 \in R_2$ such that $r = r_1 + r_2$.

Such a ring decomposition has been denoted by $R = R_1 \times R_2$ in this thesis. The subrings R_1, R_2 are called *component* rings of R.

Example The ring $R := \mathbb{F}[x]/(x^2 - x)$ decomposes as: $R = R \cdot x \times R \cdot (1 - x) \cong \mathbb{F} \times \mathbb{F}$. Here, $R \cdot x$ is a short-hand for the set $\{r \cdot x \mid r \in R\}$. Note that $R \cdot x$, $R \cdot (1 - x)$ are subrings of R and have x, $(1 - x)$ as their (multiplicative) identity elements respectively. ∎

An element $r \in R$ is called an *idempotent* if $r^2 = r$. The following lemma shows how idempotents help in decomposing a commutative ring.

Lemma A.1 *A commutative ring R decomposes iff R has an idempotent element other than $0, 1$.*

Proof: Suppose $R = R_1 \times R_2$ is a nontrivial decomposition and let the identity element 1 of R be expressible as $1 = s + t$ where $s \in R_1$, $t \in R_2$. Then we have:

$$\begin{aligned}
& 1 \cdot 1 = (s+t) \cdot (s+t) \\
\Rightarrow \quad & 1 = s^2 + t^2 \qquad [\because s \cdot t = 0] \\
\Rightarrow \quad & s + t = s^2 + t^2 \\
\Rightarrow \quad & s - s^2 = t^2 - t \\
\Rightarrow \quad & s - s^2 = 0 \qquad [\because s - s^2 \in R_1 \cap R_2 = \{0\}] \\
\Rightarrow \quad & s \text{ is an idempotent.}
\end{aligned}$$

Note that if $s = 0$ then $t = 1$ and then $R_1 = 0$ (as for all $r_1 \in R_1,\ r_1 \cdot t = 0$). Similarly, if $s = 1$ then $R_2 = 0$. As R_1, R_2 are nonzero subrings of R we deduce that $s \neq 0, 1$ and hence s is an idempotent other than $0, 1$.

Conversely, suppose that $s \neq 0, 1$ is an idempotent of R. Then consider the subrings $R \cdot s$ and $R \cdot (1 - s)$. Note that s, $(1 - s)$ are the identity elements of Rs, $R(1 - s)$ respectively. For any two elements $rs \in Rs$ and $r'(1 - s) \in R(1 - s)$: $rs \cdot r'(1 - s) = rr'(s - s^2) = 0$. If $r \in Rs \cap R(1 - s)$ then $rs = 0$ and $r(1 - s) = 0$ implying that $r = 0$. Finally, we can express any $r \in R$ as: $r = rs + r(1 - s)$. Thus, R decomposes as: $R = Rs \times R(1 - s)$. ∎

The following lemma shows that a decomposition of a ring into indecomposable rings is unique.

Lemma A.2 *Let R be a ring and $R_1, \ldots, R_k$ be indecomposable nonzero rings such that:*

$$R = R_1 \times R_2 \times \cdots \times R_k$$

Then this decomposition is unique up to ordering, i.e., if we have indecomposable nonzero S_j's such that:

$$R = R_1 \times \cdots \times R_k = S_1 \times \cdots \times S_l$$

then $k = l$ and there exists a permutation π on $[k]$ such that for all $i \in [k]$, $R_i = S_{\pi(i)}$.

Proof: Assume wlog that $k \geq l$. Let ϕ_1 be a homomorphism of the ring R such that ϕ_1 is identity on S_1 and $\phi_1(S_2) = \cdots = \phi_1(S_l) = 0$. ϕ_1 is well defined simply because $R = S_1 \times \cdots \times S_l$.

Clearly, $\phi_1(R_1), \phi_1(R_2), \cdots, \phi_1(R_k)$ are all subrings of S_1 and:

$$\phi_1(R) = \phi_1(R_1) + \phi_1(R_2) + \cdots + \phi_1(R_k) = S_1$$

Can these subrings have nontrivial intersection? Say, $s_1 \in \phi_1(R_i) \cap \phi_1(R_j)$ for some $i \neq j$ then there are some $s, s' \in S_2 + \cdots + S_l$ such that $s_1 + s \in R_i$ and $s_1 + s' \in R_j$. Let a be the (multiplicative) identity of $R_1 + \cdots + R_{i-1} + R_{i+1} + \cdots + R_k$ and b be the identity of R_i. Then:

$$\begin{aligned}
& (s_1 + s)a = 0 \text{ and } (s_1 + s')b = 0 \quad [\because R = R_1 \times \cdots \times R_k] \\
\Rightarrow \quad & (s_1 + s)a + (s_1 + s')b = 0 \\
\Rightarrow \quad & s_1(a + b) + sa + s'b = 0 \\
\Rightarrow \quad & s_1 + (sa + s'b) = 0 \quad [\because 1 = a + b] \\
\Rightarrow \quad & s_1 = (sa + s'b) = 0 \quad [\because s_1 \in S_1 \text{ and } sa, s'b \in S_2 + \cdots + S_l] \\
\Rightarrow \quad & \phi_1(R_i) \cap \phi_1(R_j) = \{0\}, \quad \text{for all } i \neq j \in [k]
\end{aligned}$$

Also, for any $r_i \in R_i$, $r_j \in R_j$, $\quad r_i r_j = 0$ implying that $\phi_1(r_i) \cdot \phi_1(r_j) = 0$. The properties above together mean that:

$$S_1 = \phi_1(R_1) \times \phi_1(R_2) \times \cdots \times \phi_1(R_k)$$

Since S_1 was assumed to be indecomposable we have that exactly one of the subrings above is nonzero. Wlog, say, $\phi_1(R_2) = \cdots = \phi_1(R_k) = 0$ and then it is implied that $\phi_1(R_1) = S_1$.

Similarly, we can define ϕ_i to be a homomorphism of the ring R such that ϕ_i is identity on S_i and $\phi_i(S_j) = 0$ for all $j \in [l] \setminus \{i\}$. Then the above argument says that there is an injective map $\tau : [l] \to [k]$ such that for all $i \in [l]$:

$$\phi_i(R_{\tau(i)}) = S_i \text{ and } \phi_i(R_j) = 0 \text{ for all } j \in [k] \setminus \{\tau(i)\} \tag{A.1}$$

Now consider an $l \times k$ matrix $D = ((\delta_{i,j}))$, where, $\delta_{i,j} = 1$ if $\phi_i(R_j) = S_i$ else $\delta_{i,j} = 0$. Equation (A.1) tells us that each row of D has exactly one 1. Now if $k > l$ then D

has more columns than rows and hence there is a zero column, say j-th, implying that $\phi_i(R_j) = 0$ for all $i \in [l]$. But this means that $R_j = 0$ which is a contradiction. Hence, $k = l$ and D has exactly one 1 in each row and column, thus making τ a permutation.

