Contents

1 Lecture 1 ................................................................. 6
   1.1 Introduction to this course ..................................... 6
      1.1.1 The Study of Structures .................................. 6
      1.1.2 What is Number Theory about? .......................... 7
   1.2 Notations .......................................................... 10
   1.3 Basic Structures over $\mathbb{Z}$ ................................. 11
   1.4 Divisibility ....................................................... 13
      1.4.1 Definitions and Properties ................................. 13

2 Lecture 2 ................................................................. 15
   2.1 Divisibility ....................................................... 15
      2.1.1 Well-ordering Axiom Implies Mathematical Induction Is Valid ....................................................... 15
      2.1.2 Greatest Common Divisor .................................. 17
   2.2 Euclidean Algorithm .............................................. 17
   2.3 Some Rules on Divisibility ..................................... 19

3 Lecture 3 ................................................................. 21
   3.1 Warm-up ........................................................... 21
   3.2 The General Structure of This Course .......................... 21
   3.3 Divisibility(Continued) ........................................ 22
   3.4 Ideals in $\mathbb{Z}$ ................................................. 23
   3.5 Prime Numbers .................................................... 24
   3.6 Some Words about the HW ..................................... 26

4 Lecture 4 ................................................................. 27
   4.1 Warm-up ........................................................... 27
   4.2 Prime Numbers(Continued) .................................... 27
      4.2.1 Some Factorization Techniques ........................... 28
      4.2.2 Euler’s Product Formula ................................... 30
   4.3 Least Common Multiple ......................................... 30
CONTENTS

5 Lecture 5 32
5.1 Warm-up .................................................. 32
5.2 Where we are heading towards ......................... 32
5.3 Linear Diophantine Equations ......................... 32
    5.3.1 The 2-variable Case ............................ 33
    5.3.2 The General Case ............................... 35

6 Lecture 6 38
6.1 Problem Session ......................................... 38
6.2 \(\mathbb{Z}/k\mathbb{Z}\) as a Commutative Ring .............. 40

7 Lecture 7 42
7.1 Motivation for Problem 10, HW2 ....................... 42
7.2 Units in \(\mathbb{Z}[(\sqrt{n})] \) .............................. 43
7.3 \(\mathbb{Z}/b\mathbb{Z}\) as a commutative ring (continued) .... 43
    7.3.1 Basic Structure on \(\mathbb{Z}/b\mathbb{Z}\) ................. 43
    7.3.2 Relation between \(\mathbb{Z}\) and \(\mathbb{Z}/b\mathbb{Z}\) ......... 44
    7.3.3 Comparing different \(\mathbb{Z}/b\mathbb{Z}\) .................. 45
    7.3.4 Relation between \(\mathbb{Z}/m\mathbb{Z}\) and \(\mathbb{Z}/n\mathbb{Z}\) when \(m|n\) .......... 46
    7.3.5 Arithmetic Tables .................................. 47
7.4 Some “Geometry” ......................................... 47

8 Lecture 8 49
8.1 Warm-up ................................................... 49
8.2 Some Geometry (continued) ............................. 49
    8.2.1 Relations among \(\mathbb{Z}, \mathbb{Z}_p\) and \(\mathbb{Z}/p\mathbb{Z}\) .... 51
8.3 Euler \(\varphi\)-function ................................... 51

9 Lecture 9 55
9.1 Problem Session .......................................... 55
9.2 Euler \(\varphi\)-function (continued) ....................... 56

10 Lecture 10 60
10.1 Problem Session ......................................... 60
10.2 Arithmetic Progression: a first look .................. 61
    10.2.1 Simplest Cases .................................. 62
    10.2.2 Less Simple Case ............................... 63

11 Lecture 11 65
11.1 Warm-up .................................................. 65
11.2 Arithmetic Progression (Continued) .................... 65
11.3 Linear Congruence Equations ......................... 66
    11.3.1 The Simple Case ................................ 66
    11.3.2 A Small Generalization ....................... 67
    11.3.3 The General Case ............................... 67
11.4 Chinese Remainder Theorem ............................ 68
CONTENTS

12 Lecture 12 70
  12.1 Warm-up ........................................ 70
  12.2 Chinese Remainder Theorem .................. 70
  12.3 Linear Algebra over $\mathbb{Z}/m\mathbb{Z}$ .......... 74

13 Lecture 13 76
  13.1 Application to Cryptography .................. 76
    13.1.1 Caesar Cipher .............................. 76
    13.1.2 Exponentiation Cipher ..................... 77
    13.1.3 RSA Cryptosystem .......................... 78
    13.1.4 The Knapsack Cipher ....................... 79
    13.1.5 Summary ................................... 81

14 Lecture 14 82
  14.1 Arithmetic Progression: a Summary .......... 82
  14.2 Euler’s Theorem, Fermat’s Little Theorem, Order of Units etc. 83
  14.3 Computations .................................. 84
  14.4 Greatest Common Divisor ...................... 85
  14.5 Divisibility Rule for $m = 29$ ............... 86

15 Lecture 15 87
  15.1 A Quick Summary ................................ 87
  15.2 High Degree Congruence Equations ............ 88
    15.2.1 $x^p - x$ over $\mathbb{Z}/p\mathbb{Z}$ .......... 89
    15.2.2 Greatest Common Divisor of Two polynomials 90
  15.3 Factorization of Polynomials .................. 91

16 Lecture 16 94
  16.1 Analogy between $\mathbb{Z}$ and $(\mathbb{Z}/p\mathbb{Z})[x]$ .... 94
  16.2 Irreducible Polynomials over $\mathbb{Z}/p\mathbb{Z}$ ........ 95
  16.3 Application 1 .................................. 96
  16.4 Application 2: Solving all Univariate Congruence Equations . 98

17 Lecture 17 99
  17.1 Warm-up ........................................ 99
  17.2 Reduction to Modulo $p$ ........................ 100

18 Lecture 18 104
  18.1 Warm-up ........................................ 104
  18.2 “Global vs Local” Revisited ................... 105
  18.3 Multiple Roots ................................. 106
  18.4 Quadratic Residues ............................. 108

19 Lecture 19 110
  19.1 Warm-up ........................................ 110
  19.2 Quadratic Residues (Continued) ............... 110
  19.3 Quadratic Reciprocity ......................... 113
## CONTENTS

### Lecture 20
- 20.1 Warm-up .......................... 115
- 20.2 Quadratic Reciprocity .......... 115
- 20.3 Proof of the Quadratic Reciprocity Law . 116

### Lecture 21
- 21.1 Warm-up .......................... 120
- 21.2 More Examples on Quadratic Congruences . . . . 122
- 21.3 More Examples .................. 123

### Lecture 22
- 22.1 More Examples .................. 125

### Problem Sets
- A.1 Problem Set 1 .................. 129
- A.2 Problem Set 2 .................. 130
- A.3 Problem Set 3 .................. 131
- A.4 Problem Set 4 .................. 131
- A.5 Problem Set 5 .................. 132
- A.6 Problem Set 6 .................. 133
- A.7 Problem Set 7 .................. 134
- A.8 Problem Set 8 .................. 134
- A.9 Problem Set 9 .................. 135

### Extra Credit Problem Sets
- B.1 Problem Set 1 .................. 136
- B.2 Problem Set 2 .................. 137
- B.3 Problem Set 3 .................. 139
- B.4 Problem Set 4 .................. 139
- B.5 Problem Set 5 .................. 140
Chapter 1

Lecture 1

1.1 Introduction to this course

Structures are the weapons of the mathematician.

Nicolas Bourbaki

1.1.1 The Study of Structures

As many of you might know, number theory is the study of integers. This is a good starting point, but what do we mean by "study"? It is certainly NOT enough to just know what all the integers are. Otherwise, the course would be over before it starts... Rather, we try to understand the structures among the integers. Here are some most important examples of structures:

1. Arithmetic, algebraic structures: one can do addition and multiplication of integers, and still get an integer; the negative of an integer is still an integer, while the reciprocal of an integer is not an integer (unless for ±1).

\[ 0 + n = n, \quad 1 \cdot n = n. \]  

Associativity, commutativity laws for addition and multiplication...

2. Distance: one has the usual absolute value \(|\cdot|\). One can say whether \(a\) and \(b\) are "close" to each other. One also has the \(p\)-adic absolute value \(|\cdot|_p\): if \(n = p^m q\), where \(q\) is relatively prime to \(p\), then \(|n|_p = (1/p)^m\); roughly, it measures "how divisible" \(n\) is, by the prime number \(p\).

3. Order: one can say an integer is larger (or smaller) than another integer.

In particular, every set of positive integers contains a smallest element.

It turns out that these structures, which we are so familiar with ever since elementary school, can lead to results about integers which are totally non-obvious and even very deep. In establishing these results, mathematicians have
developed very usual languages/formalisms/auxiliary tools, which in return have broadened the scope of mathematics itself. In the next section, we’ll list some of those results. They should give you a flavor about what kind of information mathematicians quest for in the study of number theory.

**eg.** Other structures you might have seen before:

- **topological structure:** characterizes shapes of spaces, eg. what are open/closed sets, how many holes a surface has, whether a space has boundaries. For example, in \( \mathbb{R} \), we call \((0,1)\) an *open* interval and \([0,1]\) a *closed* interval. (But we can also declare other topologies on \( \mathbb{R} \) so that \((0,1)\) is no longer open in the new sense!)

- **geometric structure:** gives measurement of angles and distances. An example you have seen before is the dot product in Calculus.

- **differentiable structure:** allows people to do calculus on certain kind of spaces.

- **complex structure:** talks about complex differentiable functions on spaces.

### 1.1.2 What is Number Theory about?

In this section, let’s take a quick glimpse on some of the important topics in number theory. The purpose of this section is only to give you a flavor of what number theory is about. We will certainly not cover everything listed here, but we will talk about some the basic results as we go along the course.

#### Factorization of Integers

One of the most important topic in elementary number theory is the factorization of integers. The following statements are all related to this topic. Related notions are “divisibility” and “congruence”. Take any integer \( n \), we can factorize them as product of prime numbers. One may imagine those primes as the building blocks of integer numbers. This phenomenon relates different integers in some way: given two integers \( a, b \), one may ask whether \( a \) is a multiple of \( b \) (*"b divides a"*), or vice versa; if neither \( a \) divides \( b \) or \( b \) divides \( a \), one can further ask “how far” \( a \) is away from dividing \( b \), using the language of congruence; similarly, one can also ask a slightly generalized question: whether \( a, b \) only differ by some power of a prime number \( p \).

Here are some results related to factorization of integers.

**Proposition 1.** Every positive integer can be uniquely written as a product of prime factors.

**Proposition 2.** \((a, b) = \min\{ma + nb | m, n \in \mathbb{Z}, ma + nb \geq 0\}\). \((a, b)\) is the greatest common divisor of \( a \) and \( b \). See the section on notations.

**Proposition 3.** \(\{ma + nb | m, n \in \mathbb{Z}\} = ((a, b))\). \(((a, b))\) is the set of all multiples of \( n \). See the section on notations.
Remark These results actually shed light on the (less obvious) algebraic structure of \( \mathbb{Z} \). In the language of abstract algebra, we say \( \mathbb{Z} \) is a “principal ideal domain”. As a consequence, it is a “unique factorization domain”, which is a fancy way to state the fact that every positive integer uniquely factorizes as a product of prime factors. Since this is a number theory class, our main focus will not be the abstract algebra language in this course.

Distribution of Prime Numbers

The distribution of primes is another important topic in number theory:

\[ 2, 3, 5, 7, 11, 13, 17, 19, 23, 29, 31, 37... \]

Is there any pattern here? At first glance, it might be even quite unclear why one can even approach such a question. Whether a number is prime or not is almost given to us as a mysterious fact, ever since humans start to count. However, mathematicians have achieved some deep understanding on this topic. Here are some results related to the distribution of primes:

Proposition 4. There are infinitely many prime numbers.

Proposition 5. For any integer \( n \) that is not prime, it has a prime factor less or equal than \( \sqrt{n} \).

Proposition 6. (Euler) \( \sum \left( \frac{1}{p} \right) = \infty \), i.e. the sum of the reciprocals of all prime numbers goes to infinity.

Proposition 7. (Chebyshev) For \( x > 2 \), the number of primes that are less than \( x \) is between \( \frac{1}{10} \left( \frac{x}{\log x} \right) \) and \( 10 \left( \frac{x}{\log x} \right) \).

Remark We are only listing some basic results here. Indeed, the distribution of primes is a very deep topic. The famous Riemann Hypothesis (see §1.2.4) is regarded as the ultimate question in this field. It remains open to this day.

The big picture is the following: if impossible to find a simple pattern to describe the infinitely many prime numbers, the next best thing to do is to study the asymptotic behavior of prime numbers, i.e. the distribution of prime numbers, when the numbers gradually approach infinity.

Diophantine Equations

An important viewpoint in the study of integers is to look for integer solutions to a given polynomial equation. These equations, for which we only allow integer solutions, are called **diophantine equations**.

Eg. The most familiar example of a Diophantine equation is probably \( x^2 + y^2 = z^2 \). Looking for integer solutions \((x, y, z)\) of this equation, is the same as looking for all triples of integers that can form the three sides of a right triangle. We refer to them as Pythagorean triples.
1.1. INTRODUCTION TO THIS COURSE

eg. The previous example generalizes easily to the following form \( x^n + y^n = z^n \) for \( n \geq 3 \). Surprisingly, this leads to one of the most famous problem in mathematical history: Fermat’s last theorem. (See §1.2.4.)

eg. Pell’s Equation: \( x^2 - ny^2 = 1 \). Notice that this expression gives a family of quadratic equations, i.e. for different \( n \), the integer solutions may behave differently. An important observation is that, the equation can be written as \((x + \sqrt{n}y)(x - \sqrt{n}y) = 1\). This motivates the study numbers of the form \( a + b\sqrt{n} \), where \( a, b \) are integers. It turns out that the set \( \mathbb{Z}[\sqrt{n}] = \{a + b\sqrt{n} | a, b \in \mathbb{Z}\} \) shares similar algebraic structure as \( \mathbb{Z} \).

Remark Another viewpoint towards Pell’s equation: take \( n = 3 \) as an example, and re-write the equation as \( \frac{x^2}{y^2} - 3 = \frac{1}{y^2} \). Suppose we found infinitely many integer solutions \((x_m, y_m)\) such that \( |y_m| \to \infty \), then \( \{\frac{x_m}{y_m}\} \) form an approximation of the irrational number \( \sqrt{3} \), by rational numbers. This might be the original motivation for ancient people to study this equation thousands years ago.

Proposition 8. (Fermat) Every positive integer can be written as the sum of the squares of 4 integers.

In other words, \( x^2 + y^2 + z^2 + w^2 = n \), where \( n > 0 \) has an integer solution.

eg. Here is a quick example showing you that algebraic question are related to geometry: consider the line \( y = \frac{1}{\pi}x \) in \( \mathbb{R}^2 \). Since the slope \( 1/\pi \) is a transcendental number, there are no integer points on this line, other than \((0,0)\).

Some Deep Results or Open Problems in Number Theory

To show the beauty of number theory, we also list some results whose difficulty are way beyond the level our class.

Conjecture 1. (The Twin Prime Conjecture) There are infinitely many pairs of consecutive primes \( p, q \) such that \( q = p + 2 \).

Maybe the most exciting progress in number theory in the year 2013 is related to the Twin Prime conjecture. Earlier last year, Yitang Zhang proved that there are infinitely pair of consecutive primes \( p, q \), such that \( q - p \leq 70 \) million. The current progress on this problem shows that this gap has been reduced to 4680. See [http://michaelnielsen.org/polymath1/index.php?title=Bounded_gaps_between_primes](http://michaelnielsen.org/polymath1/index.php?title=Bounded_gaps_between_primes) for more detail.

The story of the mathematician Yitang Zhang is very inspiring. According to Wikipedia, “he worked for several years as an accountant, a delivery worker for a New York City restaurant, in a motel in Kentucky and in a Subway sandwich shop before working as a lecturer.” Check out [http://en.wikipedia.org/wiki/Yitang_Zhang](http://en.wikipedia.org/wiki/Yitang_Zhang) for more detail.

Theorem 1. Suppose \( n \geq 3 \). There do NOT exist three positive integers \( a, b, c \) such that \( a^n + b^n = c^n \).
This was probably one of the most well-known problems in the history of mathematics. When Andrew Wiles eventually proved the conjecture in the 1990s, this problem has remained as a conjecture for more than 300 years! An important feature to the problem is that, it is extremely simple to state the result, but the proof uses very advanced machinery in modern algebraic number theory. This is also one good justification for why one should care about some seemingly esoteric algebraic structures: they really lead to something deep!

Conjecture 2. (Goldbach’s Conjecture) Every even integer greater than 2 can be written as a sum of two prime numbers.

Again, a very simple statement, but we don’t know whether it is true in general yet. The point I want to make is, in mathematics, there are lots of plausible simple statements which are difficult to prove. The fact that they are hard to prove usually suggests that there are some highly non-obvious structures hiding under the rug. It is very possible that elementary approaches will simply never solve some of the remaining open problems. Hopefully, this motivates you to climb the mountain of modern mathematics.

We shall end our list of examples with the statement of Riemann Hypothesis. If we have time, the content of this hypothesis will be explained to you later. Since this requires some complex analysis, you don’t need to worry about it at all (for HW, exam etc.).

Conjecture 3. (Riemann Hypothesis) Consider the Riemann Zeta function, defined over the complex numbers \( \mathbb{C} \) except for \( s = 1 \), which is the analytic continuation of \( \zeta(s) = \sum_{n=1}^{\infty} \frac{1}{n^s} \). The real part of any non-trivial zero of the Riemann zeta function is 1/2.

This complex-variable function is related to the distribution of prime numbers via the following identity: \( \sum_{n=1}^{\infty} \frac{1}{n^s} = \prod_{p} \frac{1}{1 - p^{-s}} \) (\( Re(s) > 1 \)), where the product on the right-hand side runs through all prime numbers.

1.2 Notations

Hereafter, we’ll fix the following set of notations. If you don’t know some of the terminologies here mentioned, that’s perfectly fine. They will be defined as the course goes along. We list them here basically for your future convenience.

1. Most frequently, we use three logic symbols: \( \forall \) means "for any XXX", \( \exists \) means "there exists XXX", and \( \in \) means "...belongs to...". \( A \cap B \) is the intersection of two sets i.e. those elements that are both in \( A \) and in \( B \); and \( A \cup B \) is the union of two sets, i.e. those elements that are either in \( A \) or in \( B \).
2. We’ll use $\mathbb{Z}$ for the set of all integers. $\mathbb{Z} = \{0, \pm 1, \pm 2, \ldots \}$. $\mathbb{N}$ stands for natural numbers, or positive integers: $\mathbb{N} = \{1, 2, 3, 4, \ldots \}$

3. $\mathbb{Q}$ will stand for the set of rational numbers, i.e. numbers of the form $a, b$ where $a, b \in \mathbb{Z}$ and $b \neq 0$. Similarly, $\mathbb{R}$ stands for real numbers, and $\mathbb{C}$ for complex numbers.

4. $\langle n \rangle = \{0, \pm n, \pm 2n, \pm 3n, \ldots \}$. For example, $\langle 2 \rangle = \{0, \pm 2, \pm 4, \ldots \}$ is the set of all even integers.

5. $p$ usually stands for a prime number, and we use other letters such as $a, b, c, n, m$ etc. for arbitrary integers, which are not necessarily prime.

6. $[x]$ is the largest integer less than or equal to a given real number $x$.

7. $(a_1, \ldots, a_n)$ stands for the greatest common divisor of the integers $a_1, \ldots, a_n$.

8. $a|b$ reads as "$a$ divides $b$", and means that there exists another integer $c$ such that $b = ac$.

9. $a \equiv b \pmod{n}$, means that $n|(a - b)$, and reads as $a$ is congruent to $b$ modulo $n$.

10. $\mathbb{Z}[i]$ will be the set of complex numbers of the form $a + bi$, where $a, b \in \mathbb{Z}$, and $i = \sqrt{-1}$.

11. $\mathbb{Z}[\sqrt{n}] = \{a + b\sqrt{n} | a, b \in \mathbb{Z}\}$.

12. As a set, $\mathbb{Z}/p\mathbb{Z}$ can be thought of as $\{0, 1, 2, \ldots, p-1\}$. As we will see in the course, $\mathbb{Z}/p\mathbb{Z}$ is related to integers modulo a prime number $p$. Naively, say $p = 2$, $\mathbb{Z}/2\mathbb{Z} = \{0, 1\}$, and can be interpreted as "every integer $n$ is either odd or even", i.e. it can be either written as $2k + 1$ or $2k$.

1.3 Basic Structures over $\mathbb{Z}$

Let’s start by reviewing some maths from elementary school.