So now we have that for any $j \in [k]$, $\quad \phi_{\tau^{-1}(j)}(R_j) = S_{\tau^{-1}(j)}$ and $\phi_i(R_j) = 0$ for all $i \in [k] \setminus \{\tau^{-1}(j)\}$. In other words for any $j \in [k]$, $\quad R_j = S_{\tau^{-1}(j)}$.

This completes the proof of unique decomposition of rings into indecomposable subrings. ∎

So what is the structure of these indecomposable rings that appear in the decomposition? Here, we sketch the form of indecomposable rings that are finite and commutative.

Lemma A.3 *Let R be a finite commutative indecomposable ring. Then,*

1) R has a prime-power characteristic, say p^m, for some prime p.

2) R can be expressed in the form:

$$R = ((\mathbb{Z}/p^m\mathbb{Z})[z]/(h(z)))\,[y_1, \ldots, y_k]/\,(y_1^{e_1}, \ldots, y_k^{e_k}, h_1(z, y_1, \ldots, y_k), \ldots, \ldots, h_\ell(z, y_1, \ldots, y_k))$$

where, $h(z)$ is irreducible over $\mathbb{Z}/p\mathbb{Z}$ and h_i's are multivariate polynomials over $\mathbb{Z}/p^m\mathbb{Z}$.

Remark: The ring $(\mathbb{Z}/p^m\mathbb{Z})[z]/(h(z))$, where $h(z)$ is irreducible over $\mathbb{Z}/p\mathbb{Z}$, is called *Galois ring*. It is a finite field if $m = 1$.

Notice that the form of R claimed in 2) above says that the generators $y_1, \ldots, y_k$ of R are *nilpotents*, i.e., they vanish when raised by a suitable integer. ∎

Proof: [1)] Suppose R is a finite commutative indecomposable ring with characteristic n. Suppose n non trivially factors as: $n = ab$, where $a, b \in \mathbb{Z}^{>1}$ are coprime, then by Euclidean-gcd algorithm we have $a', b' \in \mathbb{Z}$ such that $a'a + b'b = 1$. Then:

$$R = (a'a)R \times (b'b)R$$

(Convince yourself that this is a decomposition.) This contradiction shows that n is a prime power, say $n = p^m$. ∎

Proof: [2)] We assume $m = 1$ for simplicity of exposition. These ideas carry forward to larger m's (see [McD74]). So suppose that R is an $\mathbb{F}_p$-algebra and is given in terms of basis elements $b_1, \ldots, b_n$. Let $g_1(b_1, \ldots, b_n), \ldots, g_\ell(b_1, \ldots, b_n)$ be the multivariate polynomials that define the multiplication operation of the ring R. Thus, we have an expression for R as:

$$R \cong \mathbb{F}_p[x_1, \ldots, x_n]/(g_1(x_1, \ldots, x_n), \ldots, g_\ell(x_1, \ldots, x_n)) \tag{A.2}$$

Since R is of dimension n, $\{1, x_1, x_1^2, \ldots, x_1^n\}$ cannot all be linearly independent and, hence, there is a polynomial $f_1(z) \in \mathbb{F}_p[z]$ of degree at most n such that $f_1(x_1) = 0$ in R. Further, assume that f_1 is of lowest degree. Now if f_1 non trivially factors as: $f_1(z) = f_{11}(z)f_{12}(z)$, where f_{11}, f_{12} are coprime, then there are $a_1(z), a_2(z) \in \mathbb{F}_p[z]$ such that $a_1 f_{11} + a_2 f_{12} = 1$ and R decomposes as:

$$R \cong (a_1(x_1)f_{11}(x_1) \cdot R) \times (a_2(x_1)f_{12}(x_1) \cdot R)$$

As R is assumed to be indecomposable we deduce that f_1 is a power of an irreducible polynomial. Say, $f_1(z) = f_{11}(z)^{e_1}$ where f_{11} is an irreducible polynomial over $\mathbb{F}_p$ of degree d_1. Now we claim that there are $g'_1, \ldots, g'_\ell \in \mathbb{F}_{p^{d_1}}[x_1, \ldots, x_n]$ such that:

$$R \cong \mathbb{F}_{p^{d_1}}[x_1, \ldots, x_n]/(x_1^{e_1}, g'_1(x_1, \ldots, x_n), \ldots, g'_\ell(x_1, \ldots, x_n)) \tag{A.3}$$

To prove the above claim we need the following fact:

Claim A.0.1 *If $f(x)$ is an irreducible polynomial, of degree d, over a finite field $\mathbb{F}_q$ then*

$$S = \mathbb{F}_q[x]/(f(x)^e) \cong \mathbb{F}_{q^d}[u]/(u^e)$$

Proof of Claim A.0.1. Consider the ring $S' := (\mathbb{F}_q[x]/(f(x)))[u]/(u^e)$ isomorphic to RHS. We claim that the map $\phi : S \to S'$ which fixes $\mathbb{F}_q$ and maps $x \mapsto (x + u)$, is an isomorphism.

Note that $f(x+u)^e = 0$ in the ring S' simply because $f(x+u) - f(x) = u \cdot g(x)$ for some $g(x) \in \mathbb{F}_q[x]$. Thus, ϕ is a ring homomorphism from S to S'. Next we show

that the minimal polynomial of $\phi(x)$ over $\mathbb{F}_q$ is of degree de, thus, the dimension of $\phi(S)$ is the same as that of S' over $\mathbb{F}_q$ and hence ϕ is an isomorphism.

Suppose $g(z) := \sum_{j=0}^{d'} a_j x^j$ is the least degree polynomial over $\mathbb{F}_q$ such that $g(x+u) = 0$ in S'. This means that in S':

$$0 = g(x+u) = g(x) + u \cdot g^{(1)}(x) + u^2 \cdot \frac{g^{(2)}(x)}{2!} + \cdots + u^{e-1} \cdot \frac{g^{(e-1)}(x)}{(e-1)!}$$

where, $\frac{g^{(i)}(x)}{i!} = \sum_{j=i}^{d'} \frac{j(j-1)\cdots(j-i+1)}{i!} a_j x^{j-i}$. But since $1, u, \ldots, u^{e-1}$ are linearly independent over $\mathbb{F}_q[x]/(f(x))$. We have:

$$g(x) = g^{(1)}(x) = \cdots = g^{(e-1)}(x) = 0 \text{ over } \mathbb{F}_q[x]/(f(x))$$

Whence we get, $f(z)^e | g(z)$ which by the definition of g means that $g(z) = f(z)^e$. Thus, ϕ is an isomorphism from S to S'. □

From the above claim we now deduce:

$$\begin{aligned} R &\cong \mathbb{F}_p[x_1, \ldots, x_n]/(f_{11}(x_1)^{e_1}, g_1(x_1, \ldots, x_n), \ldots, g_\ell(x_1, \ldots, x_n)) \\ &\cong \mathbb{F}_{p^{d_1}}[u, x_2, \ldots, x_n]/(u^{e_1}, g_1'(u, x_2, \ldots, x_n), \ldots, g_\ell'(u, x_2, \ldots, x_n)) \\ &\cong \mathbb{F}_{p^{d_1}}[x_1, x_2, \ldots, x_n]/(x_1^{e_1}, g_1'(x_1, x_2, \ldots, x_n), \ldots, g_\ell'(x_1, x_2, \ldots, x_n)) \end{aligned}$$

This new ring which we obtained has x_1 as a nilpotent. We can now consider the lowest degree polynomial $f_2(z) \in \mathbb{F}_{p^{d_1}}[z]$ such that $f_2(x_2) = 0$ in R. The above process when repeated on f_2, x_2 in place of f_1, x_1 gives us that there are $d_2, e_2 \in \mathbb{Z}^{\geq 1}$ and $g_1'', \ldots, g_\ell'' \in \mathbb{F}_{p^{d_1 d_2}}[x_1, \ldots, x_n]$ such that:

$$R \cong \mathbb{F}_{p^{d_1 d_2}}[x_1, \ldots, x_n]/(x_1^{e_1}, x_2^{e_2}, g_1''(x_1, \ldots, x_n), \ldots, g_\ell''(x_1, \ldots, x_n))$$

Continuing this way we get that there is a $d \in \mathbb{Z}^{\geq 1}$ and polynomials $h_1, \ldots, h_\ell \in \mathbb{F}_{p^d}[x_1, x_2, \ldots, x_n]$ such that:

$$R \cong \mathbb{F}_{p^d}[x_1, \ldots, x_n]/(x_1^{e_1}, \ldots, x_n^{e_n}, h_1(x_1, \ldots, x_n), \ldots, h_\ell(x_1, \ldots, x_n))$$

∎

Remark: Note that the above proof can be viewed as an algorithm to decompose a finite dimensional commutative ring, given in basis form, into indecomposable

rings. It is indeed a deterministic polynomial time algorithm given oracles to integer and polynomial factorization. ∎

Let us now see a structural property of commutative indecomposable rings.

Lemma A.4 *For a field $\mathbb{F}$, consider a ring R of the form:*

$$R = \mathbb{F}[x_1, \ldots, x_n]/(x_1^{e_1}, \ldots, x_n^{e_n}, h_1(x_1, \ldots, x_n), \ldots, h_\ell(x_1, \ldots, x_n))$$

Then,

1) R is indecomposable.

2) R has a unique maximal ideal $\mathcal{M}$ and $\mathcal{M} =$ set of nilpotents of R.

Proof: [1)] Any element r of R looks like $a_0 + a_1(\overline{x})x_1 + \cdots + a_n(\overline{x})x_n$, where, $a_0 \in \mathbb{F}$ and $a_1(\overline{x}), \ldots, a_n(\overline{x}) \in \mathbb{F}[x_1, \ldots, x_n]$.

Suppose $a_0 = 0$. Since, $x_1^{e_1} = \cdots = x_n^{e_n} = 0$ we have that:

$$\begin{aligned} r^{e_1+\cdots+e_n} &= (a_1(\overline{x})x_1 + \cdots + a_n(\overline{x})x_n)^{e_1+\cdots+e_n} \\ &= 0 \end{aligned}$$

Suppose $a_0 \neq 0$. Let $r_0 := r - a_0$ and $e := e_1 + \cdots + e_n$. Then we have:

$$\begin{aligned} (a_0 + r_0)(a_0^e - a_0^{e-1}r_0 + \cdots + (-1)^{e-1}a_0 r_0^{e-1} + (-1)^e r_0^e) &= a_0^{e+1} + (-1)^e r_0^{e+1} \\ &= a_0^{e+1} \qquad [\because r_0^e = 0] \\ &\in \mathbb{F}^* \\ \Rightarrow r &\in R^* \end{aligned}$$

Thus, every element r of R is either a nilpotent or a unit depending upon whether $a_0 = 0$ or not.

Now suppose R is decomposable. By Lemma A.1 there has to be a nontrivial idempotent $t \in R$. But we have:

$$\begin{aligned} & t^2 = t \\ \Rightarrow \quad & t(t-1) = 0 \\ \Rightarrow \quad & t = 0 \text{ or } 1 \qquad [\because t \text{ or } (t-1) \text{ is a unit}] \end{aligned}$$

This contradiction shows that R is indecomposable. ∎

Proof: [2)] Define a set $\mathcal{M} := R \setminus R^*$. As shown above $\mathcal{M}$ is the set of nilpotents of R and hence is an ideal. $\mathcal{M}$ is maximal because any element outside it is a unit. $\mathcal{M}$ is unique because it contains all the non-units of R. ∎

Now we consider the special form of local rings that appear in Equation (4.7) and show how to do computations in that ring in an "efficient" way.

Lemma A.5 *Let us define a sequence of local rings, over a field $\mathbb{F}$, as:*

$$
\begin{aligned}
&S_0 := \mathbb{F} \quad \textit{having maximal ideal } \mathcal{M}_0 = 0\\
&S_1 := S_0[x_1]/(x_1^{e_1}) \quad \textit{having maximal ideal } \mathcal{M}_1 = (x_1) = x_1 \cdot S_1\\
&\quad \vdots\\
&S_k := S_{k-1}[x_k]/\left((x_k + r_{k,1}) \cdots (x_k + r_{k,e_k})\right), \quad \textit{where, } r_{k,1}, \ldots, r_{k,e_k} \in \mathcal{M}_{k-1}.\\
&\qquad \textit{Also, the maximal ideal of } S_k \textit{ is } \mathcal{M}_k = (x_1, \ldots, x_k)
\end{aligned}
$$

Define $D_i := e_1 \cdots e_i$, for all $i \in [k]$. Then the addition operation in S_k takes time $O(D_k)$ and the multiplication operation in S_k takes time $O(kD_k^2)$, assuming field operations in $\mathbb{F}$ take constant time.

Proof: Inductively we can show that S_k is a local ring. Since $(x_k + r_{k,1}) \cdots (x_k + r_{k,e_k}) = 0$ and $r_{k,1}, \ldots, r_{k,e_k} \in \mathcal{M}_{k-1}$ we have that, in the ring S_k:

$$x_k^{e_k} = r_{k-1}x_k^{e_k-1} + \cdots + r_1 x_k + r_0 \quad \text{for some } r_{k-1}, \ldots, r_0 \in \mathcal{M}_{k-1}$$

As $r_{k-1}, \ldots, r_0$ are nilpotents in S_k we deduce from the above equation that x_k is a nilpotent too and hence S_k is a local ring with the ideal of nilpotents equal to $(x_1, \ldots, x_k)$.