First of all, we have all learned how to do addition in integers. Moreover, we know the addition + satisfies the following properties:


2. Associativity: $(a + b) + c = a + (b + c)$

3. There is a "special" integer 0 such that $0 + a = a + 0 = a$ for all integers $a$.

4. For any integer $a$, there exists another integer $b$ such that $a + b = b + a = 0$. Clearly $b$ is just $-a$. 

We are all very familiar with these nice properties. They characterize some algebraic structure over \( \mathbb{Z} \). To refer to this structure, we make the following definition:

**Definition 1.** Let \( G \) be a set. Suppose there is a binary operation \(+ : G \times G \rightarrow G\) such that the following are true:

1. **Commutativity:** \( a + b = b + a \).
2. **Associativity:** \((a + b) + c = a + (b + c)\)
3. \( \exists \) some element \( 0 \) in \( G \), such that \( 0 + a = a + 0 = a \). This is usually called the zero element of \( G \).
4. For any \( a \in G \), there exists another element \( b \in G \) such that \( a + b = b + a = 0 \).

Then, \( G \) is called an abelian group with respect to the binary operation \(+\). We usually write \((G, +)\) to emphasize the binary operation we are considering.

**Remark** A “binary operation” simply means that it takes in a pair of elements from the set each time, and spits out single element in this set. Since all the defining properties are extracted from \( \mathbb{Z} \) with respect to the usual addition \( + \), \((\mathbb{Z}, +)\) is an abelian group.

One might wonder why we bother to introduce this abstract definition. The following example is motivational:

**eg.** Verify that \((\mathbb{Z}[i], +), (\mathbb{Z}[\sqrt{3}], +)\) are abelian groups. In both cases, additions are defined in the natural way.

The point is that, in number theory, \( \mathbb{Z} \) is the original object of interest. But to understand \( \mathbb{Z} \), it is often useful to understand other sets of (real or complex) numbers. See the example of Pell’s equation. Hence, it is natural that we look forward to generalizing our framework later.

**eg.** Verify that \(((\mathbb{N}), +)\) is an abelian group. Since it consists of a subset of \( \mathbb{Z} \), we call it a sub-group of \((\mathbb{Z}, +)\).

We also know how to do multiplications in \( \mathbb{Z} \). In fact, the multiplication (we’ll use \( \cdot \) instead of \( \times \), or simply omit the symbol) operation satisfies the following properties:

1. **Associativity:** \((a \cdot b) \cdot c = a \cdot (b \cdot c)\).
2. There is a special number \( 1 \), such that \( 1 \cdot a = a \cdot 1 = a \), for all integers \( a \).
1.4. DIVISIBILITY

Moreover, $+$ and $\cdot$ interact in a nice way, namely we have the distributive laws:

1. $(a + b) \cdot c = a \cdot c + b \cdot c.$
2. $a \cdot (b + c) = a \cdot b + a \cdot c.$

In short, $\mathbb{Z}$ is not only an abelian group with respect to usual addition, it has also some other good structure coming from the usual multiplication. We pack up our observation in the following definition:

**Definition 2.** Suppose $R$ is a set with two binary operations, denoted as $+$ and $\cdot$. If they satisfy the following conditions:

1. $(R, +)$ is an abelian group.
2. The other binary operation $\cdot$ satisfies associativity: $(a \cdot b) \cdot c = a \cdot (b \cdot c)$.
3. $\exists$ some element $e \in R$ such that $e \cdot a = a$, for all $a \in R$. This element is called the multiplicative unit of $R$, sometimes also denoted as $1_R$ or simply $1$.
4. The two operations satisfy the distributive laws: $(a + b) \cdot c = a \cdot c + b \cdot c$; $a \cdot (b + c) = a \cdot b + a \cdot c$.

Then, $R$ is called a ring. To emphasize the two operations, we write $(R, +, \cdot)$.

**Remark** $\mathbb{Z}$ is a ring with respect to the usual addition and multiplication.

However, we shouldn’t forget that multiplication of integers is also commutative. $a \cdot b = b \cdot a$. This further motivates the following definition:

**Definition 3.** Let $(R, +, \cdot)$ be a ring. (By definition, $(R, +)$ is an abelian group, so the operation $+$ is always commutative.) Suppose the operation $\cdot$ is also commutative, then we say $(R, +, \cdot)$ is a commutative ring.

**eg.** Verify that $\mathbb{Z}[\sqrt{3}]$ and $\mathbb{Z}[i]$ are commutative rings with respect to the usual addition and multiplication of complex numbers.

Thus, we have summarized the basic algebraic structures over $\mathbb{Z}$. Later, when we study congruences, we’ll see the convenience of the language we hereby introduced. In particular, we’ll show that $\mathbb{Z}/p\mathbb{Z} = \{0, 1, ..., p - 1\}$ is a commutative ring, and its structure encodes the information of congruence arithmetic.

**1.4 Divisibility**

**1.4.1 Definitions and Properties**

One first major topic in this course is factorization of integers. Divisibility is the preliminary language for talking about factorizations.
Definition 4. Suppose $a, b$ are two integers. We say $a$ divides $b$, and write $a|b$, if there exists another integer $c$ such that $a \cdot c = b$. In this case, we also say “$a$ is a divisor of $b$”, and “$b$ is a multiple of $a$”.
Chapter 2

Lecture 2

2.1 Divisibility

**Example.** Every non-zero integer divides 0. The only positive divisor of 1 is 1. If $a$ is a positive divisor of another positive number $b$, then $b - a \geq 0$.

Here are some simple facts:

**Lemma 1.** (1) Let $a, b, c$ be 3 integers. If $a | b$, $b | c$, then $a | c$.
(2) Let $a, b, c$ be 3 integers. Suppose $a | b$ and $a | c$, then for any two other integers $m, n$, $a | (mb + nc)$.

**Proof.** (1) By definition, $a | b$ means there is some integer $\ell$ such that $a \cdot \ell_1 = b$. Similarly, there is some integer $\ell_2$ such that $b \cdot \ell_2 = c$. Replace $b$ by $a \cdot \ell_1$, we get $a \cdot (\ell_1 \ell_2) = c$. By definition, $a | c$.
(2) Still use the definition, $\exists \ell_1, \ell_2 \in \mathbb{Z}$ such that $a\ell_1 = b, a\ell_2 = c$. Hence, $a\ell_1 m + a\ell_2 n = mb + nc$. By distributive law, $a \cdot (\ell_1 m + \ell_2 n) = mb + nc$. By definition, $a | (mb + nc)$.

2.1.1 Well-ordering Axiom Implies Mathematical Induction Is Valid

**Warning:** if you are not interested in theoretical subtlety, please skip next paragraph. However, make sure read the content of the well-ordering axiom.

Aside: When talking about the order structure, we mentioned that every set $S$ of positive integers has a smallest element. This seems an obvious statement at first sight: take any element $a \in S$, if it is not the smallest one, then there are only finitely many integers between 1 and $a$: 1, 2, ..., $a - 1$; at least one of them
is in $S$ and hence after finitely many steps we should arrive at the smallest element in $S$. However, this argument is not rigorous. In fact, we have never really defined $\mathbb{Z}!$ (In fact, axiomatically, $\mathbb{Z}$ is characterized by the Well-ordering Property, together with other properties we discussed earlier.) The upshot is that, in our convention, we treat it as an axiom. From now on, you can apply it directly in your work.

**Axiom 1.** (Well-ordering) Let $S \subset \mathbb{N}$ be any non-empty set of positive integers. Then, $\exists s_0 \in S$ such that $\forall s \in S$, either $s > s_0$ or $s = s_0$.

This axiom implies the method of mathematical induction is valid.

**Proposition 9.** Let $P_1, \ldots, P_n, \ldots$ be a sequence of statements indexed by $\mathbb{N}$. Suppose:

1. $P_1$ is true.
2. If $P_{n-1}$ is true, then $P_n$ is true.

Then, $P_i$ is true for all $i$.

**Proof.** Notice that the index set is $\mathbb{N}$, so axiom 1 applies. Suppose for some $i$, $p_i$ is not true. By axiom 1, there is a smallest number $n$ such that $p_n$ is not true. But either $n = 1$ or $p_{n-1}$ is true. This violates the given conditions. Therefore, such $p_i$ does not exist, hence all $p_i$ are true.

Another version of mathematical induction:

**Proposition 10.** Let $P_1, \ldots, P_n, \ldots$ be a sequence of statements indexed by $\mathbb{N}$. Suppose:

1. $P_1$ is true.
2. If $P_1, \ldots, P_{n-1}$ are true, then $P_n$ is true.

Then, $P_i$ is true for all $i$.

**Proof.** Similar proof applies.

**Remark** Just bear this axiom in mind. In practice, feel free to use the description $\mathbb{Z} = \{0, \pm 1, \pm 2, \ldots\}$.

After stating this axiom, we can establish a first important result in factorization of integers. It is nothing fancy, and may have already occurred to you as “division with remainder”.

**Proposition 11.** (division with remainder) Let $a, b$ be two integers, and $b > 0$. Then, there exists unique integers $q, r$ such that $a = q \cdot b + r$, and $0 \leq r < b$.\(^1\)

\(^1\)The ring of integers $\mathbb{Z}$ is the unique commutative ordered ring whose positive elements are well-ordered. Have fun.
2.2. EUCLIDEAN ALGORITHM

Proof. It turn out that \( r \) is the smallest positive number such that \( r \) can be written in the form \( a - qb \).

(Existence) We consider the set of \( S \) all positive integers of the form \( a - qb \), for some \( q \): \( S = \{a - qb | a - qb > 0, q \in \mathbb{Z} \} \). Axiom 1 tells us that this set has a smallest element. Call it \( r \), and \( r = a - qb \) for some \( q \). Clearly \( 0 \leq r < b \), otherwise \( r - b > 0 \) and \( r - b \) can also be written as \( a - (q+1)b \). So, \( r - b \in S \), but it is smaller than \( r \)! This contradicts the fact that \( r \) is the smallest element in \( S \). Hence, the pair \( (q, r) \) satisfies our requirements.

(Uniqueness) Proof by contradiction. Suppose there is another pair \( (q', r') \) which also satisfies our requirement. Then \( r' \) can be written as \( a - q'b \), and is also in \( S \). By our choice of \( r \), \( r' \geq r \). It cannot be bigger than \( r \), because \( r' - r = (q - q')b \), i.e. \( b \) is a positive divisor of another positive number \( r' - r \), i.e. \( r' - r \geq b \). But we required \( r' < b \). Contradiction! Hence, must have \( r' = r \) and therefore \( q = q' \).

Remark 1. \( r \) is just the remainder term of \( a \) divided by \( b \): e.g. take \( a = 51, b = 10 \), then \( 51 = 5 \cdot 10 + 1 \), so in this case \( q = 10 \) and \( r = 1 \).
2. The assumption that \( a, b \) are positive is just for convenience. In fact, as long as \( b \neq 0 \), we are fine. (We don’t divide a number by 0!) If \( b \) is negative, \( -b \) is positive, apply the above result to \( a, -b \), we could get \( a = q(-b) + r \), for some unique \( q \) and \( r \) where \( 0 \leq r < |b| \). So one can re-state the result as follows:

**Proposition 12.** Let \( a, b \) be two integers. Then, there exists unique integers \( q, r \) such that \( a = q \cdot b + r \), and \( 0 \leq r < |b| \).

### 2.1.2 Greatest Common Divisor

One important notion in divisibility is greatest common divisor of a finite set of integers.

**Definition 5.** Let \( a_1, ..., a_n \) be integers, and at least one of them is not zero. Then greatest common divisor of \( a_1, ..., a_n \), denoted as \((a_1, ..., a_n)\) is the largest integer that divides all of them.

**Remark** Most often, we only talk about the greatest common divisor of two non-zero integers.

**Lemma 2.** If \( a|b \), then \( (a, b) = |a| \).

**Proof.** \(|a|\) is the largest number dividing \( a \), and it also divides \( b \).

**Definition 6.** If \((a, b) = 1\), then we say “\(a, b\) are relatively prime”, or “\(a, b\) are coprime”.

### 2.2 Euclidean Algorithm

In previous section, we’ve defined the notion of greatest common divisor. The next proposition gives you a theoretic description of the the greatest common divisor.
Proposition 13. Let $a,b$ be two positive integers. Their greatest common divisor $(a,b)$ is the smallest positive number in the set $S = \{ma + nb \mid m,n \in \mathbb{Z}\}$, i.e. the set of all integers that can be written as $m \cdot a + n \cdot b$, for some $m,n \in \mathbb{Z}$. Moreover, $S = ((a,b))$.

Proof. In previous section, we have seen that $a$ can be written as $q \cdot b + r_1$, where $0 \leq r_1 < b$. Hence, $r_1 \in S$.

If $r_1 = 0$, terminate the process. This means $b|a$. In that case, it is obvious that $(a,b) = b$.

Otherwise, write $b = q_1 \cdot r_1 + r_2$, where $0 \leq r_2 < r_1$. Notice that $r_2 = b - q_1 \cdot r_1 = b - q_1 \cdot (a - q \cdot b) = (q_1 - a) \cdot (1 + q_1 \cdot q)b$, so $r_2 \in S$. If $r_2 = 0$, terminate the process.

Otherwise, run the process for $r_1$ and $r_2$ again, and get $r_1 = q_2r_2 + r_3$, where $0 \leq r_3 < r_2$...

$$a = q \cdot b + r_1, 0 \leq r_1 < b$$

$$b = q_1 \cdot r_1 + r_2, 0 \leq r_2 < r_1$$

$$\ldots$$

$$r_{n-3} = q_{n-2}r_{n-2} + r_{n-1}, 0 \leq r_{n-1} < r_{n-2}$$

$$r_{n-2} = q_{n-1} \cdot r_{n-1}$$

In general, there is some $n$ such that after doing $n$ times the same process, one gets $r_n = 0$. (Essentially, this is guaranteed by the well-ordering axiom.) We claim that $r_{n-1}$ is the smallest element in $S$, and it equals $(a,b)$.

**Step 1:** $r_{n-1}|(a,r_{n-1})$. Notice that $q_{n-1}r_{n-1} = r_{n-2}$ and $r_{n-3} = q_{n-1}r_{n-2} + r_{n-1}$ together implies that $r_{n-1}|r_{n-3}$. This further implies that $r_{n-1}|r_{n-4}$. By mathematical induction, $r_{n-1}|a,r_{n-1}b$. So, it is a common divisor of $a$ and $b$ (not yet proven to be the greatest).

**Step 2:** $r_{n-1}$ is the smallest positive number in $S$. Since $r_{n-1}|a,r_{n-1}b$, it divides any element in $S$. If there exists some positive number $r' \in S$ such that $r' < r_{n-1}$, then $r'|r_{n-1}$ will lead to contradiction.

**Step 3:** $r_{n-1} = (a,b)$. We notice that for any integer $r''$ such that $r''|a,r''|b$, must have $r''|(ma + nb)$ for all $m,n \in \mathbb{Z}$. In particular, $r''|r_{n-1}$, and hence cannot be bigger than $r_{n-1}$. This shows that $r_{n-1}$ is indeed the greatest common divisor of $a,b$.

**Step 4:** $S = (r_{n-1})$. Since $(a,b)|a,(a,b)|b$, it also divides all numbers of the form $ma + nb$. Hence, $S$ is contained in $((a,b))$. On the other hand, $(a,b)$ itself can be as $m'a + n'b$ for some $m',n'$. Hence, all the multiples of the number $(a,b)$ are of the form $ma + nb$, hence $((a,b))$ is also contained in $S$. Thus, $S = ((a,b))$. \qed
Remark 1. The process computing the greatest common divisor in the proof is called the **Euclidean Algorithm**.

2. We know how to compute the greatest common divisor for two positive integers, but in fact the assumption that \(a, b > 0\) is not necessary, because the largest positive integer dividing both \(-a\) and \(b\) is the same as the largest positive integer dividing \(a\) and \(b\). We can hence conclude \((a, b) = (|a|, |b|)\). Moreover, notice that \(\{ma + nb|m, n \in \mathbb{Z}\} = \{m|a| + n|b||m, n \in \mathbb{Z}\}\), we may drop the positivity assumption in the proposition, and the result remains true.

3. Since any positive number divides 0, the greatest common divisor of 0’s is not well-defined. However, we define the symbol \((0, 0)\) to be 0. This definition is reasonable, because in the ordinary case, \((a, b)\) is described as the smallest positive element in \(\{ma + nb|m, n \in \mathbb{Z}\}\). Yet \(\{m \cdot 0 + n \cdot 0|m, n \in \mathbb{Z}\} = \{0\}\)!

**eg.** Let’s do a concrete example:

Let \(a = 3214, b = 526\). We write down Euclidean algorithmic process for these two numbers:

Step 1: \(3214 = 6 \cdot 526 + 58\)
Step 2: \(526 = 9 \cdot 58 + 4\)
Step 3: \(58 = 14 \cdot 4 + 2\)
Step 4: \(4 = 2 \cdot 2\)

Then, \((3214, 526) = 2\).

**Corollary 1.** \(a, b\) are relatively prime iff every integer can be written as a linear combination of \(a\) and \(b\).

**Proof.** \((a, b) = 1 \iff \{ma + nb|m, n \in \mathbb{Z}\} = (1) = \mathbb{Z}.\) □

**Remark** One frequently uses the result that if \((a, b) = d\), then \(d = ma + nb\), for some \(m, n\).

### 2.3 Some Rules on Divisibility

**eg.** An integer written in decimals \(a_1a_2...a_n\) (eg. \(12345 = 10000 + 2000 + 300 + 40 + 5\)) is divisible by 2 if and only if the right-most digit is an even number. This is obvious, because \(a_1a_2...a_{n-10} = 10 \cdot a_1a_2...a_{n-1}\) is always divisible by 2.

**eg.** An integer written in decimals \(a_1a_2...a_n\) is divisible by 3 if and only if the sum of all its digits \(\sum_{i=1}^{n} a_i\) is divisible by 3. This is because \(a_1a_2...a_n = a_1 \cdot 10^{n-1} + ... + a_{n-1} \cdot 10 + a_n = [a_1 \cdot (10^{n-1} - 1) + ... + a_{n-1} \cdot (10 - 1)] + \sum a_i\). Notice that \(10^m - 1 = 9...9 = 9 \cdot 10^m - 1 = 10^m\) which is clearly divisible by 3. Hence the result follows. Similar result holds if one replaces 3 by 9.
eg. An integer written in decimals $a_1a_2...a_m$ is divisible by 4 if $a_{m-1}a_m$ is divisible by 4. Similar results holds if and only if you replace 4 by 25. This is because $a_1a_2...00$ is a multiple of 100, which is divisible by 4 and 25.

eg. What criterion can you give for divisibility by 8 or 125? More generally, $2^n$ and $5^n$?

eg. An integer written in decimals $a_1a_2...a_n$ is divisible by 5 if and only if $a_n = 0$ or 5. This is because $a_1a_2...a_{m-1}0$ is a multiple of 10, and is divisible by 5, and 5 does not divide 1,2,3,4,6,7,8,9.

eg. An integer written in decimals $a_1a_2...a_n$ is divisible by 7 if and only if $a_1a_2...a_{n-1} - 2 \cdot a_n$ is divisible by 7.

Notice that $a_1a_2...a_n = 10 \cdot (a_1a_2...a_{n-1} - 2 \cdot a_n) + 21 \cdot a_n$. Since 7 divides 21, if $7|a_1a_2...a_{n-1} - 2 \cdot a_n$, we get immediately that $7|a_1a_2...a_n$. Conversely, if $7|a_1...a_n$, then by $7|10 \cdot (a_1a_2...a_{n-1} - 2 \cdot a_n)$. Since $(7,10) = 1, 7|a_1a_2...a_{n-1} - 2 \cdot a_n$. The last step follows from the following lemma:

**Lemma 3.** Suppose $q|m \cdot n$, and $(q,n) = 1$. Then, $q|m$.

**Proof.** $(q,n) = 1$, so $1 = aq + bn$ for some $a,b$. Hence, $m = (ma)q + b(mn)$. Clearly, $a(ma)q$. And $q|mn$ was given as a condition. Then, $q((ma)q + b(mn))$, i.e. $q|m$. □

eg. An integer written in decimals $a_1a_2...a_n$ is divisible by 11 if and only if the difference between the sum of the even-indexed digits and the sum of the odd-indexed digits is divisible by 11. (This is your HW problem.)
Chapter 3

Lecture 3

3.1 Warm-up

eg. Give simple descriptions of the following sets:

1. \( S_1 = \{ 1023n + 33m | n, m \in \mathbb{Z} \} \);
2. \( S_2 = \{ 1023n + 32m | n, m \in \mathbb{Z} \} \).

Since \((1023, 33) = 33\), \( S_1 = (33) \); since \((1023, 32) = 1\), \( S_2 = (1) = \mathbb{Z} \).

3.2 The General Structure of This Course

We briefly talked about the general structure of this course. Basically, we view
the course material as being organized in the “global-local” order.

Global side: Structures over \( \mathbb{Z} \)

1. Basic structures: two operations, + and \( \cdot \), satisfying various arithmetic
   rules. In a nutshell, \((\mathbb{Z}, +, \cdot)\) is a commutative ring.
2. Greatest common divisor theory = theory of ideals of \( \mathbb{Z} \).

Subtler structures among integers are hidden in the solution (or non-existence
of solutions) to certain diophantine equations. In general, these equations are
hard to solve. (We have already seen Fermat’s equation and Pell’s equation.)
One main idea is to pass to the “local” side of the picture: to solve the cor-
responding congruence equations. Solving the latter does NOT automatically
give the solution to the original equation. However, in nice cases, if one make
sense of “every local place” and solve the equation over them, one can claim an
integer solution exists for the original equation. This motivates the second part
of this course.
CHAPTER 3. LECTURE 3

Local side: Structures over $\mathbb{Z}/d\mathbb{Z}$

1. Linear congruence equations;
2. Quadratic congruence equations.

The latter already involves quite non-trivial topics. Indeed, it is the root of some most active field of mathematical research nowadays.

3.3 Divisibility (Continued)

Corollary 2. If $(a, b) = d$, then $(a/d, b/d) = 1$.

Proof. $d = ma + nb$ for some $m, n$. Since $d|a, d|b$, we can divide $d$ on both sides, and get $1 = m(a/d) + n(b/d)$. Hence, $(a/d, b/d) = 1$. $\square$

Corollary 3. Let $a, b$ be two integers such that $b \neq 0$. The rational number $a/b$ can be re-written as $p/q$, where $(p, q) = 1$.

Proof. Follows from previous corollary. $\square$

Corollary 4. $(a + cb, b) = (a, b)$.

Proof. $(a, b) = ma + nb = m(a + cb - cb) + nb = m(a + cb) + (n - cm)b$, and hence is a multiple of $(a + cb, b)$.

Conversely, $(a + cb, b) = p(a + cb) + qb = pa + (cp + q)b$, and is a multiple of $(a, b)$.

Since $(a, b)$ and $(a + cb, b)$ are both positive, they must equal. $\square$

Lemma 4. The greatest common divisor of 3 integers $(a_1, a_2, a_3)$ equals $((a_1, a_2), a_3)$.

Proof. Denote $N = (a_1, a_2, a_3), M = ((a_1, a_2), a_3)$. By definition, $N$ divides all three integers, and $(a_1, a_2)$ is a linear combination of $a_1, a_2$. Hence, $M|(a_1, a_2)$, and further $N|M$, because $M$ is a linear combination of $(a_1, a_2)$ and $a_3$. Thus, $N \leq M$.

On the other hand, by definition, $M|a_3, M|(a_1, a_2)$; yet $(a_1, a_2)\|a_1, (a_1, a_2)\|a_2$. So, $M|a_1, M|a_2, M|a_3$, i.e. it is a positive common divisor of the three integers. But $N$ is the greatest common divisor. So, $M \leq N$.

Thus, $M = N$. $\square$

It turns out that there is a general structural result hidden behind the scene. It will lead to a systematic description of greatest common divisor of arbitrarily finitely many integers.
3.4. IDEALS IN $\mathbb{Z}$

When discussing the greatest common divisor, it took us a while to verify the greatest common divisor $(a_1, a_2)$ is the smallest positive integer in the set 

$$\{m \cdot a_1 + n \cdot a_2 | m, n \in \mathbb{Z}\}.$$ 

And we haven’t treated the general case with $n$ integers. Is there a better way to view this problem?

The answer is yes. The information we didn’t utilize is the following: a set of integers of the form $S = \{n_1 a_1 + \ldots + n_m a_m | n_1, \ldots, n_m \in \mathbb{Z}\}$ has the following special properties:

1. $(S, +)$ is an abelian subgroup of $(\mathbb{Z}, +)$. Concretely, the sum two numbers in $S$ is still in $S$, the negative of a number in $S$ is in $S$, and 0 is in $S$.
2. Take any $n \in \mathbb{Z}$, and any $m \in S$, their product $m \cdot n$ is also in $S$.

**eg.** Verify this!

This motivates the following definition:

**Definition 7.** Let $(R, +, \cdot)$ be a commutative ring. Let $I$ be a subset of $R$ satisfying the following properties:

- $(I, +)$ is an abelian subgroup of $(R, +)$;
- For any $n \in R$, and any $m \in I$, $m \cdot n$ is also in $I$.

Then, $I$ is called an ideal of $R$.

**Remark** From our discussion above, it is immediate that $S = \{n_1 a_1 + \ldots + n_m a_m | n_1, \ldots, n_m \in \mathbb{Z}\}$ is an ideal of $\mathbb{Z}$.

The following proposition reveals all the secret about greatest common divisors.

**Proposition 14.** Every ideal $I$ of $\mathbb{Z}$ is of the form $(n)$, where $n$ is some integer.

**Proof.** By the well-ordering property, $I$ contains a smallest positive integer. Call it $n$. We want to show that every number in $I$ is a multiple of $n$. If not, then there is some $n’ \in I$ such that $n$ does not divide $n’$. We can do division with remainder, and write $n’ = qn + r$, where $0 < r < n$. Since $I$ is an ideal, by definition, $r \in I$ as well. This violates the fact that $n$ is smallest positive integer in $I$. Contradiction! Hence, every element of $I$ is a multiple of $n$.

On the other hand, by the definition of an ideal, every multiple of $n$ has to be in $I$.

Therefore, $I = (n)$. \qed
Remark Here is the upshot: the Euclidean algorithm is a powerful structure over \( \mathbb{Z} \). (It secretly uses the order structure: the well-ordering axiom.) It determines the behavior of those special subsets called “ideals”. Since it is so nice, we should use it!

Proposition 15. Let \( a_1, \ldots, a_m \) be \( m \) integers, not all zero. The smallest positive integer of the set \( S = \{ n_1 a_1 + \ldots + n_m a_m | n_1, \ldots, n_m \in \mathbb{Z} \} \) is the greatest common divisor of \( a_1, \ldots, a_m \).

Proof. We know \( S \) is an ideal of \( \mathbb{Z} \). Hence, \( S = (n) \). \( n \) is the smallest positive integer of \( S \), since every other positive integer in \( S \) is a multiple of \( n \), and hence bigger. \( n \) is a common divisor of \( a_1, \ldots, a_m \), because \( a_1, \ldots, a_m \) are all in \( S \), and hence are multiples of \( n \). It is the greatest, because any number dividing \( a_1, \ldots, a_m \) must also divide \( n \), since \( n \in S \) (which means it can be written in the form \( n_1 a_1 + \ldots n_m a_m \)). \( \square \)

Remark Once we have the right language, the proof is short and clear!

This concludes our discussion of greatest common divisor.

3.5 Prime Numbers

As we have introduced in the first chapter, prime numbers are the building blocks of integers, in the sense that every integer can be factorized into a product of them. In this section, we look at the theory of prime numbers in detail.

Definition 8. An integer \( p \) is called prime if \( p > 1 \), and the only positive divisors are \( 1 \) and \( p \) itself. Similarly, if \( n > 1 \) and \( n \) is not a prime, then we say \( n \) is a composite.

Remark Here is the convention how we treat the negative integers: for the sake of uniqueness of factorization, we here define primes and composites only for positive integers. However, we sometime will also call the negative of a prime number a “prime”, or the negative of a composite a “composite”. The meaning should be clear from the context.

We are going to prove Euclid’s result on the existence infinitely many primes first.

Proposition 16. (Euclid) There are infinitely many prime numbers.

We need a lemma first.

Lemma 5. Every positive integer \( n \) greater than \( 1 \) has a prime factor.

Proof. We prove by contradiction. This essentially follows from Well-ordering property of \( \mathbb{Z} \). Assume there exists some positive integer, that does not have a prime divisor. We can define the set \( S \) of all positive integers greater than
1 with no prime divisor, which is now non-empty, and apply axiom 1 to $S$ to get a smallest element $s$. Because $s$ is assumed to have no prime divisor, $s$ is not prime itself. Hence, $\exists s'$, such that $s' > 1$ and $s'|s$. Since $s' < s$, it has a prime factor, so some prime number $p$ divides $s'$. Hence, $p$ also divides $s$. Contradiction.

Proof. (Proof of the proposition.) We shall prove by contradiction. Suppose there are only finitely many prime numbers, say $p_1, p_2, ..., p_n$. Consider the number $N = p_1 p_2 ... p_n + 1$. We claim none of the $p_i$’s divides $N$. This is because, if some $p_i$ divides $N$, then $N$ can be written as $N = p_1 p_2 ... p_n + 1 = c \cdot p_i$, for some $c$. Bringing all multiples of $p_i$ to one side, one gets $1 = (c-p_1 p_2 ... p_{i-1} p_{i+1} ... p_n)p_i$, which means $p_i$ divides 1. This is impossible! Hence, if $p_1, p_2, ..., p_n$ were all the prime numbers, $N$ couldn’t have a prime factor. This violates the previous lemma, an gives a contradiction.

eg. One might think that the proof provides a way to construct large prime numbers. This turn out to be NOT true. Let’s list the first few prime numbers: 2, 3, 5, 7, 11, 13, 17...Define $A_n = p_1 p_2 ... p_n + 1$. One can compute:

- $A_1 = 2 + 1 = 3$.
- $A_2 = 2 \cdot 3 + 1 = 7$.
- $A_3 = 2 \cdot 3 \cdot 5 + 1 = 31$.
- $A_4 = 2 \cdot 3 \cdot 5 \cdot 7 + 1 = 211$.
- $A_5 = 2 \cdot 3 \cdot 5 \cdot 7 \cdot 11 + 1 = 2311$.
- $A_6 = 2 \cdot 3 \cdot 5 \cdot 7 \cdot 11 \cdot 13 + 1 = 30031 = 59 \cdot 509$. And it is not prime!

Let’s re-formulate the definition of prime numbers and composite numbers.

Lemma 6. (1) Suppose $p \neq 0$. Then, $p$ is a prime number if and only if whenever $a \cdot b \in (p)$, we have either $a \in (p)$ or $b \in (p)$.

(2) $n$ is a composite if and only if there are two integers $a, b$ such that neither of $a, b$ is in $(n)$, but $a \cdot b$ is in $(n)$.

Proof. (1) $a \cdot b \in (p)$ means that there is some $c \in \mathbb{Z}$ such that $c \cdot p = a \cdot b$, i.e. $p|ab$. Suppose $p$ is prime, and $p$ does not divide $a$. Since the only positive factors of $p$ are 1 and $p$, $(p, a) = 1$. This means $1 = ma + np$, and hence $b = mab + nbp = (mc + nb)p$, i.e. $p|b$. Hence, either $p|a$ or $p|b$.

Conversely, suppose whenever $a \cdot b \in (p)$, we have either $a \in (p)$ or $b \in (p)$. This means $p$ has no other positive factor than 1, $p$. Otherwise, we could get some $a|p$, such that both $a$ and $b = p/a$ are not in $(p)$. Hence, $p$ is prime. The second part is just the complement of the first statement, so follows directly.

Remark Hereafter, I’ll write “iff” for “if and only if”.

The above lemma says that $(n)$ behaves quite differently, depending on whether $n$ is prime or not, and hence motivates the following definition:
Definition 9. Let $I$ be a proper ideal (i.e. $I \neq R$) of a commutative ring $R$. Suppose $I$ satisfies the following property: “if $a \cdot b \in I$, then either $a \in I$ or $b \in I$”. Then $I$ is called a prime ideal of $R$.

By the previous lemma, we can conclude immediately $(p)$ is a prime ideal of $\mathbb{Z}$, where $p$ is a prime number. However, we know also another basic fact: the product of two non-zero integers cannot be zero. Hence, $(0)$ is also a prime ideal of $\mathbb{Z}$. In fact, by the previous lemma, these are all the prime ideal of $\mathbb{Z}$.

A common phenomenon in mathematics: sometimes a same concept can be defined in various ways. Some definition may be easier for computational purposes, while some other definition may be generalized to broader situations more easily.

We are now ready to demonstrate the first important theorem regarding factorization of integers.

Theorem 2. (Fundamental Theorem of Arithmetic) Every positive integer $n$ greater than 1 can be written uniquely as a product of prime factors $n = p_1^{n_1} \cdots p_k^{n_k}$. (Here, "unique" means the prime numbers that appear are fixed, and their powers are fixed. You can at most change the order you write those factors.)

3.6 Some Words about the HW

We also mentioned the following ring-theoretic notion in class:

Definition 10. Let $R$ be a ring. $r \in R$ is called a unit, if there exists some $s \in R$ such that $rs = 1$, where 1 is the multiplicative unit of the ring $R$.

eg. In the ring $\mathbb{Z}$, the only units are 1,−1.

Remark In the definition of an ideal, we certainly don’t require the existence of all multiplicative reciprocals. In fact, it is pretty hard to get reciprocal for elements in $\mathbb{Z}$.

eg. The following example shows why the definition of a unit is useful more generally: consider Pell’s equation $x^2 - 3y^2 = 1$, which we can re-write as $(x + \sqrt{3}y)(x - \sqrt{3}y) = 1$. Finding integer solutions of the equation is equivalent to looking for units in the ring $(\mathbb{Z}[^\sqrt{3}],[+,\cdot])$. 

Chapter 4

Lecture 4

4.1 Warm-up

e.g. Consider $a = -2014, b = 702$. Find $(a, b)$ and write it as an integer linear combination of $a$ and $b$.

$$-2014 = (-3) \cdot 702 + 92$$
$$702 = 7 \cdot 92 + 58$$
$$92 = 58 + 34$$
$$58 = 34 + 24$$
$$34 = 24 + 10$$
$$24 = 2 \cdot 10 + 4$$
$$10 = 4 \cdot 2 + 2$$
$$4 = 2 \cdot 2$$

So, $(-2014, 702) = 2$. Now, use these equations, we get

$$2 = 10 - 4 \cdot 2 = 10 - (24 - 2 \cdot 10) \cdot 2 = 5 \cdot 10 + (-2) \cdot 24$$
$$= 5 \cdot (34 - 24) + (-2) \cdot 24 = 5 \cdot 34 + (-7) \cdot 24 = 5 \cdot 34 + (-7) \cdot (58 - 34)$$
$$= 12 \cdot 34 + (-7) \cdot 58 = 12 \cdot (92 - 58) + (-7) \cdot 58 = 12 \cdot 92 + (-19) \cdot 58$$
$$= 12 \cdot 92 + (-19) \cdot (702 - 7 \cdot 92) = 145 \cdot 92 + (-19) \cdot 702$$
$$= 145 \cdot (-2014 + 3 \cdot 702) + (-19) \cdot 702$$
$$= 145 \cdot (-2014) + 416 \cdot 702$$

4.2 Prime Numbers (Continued)

**Theorem 3.** (*Fundamental Theorem of Arithmetic*) Every positive integer $n$ greater than 1 can be written uniquely as a product of prime factors $n = p_1^{n_1} \cdots p_k^{n_k}$. Here, $p_1, \ldots, p_k$ are distinct primes and $n_1, \ldots, n_k > 0$. 
Remark Here, we emphasize \( n_1, \ldots, n_k > 0 \), because one can write an arbitrary positive integer as \( n = \prod_{p \text{ prime}} p^{n(p)} \), where \( n(p) \geq 0 \) and \( n(p) = 0 \) for all but finitely many \( p \)'s. Although formally it seems that we are writing out a product of infinitely many prime numbers, but most of the factors are actually of the form \( p^0 = 1 \), so the symbol makes sense.

Proof. Recall that every positive integer greater than 1 has a prime factor. Given any \( n > 1 \), we can write \( n = p_1 \cdot q \), where \( p_1 \) is prime. If \( q = 1 \), we are done. Otherwise, \( q \) has a prime factor and we can get \( q = p_2 \cdot q' \) and hence \( n = p_1 \cdot p_2 \cdot q' \), where \( p_2 \) is also prime. This process terminates in finitely many steps, since every prime is \( \geq 2 \), yet \( n \) is a finite number.

Uniqueness is easy to see. Suppose \( n = p_1^{r_1} \cdots p_m^{r_m} = q_1^{s_1} \cdots q_K^{s_K} \) are two prime factorizations. Every \( p_i \) must divide one of the factors, say \( q_j \), on the right hand side. Since they are both primes, they must equal. So, one can divide \( p_1 \) from the left-hand side and \( q_j \) from the right-hand side and the equality still holds. Repeatedly using this argument, we see the prime factors on both sides must precisely cancel out. This concludes the proof. \( \square \)

e.g. \( \sqrt{4} \) is irrational. Otherwise, \( \sqrt{4} = a/b \), where \( (a, b) = 1 \). Hence, \( 4b^3 = a^3 \), and in particular \( 2|a^3 \), so \( 2|a \). Then, \( 8|a^3 = 4b^3 \), so \( 2|b^3 \) and hence \( 2|b \). This implies that \( (a, b) \geq 2 \). Contradiction.

4.2.1 Some Factorization Techniques

Imagine that you want to write a computer program to let the computer factorize an arbitrary integer into a product of prime numbers. What would be the algorithm?

A dumb algorithm

Suppose \( n \) is the integer you want to factorize. One algorithm would be to let the computer try out prime numbers among 1, 2, \ldots, \( n - 1 \) and see which of them divides \( n \). This works, but may not be very efficient.

Sieve of Eratosthenes

Based on the following proposition, this method reduces the workload of prime factorization.

**Proposition 17.** Suppose \( n \) is a composite. Then \( n \) has a prime factor less than or equal to \( \sqrt{n} \).

\(^1\)This uses a previous lemma: if \( c|m \cdot n \) and \( (c, m) = 1 \), then \( c|n \).
Proof. By the fundamental theorem of arithmetic, \( n = p_1^{m_1} \cdots p_K^{m_K} \). Since \( n \) is assumed to be composite, it is the product of at least two prime numbers (whether distinct or not). If all of the prime factors are larger than \( \sqrt{n} \), right hand side will be strictly bigger than left hand side, which leads to a contradiction.

Algorithmically, this is saying that to find prime factors of a composite, you can make the computer check prime numbers among \( 1, 2, \ldots, \lceil \sqrt{n} \rceil \). One you find a factor, take a quotient and run the same procedure for the quotient. If the procedure does not return anything, this means the number is already prime itself.

eg. Say we want to factorize 2014. It is easy to see this number is even so \( 2014 = 2 \cdot 1007 \). Since \( \sqrt{1007} < 32^2 \), it suffices to test for divisibility by 3, 5, 7, 11, 13, 17, 19, 23, 29, 31. Given the divisibility rules we developed in class, one can skip 3, 5, 7, 11 and start from 13. After trying, you will see 19 divides 1007 and \( 2014 = 2 \cdot 19 \cdot 53 \). This is the prime factorization: to see 53 is prime, it suffices to check for divisibility by 2, 3, 5, 7, which can be done using out divisibility rules.

**Fermat Factorization**

This method take a different approach towards prime factorization and is based on the following lemma:

**Lemma 7.** Let \( n > 0 \) be an odd integer. Then, there is a 1-1 correspondence between pairs of positive integers \( (a, b) \) and \( (s, t) \) such that \( a > b, s > t \) and \( n = a \cdot b = s^2 - t^2 \).

**Proof.** In one direction, \( (a, b) \mapsto \left( \frac{a+b}{2}, \frac{a-b}{2} \right) \). It’s inverse is \( (s, t) \mapsto (s+t, s-t) \). \( \square \)

Algorithmically, given a positive odd integer \( n \), start from the first integer \( t \) greater than \( \sqrt{n} \), compute \( t^2 - n, (t+1)^2 - n, \ldots \) until you reach the first perfect square. Say, \( s^2 - n = q^2 \). Then, \( n = (s-q)(s+q) \), and you have a factorization. However, \( s - q \) and \( s + q \) need not be prime and in general you have to further apply this method to them to eventually obtain a prime factorization of \( n \).

eg. As a concrete example from your textbook, (p131, sixth edition): given \( n = 6077 \), notice that \( \sqrt{6077} \) is between 77 and 78, compute:

\[
\begin{align*}
78^2 - 6077 &= 7 \\
79^2 - 6077 &= 164 \\
80^2 - 6077 &= 323 \\
81^2 - 6077 &= 484 = 22^2 \\
\end{align*}
\]

Hence, \( 6077 = (81 + 22)(81 - 22) = 103 \cdot 59 \).

**Remark** The restriction to odd numbers is necessary for stating the lemma correctly. However, in practice this does not cause too much obstruction, since one can first factor a power of 2 from \( n \) and then apply the method.
CHAPTER 4. LECTURE 4

4.2.2 Euler’s Product Formula

The Fundamental Theorem of Arithmetic also leads to a proof of the identity
\[ \sum_{n=1}^{\infty} \frac{1}{n^s} = \prod \left( 1 - \frac{1}{p_i^{\alpha_i}} \right), \]
where \( \text{Re}(s) > 1 \), which shows why Riemann Zeta function is related to theory of prime numbers.