Assume that the addition operation in S_{k-1} takes time: $O(D_{k-1})$. Let $r := (\alpha_{e_k-1}x_k^{e_k-1} + \cdots + \alpha_1 x_k + \alpha_0)$ and $r' := (\alpha'_{e_k-1}x_k^{e_k-1} + \cdots + \alpha'_1 x_k + \alpha'_0)$ be two elements in S_k such that for all $0 \le i \le e_k - 1, \quad \alpha_i, \alpha'_i \in S_{k-1}$. Now the addition operation: $r + r'$ entails computing e_k additions (of the form $\alpha_i + \alpha'_i$) in S_{k-1}. Thus, addition in S_k takes time: $e_k \cdot O(D_{k-1}) = O(D_k)$.

Assume that the multiplication operation in S_{k-1} takes time: $O((k-1)D_{k-1}^2)$. Then the multiplication operation: $r \cdot r'$ entails e_k^2 multiplications (of the form $\alpha_i \cdot \alpha'_j$) in the ring S_{k-1} and those many additions. Hence, the time taken is:

$$\begin{aligned} e_k^2 O((k-1)D_{k-1}^2) + e_k^2 O(D_{k-1}) &= O((k-1)D_k^2) + e_k O(D_k) \\ &= O(kD_k^2) \end{aligned}$$

∎

The next lemma gives an important property of the multiplicative group of the ring: $\mathbb{Z}/p^s\mathbb{Z}$.

Lemma A.6 *Let p be a prime and $G := (\mathbb{Z}/p^s\mathbb{Z})^*$ be the multiplicative group of invertible elements modulo p^s. Then,*

- *If $p = 2$ then G is a cyclic group only if $s \in \{1, 2\}$.*
- *If $p \geq 2$ then G is always a cyclic group.*

Proof: See [NZM91]. ∎

It is easy to see that a finite field $\mathbb{F}_q$ has to be of size $q = p^m$, for some prime p. The following lemma describes some more interesting properties of finite fields.

Lemma A.7 *Let $\mathbb{F}_q$ be a finite field. Then,*

- *$\mathbb{F}_q^*$ is a cyclic group of size $(q-1)$.*
- *The automorphism group of the ring $\mathbb{F}_q$ is generated by the Frobenius map $\sigma_q : \alpha \mapsto \alpha^q$, i.e., $Aut(\mathbb{F}_q) = (\sigma_q)$.*
- *For any r coprime to q, the polynomial $(x^r - 1)$ factorizes into irreducible polynomials over $\mathbb{F}_q$ as:*

$$(x^r - 1) = \prod_{d_i | r} \prod_{j=1}^{\frac{\phi(d_i)}{o_{d_i}(q)}} f_{i,j}(x), \quad \textit{where, } f_{i,j} \textit{ is of degree } o_{d_i}(q)$$

Proof: See [LN86] for the proofs. ∎

Suppose we are given a multivariate polynomial $f \in \mathbb{F}[x_1, \ldots, x_n]$ having total degree d. Then there exists a linear transformation τ on the variables $x_1, \ldots, x_n$ that transforms f to a multivariate polynomial $\tau(f)$ having a nonzero term x_1^d. This observation was useful in the proofs of chapter 5.

Lemma A.8 *Let $f(x_1, \ldots, x_n) \in \mathbb{F}[x_1, \ldots, x_n]$ has total degree d. Then there is an invertible linear transformation $\tau : \overline{\mathbb{F}}^n \to \overline{\mathbb{F}}^n$ such that $f(\tau(x_1), \cdots, \tau(x_n))$ has a nonzero coefficient of x_1^d. ($\overline{\mathbb{F}}$ is the algebraic closure of $\mathbb{F}$.)*

Proof: Collect the degree d terms of f in the polynomial:

$$f_d(x_1, \ldots, x_n) := \sum_{i_1+\cdots+i_n=d} a_{i_1,\ldots,i_n} x_1^{i_1} \cdots x_n^{i_n}, \quad \text{where, } a_{i_1,\ldots,i_n}\text{'s } \in \mathbb{F}$$

By the hypothesis, $f_d \neq 0$. If we apply a linear transformation τ on f such that:

$$\tau(x_i) = \sum_{1 \leq j \leq n} \tau_{i,j} x_j, \quad \text{where, } \tau_{i,j} \in \overline{\mathbb{F}}$$

Then the coefficient of x_1^d in the polynomial $f(\tau(x_1), \cdots, \tau(x_n))$ is:

$$\sum_{i_1+\cdots+i_n=d} a_{i_1,\ldots,i_n} \tau_{1,1}^{i_1} \cdots \tau_{n,1}^{i_n}$$

which is nothing but $f_d(\tau_{1,1}, \ldots, \tau_{n,1})$. By the Schwartz-Zippel lemma we have that there are values for $\tau_{1,1}, \ldots, \tau_{n,1} \in \overline{\mathbb{F}}$ such that $f_d(\tau_{1,1}, \ldots, \tau_{n,1}) \neq 0$ and, hence, the coefficient of x_1^d in the polynomial $f(\tau(x_1), \cdots, \tau(x_n))$ is nonzero. ∎

Suppose R is a ring, $\mathcal{I}$ is an ideal of R and $f \in R[z]$. Then a factorization of $f(z)$ modulo $\mathcal{I}$ can be "lifted" to one modulo $\mathcal{I}^2$ by a well known trick in algebra called *Hensel's Lifting*. This is a useful trick in many situations, for example, given a root of $f(x)$ modulo p we can lift it to a root of $f(x)$ modulo p^2.

Lemma A.9 (Hensel's Lifting) *Let R be a ring and $\mathcal{I}$ be an ideal. Let $f(z) \in R[z]$ and $f = gh \ (mod\, \mathcal{I})$ be a factorization of f over $R/\mathcal{I}$ such that there exists $a, b \in R[z]$, $ag + bh = 1 \ (mod\, \mathcal{I})$. Then,*

- *There are easily computable* $g^*, h^*, a^*, b^* \in R[z]$ *satisfying:*

$$f = g^* h^* \ (mod\ \mathcal{I}^2)$$
$$g^* = g \ (mod\ \mathcal{I}) \ and \ h^* = h \ (mod\ \mathcal{I})$$
$$a^* g^* + b^* h^* = 1 \ (mod\ \mathcal{I}^2)$$

- *Also,* g^*, h^* *above are unique in the sense that for any other* g', h' *satisfying the above conditions we have some* $u \in \mathcal{I}$ *such that:*

$$g' = g^*(1+u) \ (mod\ \mathcal{I}^2)$$
$$h' = h^*(1-u) \ (mod\ \mathcal{I}^2)$$

Proof: See [LN86] for the proof. ∎

We can define the *ring of fractions* S^{fr} of a ring S as the set of elements $\frac{u}{v}$, where, $u, v \in S$ and v is not a zero divisor of S. Clearly, S^{fr} is also a ring. We will be considering polynomials over rings S and S^{fr}. A polynomial $f(z) \in S[z]$ is called *monic* if its leading coefficient is a unit of S. The following is a well known lemma that relates polynomial factorization over the ring S to its ring of fractions S^{fr}.