**eg.** Since some of you haven’t taken complex analysis yet, I will only cover the real case, namely, \( \sum_{n=1}^{\infty} \frac{1}{n^s} = \prod \left( 1 - \frac{1}{p_i^{\alpha_i}} \right), \) where \( s \in \mathbb{R} \) and \( s > 1 \).

By the convergence test for \( p \)-series, we know \( \sum_{n=1}^{\infty} \frac{1}{n^s} < \infty \).

By the summation formula for convergent geometric series, we know \( \sum_{k=0}^{\infty} \frac{1}{p^k} = \frac{1}{1 - \frac{1}{p^s}} \).

Just as one uses partial sums to approximate an infinite series, we use partial product \( \prod_{p \leq N} \left( \sum_{k=0}^{\infty} \frac{1}{p^k} \right)^s \) to approximate the infinite product on the right-hand side.

It is important to observe that \( \prod_{p \leq N} \left( \sum_{k=0}^{\infty} \frac{1}{p^k} \right)^s \) is a product of primes \( \leq N \) for \( \frac{1}{n^s} \).

Suppose \( p_1, p_2, ..., p_n \) are all the primes \( \leq N \). Then, the product is \( \sum_{k_1,...,k_n \geq 0} \left( \prod_{i=1}^{n} \frac{1}{p_1^{k_1} \cdots p_n^{k_n}} \right)^s \).

The Fundamental Theorem of Arithmetic then tells us that indeed the denominators run through all integers whose primes factors are \( \leq N \), and reach each of them precisely once.

Moreover, \( 0 < \sum_{n=1}^{\infty} \frac{1}{n^s} - \prod_{p \leq N} \left( \sum_{k=0}^{\infty} \frac{1}{p^k} \right)^s < \sum_{n>N} \frac{1}{n^s} \), since any integer less than or equal to \( N \) cannot have a prime factor bigger than \( N \).

As \( N \to \infty \), \( \sum_{n>N} \frac{1}{n^s} \to 0 \), since \( \frac{1}{n^s} \) converges. This concludes the proof.

**Remark** For the complex case, \( \zeta(s) \) the analytic continuation of \( \sum_{n=1}^{\infty} \frac{1}{n^s} \). When \( \text{Re}(s) > 1 \), to show the formula, one needs to do some modulus estimate and adopt the notion of absolute convergence, but the consideration is similar.

4.3 Least Common Multiple

Related to the greatest common divisor and prime decomposition, the notion of least common multiple is introduced here.

**Definition 11.** Let \( a, b \) be two integers, not both zero. Define the least common multiple of \( a, b \), denoted as \([a, b]\), to be the smallest positive integer which is both a multiple of \( a \) and a multiple of \( b \).

Following out remark after the fundamental theorem of arithmetic, we can write \( |a| = \prod p_i^{\alpha_i}, \ |b| = \prod p_i^{\beta_i} \). Consequently, we can write \((a, b)\) and \([a, b]\) in the following way:

**Lemma 8.**

\[
(a, b) = \prod p_i^{\min\{\alpha_i, \beta_i\}};
\]

\[
[a, b] = \prod p_i^{\max\{\alpha_i, \beta_i\}}.
\]
4.3. LEAST COMMON MULTIPLE

With this lemma we conclude immediately:

**Proposition 18.** \([a, b] = \frac{|ab|}{(a, b)}\)

**Proof.** Notice that \(|ab| = \prod p_i^{\alpha_i + \beta_i}\) and that \(\alpha_i + \beta_i = \max\{\alpha_i, \beta_i\} + \min\{\alpha_i, \beta_i\}\).

Hence,

\[
|ab| = \prod p_i^{\alpha_i + \beta_i} = \prod p_i^{\max\{\alpha_i, \beta_i\} + \min\{\alpha_i, \beta_i\}} = (\prod p_i^{\min\{\alpha_i, \beta_i\}})(\prod p_i^{\max\{\alpha_i, \beta_i\}})
\]

Thus, \([a, b] = \frac{|ab|}{(a, b)}\).

**eg.** Can you find a pair of integers such that \((a, b) = 10\) and \([a, b] = 55\)? No. Because \((a, b)|a\) and \(a|[a, b]\), so \((a, b)[a, b]!\)
Chapter 5

Lecture 5

5.1 Warm-up

eg. Let \((R, +, \cdot)\) be a commutative ring. Recall an ideal \(I\) of \(R\) is a subset of \(R\) satisfying the following properties:

- \((I, +)\) is an abelian subgroup of \((R, +)\);
- For any \(n \in \mathbb{Z}\), and any \(m \in I\), \(m \cdot n\) is also in \(I\).

We now define the \textit{sum} of two ideals \(I, J\) to be \(I + J = \{x + y|x \in I, y \in J\}\).

Consider \(R\) to be the ring of integers \(\mathbb{Z}\), describe the sum of two ideals in \(\mathbb{Z}\).

From previous lectures, we know that ideals in \(\mathbb{Z}\) are all of the form \((n)\), i.e. the set of all multiples of a fixed integer \(n\).

By definition, \((a) + (b) = \{ma + nb|m, n \in \mathbb{Z}\} = ((a, b))\). This is a re-emphasis of the theory of greatest common divisor.

5.2 Where we are heading towards

So far, we have discussed some of the basic structures over \(\mathbb{Z}\): the ring structures, greatest common divisor theory (a.k.a ideals of \(\mathbb{Z}\)), fundamental theorem of arithmetic. There are subtler structures/results over \(\mathbb{Z}\), which we shall see when solving for integer solutions of some polynomial equation. The simplest kind of polynomial equation is certainly linear equations. Solving them is a direct application of the greatest common divisor theory, which will be our next topic.

5.3 Linear Diophantine Equations

In the motivation part, we introduced the notion of Diophantine Equation, namely, the study of integer solutions to some polynomial equation. Since we
5.3. LINEAR DIOPHANTINE EQUATIONS

have discussed greatest common divisor, we are ready to discuss the simplest case of diophantine equations.

**Definition 12.** A linear equation \( a_1 x_1 + \ldots + a_n x_n = c \) is called a linear Diophantine equation if \( a_1, \ldots, a_n, c \) are all integers. We seek integer solutions to this equation.

**Remark** The reason to consider polynomial equations is the following: ultimately, we are interested in the algebraic structures on \( \mathbb{Z} \). We therefore only allow “+” and “⋅”. In fact, if you take operations such as logarithm or exponential to integers, you do not even get integers as outputs! These operations are more related to real analysis than to number theory.

**eg.** \( y = \frac{1}{2} x + \frac{1}{2} \) has no integer solution at all, since you can never produce \( \pi \) out of sums of products of integers. This suggests that the existence of solutions (for the two-variable case) should be related to the slope of the line that the equation defines and it is more wise to consider integer coefficients; otherwise, we may be asking the wrong question.

**eg.** However, compare \( 2x + 2y = 1 \) with \( 2x + 2y = 2 \). The former does not have an integer solution, but the latter has infinitely many integer solutions: \((x, y) = (-n, n+1)\) are all solutions. So, we need more information than just the slope.

### 5.3.1 The 2-variable Case

Here is a simple observation:

**eg.** We have seen before that if \((a, b) = 1\), any integer \( c \) can be written as a linear combination of \( a, b \). Put it in the language of Diophantine equations: if \((a, b) = 1\), the equation \( ax + by = c \) always has an integer solution, no matter what \( c \) is.

Moreover, suppose \((x, y)\) is a solution to the linear equation. Then, \( x + tb, y - ta \) (\( t \) being any integer) is also a solution:

\[
ax + by = c
\]

Conversely, if \((x', y')\) is any other solution to the linear equation,

\[
(ax + by) - (ax' + by') = c - c = 0,
\]

i.e. \( a(x - x') = b(y' - y) \), i.e. \( x - x' = b \left( \frac{y' - y}{a} \right) \) and \( y - y' = a \left( -\frac{x' - x}{b} \right) \). We then denote \( t = \frac{x' - x}{b} = \frac{y' - y}{a} \in \mathbb{Z} \), and get \( x' = x + tb, y' = y - ta \).

With the above argument, we arrive at the following result:
CHAPTER 5. LECTURE 5

**Proposition 19.** Let $a, b, c$ be 3 integers such that $(a, b) = 1$. Then, the equation $ax + by = c$ always has solutions. Moreover, the solution set $S = \{(x = x_0 + tb, y = y_0 - ta) | t \in \mathbb{Z}\}$ where $x_0, y_0$ is some particular solution to the equation.

e.g. The solution set to the equation $3x + 4y = 11$ is $\{1 + 4t, 2 - 3t | t \in \mathbb{Z}\}$.

For the general case, let’s drop the assumption that $a, b$ are co-prime. Since $((a, b)) = \{na + mb | n, m \in \mathbb{Z}\}$, if $c$ is not a multiple of $(a, b)$, the equation $ax + by = c$ has no solution.

If $c$ is a multiple of $(a, b)$, one can divide both sides of the equation by $(a, b)$. Then, we get $(\frac{a}{(a, b)}, \frac{b}{(a, b)}) = 1$, and we have successfully reduced to the simple case, discussed earlier. We can therefore state the general result:

**Proposition 20.** Let $a, b, c$ be three integers such that $(a, b)|c$. The solution set to the equation $ax + by = c$ is $S = \{(x = x_0 + t\frac{b}{(a, b)}, y = y_0 - t\frac{a}{(a, b)}) | t \in \mathbb{Z}\}$.

**Remark** An annoying issue about notation: for most of the cases we use $(a, b)$ for the greatest common divisor of $a, b$. However, when describing a solution to some diophantine equation, we use $(x, y)$ for a pair of integers, which satisfies the given equation.

Therefore, to explicitly solve a linear Diophantine equation, it suffice to find one particular solution to the given equation. Luckily, with the Euclidean algorithm, this is not hard to do.

Step 1: Run the Euclidean Algorithm for $a, b$, and find $(a, b)$.

Step 2: Check whether $c$ divides $(a, b)$.

Step 3: From step 2 you get

$$a = q_0 \cdot b + r_1, 0 \leq r_1 < b$$

$$b = q_1 \cdot r_1 + r_2, 0 \leq r_2 < r_1$$

$$......$$

$$r_{n-3} = q_{n-2}r_{n-2} + (a, b)$$

This implies $(a, b)$ is a linear combination of $r_{n-2}$ and $r_{n-3}$; yet $r_{n-2}$ is also a linear combination of $r_{n-3}$ and $r_{n-4}$. Hence, one can write $(a, b)$ as a linear combination of $r_{n-3}$ and $r_{n-4}$. By induction, one can eventually write $(a, b)$ as a linear combination of $a$ and $b$, say $ma + nb$. (See previous lecture notes for an example.)

Step 4: $(md, nd)$ is a particular solution to the equation $ax + by = c$.

Step 5: Write down the solution set.

\footnote{$a, b \neq 0$, otherwise the situation will be trivial.}
eg. Find the solution set to the equation $702x - 2012y = 16$.

$$-2012 = (-3) \cdot 702 + 94$$
$$702 = 7 \cdot 94 + 44$$
$$94 = 2 \cdot 44 + 6$$
$$44 = 7 \cdot 6 + 2$$
$$6 = 3 \cdot 2$$

So,

$$(-2012, 702) = 2 = 44 - 7 \cdot 6 = 44 - 7(94 - 2 \cdot 44)$$
$$= 15 \cdot 44 - 7 \cdot 94 = 15(702 - 7 \cdot 94) - 7 \cdot 94$$
$$= -112 \cdot 94 + 15 \cdot 702$$
$$= -112 \cdot (-2012 + 3 \cdot 702) + 15 \cdot 702$$
$$= -321 \cdot 702 - 112 \cdot (-2012)$$

Since $16 = 2 \cdot 8$, get $(-2568, -896)$ is a particular solution to the linear equation.

Hence, the solution set to this equation is

$$\{(x = -2568 - 1006t, y = -896 - 351t)| t \in \mathbb{Z}\}.$$

### 5.3.2 The General Case

Now consider the case with $n$ variables:

$$a_1x_1 + \ldots + a_nx_n = c.$$ 

Using the general result about greatest common divisor, this equation has a solution iff $c$ is a multiple of $(a_1, \ldots, a_n)$.

The general idea to solving for a particular solution is to reduce an $n$-variable equation to an $(n-1)$-variable equation, using the property of greatest common divisor: $a_{n-1}x_{n-1} + a_nx_n = (a_{n-1}, a_n)y$.

Thus, by introducing a new variable $y$, one can equivalently solve

$$a_1x_1 + \ldots + a_{n-2}x_{n-2} + (a_{n-1}, a_n)y = c.$$ 

If you re-look at how we solved the equation with two variables, essentially we reduce the equation to $(a, b)u = c$, so it is the same idea.

Moreover, two solutions $(x_1, \ldots, x_n)$ and $(x'_1, \ldots, x'_n)$ satisfy the relation

$$a_1(x_1 - x'_1) + \ldots + a_n(x_n - x'_n) = 0.$$ 

One can denote $y_i = x_i - x'_i$ for $i = 1, \ldots, n$, and write $(a_1, \ldots, a_n) \cdot (y_1, \ldots, y_n) = 0$ (dot product). This means, any two solutions differ by an integer point (i.e. all the entries are integers) in the hyperplane \footnote{A hyperplane can be thought of an \((n-1)\)-dimensional flat subset of \(\mathbb{R}^n\). When \(n = 3\), it is just a plane in the usual sense.} in \(\mathbb{R}^n\) defined by the equation

$$a_1x_1 + \ldots + a_nx_n = 0.$$
i.e. \((a_1, ..., a_n)\) is a normal vector to this hyperplane. So, in general the solution set looks like the following:

\[
S = \{(x_1, ..., x_n) + (y_1, ..., y_n) | (y_1, ..., y_n) \cdot (a_1, ..., a_n) = 0\}
\]

where \((x_1, ..., x_n)\) is a particular solution.

![Two-variable Case](image1)

![Three-variable Case](image2)

**Remark** These pictures describe what’s going on geometrically:

When \(n = 2\), we are trying to find all integer points on a given line. We start with find one integer point and then move by a smallest feasible step each time. Another way to put it, we move by distance \(d = \sqrt{\frac{a^2 + b^2}{(a, b)^2}}\) along the line each time. We do not want to move by a bigger distance, since otherwise we will miss some solutions.

When \(n = 3\), solutions are integer points on a plane. Again, we start with one integer point, but now we have two directions to move into.

These integer points are sometimes called “lattice points”.

**eg.** In the 3 variable case, we can understand the solution set more concretely; namely, we can describe the condition

\[
a_1 x_1 + a_2 x_2 + a_3 x_3 = 0
\]

more explicitly. We start with the assumption that \((a_1, a_2, a_3) = 1\). We claim that one can write any integer points in this plane as a linear combination of two integer vectors.

Suppose \(a_1 n + a_2 m = (a_1, a_2)\), let \(v_1 = (a_3 n, a_3 m, -(a_1, a_2))\), \(v_2 = (\frac{a_2}{(a_1, a_2)}, -\frac{a_1}{(a_1, a_2)}, 0)\).

We claim that \(v_1, v_2\) are points in the plane defined by \(a_1 x_1 + a_2 x_2 + a_3 x_3 = 0\) and any other integer point in this plane can be written as an integer linear combination of these two points.

One can check by plugging them into the equation that \(v_1, v_2\) are in the plane defined by \(a_1 x_1 + a_2 x_2 + a_3 x_3 = 0\).
Suppose \((x', y', z')\) is any integer point in the plane, then

\[ a_1 x' + a_2 y' = N \cdot (a_1, a_2) = (N n) a_1 + (N m) a_2 = -a_3 z', \]

where \(N\) is a multiple of \(a_3\), so that \(ax' + by' = (Mcn) a + (Mcm) b\). But this implies \(-a_1(x' - Ma_3 n) = a_2(y' - Ma_3 m),\) and further

\[ \frac{-a_1}{(a_1, a_2)} (x' - Ma_3 n) = \frac{a_2}{(a_1, a_2)} (y' - Ma_3 m). \]

Thus, \(x' - Ma_3 n = \frac{a_2 t}{(a_1, a_2)}, y' - Ma_3 m = \frac{-a_1 t}{(a_1, a_2)},\) for some \(t \in \mathbb{Z}\).

Thus,

\[
(x', y', z') = (Ma_3 n, Ma_3 m, z') + (x' - Ma_3 n, y' - Ma_3 m, 0)
= M(a_3 n, a_3 m, -(a_1, a_2)) + t\left(\frac{a_2}{(a_1, a_2)} - \frac{a_1}{(a_1, a_2)}, 0\right)
= Mv_1 + tv_2
\]

This tells us that the solution set to \(a_1 x_1 + a_2 x_2 + a_3 x_3 = c\) is of the form:

\[
S = \{(x_0, y_0, z_0) + t_1 v_1 + t_2 v_2 | t_1, t_2 \in \mathbb{Z}\}
\]

where \((x_0, y_0, z_0)\) is a particular solution.

**Remark** It is important to start with the assumption that \((a_1, a_2, a_3) = 1\). Otherwise, the formula here is not correct. We run into the problem of missing some solutions.

---

\(^3\)Here, we use the assumption \((a_1, a_2, a_3) = ((a_1, a_2), a_3) = 1.\) Since \(a_3|N \cdot (a_1, a_2)\) and it is co-prime to \((a_1, a_2),\) it must divide \(N.\)
Chapter 6

Lecture 6

6.1 Problem Session

eg. Pell’s equation re-visited.

As we’ve discussed earlier, when we solve for solutions to the equation \( x^2 - ny^2 = 1 \), where \( n > 0 \) is not a perfect square, we are studying units in the ring \((\mathbb{Z}[\sqrt{n}], +)\).

Definition 13. Let \( R \) be a ring. \( r \in R \) is called a unit, if there exists some \( s \in R \) such that \( r \cdot s = 1 \).

Remark This is just a name for elements which have multiplicative reciprocals.

It turns out that the units \( a + \sqrt{n}b \) of this ring correspond to \((a, b)\) that are solutions to equations \( x^2 - ny^2 = \pm 1 \). This can be seen in an elementary way:

If \((a + \sqrt{n}b)(c + \sqrt{n}d) = (ac + nbd) + (ad + bc)\sqrt{n} = 1\), then since \( \sqrt{n} \) is irrational, one must have

\[
\begin{align*}
ad + bc &= 0 \\
ac + nbd &= 1
\end{align*}
\]

We have two possibilities for \( d \).

Case 1: \( d = 0 \). In this case, \( c \neq 0 \) and \( b = 0 \). Thus, \( ac = 1 \), so \( a, c = \pm 1 \).

Case 2: \( d \neq 0 \). We then get \( a = -\frac{b}{d} = t \), i.e. \( a = ct, b = -dt \). Plug this into \( ac + nbd = 1 \), we get \( t(c^2 - nd^2) = 1 \). We claim \( t \) must be an integer. Suppose not. Then, \((a + b\sqrt{n})(a - b\sqrt{n}) = t(a + b\sqrt{n})(c + d\sqrt{n})\) is a rational number which is not an integer. This is impossible, because if you multiply two numbers in \( \mathbb{Z}[\sqrt{n}] \), the product is still in \( \mathbb{Z}[\sqrt{n}] \). ((\( \mathbb{Z}[\sqrt{n}], +, \cdot \)) is a ring, which is closed under multiplication.)
Therefore, \( t = \pm 1 \). So either \( c + \sqrt{nd} = a - \sqrt{nb} \), or \( c + \sqrt{nd} = -a + \sqrt{nb} \). In the first case, \((a, b)\) is a solution to \( x^2 - ny^2 = -1 \); in the latter case, it is a solution to \( x^2 - ny^2 = -1 \). We shall refer to them as units of type I and units of type II.

Back to Pell’s equation, we are only interested in units of type I.

**Here is a relation between units of type I and units of type II:** if \((a, b)\) satisfies the equation \( a^2 - nb^2 = -1 \), then \( a' + b'\sqrt{\alpha} = (a + b\sqrt{\alpha})^2 \) is a unit of the first type; in other words, \((a', b')\) is a solution to the Pell’s equations. (Why?) In conclusion, units in \( \mathbb{Z}[\sqrt{\alpha}] \) are either \( a + b\sqrt{\alpha} \), where \((a, b)\) is a solution to Pell’s equation; or, a square root of \( a + b\sqrt{\alpha} > 0 \), where \((a, b)\) is a solution to Pell’s equation.

The logic to solving the HW problem: find a smallest positive solution \((x_0, y_0)\) via trial. Then, show that every solutions is of the form \( \pm (x_0, y_0)^m \).

This can be done via proof by contradiction. You need to argue that if this is not true, then there is a positive solution \((x', y')\) such that \( x_0 + \sqrt{\alpha} < x' + \sqrt{\alpha} \) \( (x_0 + \sqrt{\alpha})^M < (x_0 + \sqrt{\alpha})^M + 1 \). Then, you show that \( (x' + \sqrt{\alpha}) (x_0 + \sqrt{\alpha})^{-M} \) still corresponds to a positive solution, which is smaller than the smallest positive solution. This leads to a contradiction.