Lemma A.10 (Gauss' Lemma) *Suppose* $f, g \in S[z]$ *and* $h \in S^{fr}[z]$ *such that* $f = gh$. *If* g *is monic then* $h \in S[z]$.

Proof: A proof for the case of $S = \mathbb{Z}$ can be found in any algebra text, eg., [NZM91]. The proof for general S is similar in spirit. ∎

It was shown by Hendrik Lenstra, Jr. that if Fermat's little test modulo n passes for all $a \leq 4(\log^2 n)$ then n has to be square-free.

Lemma A.11 (Lenstra) *Let* n *be a positive integer. If* $a^n = a \ (mod\ n)$, *for all* $1 \leq a \leq 4(\log^2 n)$, *then* n *is square-free.*

Proof: Suppose $n = p^k m$ where prime p does not divide m. Suppose $k \geq 2$. We have that:

$$\begin{aligned} a^{p^k m} &= a \pmod{n} \\ \Rightarrow a^{p^k m} &= a \pmod{p^2} \\ \Rightarrow a^{pm} &= a \pmod{p^2} \end{aligned}$$

$$[\because a^{p^2-p} = a^{\phi(p^2)} = 1 \pmod{p^2}. \text{ Thus, } a^p = a^{p^2} = \cdots = a^{p^k} \pmod{p^2}.]$$

Now the above gives us that $a^{pm} = a \pmod{p}$ implying that $a^m = a \pmod{p}$. Thus, there is an integer b such that $a^m = a + bp$ and now raising both sides by p gives us that: $a^{pm} = a^p \pmod{p^2}$. Thus,

$$a^p = a \pmod{p^2}, \quad \text{for all } 1 \leq a \leq 4(\log^2 n)$$

The equation $x^p = x \pmod{p^2}$ can have at most p distinct solutions. Since all the $1 \leq a \leq 4\log^2 p$ numbers are its solution, so will be their products. But the bound in [CEG83] shows that the $(4\log^2 p)$-smooth numbers smaller than p^2 are more than p, which gives us a contradiction. Thus, $k = 1$ and n is square-free. ∎

We give below some interesting identities in $\mathbb{Q}(\zeta_r)$, where ζ_r is a primitive r-th root of unity. Note that $Q_r(y) := \frac{y^r-1}{y-1}$ is a polynomial having ζ_r as a root.

Lemma A.12 *Let n be an odd integer and r be an odd prime not dividing n. Let $B = 16^{-1} (mod\ r)$. Then,*

1) $(1-x)(1-x^2)\cdots(1-x^{r-1}) = r\ (mod\ Q_r(x))$.

2) $\left(x^B(1-x)(1-x^2)\cdots(1-x^{\frac{r-1}{2}})\right)^2 = (-1)^{\frac{r-1}{2}} \cdot r\ (mod\ Q_r(x))$.

3) $x^{Bn}(1-x^n)(1-x^{2n})\cdots(1-x^{\frac{r-1}{2}n}) = \left(\frac{n}{r}\right)x^B(1-x)(1-x^2)\cdots(1-x^{\frac{r-1}{2}})\ (mod\ Q_r(x))$.

Proof: [1)] Since x is an r-th primitive root of unity we have that $Q_r(y)$ factorizes as:

$$Q_r(y) = (y-x)\cdots(y-x^{r-1}) \pmod{Q_r(x)}$$

Substituting $y = 1$ above we get: $(1-x)(1-x^2)\cdots(1-x^{r-1}) = r \pmod{Q_r(x)}$.

∎

Proof: [2)] Starting from the identity we got above, we deduce:

$$\begin{aligned}
& (1-x)(1-x^2)\cdots(1-x^{r-1}) \;=\; r \pmod{Q_r(x)} \\
\Rightarrow\quad & (1-x)(1-x^2)\cdots(1-x^{\frac{r-1}{2}})(1-x^{\frac{r+1}{2}})\cdots(1-x^{r-1}) \;=\; r \pmod{Q_r(x)} \\
\Rightarrow\quad & (1-x)(1-x^2)\cdots(1-x^{\frac{r-1}{2}})x^{\frac{r+1}{2}}(x^{\frac{r-1}{2}}-1)\cdots x^{r-1}(x-1) \;=\; r \pmod{Q_r(x)} \\
\Rightarrow\quad & (-1)^{\frac{r-1}{2}}\cdot x^{(\frac{r+1}{2})+\cdots+(r-1)}\left((1-x)(1-x^2)\cdots(1-x^{\frac{r-1}{2}})\right)^2 \;=\; r \pmod{Q_r(x)} \\
\Rightarrow\quad & \left(x^B(1-x)(1-x^2)\cdots(1-x^{\frac{r-1}{2}})\right)^2 \;=\; (-1)^{\frac{r-1}{2}}\cdot r \pmod{Q_r(x)}
\end{aligned}$$

∎

Proof: [3)] Consider the set $T := \{1\cdot n,\ 2\cdot n, \ldots,\ \frac{r-1}{2}\cdot n\}$. Let $s_1, \ldots, s_u \in T$ be the numbers that are congruent (modulo r) to a number between 1 and $\frac{r}{2}$. Let $l_1, \ldots, l_v$ be the numbers that are congruent (modulo r) to a number between $\frac{r}{2}$ and $(r-1)$. It is easy to show that the set $\{s_1, \ldots, s_u, (r-l_1), \ldots, (r-l_v)\}$(modulo r) has all the elements $\left[1..\frac{r-1}{2}\right]$. We now have:

$$\begin{aligned}
\frac{r-1}{2}! &= s_1\cdots s_u\cdot(r-l_1)\cdots(r-l_v) \pmod{r} \\
&= (-1)^v\cdot s_1\cdots s_u\cdot l_1\cdots l_v \pmod{r} \\
&= (-1)^v\cdot n\cdot 2n\cdots\frac{r-1}{2}n \pmod{r} \\
&= (-1)^v\cdot n^{\frac{r-1}{2}}\cdot\left(\frac{r-1}{2}!\right) \pmod{r} \\
&= (-1)^v\cdot\left(\frac{n}{r}\right)\cdot\left(\frac{r-1}{2}!\right) \pmod{r}
\end{aligned}$$

Finally, we get that:

$$(-1)^v \;=\; \left(\frac{n}{r}\right) \tag{A.4}$$

Also, $s_1+\cdots+s_u+l_1+\cdots+l_v = n+2n+\cdots+\frac{r-1}{2}n = \frac{(r+1)(r-1)}{8}n = -2Bn \pmod{r}$. And $s_1+\cdots+s_u-l_1-\cdots-l_v = 1+2+\cdots+\frac{r-1}{2} = \frac{(r+1)(r-1)}{8} = -2B \pmod{r}$. By taking the difference of the two equations we get that:

$$l_1+\cdots+l_v \;=\; (-Bn+B) \pmod{r} \tag{A.5}$$

Now we have enough 'machinery' to tackle the original problem:

$$
\begin{aligned}
& x^{Bn}(1-x^n)(1-x^{2n})\cdots(1-x^{\frac{r-1}{2}n}) \\
\equiv\ & x^{Bn}(1-x^{s_1})\cdots(1-x^{s_u})(1-x^{l_1})\cdots(1-x^{l_v}) \pmod{Q_r(x)} \\
\equiv\ & (-1)^v \cdot x^{Bn+l_1+\cdots+l_v}(1-x^{s_1})\cdots(1-x^{s_u})(1-x^{r-l_1})\cdots(1-x^{r-l_v}) \pmod{Q_r(x)} \\
\equiv\ & \left(\frac{n}{r}\right)x^B(1-x)(1-x^2)\cdots(1-x^{\frac{r-1}{2}}) \pmod{Q_r(x)} \quad \text{[by Equations (A.4) and (A.5)]}
\end{aligned}
$$

∎

The following lemma states the famous Quadratic Reciprocity Law which gives a beautiful relation between the Jacobi symbols: $\left(\frac{m}{n}\right)$ and $\left(\frac{n}{m}\right)$.

Lemma A.13 (Quadratic Reciprocity) *If m and n are two odd and coprime positive integers then,*

$$\left(\frac{m}{n}\right)\cdot\left(\frac{n}{m}\right) = (-1)^{\frac{m-1}{2}\cdot\frac{n-1}{2}}$$

Proof: See [NZM91]. ∎

The following lemma lists two useful results regarding the polynomial hierarchy (PH): BPP is low for Σ_2 and the Swapping lemma.

Lemma A.14 *1) $BPP \in \Sigma_2 \cap \Pi_2$.*

2) Let M be a polynomial time deterministic Turing machine and c be a positive constant then there is a polynomial time deterministic Turing machine M' and a positive c' such that:

$$
\begin{aligned}
L &= \left\{ x \mid \ (\exists y \in \{0,1\}^c) Prob_{z\in\{0,1\}^c}[M(x,y,z) \ accepts] \geq \frac{2}{3} \right\} \\
&= \left\{ x \mid \ Prob_{z\in\{0,1\}^{c'}}[(\exists y \in \{0,1\}^c)\ M'(x,y,z) \ accepts] \geq \frac{2}{3} \right\}
\end{aligned}
$$

Proof: See [Sch88]. ∎

References

[AB99] M. Agrawal and S. Biswas. *Primality and identity testing via Chinese remaindering.* Proceedings of Annual IEEE Symposium on Foundations of Computer Science, 202–209, 1999.

[AGP94] W. R. Alford, Andrew Granville and Carl Pomerance. *There are infinitely many Carmichael numbers.* Annals of Mathematics, 139:703–722, 1994.

[AH92] L. M. Adleman and M.-D. Huang. *Primality testing and two dimensional Abelian varieties over finite fields.* Lecture Notes in Mathematics, 1512, 1992.

[AKS02] Manindra Agrawal, Neeraj Kayal and Nitin Saxena. *PRIMES is in P.* Preprint (http://www.cse.iitk.ac.in/users/manindra/primality.ps), August 2002.

[AKS04] Manindra Agrawal, Neeraj Kayal and Nitin Saxena. *Primes is in P.* Annals of Mathematics, 160(2):781–793, 2004.

[ALM+98] S. Arora, C. Lund, R. Motwani, M. Sudan and M. Szegedy. *Proof verification and the hardness of approximation problems.* Journal of the Association for Computing Machinery, 45(3):501–555, 1998. (preliminary version in FOCS 1992).

[Ank52] N. C. Ankeny. *The least quadratic non-residue.* Annals of Mathematics, 55:65–72, 1952.

[Apo97] T. M. Apostol. *Introduction to Analytic Number Theory.* Springer-Verlag, 1997.

[APR83] L. M. Adleman, C. Pomerance and R. S. Rumely. *On distinguishing prime numbers from composite numbers.* Annals of Mathematics, 117:173–206, 1983.

[AS97] S. Arora and M. Sudan. *Improved low-degree testing and its applications.* Proceedings of the Twenty-Ninth Annual ACM Symposium on Theory of Computing, 485–495, 1997.

[AS98] S. Arora and S. Safra. *Probabilistic checking of proofs: A new characterization of NP.* Journal of the Association for Computing Machinery, 45(1):70–122, 1998. (preliminary version in FOCS 1992).

[AS05] M. Agrawal and N. Saxena. *Automorphisms of Finite Rings and Applications to Complexity of Problems.* STACS'05, Springer LNCS 3404, 1–17, 2005.

[AS06] M. Agrawal and N. Saxena. *Equivalence of $\mathbb{F}$-algebras and cubic forms.* STACS'06, Springer LNCS 3884, 115–126, 2006.

[Atk86] A. O. L. Atkin. Lecture notes of a conference, Boulder (Colorado). Manuscript, August 1986.

[AT04] V. Arvind and Jacobo Toran. *Solvable group isomorphism is (almost) in $NP \cap coNP$.* Proc. 19$^{\text{th}}$ IEEE Conference on Computational Complexity, 91–103, 2004.

[Bac96] Eric Bach. *Weil bounds for singular curves.* Applicable Algebra in Engineering, Communication and Computing **7**:289–298, 1996.

[BCW80] M. Blum, A. K. Chandra and M. N. Wegman. *Equivalence of free Boolean graphs can be tested in polynomial time.* Information Processing Letters, 10:80–82, 1980.

[Berl70] E. R. Berlekamp. *Factoring Polynomials over Large Finite Fields.* Mathematics of Computation **24**:713–735, 1970.

[Bern] Dan Bernstein. *http://cr.yp.to/primetests.html.*

[Berr03] Pedro Berrizbeitia. *Sharpening PRIMES is in P for a large family of numbers.* Preprint, 2003. Available from *http://arxiv.org/abs/math.NT/0211334.*

[BH96] R. C. Baker and G. Harman. *The Brun-Titchmarsh Theorem on average.* Proceedings of a conference in Honor of Heini Halberstam, Volume 1:39–103, 1996.

[BK95] M. Blum and S. Kannan. *Designing programs that check their work.* Journal of the Association for Computing Machinery, 42:269–291, 1995.

[BP01] Rajat Bhattacharjee and Prashant Pandey. *Primality testing.* Technical report, IIT Kanpur, 2001. Available at http://www.cse.iitk.ac.in/research/btp2001/primality.html.

[Bro87] W. D. Brownawell. *Bounds for the degrees in the Nullstellensatz.* Annals of Maths, 126, 1987, 577-591.

[BS84] L. Babai and E. Szemeredi. *On the complexity of matrix group problems.* Proc. 25$^{\text{th}}$ IEEE Foundations of Computer Science, 1984, 229-240.

[BS96] Eric Bach and Jeffrey Shallit. *Algorithmic Number Theory.* MIT Press, Cambridge, 1996.

[Car10] R. D. Carmichael. *Note on a number theory function.* Bulletin of American Mathematical Society, 16:232-238, 1910.