As a concrete example, from the extra credit homework, we know \( (3, 2) \) is the smallest positive solution to

\[
x^2 - 2y^2 = 1.
\]

Moreover, any solution is of the form \( \pm (3, 2)^m \), i.e., \( \pm (3 + 2\sqrt{2})^m \) are units of \( \mathbb{Z}[\sqrt{2}] \) of type I. Notice that \( (1 + \sqrt{2})^2 = 3 + 2\sqrt{2} \) and \( 1 + \sqrt{2} \) is a unit of type II. In particular, square roots of \( (3 + 2\sqrt{2})^m \) can be written as \( \pm (1 + \sqrt{2})^m \). We have thus determined all units in \( \mathbb{Z}[\sqrt{2}] \).

**eg.** True of false: \( a^2 | b^3 \Rightarrow a | b \). This is false: \( a = 8, b = 4 \) is a counter-example. More generally, let \( p \) be any prime number. Then, \( p^6 = p^{3^2} = p^3 \), but \( p^3 \) does not divide \( p^6 \). Compare this with a similar problem in HW2.

**Remark** Also notice that \( a \) does not divide \( b \) is not the same as \( (a, b) = 1 \).

**eg.** Show that for any \( n > 0 \), there are \( n \) consecutive integer which are all composite.

Take \( (n + 1)! + 2, ..., (n + 1)! + (n + 1) \). They are clearly consecutive, and since \( i | (n + 1)! + i \) for \( i = 2, ..., n + 1 \), they are all composites.

**Remark** Clearly, the only consecutive primes are 2, 3. We have also seen that the only triple of primes, \( p, p + 2, p + 4 \), are 3, 5, 7. And we mentioned that whether there are infinitely many pairs of twin primes, \( p, p + 2 \), remains an extremely difficult open question.

This example shows that in one aspect composites are much simpler than primes: you can easily get long chains of consecutive composites.

**eg.** Find the solution set to \( 12x + 30y - 45z = 9 \).
First, find a particular solution: since \(12 = 2^2 \cdot 3, 30 = 2 \cdot 3 \cdot 5\), we get \((12, 30) = 6\); one can write \(6 = (-2) \cdot 12 + 30\). Then, \((6, -45) = 3\) and \(3 = 6 \cdot (-7) + (-1) \cdot (-45)\).

So,
\[
3 = 6 \cdot (-7) + (-1) \cdot (-45) = 14 \cdot 12 + (-7) \cdot 30 + (-1) \cdot 45.
\]

Since \(9 = 3 \cdot 3\), \((42, -21, -3)\) is a particular solution to the equation.

Second, consider \(12x + 30y - 45z = 0\). We want to use formula from yesterday to find two integer vectors \(v_1, v_2\). **Warning:** one cannot the formula directly, since \(12, 30, -45\) are not co-prime. Dividing out by \((12, 30, -45)\), we get
\[
4x + 10y - 15z = 0.
\]

Notice that \(2 = 4 \cdot (-2) + 10 \cdot 1\). By our given formula, \(v_1 = (30, -15, -2)\), \(v_2 = (5, -2, 0)\). Hence, the solution set is
\[
S = \{(42, -21, -3) + t_1(30, -15, -2) + t_2(5, -2, 0) | t_1, t_2 \in \mathbb{Z}\}.
\]

**Remark** We gave the following formula for \(v_1, v_2\) yesterday:
\[
\begin{align*}
  v_1 &= (a_3 n, a_3 m, -(a_1, a_2)) \\
  v_2 &= \left(\frac{a_2}{(a_1, a_2)}, \frac{-a_1}{(a_1, a_2)}, 0\right)
\end{align*}
\]

This was based on the assumption \((a_1, a_2, a_3) = 1\). They are indeed the “smallest” integer points in the plane defined by \(a_1 x_1 + a_2 x_2 + a_3 x_3 = 0\), spanning all integer points in it.

First, \(\left(\frac{a_2}{(a_1, a_2)}, \frac{-a_1}{(a_1, a_2)}\right) = 1\), and \(\left|\frac{a_1 a_2}{(a_1, a_2)}\right|\) is the least common multiple of \(a_1, a_2\). Therefore, fixing \(z = 0\), \(\left(\frac{a_2}{(a_1, a_2)}, \frac{-a_1}{(a_1, a_2)}, 0\right)\) is the solution to \(a_1 x_1 + a_2 x_2 + a_3 x_3 = 0\) such that \(|a_1 x_1|\) is minimal.

But then we also need to take into account freedom in the \(z\)-variable. This is taken care of by \(v_2\). Note that \((a_1, a_2, a_3) = 1\), so the least common multiple of \((a_1, a_2)\) and \(a_3\) is \((a_1, a_2) \cdot a_3\). Therefore, \((a_3 n, a_3 m, -(a_1, a_2))\) is the solution to \(a_1 x_1 + a_2 x_2 + a_3 x_3 = 0\) such that \(|a_1 x_1 + a_2 x_2| = |a_3 x_3|\) is minimal.

Recall the picture from last lecture: when considering those integer points in the plane \(a_1 x_1 + a_2 x_2 + a_3 x_3 = c\), starting from some integer point, we want to make sure that each time we travel only one step in the direction of \(v_1\) or \(v_2\) (or, \(-v_1\) or \(-v_2\)). In that way, we won’t miss any possible solution in the solution set.

### 6.2 \(\mathbb{Z}/k\mathbb{Z}\) as a Commutative Ring

> What is crucial is there be laws.

Andre Weil

In previous lectures, we have talked about division with remainder. Fix \(b > 0\), every integer \(a\) can be written uniquely as \(a = qb + r\), where \(0 \leq r < b - 1\). A different way to look at it: given any \(b > 0\), \(\mathbb{Z} = \bigcup_{i=0}^{b-1} \{qb + ri | q \in \mathbb{Z}\}\).
6.2. $\mathbb{Z}/K\mathbb{Z}$ AS A COMMUTATIVE RING

From now on, we will study congruence arithmetic systematically. By the division with remainder, we know that any integer $n$ is a sum of some multiple of $b$ with an integer $r$ between 0 and $b-1$. Moreover, that integer $r$ between 0 and $b-1$ is unique (and hence $n-r$ is also uniquely determined.) This motivates the following definition:

**Definition 14.** Two integers $a$ and $c$ are said to be congruent modulo $b$, if $b|a-c$, i.e. $a$ and $c$ differ by a multiple of $b$. We write $a \equiv c \pmod{b}$, and read “$a$ is congruent to $c$ modulo $b$”.

A finite set of $b$ integers $\{a_1, \ldots, a_b\}$ is called a complete residue system modulo $b$, if for any integer $n$, $n$ is congruent to exactly one $a_i$ modulo $b$.

**Remark** As we mentioned in earlier lectures, we sometimes pass to congruence (polynomial) equations, when the original integer polynomial equation is hard to solve. This may provide us with partial information about the original equation. The main benefit is that a complete residue system has only finitely many elements and we only seek solution inside this finite set.

**eg.** Complete residue systems modulo $b$ is not unique. $\{0, 1, \ldots, b-1\}$ is a complete residue system modulo $b$. So is $\{-1, 0, \ldots, b-2\}$.

For the moment, we define $\mathbb{Z}/b\mathbb{Z} = \{0, 1, \ldots, b-1\}$. Notice that we put a line over 0, 1, ..., $k-1$ to emphasize the difference between an element in $\mathbb{Z}/k\mathbb{Z}$ and an element in $\mathbb{Z}$. In the following lectures, we fix a choice of a complete residue system modulo $b$: $\{0, 1, \ldots, b-1\}$, unless otherwise mentioned. So it makes sense to say “the remainder of a modulo $b$”. Our first goal is to describe a ring structure on $(\mathbb{Z}/b\mathbb{Z}, +, \cdot)$. This is saying nothing else than the fact that we can add two remainders and multiply two remainders; $\mathbb{Z}/b\mathbb{Z}$ with these two operations shares the nine properties of $\mathbb{Z}$ with the usual addition and multiplication

1. Define a binary operation $+: \mathbb{Z}/b\mathbb{Z} \times \mathbb{Z}/b\mathbb{Z} \to \mathbb{Z}/b\mathbb{Z} : (\bar{i}, \bar{j}) \mapsto \bar{i} + \bar{j}$, where $\ell$ is the remainder of $i + j$ modulo $b$, chosen from $\{0, \ldots, b-1\}$.

2. Define another binary operation $\cdot : \mathbb{Z}/b\mathbb{Z} \times \mathbb{Z}/b\mathbb{Z} \to \mathbb{Z}/b\mathbb{Z} : (\bar{i}, \bar{j}) \mapsto \bar{i} \cdot \bar{j}$, where $\ell$ is the remainder of $i \cdot j$ modulo $b$, chosen from $\{0, \ldots, b-1\}$.

**eg.** Let $b = 6$. Then $\bar{3} + \bar{4} = \bar{1}$.

**eg.** Let $b = 6$. Then $\bar{3} \cdot \bar{4} = \bar{0}$. We see that although $(\mathbb{Z}/b\mathbb{Z}, +, \cdot)$ shares a lot of similarities with $(\mathbb{Z}, +, \cdot)$, there are interesting differences as well. In general, one can have two non-zero elements multiplying together and get the zero element. This certainly cannot happen with the integers.
Chapter 7

Lecture 7

7.1 Motivation for Problem 10, HW2

Prime numbers are the central objects of study in number theory. However, it is hard to write down a simple formula that produces lots of prime numbers. Euler produces the following quadratic formula: \( f(x) = x^2 - x + m \). Among all \( m \)'s, 41 is a special one.

**eg.** \( x^2 - x + 41 \) is prime for \( x : 0 \leq x \leq 40 \). When \( x = 41 \), it is composite: \( 41|41^2 - 41 + 41 \). However, you may get other primes for \( x > 41 \). In general, if \( p \in \mathbb{Z} \) satisfies the property that \( x^2 - x + p \) is prime for all \( x : 0 \leq x \leq p - 1 \), \( p \) is called Euler’s lucky number.

We shall turn Problem 10 into an extra credit HW. Here is a rough sketch as for how one solves the problems with more mathematical insight rather than by brute force. \( f(x) = x^2 - x + 41 \) is a quadratic polynomial.

We define the discriminant of \( ax^2 + bx + c \) to be \( \Delta := b^2 - 4ac \). \( \Delta(f) = (-1)^2 - 4 \cdot 41 = -163 \).

Suppose \( n^2 - n + 41 = s \cdot t \) such that \( 1 < s, t < 41 \), for some \( n : 0 \leq n \leq 40 \). If we take the false assumption, one can define a related polynomial \( g(y) = sy^2 - (2n - 1)y + t = 0 \). Then,

\[
\Delta(g) = (4n^2 - 4n + 1) - 4s \cdot t \\
= (4n^2 - 4n - 4s \cdot t) + 1 \\
= 4 \cdot (-41) + 1 = -163 \\
= \Delta(f)
\]

This is not a random coincidence and we shall use this fact to prove the statement. For more detail, please see Extra Credit HW2.

**Remark** If you have time for some fun summer reading, check out “imaginary quadratic field” in google.
7.2 Units in \( \mathbb{Z}[\sqrt{n}] \)

We mentioned yesterday that the units (a.k.a. elements that have multiplicative inverse) in \((\mathbb{Z}[\sqrt{n}], +, \cdot)\) can be classified into two types:

1. type I: \(c + \sqrt{n}d\) such that \(c^2 - nd^2 = 1\);

2. type II: \(c + \sqrt{n}d\) such that \(c^2 - nd^2 = -1\).

\textbf{eg.} When \(n = 2\), from your Extra Credit HW1, all units of type I in \((\mathbb{Z}[\sqrt{2}], +, \cdot)\) are of the form \(\pm(3 + 2\sqrt{2})m, \; m \in \mathbb{Z}\).

Suppose \(c + \sqrt{n}d\) is a type II unit. Then, \((c + \sqrt{n}d)(c - \sqrt{n}d) = -1\); therefore, \((c + \sqrt{n}d)^2(c - \sqrt{n}d)^2 = (-1)^2 = 1\). In particular,

\[(c + \sqrt{n}d)^2 = (c^2 + nd^2) + (2cd)\sqrt{n}\]

is a unit of type I. In other words, \((c^2 + nd^2, 2cd)\) is a solution to Pell’s equation:

\[x^2 - ny^2 = 1\]

\textbf{eg.} When \(n = 2\), since \((1 + \sqrt{2})^2 = 3 + 2\sqrt{2}\). Therefore, all units of type II are of the form \(\pm(1 + \sqrt{2})m\).

7.3 \(\mathbb{Z}/b\mathbb{Z}\) as a commutative ring (continued)

Yesterday, we defined addition and multiplication of remainders modulo some integer \(b\), i.e. + and \(\cdot\) on \(\mathbb{Z}/b\mathbb{Z}\).

7.3.1 Basic Structure on \(\mathbb{Z}/b\mathbb{Z}\)

By the end of the day, we want to say “\((\mathbb{Z}/b\mathbb{Z}, +, \cdot)\)” is a commutative ring. This is just a concise way of summarizing the nine properties that \((\mathbb{Z}/b\mathbb{Z}, +, \cdot)\) possesses. Let’s compare \(\mathbb{Z}/b\mathbb{Z}\) with \(\mathbb{Z}\):

<table>
<thead>
<tr>
<th></th>
<th>(\mathbb{Z})</th>
<th>(\mathbb{Z}/b\mathbb{Z})</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>associativity for usual addition</td>
<td>((\tilde{i} + \tilde{j}) + \tilde{l} = \tilde{i} + (\tilde{j} + \tilde{l}))</td>
</tr>
<tr>
<td>2</td>
<td>commutativity for usual addition</td>
<td>(\tilde{i} + \tilde{j} = \tilde{j} + \tilde{i})</td>
</tr>
<tr>
<td>3</td>
<td>it contains 0</td>
<td>(\tilde{0} + \tilde{i} = \tilde{i} + \tilde{0} = \tilde{0})</td>
</tr>
<tr>
<td>4</td>
<td>every integer has its negative</td>
<td>(\tilde{i} + \tilde{b} = \tilde{0})</td>
</tr>
<tr>
<td>5</td>
<td>associativity for usual multiplication</td>
<td>((\tilde{i} \cdot \tilde{j}) \cdot \tilde{l} = \tilde{i} \cdot (\tilde{j} \cdot \tilde{l}))</td>
</tr>
<tr>
<td>6</td>
<td>commutativity for usual multiplication</td>
<td>(\tilde{i} \cdot \tilde{j} = \tilde{j} \cdot \tilde{i})</td>
</tr>
<tr>
<td>7</td>
<td>it contains 1</td>
<td>(\tilde{1} \cdot \tilde{i} = \tilde{i} \cdot \tilde{1} = \tilde{1})</td>
</tr>
<tr>
<td>8</td>
<td>distributive law 1</td>
<td>((\tilde{i} + \tilde{j}) \cdot \tilde{l} = \tilde{i} \cdot \tilde{l} + \tilde{j} \cdot \tilde{l})</td>
</tr>
<tr>
<td>9</td>
<td>distributive law 2</td>
<td>(\tilde{i} \cdot (\tilde{j} + \tilde{l}) = \tilde{i} \cdot \tilde{j} + \tilde{i} \cdot \tilde{l})</td>
</tr>
</tbody>
</table>
Put in fancier terminologies:

**Lemma 9.** \((\mathbb{Z}/b\mathbb{Z},+)\) is an abelian group.

**Proposition 21.** \((\mathbb{Z}/b\mathbb{Z},+,:)\) is a commutative ring.

We can now explain the notation \(\mathbb{Z}/b\mathbb{Z}\): \(b\mathbb{Z}\) stands for the ideal \((b)\) of \(\mathbb{Z}\) (you can think of \(b\mathbb{Z}\) as multiplying every integer by \(b\), which gives you precisely \((b)\)). We write “\(\mathbb{Z}/b\mathbb{Z}\)” to denote the ring obtained from \(\mathbb{Z}\) modulo out the ideal \((b)\). More rigorously, we have the following notion (for fun):

**Definition 15.** Let \(R\) be a commutative ring, and \(I\) be an ideal. The quotient ring \(R/I\) is defined to be the set of “remainders” of \(R\) modulo \(I\):

\[R/I = \{\text{equivalence classes of elements in } R \text{ modulo the equivalence relation } \sim\},\]

where \(a \sim b\) if and only if \(a - b \in I\). The two binary operations for \(R/I\) are just the operations of \(R\), applied modulo \(I\).

**Remark** In our case, we identify two integers \(a, c\) inside the ring \(\mathbb{Z}/b\mathbb{Z}\) iff they have the same remainder modulo \(b\). Another way to put it, anything in \((b)\) is regarded as “zero”. And the two operations we call + and \(\cdot\) are precisely coming from the usual addition and multiplication (as suggested by the above more general definition.) It is no wonder that they have nice properties.

### 7.3.2 Relation between \(\mathbb{Z}\) and \(\mathbb{Z}/b\mathbb{Z}\)

The relation between \(\mathbb{Z}\) and \(\mathbb{Z}/b\mathbb{Z}\) is described through a natural map:

\[q_b : \mathbb{Z} \to \mathbb{Z}/b\mathbb{Z} : a \mapsto \text{the reminder of } a \text{ modulo } b \text{ chosen from } \{0, ..., b - 1\}\]

This map has the following nice properties:

1. \(q_b(a + c) = q_b(a) + q_b(c)\);
2. \(q_b(a \cdot b) = q_b(a) \cdot q_b(b)\);
3. \(q_b(1) = 1\);
4. \(q_b\) is surjective.

Again, a fancier way to put it (for fun):

**Proposition 22.** There is a surjective ring homomorphism\(^1\) \(q_b : \mathbb{Z} \to \mathbb{Z}/b\mathbb{Z}\) sending every integer \(n\) to its remainder modulo \(b\).

\(^1\)Let \(R_1, R_2\) be two rings. A ring homomorphism from \(R_1\) to \(R_2\) is a map \(f : R_1 \to R_2\) such that:

1. \(f(a + b) = f(a) + f(b)\).
2. \(f(a \cdot b) = f(a) \cdot f(b)\).
3. \(f(1_{R_1}) = 1_{R_2}\).
7.3. \( \mathbb{Z}/b\mathbb{Z} \) AS A COMMUTATIVE RING (CONTINUED) 45

Remark A side remark: In different mathematical context, one should require maps to have different properties.

1. In the context of sets, there is no extra structure at all. Hence, maps between sets only need to tell us where each element is sent to.

2. In the context of topological spaces, we are interested in “continuous” maps. They need to relate open sets in both spaces.

3. In the context where we do calculus, we talk about “smooth” maps. They need to relate the calculus one does on one space to the calculus on the other.

4. In the context where we do complex analysis, we look at “holomorphic” maps. They need to relate the complex analysis one does on one space to the complex analysis on the other.

5. In the context of rings, we want maps to preserve the ring structure. If a map doesn’t preserve these algebraic structures, it really doesn’t tell us much about the relation between the two rings. i.e. they forget the algebraic information and are no so useful.

To make long stories short, when we encounter different structures, we are always interested in maps preserving these structures.

The following corollary is very useful in computation:

**Corollary 5.** Let \( f(x_1, \ldots, x_n) \) be an integer-coefficient polynomial. Suppose \( a_1 \equiv b_1 ( \text{mod} \ b) \), \( a_n \equiv b_n ( \text{mod} \ b) \), then \( f(a_1, \ldots, a_n) \equiv f(b_1, \ldots, b_n) ( \text{mod} \ b) \).

**Proof.** A polynomial \( f(x_1, \ldots, x_n) \) can be written as \( \sum a_{m_1, \ldots, m_n} x_1^{m_1} \cdots x_n^{m_n} \).

\[
q_b(f(a_1, \ldots, a_n)) = q_b(\sum a_{m_1, \ldots, m_n} a_1^{m_1} \cdots a_n^{m_n}) \\
= \sum q_b(a_{m_1, \ldots, m_n}) \cdot q_b(a_1)^{m_1} \cdots q_b(a_n)^{m_n} \\
= \sum q_b(a_{m_1, \ldots, m_n}) \cdot q_b(b_1)^{m_1} \cdots q_b(b_n)^{m_n} \\
= q_b(\sum b_{m_1, \ldots, m_n}) \\
= q_b(f(b_1, \ldots, b_n))
\]

Therefore, \( f(a_1, \ldots, a_n) \equiv f(b_1, \ldots, b_n) (\text{mod} \ k) \). \( \square \)

7.3.3 Comparing different \( \mathbb{Z}/b\mathbb{Z} \)

Yesterday, we have seen that \( \mathbb{Z}/6\mathbb{Z} \) is \( 0 \) in \( \mathbb{Z}/6\mathbb{Z} \), i.e. two non-zero elements multiply to zero. This cannot happen when \( b \) is some prime number \( p \).

First of all, we have seen that if \( a \cdot c \equiv q \cdot p \) then at least one of \( a|p, b|p \) must be true. Hence, \( i \cdot j = 0 \) iff at least one of \( i, j \) is zero. This motivates the following definition:
Definition 16. Let $R$ be a commutative ring such that $\forall a, c \in R$, $a \cdot c = 0$ iff at least one of $a, c$ is the zero element, then $R$ is called an integral domain.

eg. $\mathbb{Z}/b\mathbb{Z}$ is an integral domain if and only if $b$ is prime. When $b$ is not prime, one can write $b = c \cdot d$, where $1 < c, d < b$. Then, $q_b(c), q_b(d)$ are not zero in $\mathbb{Z}/b\mathbb{Z}$, but their product is $0$.

When $p$ is prime, we know that for any positive integer $a$ smaller than $p$, the greatest common divisor $(a, p) = 1$. Hence, $\exists m, n$ such that $an + pm = 1$. Therefore, $q_p(n) \cdot \overline{p} = 1$. This shows that every non-zero element in $\mathbb{Z}/p\mathbb{Z}$ has a reciprocal, i.e. it is a unit.

Definition 17. Let $R$ be an integral domain. If every non-zero element of $R$ is a unit, then $R$ is called a field.

eg. You have heard of $\mathbb{Q}, \mathbb{R}, \mathbb{C}$ as fields. Now you have one more example: $\mathbb{Z}/p\mathbb{Z}$ ($p$ prime). It is different from the previous three in a fundamental way: in the previous, $1 + 1 + 1 + \ldots + 1$ is never zero. But in $\mathbb{Z}/p\mathbb{Z}$, $1 + \ldots + 1 = 0$.

This seemingly pathological structure can be extremely powerful! If you want to read some enlightening stories, you can google “Shigefumi Mori”, “bend and break” and etc.

7.3.4 Relation between $\mathbb{Z}/m\mathbb{Z}$ and $\mathbb{Z}/n\mathbb{Z}$ when $m | n$.

It turns out there is also a surjection from $\mathbb{Z}/n\mathbb{Z}$ to $\mathbb{Z}/m\mathbb{Z}$ when $m | n$. We describe this by a concrete example:

eg. We know that $\mathbb{Z}/6\mathbb{Z}$. Hence, $(6) \subset (2)$. As a consequence, we have a natural map $q_{2,6} : \mathbb{Z}/6\mathbb{Z} \rightarrow \mathbb{Z}/2\mathbb{Z} : \overline{i} \mapsto \text{(remainder of } i \text{ modulo } 2)$. More concretely, $0 \rightarrow 0, \overline{1} \rightarrow \overline{1}, \overline{2} \rightarrow \overline{0}, \overline{3} \rightarrow \overline{1}, \overline{4} \rightarrow \overline{0}, \overline{5} \rightarrow \overline{1}$. Moreover, one can check directly that the map $q_2 : \mathbb{Z} \rightarrow \mathbb{Z}/2\mathbb{Z}$ we defined earlier factors as $\mathbb{Z} \xrightarrow{q_6} \mathbb{Z}/6\mathbb{Z} \xrightarrow{q_{2,6}} \mathbb{Z}/2\mathbb{Z}$, i.e. $q_2 = q_{2,6} \circ q_6$. This generalizes:

Proposition 23. Suppose $m | n$. Then, there is a surjective ring homomorphism $q_{m,n} : \mathbb{Z}/n\mathbb{Z} \rightarrow \mathbb{Z}/m\mathbb{Z}$, such that $\overline{i}$ is mapped to the remainder of $i$ modulo $m$. The kernel is the ideal $(\overline{m}) \subset \mathbb{Z}/n\mathbb{Z}$. Moreover, the map $q_m : \mathbb{Z} \rightarrow \mathbb{Z}/m\mathbb{Z}$ factors as

\[\begin{array}{ccc}
\mathbb{Z} & \xrightarrow{q_n} & \mathbb{Z}/n\mathbb{Z} \\
\downarrow{q_m} & & \downarrow{q_{m,n}} \\
\mathbb{Z}/m\mathbb{Z} & \end{array}\]

Proof. To check $q_{m,n}$ is a surjective ring homomorphism is essentially the same argument as the one for $q_n$, replacing $\mathbb{Z}$ with $\mathbb{Z}/n\mathbb{Z}$. (You will check some of the properties of the map $q_6$ in the next homework.)

As for the factorization of $q_m$, just notice that if $a = qn + r$, then $a$ can also be written as $a = q \cdot \frac{a}{m} + m + r$. Therefore, $a \equiv r \pmod{n}$ and $a \equiv r \pmod{m}$.  \[\square\]
7.3.5 Arithmetic Tables

Let’s work out the arithmetic table of \( \mathbb{Z}/k\mathbb{Z} \) for some small \( k \).

1) \( k = 2 \). \( \mathbb{Z}/2\mathbb{Z} \) has only two elements, which we denote as 0, 1. The arithmetic tables are as follows:

\[
\begin{array}{c|cc}
* & 0 & 1 \\
\hline
0 & 0 & 1 \\
1 & 1 & 0
\end{array}
\quad \quad \quad
\begin{array}{c|cc}
\circ & 0 & 1 \\
\hline
0 & 0 & 0 \\
1 & 0 & 1
\end{array}
\]

2) \( k = 3 \). \( \mathbb{Z}/3\mathbb{Z} \) has 3 elements, which we denote as 0, 1, 2.

\[
\begin{array}{c|ccc}
* & 0 & 1 & 2 \\
\hline
0 & 0 & 1 & 2 \\
1 & 1 & 2 & 0 \\
2 & 2 & 0 & 1
\end{array}
\quad \quad \quad
\begin{array}{c|ccc}
\circ & 0 & 1 & 2 \\
\hline
0 & 0 & 0 & 0 \\
1 & 0 & 1 & 2 \\
2 & 0 & 2 & 1
\end{array}
\]

3) \( k = 4 \). \( \mathbb{Z}/4\mathbb{Z} \) has 4 elements, which we denote as 0, 1, 2, 3.

\[
\begin{array}{c|cccc}
* & 0 & 1 & 2 & 3 \\
\hline
0 & 0 & 1 & 2 & 3 \\
1 & 1 & 2 & 3 & 0 \\
2 & 2 & 3 & 0 & 1 \\
3 & 3 & 0 & 1 & 2
\end{array}
\quad \quad \quad
\begin{array}{c|cccc}
\circ & 0 & 1 & 2 & 3 \\
\hline
0 & 0 & 0 & 0 & 0 \\
1 & 0 & 1 & 2 & 3 \\
2 & 0 & 2 & 0 & 2 \\
3 & 0 & 3 & 2 & 1
\end{array}
\]

7.4 Some “Geometry”

This section is mainly motivational. We shall see that one can create useful geometric structures for studying integers.

In previous lectures, we’ve seen that the prime ideals of \( \mathbb{Z} \) are just \((0)\) and \((p)\) where \( p \) is any prime number. Let’s draw all of them on a line, as follows:

\[
\text{Q} \quad \mathbb{Z}/2\mathbb{Z} \quad \mathbb{Z}/3\mathbb{Z} \quad \mathbb{Z}/5\mathbb{Z} \quad \mathbb{Z}/7\mathbb{Z} \quad \cdots
\]

For each of the small dots on the line, we have associated a field \( \mathbb{Z}/p\mathbb{Z} \); for the big dot, we associate \( \mathbb{Q} \).

Just as you can consider a univariate function over \( \mathbb{R} \), we can consider “functions” over this space. Notice however, for a real-valued function \( f(x) \), if you evaluate at every \( x \in \mathbb{R} \), the function value \( f(x) \) is always in \( \mathbb{R} \). However, in our
case, at each \((p)\), the field we consider is different: it is \(\mathbb{Z}/p\mathbb{Z}\), which varies with respect to \(p\). ((0) is different from the other points, which we drew bigger. We’ll come to that later.)

In particular, we consider integers themselves to be functions defined globally over this space. The **values of \(n\) at the point \((p)\) is just \(q_p(n)\).**

eg. One can regard \(12\) as a global function, whose value at \((2)\) is \(\overline{5}\), at \((3)\) is \(\overline{0}\), at \((5)\) is \(\overline{2}\), at \((7)\) is \(\overline{5}\), and so on.

eg. One can also regard \(151\) as a global function, whose value at \((2)\) is \(\overline{1}\), at \((3)\) is \(\overline{1}\), at \((5)\) is \(\overline{1}\), at \((7)\) is \(\overline{4}\), and so on.

Over \(\mathbb{R}\), one not only has globally defined functions, but also locally defined functions. So you might ask, what are the locally defined functions in our context? They are the rational numbers. A rational number \(\frac{a}{b}\) can be thought of defined at \((p)\) if \(p\) does not divide \(b\), i.e. the denominator is not zero in \(\mathbb{Z}/p\mathbb{Z}\). The value of \(a/b\) at \((p)\) is just \(q_p(a) \cdot q_p(b)^{-1}\).

**Remark** Notice the analogy to locally defined functions in calculus: \(\frac{1}{1-x}\) is not defined at \(x = 1\) because the denominator is 0.

eg. \(\frac{3}{5}\) is not defined at \((5)\), but is defined everywhere else. It has value \(\overline{1}\) at \((2)\), and value \(\overline{0}\) at \((3)\), and value \(\overline{2}\) at \((7)\). For example, \(q_7(3) \cdot q_7(5)^{-1} = \overline{3} \cdot \overline{3} = \overline{2}\). The multiplicative inverse of \(\overline{5}\) in \(\mathbb{Z}/7\mathbb{Z}\) is \(\overline{3}\), since \(3 \cdot 5 = 7 \cdot 2 + 1\).

\(^2\mathbb{Z}/p\mathbb{Z}\) is also called the residue field at \((p)\).
Chapter 8

Lecture 8

8.1 Warm-up

e.g. Find \( q_7(6!) \), \( q_7(6^6) \) and \( q_7(123456789) \).

1. \( q_7(6!) = q_7((-1)(-2)(-3) \cdot 2 \cdot 1) = q_7((-1)(-1)(-1) \cdot 1) = q_7(-1) = \overline{6}. \)

(In the future lectures, we will see that \( x^2 - 1 = 0 \) has exactly two solutions over \( \mathbb{Z}/7\mathbb{Z} \). It is not hard to see 6 and 1 are the solutions. All other non-zero elements pair up to give 1, since each non-zero element has a multiplicative inverse. Using this fact, one can conclude immediately that \( q_7(6!) = q_7((-1) \cdot 1 \cdot \ldots \cdot 1 \cdot 1) = \overline{6}. \))

2. \( q_7(6^6) = q_7((-1)^6) = q_7(1) = \overline{1}. \)

123456789 = 1 \cdot 10^8 + 2 \cdot 10^7 + \ldots + 9 \cdot 10^0. \) Notice that \( q_7(10) = \overline{3}, q_7(10^2) = \overline{3^2} = \overline{2}, q_7(10^3) = \overline{3 \cdot 2} = \overline{6}, q_7(10^4) = \overline{3 \cdot 6} = \overline{4}, q_7(10^5) = \overline{3 \cdot 4} = \overline{5}, q_7(10^6) = \overline{3 \cdot 5} = \overline{1}. \)

Thus, \( q_7(10^7) = q_7(10) = \overline{3}, q_7(10^8) = q_7(10^2) = \overline{2}. \)

Hence,

\[
q_7(1 \cdot 10^8 + 2 \cdot 10^7 + \ldots + 9 \cdot 10^0) \\
= \overline{1} \cdot \overline{2} + \overline{3} + \overline{3} \cdot \overline{1} + \overline{4} \cdot \overline{5} + \overline{5} \cdot \overline{4} + \overline{6} + \overline{6} + \overline{2} \cdot \overline{3} + \overline{2} \cdot \overline{1} \\
= \overline{2} + \overline{6} + \overline{3} + \overline{6} + \overline{1} + \overline{6} + \overline{6} + \overline{2} \\
= \overline{3}
\]

8.2 Some Geometry (continued)

Yesterday, we talked about functions on the following space:
1. Global functions are integers;

2. the set of functions locally defined at \((p)\) is denoted as \(\mathbb{Z}_p\); elements in \(\mathbb{Z}_p\) are roughly rational numbers whose denominator is not divisible by \(p\).

Notice the following phenomenon: 1 can be written as \(\frac{7}{7}\) in \(\mathbb{Q}\). But \(\frac{7}{7}\) shouldn’t be allowed as something equivalent to 1, when we consider “functions locally defined at \((7)\)”. We now make this idea into rigorous formal language (for fun).

**Definition 18.** Let \(p\) be a prime number. Define \(\mathbb{Z}_p = \{(a, b) | a, b \in \mathbb{Z}, b \notin (p)\}/\sim\), where the equivalence relation \(\sim\) identifies \((a, b)\) with \((a', b')\), if \(b'a - ba' = 0 \in \mathbb{Z} \).

**Remark** The intuition should be clear. You should think of \((a, b)\) as just the rational number \(\frac{a}{b}\). The reason for writing this set of rational numbers in form of some set modulo some equivalence relation is precisely to rule out those illegal symbols.

Since one can add and multiple rational numbers, it is not hard to show the following:

**Proposition 24.** \((\mathbb{Z}_p, +, \cdot)\) is an integral domain.

**Remark** Clearly, \(\mathbb{Z}_p\) is not a field, because \(p\) is an element in \(\mathbb{Z}_p\) which is non-zero, but does not have a reciprocal in \(\mathbb{Z}_p\).

One can think of \(\mathbb{Q}\) as the collection of all “functions” over \(\mathbb{Z}\), some of which are globally defined, some of which are locally defined.

**Remark** In earlier lectures, we mentioned the terminology “to solve a diophantine equations locally”. We can now be more precise about what we mean. It means to mod out the diophantine equation by every prime number \(p\) and solve the congruence equation inside a field related to \(\mathbb{Z}/p\mathbb{Z}\) moreover, one also needs to solve the equation over \(\mathbb{R}\). Yet it is NOT always true that if the equation is everywhere locally solvable then it is globally solvable (a.k.a the original diophantine equation admits integer solutions). However, the celebrated Hasse-Minkowski Theorem says this is true for quadratic diophantine equations.

This philosophy in number theory is referred to as the **Local-global Principle**.

---

\(^1\)More precisely, it is the completion of \(\mathbb{Q}\) with respect to the \(p\)-adic absolute value we introduced in the motivation lecture.
8.2.1 Relations among \( \mathbb{Z}, \mathbb{Z}_p \) and \( \mathbb{Z}/p\mathbb{Z} \)

Let’s summarize the relations among the three rings \( \mathbb{Z}, \mathbb{Z}_p, \mathbb{Z}/p\mathbb{Z} \).

**Proposition 25.**

1. There is a ring inclusion \( i_p : \mathbb{Z} \to \mathbb{Z}_p : a \mapsto \frac{a}{1} \).

2. There is a surjective ring homomorphism \( \mathbb{Z}_p \to \mathbb{Z}/p\mathbb{Z} : \frac{a}{c} \mapsto q_p(a) \circ q_p(c)^{-1} \).

Moreover, \( q_p \) factorizes as follows:

\[
\begin{array}{ccc}
\mathbb{Z} & \xrightarrow{i_p} & \mathbb{Z}_p \\
& \searrow & \downarrow q_p \\
& i_p & \downarrow q_p \\
\mathbb{Z}_p & \rightarrow & \mathbb{Z}/p\mathbb{Z}
\end{array}
\]

**Proof.** One can check these results from the definitions directly.

**Remark**

The first can be interpreted as saying globally defined function are naturally locally defined functions. The second one can be thought of as saying one can evaluate locally defined functions (which includes all globally defined functions).

Lastly, similar to \( \mathbb{Z}_p \), one gets \( \mathbb{Z}_0 \), which is just \( \mathbb{Q} \)!

One might ask how this kind of thinking could be helpful. Our first example would be a preparation for the Chinese Remainder Theorem.

Consider the prime ideals of \( \mathbb{Z}/m\mathbb{Z} \). The prime ideals of \( \mathbb{Z}/m\mathbb{Z} \) corresponds to precisely the prime factors of \( m \). (Note: if \( m \) is prime, \( 0 \) is the only prime ideal of \( \mathbb{Z}/m\mathbb{Z} \); otherwise, it is not a prime ideal.)

**eg.** Consider \( R = \mathbb{Z}/6\mathbb{Z} \). Clearly, it’s prime ideals are just \( (2), (3) \). So we have a space of two points, and we associate the fields \( \mathbb{Z}/2\mathbb{Z} \) and \( \mathbb{Z}/3\mathbb{Z} \) to them respectively. Again, we think of elements in \( \mathbb{Z}/6\mathbb{Z} \) as global functions, which are taking values at both points. Intuitively, one would think that a function over a space of two points should be just determined by its values at the two points. It turns out that this is indeed the case. A fancier to say this: one gets an isomorphism of rings \( \mathbb{Z}/6\mathbb{Z} \cong \mathbb{Z}/2\mathbb{Z} \times \mathbb{Z}/3\mathbb{Z} \).

### 8.3 Euler \( \varphi \)-function

When \( b \) is not necessarily prime, we can classify all elements in \( \mathbb{Z}/b\mathbb{Z} \) in the following way:

1. The zero element \( \bar{0} \);
2. units in \( \mathbb{Z}/b\mathbb{Z} \), i.e. elements that admit multiplicative inverses;
3. non-units in \( \mathbb{Z}/b\mathbb{Z} \), i.e. elements that do not admit multiplicative inverses.

**eg.** Take \( b = 6 \). \( \mathbb{Z}/6\mathbb{Z} \) has elements \( \bar{0}, \bar{1}, \bar{2}, \bar{3}, \bar{4}, \bar{5} \).
1. $\bar{0}$ is the zero element;
2. $\bar{1}$ and $\bar{5}$ are units;
3. $\bar{2}, \bar{3}, \bar{4}$ are zero-divisors.

A natural question to ask is how many units are there in $\mathbb{Z}/b\mathbb{Z}$? This can be translated into the following question about integers: how many integers between 0 and $b-1$ are coprime to $b$?

**Example.** $i \cdot j = \bar{1}$ if and only if $i \cdot j = c \cdot b + 1$ for some $c \in \mathbb{Z}$. By the theory of greatest common divisors, the latter is equivalent to the condition $(i, b) = 1$.

Euler defined the $\varphi$ function to be the answer to this question:

**Definition 19.** Euler’s $\varphi$-function associates to every integer $n$ greater than 1 the number of positive integers smaller than $n$, which are co-prime to $n$.

For future reference, we also introduce the following notation:

**Definition 20.** Denote $U_n$ to be the set of all units in $\mathbb{Z}/n\mathbb{Z}$.

Then,

$$\varphi(n) = |\{m|0 \leq m \leq n-1, (m, n) = 1\}|$$

= number of units in $\mathbb{Z}/n\mathbb{Z}$

= $|U_n|$

Our first goal is to prove a computational formula for $\varphi(n)$: $\varphi(n) = n \cdot \prod_{p|n} (1 - \frac{1}{p})$.

We start with a preliminary result.

**Proposition 26.** Suppose $m, n > 1$, such that $(m, n) = 1$. Then $\varphi(mn) = \varphi(n) \varphi(m)$.

Before we prove this proposition, we need two lemmas about complete residue systems.

**Lemma 10.** Let $\{a_1, ..., a_b\}$ be a complete residue system modulo $b$. Then,

1. $\{a_1 + m, ..., a_b + m\}$ is a complete residue system modulo $b$.
2. If $(n, b) = 1$, then $\{n \cdot a_1, ..., n \cdot a_b\}$ is a complete residue system modulo $b$.

**Proof.** (1) For every $i$, $q_b(a_i + m) = q_b(a_i) + q_b(m)$. As sets,

$$\{q_b(a_1) + q_b(m), ..., q_b(a_b) + q_b(m)\} = \{\bar{0} + q_b(m), ..., \bar{b-1} + q_b(m)\} = \{\bar{0}, ..., \bar{b-1}\}.$$

(2) As sets,

$$\{q_b(n) \cdot q_b(a_1), ..., q_b(n) \cdot q_b(a_b)\} = \{q_b(n) \cdot \bar{0}, ..., q_b(n) \cdot \bar{b-1}\} = \{\bar{0}, ..., \bar{b-1}\}.$$

This is because any element $\bar{i}$ in $\{\bar{0}, ..., \bar{b-1}\}$ can be written as a product of the form

$$\bar{i} = (i \cdot q_b(n)^{-1}) \cdot q_b(n).$$

\qed
8.3. EULER $\varphi$-FUNCTION

Lemma 11. Suppose $\ell, m, n > 0$. $(\ell, m \cdot n) = 1$ if and only if $(\ell, m) = 1$ and $(\ell, n) = 1$.

Proof. This follows from Fundamental Theorem of Arithmetic, by comparing the primes factors of $\ell, m$ and $n$.

Now we can prove the proposition.