[CDGK91] M. Clausen, A. Dress, J. Grabmeier and M. Karpinsky. *On zero-testing and interpolation of k-sparse multivariate polynomials over finite fields.* Theoretical Computer Science, 84(2):151–164, 1991.

[CEG83] E. R. Canfield, P. Erdos and A. Granville. *On a problem of Oppenhein concerning "Factorisatio Numerorum".* Journal of Number Theory, 17:1–28, 1983.

[CGP98] N. Courtois, L. Goubin and J. Patarin. *Improved Algorithms for Isomorphism of Polynomials.* Eurocrypt'98, Springer LNCS 1403, 1998, 184-200.

[Chen03] Q. Cheng. *Primality proving via one round of ECPP and one iteration in AKS.* Crypto 2003, Santa Barbara. Available from *http://www.cs.ou.edu/ qcheng/.*

[Chr05] Christiaan van de Woestijne. *Computing zeros of diagonal forms over finite fields.* PhD thesis, Universiteit Leiden, 2005.

[CK97] Zhi-zhong Chen and Ming Yang Kao. *Reducing Randomness via irrational numbers.* Proceedings of the 29th annual ACM Symposium on Theory of Computing, ACM Press, 1997, 200-209.

[DH88] J. Davenport and J. Heintz. *Real Quantifier Elimination Is Doubly Exponential.* Journal of Symbolic Computation, 5, 1988, 29-35.

[DS05] Zeev Dvir and Amir Shpilka. *Locally Decodable Codes with 2 queries and Polynomial Identity Testing for depth 3 circuits.* Proceedings of the 37th annual ACM Symposium on the Theory of Computing, ACM Press, 2005.

[Evd94] S. A. Evdokimov. *Factorization of polynomials over finite fields in subexponential time under GRH.* Proceedings of the 1994 Algorithmic Number Theory Symposium, 209–219, 1994.

[Fou85] E. Fouvry. *Theoreme de Brun-Titchmarsh; application au theoreme de Fermat.* Inventiones Math., 79:383–407, 1985.

[Gal] E. Galois. *Oeuvres mathmatiques.* Journal des mathmatiques pures et appliques, 11:381-444, 1846.

[GK86] S. Goldwasser and J Kilian. *Almost all primes can be quickly certified.* Proceedings of Annual ACM Symposium on the Theory of Computing, 316–329, 1986.

[GK98] Dima Grigoriev and Marek Karpinski. *An exponential lower bound for depth* 3 *arithmetic circuits.* Proceedings of the 30th annual ACM Symposium on the Theory of Computing, 577–582, 1998.

[GM84] R. Gupta and M. Ram Murty. *A remark on Artin's conjecture.* Inventiones Math., 78:127–130, 1984.

[GMM85] R. Gupta, V. Kumar Murty and M. Ram Murty. *The Euclidean algorithm for S integers.* CMS Conference Proceedings, 189–202, 1985.

[GMR85] S. Goldwasser, S. Micali and C. Rackoff. *Knowledge complexity of interactive proofs.* Proceedings of the 17$^{\text{th}}$ Annual ACM Symposium on Theory of Computing, 291–304, 1985.

[Gran] Andrew Granville. *It is easy to determine whether a given integer is prime.* Bulletin (New Series) of the American Mathematical Society, 42(1): 3-38, 2004.

[GW02] O. Goldreich and A. Wigderson. *Derandomization That Is Rarely Wrong from Short Advice That Is Typically Good.* Proceedings of the 6$^{\text{th}}$ International Workshop on Randomization and Approximation Techniques, 209–223, 2002.

[Har75] D. Harrison. *A Grothendieck ring of higher degree forms.* Journal of Algebra, 35:123–128, 1975.

[Has21] Helmut Hasse. *Über die Darstellbarkeit von Zahlen durch quadratische Formen im Körper der rationalen Zahlen.* PhD thesis, 1921.

[Her26] Grete Hermann. *Die Frage der endlich vielen Schritte in der Theorie der Polynomideale.* Math. Ann. 95:736-788, 1926.

[HB86] D. R. Heath-Brown. *Artin's conjecture for primitive roots.* Quart. J. Math. Oxford, (2) 37:27–38, 1986.

[HP88] D. K. Harrison and B. Pareigis. *Witt rings of higher degree forms.* Communications in Algebra, 16(6):1275–1313, 1988.

[IK03] Russell Impaggliazzo and Valentine Kabanets. *Derandomizing Polynomial Identity Testing means proving circuit lower bounds.* Proceedings of the ACM 35$^{\text{th}}$ annual Symposium on Theory of Computing, 355–364, 2003.

[JS80] Mark Jerrum and Marc Snir. *Some exact complexity results for straight-line computations over semi-rings.* Technical Report CRS-58-80, University of Edinburgh, 1980.

[Kay06] Neeraj Kayal. *Derandomizing some Number-theoretic and Algebraic Problems.* PhD Thesis, IIT Kanpur, 2006.

[Keet93] Arnold Keet. *Higher degree hyperbolic forms.* Quaestiones Mathematicae, 16(4):413–442, 1993.

[Klap89] A. Klapper. *Generalized lowness and highness and probabilistic classes.* Mathematical Systems Theory, 22:37–45, 1989.

[KS01] Adam Klivans and Daniel Spielman. *Randomness efficient identity testing of multivariate polynomials.* Proceedings of the ACM 33$^{\text{rd}}$ annual Symposium on Theory of Computing, 216–223, 2001.

[KS02] Neeraj Kayal and Nitin Saxena. *Towards a deterministic polynomial-time test.* Technical report, IIT Kanpur, 2002. Available at http://www.cse.iitk.ac.in/research/btp2002/primality.html.

[KS05] N. Kayal and N. Saxena. *On the Ring Isomorphism and Automorphism Problems.* IEEE Conference on Computational Complexity, 2–12, 2005.

[KS06] N. Kayal and N. Saxena. *Polynomial Identity Testing for Depth* 3 *Circuits.* IEEE Conference on Computational Complexity, 2006.

[Lang] S. Lang. *Algebra.* 3$^{\text{rd}}$ edition, Addison Wesley.

[Lee90] J. V. Leeuwen, editor. *Handbook of Theoretical Computer Science, Volume A.* Elsevier, 1990.

[Len87] Hendrik W. Lenstra Jr. *Factoring integers with elliptic curves.* Annals of Mathematics **126** (2) 1987, 649-673.

[Len02] H. W. Lenstra Jr. *Primality testing with cyclotomic rings.* Unpublished (http://cr.yp.to/papers.html#aks has an exposition of Lenstra's argument), August 2002.

[Len04] Hendrik Lenstra, Jr. *Private communication.* 2004.

[LFKN92] C. Lund, L. Fortnow, H. Karloff and N. Nisan. *Algebraic methods for interactive proof systems.* Journal of the Association for Computing Machinery, 39(4):859–868, 1992.