Proof. (Proof of the proposition)

One can write $0, 1, \ldots, mn$ in the following way:

\[
\begin{array}{cccccccc}
0 & 1 & 2 & 3 & \ldots & m - 2 & m - 1 \\
0 + m & 1 + m & 2 + m & 3 + m & \ldots & 2m - 2 & 2m - 1 \\
\vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\
0 + (n - 1)m & 1 + (n - 1)m & 2 + (n - 1)m & 3 + (n - 1)m & \ldots & nm - 2 & nm - 1
\end{array}
\]

By Lemma 10 part (1), each row is a complete residue system modulo $m$.
By Lemma 10 part (1) and (2), each column is a complete residue system modulo $n$. (Here we have to use the assumption $(m, n) = 1$.)
By Lemma 11 an integer is coprime to $mn$ iff it is coprime to $n$ and to $m$. To get $\varphi(mn)$, we can count the number of positive integers that are both coprime to $m$ and coprime to $n$.
Notice that in the first row, there are precisely $\varphi(m)$ many integers coprime to $m$, by the definition of $\varphi(m)$. Moreover, all integers in the same column are congruent modulo $m$. Therefore, among all the $n$ columns there are precisely $\varphi(m)$ of them consisting of integers coprime to $m$.
Since every one of the $\varphi(m)$ columns is a complete residue system modulo $n$, each of them contains precisely $\varphi(n)$ many of integers which are also coprime to $n$.
Thus, $\varphi(mn) = \varphi(m) \varphi(n)$.

\[
\text{eg. A concrete example of the argument above: } n = 3, m = 4:
\]

\[
\begin{array}{cccc}
0 & 1 & 2 & 3 \\
4 & 5 & 6 & 7 \\
8 & 9 & 10 & 11
\end{array}
\]

The numbers in the second and fourth columns are all coprime to $m = 4$. Inside the second column, 1 and 5 are also coprime to $n = 3$; inside the last column, 7 and 11 are coprime to 3.
$\varphi(12) = \varphi(3) \varphi(4) = 2 \cdot 2 = 4$.

Proposition 27. For every $n > 1$, $\varphi(n) = n \cdot \prod_{p|n} (1 - \frac{1}{p})$.

Proof. By the fundamental theorem of arithmetic, $n$ has a unique prime factorization

\[
n = p_1^{m_1} \cdots p_k^{m_k}.
\]

By Proposition 26 $\varphi(n) = \prod_{p|n} \varphi(p_i^{m_i})$, since $(p_i^{m_i}, p_j^{m_j}) = 1$ when $p_i \neq p_j$.
Thus, it suffices to show that $\varphi(p^s) = p^s(1 - \frac{1}{p})$ for any prime number $p$.

The integers between 0 and $p^s - 1$ which are NOT coprime to $p^s$ are the integers in this range divisible by $p$:

<table>
<thead>
<tr>
<th>0</th>
<th>$p$</th>
<th>$2p$</th>
<th>$3p$</th>
<th>...</th>
<th>$(p-2)p$</th>
<th>$(p-1)p$</th>
</tr>
</thead>
<tbody>
<tr>
<td>0 + $p^2$</td>
<td>$p + p^2$</td>
<td>$2p + p^2$</td>
<td>$3p + p^2$</td>
<td>...</td>
<td>$(p-2)p + p^2$</td>
<td>$(p-1)p + p^2$</td>
</tr>
</tbody>
</table>

One can check these numbers are precisely integers of the form $a_1 p + a_2 p^2 + ... + a_{s-1} p^{s-1}$, where $a_i : 0 \leq a_i \leq p - 1$. Since there are $p$ choices for each $a_i$, there are $p^{s-1}$ many of such integers.

Hence, we get $\varphi(p^s) = p^s - p^{s-1} = p^s(1 - \frac{1}{p})$. This concludes the proof. $\square$

**Example.** When $p = 3$, $s = 2$, integers between 0 and $3^2 - 1$ which are not coprime to $3^2$ are 0, 3 and 2·3. One can see they are precisely all integers of the form $a_1 \cdot 3$, with $a_1 : 0 \leq a_1 \leq 3 - 1$. 

eg. When $p = 3$, $s = 2$, integers between 0 and $3^2 - 1$ which are not coprime to $3^2$ are 0, 3 and 2·3. One can see they are precisely all integers of the form $a_1 \cdot 3$, with $a_1 : 0 \leq a_1 \leq 3 - 1$. 

......
Chapter 9

Lecture 9

9.1 Problem Session

eg. Find \( q_7(9^{10000}) \).

Since \( q_7(9) = 2 \) and \( 2^3 = \bar{1} \), we get \( q_7(9^{10000}) = \bar{1}^{3333} \cdot 2 = \bar{2} \).

eg. Find a positive integer \( n \) such that \( 3^n = \bar{1} \) in \( \mathbb{Z}/8\mathbb{Z} \).

In \( \mathbb{Z}/8\mathbb{Z} \), \( 3 \cdot 3 = \bar{1} \); so one can take \( n = 2 \).

eg. Find a positive integer \( m \) such that \( 2^m = \bar{1} \) in \( \mathbb{Z}/21\mathbb{Z} \).

In \( \mathbb{Z}/21\mathbb{Z} \), \( \bar{2}^2 = \bar{4} \), \( \bar{2}^3 = \bar{8} \), \( \bar{2}^4 = \bar{16} \), \( \bar{2}^5 = \bar{1} \), \( \bar{2}^6 = \bar{1} \); so one can take \( m = 6 \).

Remark It turns out that for every unit \( \bar{i} \) in \( \mathbb{Z}/b\mathbb{Z} \), there is always some positive integer \( n \) such that \( \bar{i}^n = \bar{1} \). We will see this in today’s lecture.

eg. How many integers between 1 and 119 are coprime to 120?

\( 120 = 2^3 \cdot 3 \cdot 5 \). \( \varphi(120) = 120 \cdot \left(1 - \frac{1}{2}\right) \left(1 - \frac{1}{3}\right) \left(1 - \frac{1}{5}\right) = 32 \). So there are 32 such integers.

Lemma 12.

\[
\varphi(n) = n \cdot \prod_{p \mid n} \left(1 - \frac{1}{p}\right)
\]

\[
= \prod_{p_i \mid n} p_i^{m_i} \left(1 - \frac{1}{p_i}\right)
\]

\[
= \prod_{p_i \mid n} p_i^{m_i - 1} (p_i - 1).
\]
where $p_i$ are the prime factors of $n$.

**Remark** Three useful facts from these equivalent forms of the formula:

1. If $p|n$, then $p - 1|\varphi(n)$;
2. if $p^s|n$, then $p^{s-1}|\varphi(n)$; in particular, if $s \geq 2$, $p|\varphi(n)$;
3. $\varphi(p) = p - 1$, when $p$ is prime.

**eg.** We now demonstrate that Euler-\(\varphi\) function provides partially information about the integer. For which positive integers $n$ is $\varphi(n) = 6$?

First, $n$ cannot have a prime factor $p$ strictly greater than 7. Otherwise $\varphi(n) \geq p - 1 > 6$. Therefore, it suffices to look at $N = 2^a \cdot 3^b \cdot 5^c \cdot 7^d$.

A first observation is $n_1 \leq 1$, otherwise $\varphi(n) \geq 42$. Also if $5|n$, one would have $4|\varphi(n)$, which is absurd. Hence, it remains to look at integers of the form $2^a \cdot 3^b \cdot 7^d$.

When $n_4 = 1$, $n_3 = 0$; otherwise, $\varphi(n) \geq \varphi(3) \cdot 6 = 12$ and one gets a contradiction. Hence, $\varphi(7) = \varphi(14) = 6$ are the only possible solutions in this case. (Notice that $\varphi(2) = 2^0 \cdot (2 - 1) = 1$.)

We now turn to the case $n_4 = 0$. It remains to look at integers of the form $2^a 3^b$.

We claim that $b = 2$. This is because, $b > 2$ will make $\varphi(n) > 6$, and $b < 2$ would imply $\varphi(n)$ is a power of 2, which is not the case. Then, $n = 9, 18$ are the only possibilities in this case.

Hence, we conclude that $n = 7, 9, 14, 18$ are all the possibilities for $n$.

**Remark** Conventionally, we define $\varphi(1) = 1$. In all the formulas so far for $\varphi(n)$, we assumed $n > 1$. Notice that when we define $\varphi(n)$, we could also say “the number of all positive integers less than or equal to $n$ that are coprime to $n$”. This would not change anything, since for $n > 1$, $n$ is never coprime to itself. But for the number 1, it then makes sense to say $\varphi(1) = 1$, because $(1, 1) = 1$.

Notice also $\varphi(n) = \varphi(n \cdot 1) = \varphi(1) \cdot \varphi(n) = 1 \cdot \varphi(n)$, so it really does no harm.

### 9.2 Euler $\varphi$-function (continued)

We defined $U_b$ to be the set of all units of $\mathbb{Z}/b\mathbb{Z}$. Here is a first structural result about $U_b$.

**Proposition 28.** $(U_b, \cdot)$ is an abelian group.

**Remark** Since this involves multiplication rather than addition, we call it a multiplicative sub-group of $\mathbb{Z}/b\mathbb{Z}$, to emphasize the difference.

**Proof.** This statements breaks down into four sentences:

1. First of all, if $i, j$ are units, then $ij$ is also a unit: $(ij) \cdot (ij^{-1}j^{-1}) = 1$.

Hence, $U_b$ is closed under multiplication.
(2) Associativity and commutativity of \( \circ \) is inherited from \( \mathbb{Z}/b\mathbb{Z} \) directly.

(3) \( \overline{1} \in \mathbb{Z}/b\mathbb{Z} \) is always a unit and \( \overline{1} \cdot \overline{1} = \overline{1} \).

(4) Every element has inverse with respect to \( \cdot \), which follows from the definition of a unit.

In the first three problems of today, we have seen some examples of powers of units being equal to \( \overline{1} \). This is true in general, known as Euler’s theorem, which we now state and prove. One cannot over-estimate its importance in the current chapter.

**Theorem 4.** If \( \overline{i} \in U_b \), then \( \overline{i}^{\varphi(b)} = \overline{1} \).

**Proof.** We first show that there is some number \( N > 0 \) such that \( \overline{i}^N = \overline{1} \). We have shown \( (U_b, \cdot) \) is an abelian group, and it has finitely many elements. Hence, \( \overline{i}^m = \overline{i}^n \) for some \( m, n \) such that \( m > n \); the group is finite, so a large enough power of \( \overline{i} \) must equal some smaller power of \( \overline{i} \). Therefore, multiply \( (\overline{i}^m)^{-1} \) on both sides of the identity, conclude \( \overline{i}^{n-m} = \overline{1} \), and \( n - m > 0 \).

Denote \( S \) to be the set of all positive integers \( N \) such that \( \overline{i}^N = \overline{1} \). We have shown that \( S \) is non-empty. By the Well-ordering Axiom, \( S \) has a smallest element \( n_0 \).

We next show that \( n_0 \mid \varphi(b) \). Notice that if every element of \( U_b \) is a power of \( \overline{i} \), then \( n_0 = \varphi(m) \), and we are done. Otherwise, we can find some \( \overline{j} \in U_b \) such that \( \overline{j} \neq \overline{i}^m \) for any \( m \). Then, \( \overline{j} \cdot \overline{i}, \overline{j} \cdot \overline{i}^2, \ldots, \overline{j} \cdot \overline{i}^{n_0} \) are \( n_0 \) distinct elements in \( U_b \). Moreover, one can check that \( \overline{i}, \ldots, \overline{i}^{n_0}, \overline{j} \cdot \overline{i}, \ldots, \overline{j} \cdot \overline{i}^{n_0} \) are \( 2n_0 \) many distinct elements all together: if \( \overline{j} \cdot \overline{i}^m = \overline{i}^n \), \( \overline{j} \) would be a power of \( \overline{i} \).

By induction, every time there is some \( \overline{l} \in U_b \) not among the previously found units, one can find \( n_0 \) many distinct elements, \( \overline{l}, \ldots, \overline{l} \cdot \overline{i}^{n_0} \) in \( U_b \), that are different from the previously found ones. Since \( U_b \) has finitely many elements, this process terminates in finitely many steps. Since at each step we pick out \( n_0 \) many distinct elements, the number \( \varphi(b) \), which is the size of \( U_b \), is a multiple of \( n_0 \). Say, \( \varphi(b) = d \cdot n_0 \).

Hence we get \( \overline{i}^{\varphi(b)} = (\overline{i}^{n_0})^d = \overline{1}^d = \overline{1} \).

**Corollary 6.** (Fermat’s Little Theorem) Suppose \( p \) is a prime number, and \( \overline{i} \neq \overline{0} \) in \( \mathbb{Z}/p\mathbb{Z} \). Then, \( \overline{i}^{p-1} = \overline{1} \).

**Proof.** \( \varphi(p) = p - 1 \).

It turns out that the \( n_0 \) in the proof of Euler’s theorem is an important number associated to a element in \( U_b \):

**Definition 21.** Let \( \overline{i} \in U_b \). The smallest positive integer \( n_0 \) such that \( \overline{i}^{n_0} = \overline{1} \in \mathbb{Z}/b\mathbb{Z} \) is called the order of \( \overline{i} \) in \( U_b \), and is denotes as \( \text{ord}_b(\overline{i}) \).
Proposition 29. Let $\bar{i} \in U_b$. Suppose $N$ is a positive number such that $\bar{i}^{-N} = 1$, then $\text{ord}_b(\bar{i}) | N$.

Proof. We claim $\mathcal{N} = \{ N | \bar{i}^N = 1 \}$ is an ideal of $\mathbb{Z}$.

We first check $(\mathcal{N}, +)$ is an abelian group. If $N_1, N_2 \in \mathcal{N}$, then $\bar{i}^{N_1 + N_2} = \bar{i}^{N_1} \cdot \bar{i}^{N_2} = 1 \cdot 1 = 1$. This shows that $\mathcal{N}$ is closed under addition. $\bar{i}^{-N} \cdot \bar{i}^{-N} = \bar{i}$ implies $\bar{i}^{-N} = 1$, so the negative of any $N \in \mathcal{N}$ is still in $\mathcal{N}$. (In particular, $1^0 = 1$ and $0 \in \mathcal{N}$.) Thus, $\mathcal{N}$ is an (additive) subgroup of $\mathbb{Z}/b\mathbb{Z}$.

Secondly, take any $d \in \mathbb{Z}$ and any $N \in \mathcal{N}$, $\bar{i}^{dN} = (\bar{i}^N)^d = 1^d = 1$; so, $dN \in \mathcal{N}$. We have thus showed $\mathcal{N}$ is an ideal of $\mathbb{Z}$.

We know that all ideal of $\mathbb{Z}$ are of the form $(m)$, where $m$ is the smallest element in the ideal. (This is just theory of g.c.d! Since $\text{ord}_b(\bar{i})$ is the smallest element in $\mathcal{N}$. Hence, $\mathcal{N} = \{ \text{ord}_b(\bar{i}) \}$. \hfill \qed

Remark When $N < 0$, $\bar{i}^N = (\bar{i}^{-1})^{-N}$, so this operation makes sense.

The only reason to include negative power into consideration is to utilize the nice result about ideals of $\mathbb{Z}$.

We’ll end this section with another property of the $\varphi$ function, which we shall refer to in later lectures.

Proposition 30. For $n > 0$, $\sum_{d|n} \varphi(d) = n$.

Proof. We basically just regroup the elements of the set $S = \{ 1, 2, \ldots, n \}$. First of all, let $d$ be a positive factor (not necessarily prime) of $n$. Denote the subset $C_d \subset S$ to be the set of integers $a$ between 1 and $n$ such that the greatest common divisor $(a, n) = d$. Since $(a, n)$ is a determined number, $C_d \cap C_{d'} = \emptyset$ if $d \neq d'$. Moreover, for any number $a$ between 1 and $n$, $(a, n)|n$ by definition. Therefore

$$n = |\{ 1, 2, \ldots, n \}| = \sum_{d|n} |C_d|$$

It remains to compute size of $C_d$. We’ve seen in a previous lecture that $(a, n) = d$ if and only if $(\frac{a}{d}, \frac{n}{d}) = 1$. This tells us that when counting the number of elements in $C_d$, we can equivalently count all the integers between 1 and $\frac{n}{d}$ which are coprime to $\frac{d}{d}$. The latter number is just $\varphi(\frac{d}{d})$, by definition.

A last observation: $\sum_{d|n} \varphi(\frac{n}{d}) = \sum_{d|n} \varphi(d)$. This is just saying, when $d$ runs through all the factors of $n$, $\frac{n}{d}$ also runs through all the factors of $n$. (See the following example.) Hence we are done. \hfill \qed

eg. The factors of $n = 24$ are $d = 1, 2, 3, 4, 6, 8, 12, 24$. Meanwhile, $\frac{n}{d} = 24, 12, 8, 6, 4, 3, 2, 1$.

eg. Let’s do a concrete example of this result. Take $n = 44$.

---

$|A|$ denotes the size of the set $A$. 

The factors of 44 are 1, 2, 4, 11, 22, 44.

\[\varphi(1) = 1, \varphi(2) = 1, \varphi(4) = 4 - 2 = 2, \varphi(11) = 10, \varphi(22) = 10, \varphi(44) = 2 \cdot 10 = 20.\]

\[1 + 1 + 2 + 10 + 10 + 20 = 44.\]
Chapter 10

Lecture 10

10.1 Problem Session

eg. Compute $\varphi(p) + \varphi(p^2) + ... + \varphi(p^n)$, where $p$ is a prime number.

$\varphi(p) + \varphi(p^2) + ... + \varphi(p^n) = (p - 1) + (p^2 - p) + ... + (p^n - p^{n-1}) = p^n - 1$.

Alternatively, since $1, p, ..., p^n$ are all the positive factors of $p^n$, one proposition from yesterday’s lecture says

$$\sum_{i=0}^{n} \varphi(p^i) = p^n.$$  

Thus, $\sum_{i=1}^{n} \varphi(p^i) = p^n - 1$.

eg. Determine when is $\varphi(n)$ a prime number.

Let $n = p_1^{m_1} \cdots p_k^{m_k}$ be the unique prime factorization of $n$.

Notice that $p - 1 | \varphi(p^i)$ for all primes. $\varphi(n) = \prod \varphi(p_i^{m_i})$. If $n$ has three distinct prime factors, then at least two of them are larger than 2 and hence $\varphi(p_i^{m_i}) > 1$ for at least two $i$’s. This implies that $\varphi(n)$ is a composite.

Thus, it suffices to look at cases where $k \leq 2$, i.e. $n$ has at most two prime factors.

Notice also if $p | n$ and $p > 3$, then $p - 1 | \varphi(n)$ is an even number $> 2$ and again $\varphi(n)$ is a composite. So it suffices to look at $n = 2^s 3^t$.

1. When $k = 2$, one must have $s \leq 2$, otherwise $4 | \varphi(n)$. When $s = 2$, $n = 2^2 = 4$ and cannot have other factors, so we get a contradiction; when $s = 1$, we
10.2. ARITHMETIC PROGRESSION: A FIRST LOOK

get \( n = 6 \) is the only possibility.

2. When \( k = 1 \), we get \( n = 4 \) or \( n = 3 \).

In conclusion, \( n = 2, 4, 6 \) are the only possibilities.

**eg.** Can \( \mathbb{Z}/23\mathbb{Z} \) has a unit of order 4? No, because 4 does not divide 22!

With more detail, 23 is prime and \( \varphi(23) = 23 - 1 = 22 \).

By Fermat’s little theorem, \( i^{22} = 1 \); hence, \( ord_{23}(i)|22 \).

Since 4 does not divide 22, there does not exist a unit of order 4 in \( \mathbb{Z}/23\mathbb{Z} \).

Alternatively, you can do a proof by contradiction. Suppose there is a unit \( i \) of order 4, then \( i^2 = 1 \) also holds, since \( 2 = 22 - 4 \cdot 5 \) and \( i^2 = i^{22} \cdot (i^4)^{-5} \). This gives a contradiction.

**eg.** Does \( \mathbb{Z}/b\mathbb{Z} \) has a unit of order 2 for every \( b > 1 \)? Almost, since \( b - 1 \cdot 2 = 1 \) always holds. But be careful, \( \mathbb{Z}/2\mathbb{Z} \) does not have a unit of order 2: \( \bar{1} \) is the only unit and has order 1; another way to put it, \( \varphi(2) = 1 \).

**Remark** Warning: The following is a common mistake: “Since \( \bar{i}^N = \bar{1} \), \( ord_k(\bar{i}) = N \)”. **NO. THIS IS NOT RIGHT.** You can only conclude \( ord_k(\bar{i})|N \).