[Lov79] L. Lovasz. *On determinants, matchings and random algorithms.* Fundamentals of Computing Theory, edited by L. Budach. Akademia-Verlag, Berlin, 1979.

[Luk82] E. M. Luks. *Isomorphism of Graphs of Bounded Valence Can Be Tested in Polynomial Time.* Journal of Computer and System Sciences, 25:42–49, 1982.

[LL93] Arjen K. Lenstra and Hendrik W. Lenstra Jr. (Eds.) *The development of the number field sieve.* Lecture Notes in Math. 1554. Springer-Verlag, 1993.

[LN86] R. Lidl and H. Niederreiter. *Introduction to finite fields and their applications.* Cambridge University Press, 1986.

[LP03] H.W. Lenstra, Jr. and Carl Pomerance. *Primality testing with Gaussian periods.* Preprint, 2003.

[LP03b] H. W. Lenstra, Jr. and Carl Pomerance. *Remarks on Agrawals conjecture.* Unpublished (http://www.aimath.org/WWN/primesinp/articles/html/50a/), March 2003.

[LV98] Daniel Lewin and Salil Vadhan. *Checking polynomial identities over any field: towards a derandomization?* Proceedings of thirtieth annual ACM Symposium on Theory of Computing, ACM Press, 1998, 438-447.

[Mas84] R. C. Mason. *Diophantine Equations Over Function Fields.* London Mathematical Society Lecture Note Series, 96, Cambridge University Press, 1984.

[MA76] K. Manders and L. Adleman. *NP-complete decision problems for quadratic polynomials.* Proceedings of the 8$^{\text{th}}$ ACM Symposium on Theory of Computing, 1976, 23-29.

[MA04] Preda Mihailescu and Roberto Avanzi. *Efficient quasi - deterministic primality test improving AKS.* Preprint, 2004.

[McD74] Bernard R. McDonald. *Finite Rings with Identity.* Marcel Dekker, Inc., 1974.

[MH74] Y. I. Manin and M. Hazewinkel. *Cubic forms: algebra, geometry, arithmetic.* North-Holland Publishing Co., Amsterdam, 1974.

[Mil76] G. L. Miller. *Riemann's hypothesis and tests for primality.* J. Comput. Sys. Sci., 13:300–317, 1976.

[Minkow] Hermann Minkowski. *Untersuchungen über quadratische Formen, Bestimmung der Anzahl verschiedener Formen, welche ein gegebenes Genus enthält.* PhD thesis, Königsberg, 1885.

[MM82] E. Mayr, A. Meyer. *The complexity of the word problems for commutative semigroups and polynomial ideals.* Adv. Math. 46(3):305–329, 1982.

[MVV87] K. Mulmuley, U. Vazirani and V. Vazirani. *Matching is as easy as matrix inversion.* Combinatorica, 7(1):105–113, 1987.

[NW94] N. Nisan and A. Wigderson. *Hardness vs. Randomness.* J. Comput. System Sci., 49:149–167, 1994.

[NZM91] I. Niven, H. Zuckerman and H. Montgomery. *An Introduction to the Theory of Numbers.* John Wiley & Sons, Inc., 5th edition, 1991.

[Pal93] R. E. A. C. Paley. *Theorems on Polynomials in a Galois Field.* Quarterly Journal of Math., 4:52-63, 1993.

[Pat96] J. Patarin. *Hidden Field Equations (HFE) and Isomorphisms of Polynomials (IP): Two new families of asymmetric algorithms.* Eurocrypt'96, Springer LNCS 1070, 1996, 33-48.

[Pud94] Pavel Pudlak. *Communication in bounded depth circuits.* Combinatorica, 14(2):203-216, 1994.

[Rab80] M. O. Rabin. *Probabilistic algorithm for testing primality.* Journal of Number Theory, 12:128–138, 1980.

[Ron88] L. Ronyai. *Factoring polynomials over finite fields.* Journal of Algorithms, 9:391–400, 1988.

[Ros95] M. I. Rosen. *Niels Henrik Abel and the equation of the fifth degree.* American Mathematics Monthly, 102:495–505, 1995.

[RS01] Ran Raz and Amir Shpilka. *Lower bounds for matrix product, in bounded depth circuits with arbitrary gates.* Proceedings of the 33rd annual ACM Symposium on Theory of Computing, ACM Press, 409–418, 2001.

[RS03] M. O'Ryan and D. B. Shapiro. *Centers of higher degree forms.* Linear Algebra Applications, 371:301–314, 2003.

[RS04] Ran Raz and Amir Shpilka. *Deterministic Polynomial identity testing in noncommutative models.* Conference on Computational Complexity, 2004.

[Rup03] C. Rupprecht. *Cohomological invariants for higher degree forms.* PhD Thesis, Universität Regensburg, 2003.

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

[Sch88] U. Schoning. *Graph isomorphism is in the low hierarchy.* Journal of Computer and System Science, **37**, 1988, 312-323.

[Serre] Jean-Pierre Serre. *A course in arithmetic.* Springer-Verlag, New York, NY, 1973.

[Sha92] Adi Shamir. *IP=PSPACE.* Journal of the Association for Computing Machinery, 39(4):869–877, 1992.

[Shou92] Victor Shoup. *Searching for primitive roots in finite fields.* Mathematics of Computation, 58:369–380, 1992.

[ShS77] Eli Shamir and Marc Snir. *Lower bounds on the number of multiplications and the number of additions in monotone computations.* Research Report RC6757, IBM Thomas J. Watson Research Center, Yorktown Heights, N.Y., 1977.

[ShS80] Eli Shamir and Marc Snir. *On the depth complexity of formulas.* Mathematical Systems Theory, 13:301-322, 1980.

[SoS77] R. Solovay and V. Strassen. *A fast Monte-Carlo test for primality.* SIAM Journal on Computing, 6:84–86, 1977.

[Sto81] W. W. Stothers. *Polynomial identities and Hauptmoduln.* Quarterly Journal of Mathematics, Oxford Second Series, 32:349-370, 1981.

[SW99] Amir Shpilka and Avi Wigderson. *Depth-3 arithmetic formulae over fields of characteristic zero.* Proceedings of the 14^{th} annual IEEE Conference on Computational Complexity, 87–96, 1999.

[TT94] Prasoon Tewari and Martin Tompa. *A direct version of Shamir and Snir's lower bounds on monotone circuit depth.* Information Processing Letters, 49(5):243-248, 1994.

[vzGG99] Joachim von zur Gathen and Jürgen Gerhard. *Modern Computer Algebra.* Cambridge University Press, 1999.

[Witt] Ernst Witt and Ina Kersten (Editor). *Collected Papers - Gesammelte Abhandlungen.* Springer, 1st edition, October 1, 1998.

[Zip79] Richard E. Zippel. *Probabilistic algorithms for sparse polynomials.* Proceedings of the International Symposium on Symbolic and Algebraic Computation, Springer Verlag, 216–226, 1979.

Index