10.2 Arithmetic Progression: a first look

Earlier, we have shown the result that there exist infinitely many prime numbers. However, one can say even more about the distribution of prime numbers. Our next topic will be called “arithmetic progression”, which basically studies integers of the form \( q \cdot a + r \), for fixed \( a \) and \( r \). One way to think of it is the following: we can divide \( \mathbb{Z} \) into a disjoint union of subsets

\[
\mathbb{Z} = \bigcup_{r=0}^{a-1} \{ q \cdot a + r | q \in \mathbb{Z} \} = \bigcup_{r=0}^{a-1} q^{-1}(r),
\]

and we want to say something about integers in \( q^{-1}(r) \).

A weak form of the famous Dirichlet’s theorem states the following:

**Theorem 5.** Suppose \( (a, r) = 1 \). Then, there are infinitely many prime numbers of the form \( q \cdot a + r \).

The proof of this result involves some advanced machinery called Dirichlet’s L-series. However, with what we’ve seen so far, we can prove some special case of this theorem. In particular, Euler’s theorem and the notion of order are very useful.

In the previous section, we have defined the order of an element \( \bar{i} \) in \( \mathbb{Z}/b\mathbb{Z} \), that is, the smallest positive integer \( n \) such that \( \bar{i}^n = \bar{1} \). We look at some examples of small \( b \).
10.2.1 Simplest Cases

eg. When $b = 2$, $\mathbb{Z}/2\mathbb{Z}$ has only one unit, namely $1$, which has order 1. Dirichlet’s theorem is trivial in this case (i.e. “there are infinitely many prime numbers of the form $2n + 1$”, or “there are infinitely many odd primes”), since we have showed previously that there are infinitely many prime numbers, and clearly only one of them is even.

eg. When $b = 3$, $1, \overline{2}$ are the units in $\mathbb{Z}/3\mathbb{Z}$. $1$ has order 1, and $\overline{2}$ has order 2. Since $1 \cdot \ldots \cdot 1 = 1$, this basically says $3n + 1$’s can be separated from $3n + 2$’s in some sense: a product of numbers of the form $3n + 1$ is never of the form $3n + 2$.

Using this observation, the following case may be the simplest one of Dirichlet’s Theorem.

**Proposition 31.** There are infinitely many prime numbers of the form $3n + 2$.

**Proof.** Clearly, 5 is a prime of this form. Suppose you have obtained $p_1, \ldots, p_n$ which are prime numbers of this form. (We start with 5.)

Let $Q = 3 \cdot p_1 \cdot p_2 \cdot \ldots \cdot p_n + 2$. We consider the prime factors of $Q$. None of these prime factors can be $p_1, \ldots, p_n$, because otherwise, this $p_i$ would divide 2, which is impossible. Nor could 2 be a factor, since $Q$ is odd. Then, we claim that $Q$ has at least one prime factor of the form $3n + 2$. This follows directly from the observation $\overline{1}^m = \overline{1}$, i.e. a product of numbers of the form $3n + 1$ can never be of the form $3n + 2$. Thus we have found one more prime number of the form $3n + 2$. By induction, there are infinitely many prime numbers of this form.

As you can see, the same argument does NOT apply to show “there are infinitely many prime numbers of the form $3n + 1$”, precisely because $(3n+2)(3m+2) = 3(3mn+2m+2n+1)+1$. However, the argument does apply to the following case:

eg. $k = 4$. Only $1, \overline{3}$ are units. By the same argument as before, we get:

**Proposition 32.** There are infinitely many prime numbers of the form $4n + 3$.

**Proof.** One replaces the previous construction of $Q$ by $Q = 4p_1p_2\ldots p_n + 3$, and the same argument applies.

As you might have observed, these argument are relatively simple, because the other units cannot generate the particular unit we have chosen. This is certainly not always the case. Otherwise, Dirichlet’s theorem wouldn’t be a hard statement.

eg. $b = 5$. We compute the table of powers of units in $\mathbb{Z}/5\mathbb{Z}$.
As we see from the table, the powers of 2 or 3 run through all units of $\mathbb{Z}/5\mathbb{Z}$.
Hence, it is impossible to apply the previous argument to any of the cases $5n + 1, ..., 5n + 4$.

10.2.2 Less Simple Case

It turns out that the case of $4n + 1$ is the next simplest case. It bespeaks the power of Euler’s theorem.

**Proposition 33.** There are infinitely many prime numbers of the form $4n + 1$.

We first prove a lemma.

**Lemma 13.** Let $a > 1$ be an integer. The prime factors of $a^2 + 1$ that are greater than 2 must be of the form $4n + 1$.

**Proof.** $p|(a^2 + 1)$ implies that $a^2 = p \cdot q - 1$ for some $q$, i.e. $(q_1(a))^2 + 1$ is a prime. Hence, $(q_1(a))^2 = \overline{1}$. This implies that 4 should be a multiple of $\text{ord}_p(q_1(a))$, and hence $\text{ord}_p(q_1(a))$ can only be 1, 2 or 4.

But $q_1(a)^2 + 1 = \overline{0}$, and we assumed $p > 2$, which means $-1 \neq \overline{1}!$. Hence, the order cannot be 2. It also cannot be 1, since that would mean $q_1(a) = \overline{1}$, which would imply $(q_1(a))^2 + 1 = 2 \neq \overline{0}$. ($p > 2$)

Hence, $\text{ord}_p(q_1(a))$ must be 4.

Since $\varphi(p) = p-1$, as we have seen in previous section, $\text{ord}_p(a)|\varphi(p)$, conclude $p - 1 = 4 \cdot n$ for some $n$, i.e. $p = 4n + 1$.

Now the proof of the proposition:

**Proof.** (Proof of the proposition) Clearly, 5=4+1. In general, suppose $p_1, p_2, ..., p_n$ are all primes of the form $4n + 1$, take $Q = (2p_1 \cdots p_n)^2 + 1$, then $Q$ must contain some prime factor, which cannot be any of the $p_i$'s, or 2; otherwise, 2|1 or $p_i|1$ which is absurd. By induction, one can conclude that there are infinitely many prime numbers of the form $4n + 1$.

In fact one can generalize this result a little bit:

**Proposition 34.** For any $r \geq 1$, there are infinitely many prime numbers of the form $2^r \cdot n + 1$.

**Proof.** The proof is more or less the same as for the previous case.

We claim any odd prime factors of $a^{2^r-1} + 1$ must be of the form $2^r \cdot n + 1$ ($a > 1$).
This is because $(q_1(a))^{2^r-1} + 1 = \overline{0}$, and hence $(q_1(a))^{2^r} = \overline{1}$. Thus, the order
must be \(2^i\) for some \(i \leq r\). But if \(i < r\), one would get \((q_p(a))^2^{2^{r-i-1}} + 1 = 2 \neq 0\). So, \(ord_p(q_p(a)) = 2^r\). Hence, \(2^r \mid p - 1\), and \(p = 2^r \cdot n + 1\).

By this claim, take \(a = 2\), we immediately see that \(2^{r+1}\) has an odd prime factor of the desired form, which we can take as our \(p_1\).

Eventually, let \(Q = (2^{p_1}p_2 \cdots p_n)^{2^r} + 1\), where \(p_i's\) are prime numbers of the form \(2^r \cdot n + 1\). We know that \(Q\) is odd and none of \(p_i\) divides \(Q\). So, by the above claim, \(Q\) must have a prime factor of the form \(2^r \cdot n + 1\) other than \(p_1, \ldots, p_n\). By induction, we’ve proved the result.

\textbf{eg.} In your next HW, you will be asked to prove there are infinitely many prime numbers of the form \(5n + 1\). More generally, can you prove there are infinitely many prime numbers of the form \(qn + 1\), where \(q\) is some fixed \textbf{prime} number?
Chapter 11

Lecture 11

11.1 Warm-up

eg. Show that there are infinitely many prime numbers of the form \( 6n + 5 \).

Proof. First of all, \( \mathbb{Z}/6\mathbb{Z} \) are the only units in \( \mathbb{Z}/6\mathbb{Z} \). So an odd prime number greater than 3 is either of the form \( 6n + 1 \) or of the form \( 6n + 5 \). Since \( \mathbb{T}^N \neq 5 \) and 2, 3 do not divide \( 6n + 5 \), we conclude that an integer of the form \( 6n + 5 \) must have a prime factor of the same form.

Start with \( p_1 = 11 \) and construct \( Q_n = 6P_1p_2...p_{n-1} + 5 \). By the above observation, \( Q_n \) has a prime factor of the form \( 6q + 5 \). Since each of \( p_1, ..., p_{n-1} \) does not divide \( Q_n \) (otherwise, it would divide 5, which is absurd), this prime factor must be different from \( p_1, ..., p_{n-1} \), call it \( p_n \). \((p_n \neq 5 \text{ since we start the induction with } p_1 = 11, \text{ so that } p_1, ..., p_{n-1} > 5, \text{ and } 5 \text{ cannot divide } Q_n.\)

By induction, we arrive at the conclusion.

eg. Show that there are infinitely many prime numbers of the form \( 8n + 1 \).

Proof. First, we claim that if \( p \) is a prime factor of \( a^4 + 1 \), where \( a > 1, \) \( p \) is of the form \( 8n + 1 \). This is so, because \( q_p(a)^4 + 1 \equiv 0 \) implies \( (q_p(a))^8 = 1 \). From a property of \( \text{ord}_p(q_p(a)) \), we know \( \text{ord}_p(q_p(a))|8 \), so it can only be 1, 2, 4 or 8.

We further see that it must be 8, since otherwise \( (q_p(a))^4 = \mathbb{T} \neq -\mathbb{T} \) and we get a contradiction. By Fermat's Little Theorem, \( 8|\varphi(p) = p - 1 \), i.e. \( p \) is of the form \( 8n + 1 \).

Therefore, start with \( p_1 = 17 \) and let \( Q_n = (p_1p_2...p_{n-1})^4 + 1 \). By the above claim, \( Q_n \) has a prime factor of the form \( 8m + 1 \), which is different from \( p_1, ..., p_{n-1} \) (otherwise one of them would divide 1, which is absurd). Call it \( p_n \).

By induction, we arrive at the conclusion.

11.2 Arithmetic Progression (Continued)

Here is one more example regarding arithmetic progression.
Proposition 35. There are infinitely many prime numbers of the form $3q + 1$.

Suppose $p$ is a prime number. If we can find an element in $\mathbb{Z}/p\mathbb{Z}$ of order 3, then we know $3|\varphi(p) = p - 1$, i.e. $p = 3n + 1$ for some $n$. To find such a prime $p$, based on our previous experience, one might pose the condition that $p|a^3 - 1$ for some $a > 0$. However, since $a^3 - 1 = (a - 1)(a^2 + a + 1)$, this condition could be achieved by some $a$ of the form $pn + 1$ as well. In order to rule it out, we do the following:

Lemma 14. Suppose $p > 3$ is a prime number and $p|a^2 + a + 1$ for some $a > 1$. Then, $p = 3n + 1$ for some $n > 0$.

Proof. First, $q_p(a^3 - 1) = q_p(a - 1) \cdot q_p(a^2 + a + 1) = T$, i.e. $q_p(a)^3 = T$. So the order of $q_p(a)$ has to be 1 or 3.

Since $q_p(a^2 + a + 1) = T$ and $p > 3$, we first conclude $a$ is not congruent to 1 modulo $p$, because otherwise $q_p(a^2 + a + 1)$ would equal 3.

Therefore, $ord_p(q_p(a)) = 3$. Hence, $3|p - 1$, and hence $p = 3n + 1$. \hfill \square

Proof. (Proof of the proposition.) Obviously, $7 = 3 \cdot 2 + 1$. Suppose one has obtained prime numbers $p_1, ..., p_{n-1}$ of the form $3q + 1$. Define $Q_n = (3p_1 \cdot \ldots \cdot p_{n-1})^2 + (3p_1 \cdot \ldots \cdot p_{n-1}) + 1$. Clearly, $Q_n$ is an odd number and 3 does not divide it (otherwise 3 would divide 1). Hence, it has a prime factor greater than 3, which cannot be among $p_1, ..., p_{n-1}$. By the previous lemma, this prime factor is also of the form $3q + 1$. Call is $p_n$.

By induction, we are done. \hfill \square

11.3 Linear Congruence Equations

A (univariate) linear congruence equation modulo $m$ is an equation of the form

$$ax \equiv b \pmod{m}.$$

In this section, we shall discuss how to solve for the unknown $x$. Before we do so, let’s make one observation.

It is easy to see that in the equation $ax \equiv b \pmod{m}$ the coefficients should really be thought of as elements in $\mathbb{Z}/m\mathbb{Z}$. In other words, if $a \equiv a' \pmod{m}$ and $b \equiv b' \pmod{m}$, then $ax - a'x \equiv 0 \pmod{m}$, so the two equations $ax \equiv b \pmod{m}$ and $a'x \equiv b' \pmod{m}$ are equivalent. Moreover, it suffices to solve for solution(s) not in $\mathbb{Z}$, but in $\mathbb{Z}/m\mathbb{Z}$. Once we have obtained all the solutions in $\mathbb{Z}/m\mathbb{Z}$, to translate them into integer solution is almost trivial. Thus, a linear congruence equation modulo $m$ is just a linear equation over $\mathbb{Z}/m\mathbb{Z}$.

11.3.1 The Simple Case

The simplest examples are linear congruence equation modulo $p$, where $p$ is a prime number. This is because $\mathbb{Z}/p\mathbb{Z}$ is a field and every non-zero element in it
has a multiplicative inverse! See the following example:

**eg.** Solve $6x \equiv 11 \pmod{17}$. This is the same as $6 \cdot x = 11 \pmod{17}$. Since $6 \cdot 3 = 17 + 1$, the reciprocal of $6$ in $\mathbb{Z}/17\mathbb{Z}$ is $3$. Therefore, multiply $3$ on both sides of the equation, one gets $x = 3 \cdot 11 = 16$ is the solution. One can also say “$x \equiv 16 \pmod{17}$” is the solution; or, the set of integer solutions is $S = \{16 + 17t | t \in \mathbb{Z}\}$.

### 11.3.2 A Small Generalization

The way we solved the previous example generalizes to equations of the form $ax \equiv b \pmod{m}$, where $q_m(a) \in U_m$, even if $m$ is not necessarily prime. As long as $q_m(a)$ has a multiplicative inverse in $\mathbb{Z}/m\mathbb{Z}$, the equation has a unique solution over $\mathbb{Z}/m\mathbb{Z}$.

**Remark** These cases are very similar to solving linear equations over real numbers: $2x = 3$ implies $x = 1.5$. One just multiply $\frac{1}{2}$, which is the multiplicative inverse of 2, on both sides of the equation.

### 11.3.3 The General Case

When $q_m(a)$ is not a unit in $\mathbb{Z}/m\mathbb{Z}$, the situation is subtler. Nevertheless, we know that $ax \equiv b \pmod{m}$ implies

$$ax + my = b$$

(11.1)

This is a linear diophantine equation in $x$ and $y$! From our knowledge about linear diophantine equations, (1) has a solution iff $(a,m) | b$. We also know that the solution set to $ax + my = b$ is

$$S = \{(x_0, y_0) + t \cdot (\frac{m}{(a,m)}, -\frac{a}{(a,m)})\},$$

where $(x_0, y_0)$ is some particular solution.

From the viewpoint of the linear congruence equation, we are only interested in the $x$-values: $x_0 + t \cdot \frac{m}{(a,m)}$. By collecting all the $x$-values, we have obtained all the integer solutions to the congruence equation we started with. But one can say more.

From earlier discussion, when we solve a congruence equation modulo $m$, we are more or less only interested in solutions over $\mathbb{Z}/m\mathbb{Z}$, since the equations themselves are nothing but linear equations over $\mathbb{Z}/m\mathbb{Z}$. When $q_m(a)$ is not a unit, the equation has more than one solutions. In fact, one can count the number of solutions over $\mathbb{Z}/m\mathbb{Z}$:

Two integer solutions $x_0 + t_1 \cdot \frac{m}{(a,m)}, x_0 + t_2 \cdot \frac{m}{(a,m)}$ are congruent modulo $m$ iff

$$m | (t_1 - t_2) \cdot \frac{m}{(a,m)}.$$
This means \( qm = \frac{t_2 - t_1}{(a,m)} m \), for some \( q \), i.e. \( q \cdot (a,m) = t_1 - t_2 \). Hence, the two solutions are congruent iff \( (a,m)|(t_1 - t_2) \), i.e. \( t_1 \equiv t_2 \pmod{(a,m)} \). Therefore, start with any particular solution \( x_0 \),

\[
qm(x_0), qm(x_0 + \frac{m}{(a,m)}), qm(x_0 + \frac{2m}{(a,m)}), \ldots, qm(x_0 + \frac{m}{(a,m) \cdot (a,m) - 1})
\]

are all the distinct solutions over \( \mathbb{Z}/m\mathbb{Z} \). (Notice, when \( (a,m) = 1 \), there is a unique solution over \( \mathbb{Z}/m\mathbb{Z} \! \))

Therefore, we can draw the following conclusion:

**Proposition 36.** The equation \( ax \equiv b \pmod{m} \) has a solution iff \( (a,m)|b \). In this case, it has \( (a,m) \) many solutions over \( \mathbb{Z}/m\mathbb{Z} \).

**eg.** Solve \( 12x \equiv 34 \pmod{10} \).

Notice that \( 12 \equiv 2 \pmod{10}, \) \( 34 \equiv 4 \pmod{10} \), we can equivalently solve the equation \( 2x \equiv 4 \pmod{10} \).

Since \( (2,10) = 2 \) and divides 4, the equation has a solution.

To find a particular solution is easy, since \( 2 \cdot 2 \equiv 4 \pmod{10} \).

So, take \( x_0 = 2 \). By the above proposition, we know there should two solutions over \( \mathbb{Z}/10\mathbb{Z} \), namely \( q_{10}(2 + 0 \cdot \frac{10}{2}) = 2, q_{10}(2 + 1 \cdot \frac{10}{2}) = 7 \).

We can also say that the set of integer solutions is

\[
S = \{2 + 10q|q \in \mathbb{Z}\} \cup \{7 + 10q|q \in \mathbb{Z}\}.
\]

**Remark** We point out the following counter-intuitive fact: when \( m \) is not prime, a polynomial equation over \( \mathbb{Z}/m\mathbb{Z} \) may have more solutions than the degree of the polynomial. In the previous example, the equation is linear, i.e. the degree of the polynomial is 1, but it has two solutions over \( \mathbb{Z}/10\mathbb{Z} \! \)!

### 11.4 Chinese Remainder Theorem

We are now ready to state the Chinese remainder theorem.

**Theorem 6.** Suppose \( m_1, \ldots, m_n \) are integers such that \( (m_i, m_j) = 1, \forall i \neq j \). Given \( b_1, \ldots, b_n \in \mathbb{Z} \), the system:

\[
\begin{align*}
x & \equiv b_1 \pmod{m_1} \\
x & \equiv b_2 \pmod{m_2} \\
& \quad \ldots \ldots \ldots \\
x & \equiv b_n \pmod{m_n}
\end{align*}
\]

has a unique solution over \( \mathbb{Z}/m\mathbb{Z} \), where \( m = m_1 \cdot m_2 \cdot \ldots \cdot m_n \).

Before prove of the theorem, let’s try to appreciate the intuition. Consider the special case where \( m_i = p_i \) are distinct prime numbers. Recall the space we constructed using prime ideals of \( \mathbb{Z} \), and how we interpreted integers as functions
over this space. In that context, this theorem merely says, if I know the values of a certain “function” at various given points, the function is determined over the subspace containing just these points, which naively seems tautological. (A function is determined by its values at all points in the domain.)

In the general case, one can think of the condition \((m_i, m_j) = 1\) as saying that we are gathering information of this function over various subsets that do not intersect, hence there should be no interference of local information.

This intuition will become more helpful, when we actually prove o the theorem. In the proof, we use an idea similar to the idea of characteristic functions in analysis: we get a number \(f_i\) which is \(\bar{i}\) in \(\mathbb{Z}/m_i\mathbb{Z}\), but is \(\bar{0}\) in \(\mathbb{Z}/m_j\mathbb{Z}\), for all other \(j\)’s. Then, if you want to find some \(f\) such that \(q_{m_i}(f) = q_{m_i}(b_i)\), for all \(i\), you can simply take \(f = \sum b_i f_i\). This also gives you an algorithm to solve the system of linear congruence equations.

First consider the case \(n = 2\).

**Lemma 15.** Suppose \((m_1, m_2) = 1\). One can always find \(a \in \mathbb{Z}\) such that \(a \equiv 1 \pmod{m_1}\) and \(a \equiv 0 \pmod{m_2}\).

**Proof.** \((m_1, m_2) = 1\), so \(\exists q_1, q_2\) such that \(q_1 m_1 + q_2 m_2 = 1\). Take \(a = 1 - q_1 m_1 = q_2 m_2\), it is easy to check \(a\) satisfies the conditions we want. \(\square\)
Chapter 12

Lecture 12

12.1 Warm-up

eg. Solve $7x \equiv 10 \pmod{11}$.

$7^{-1} = \bar{8}$ in $\mathbb{Z}/11\mathbb{Z}$, so multiply 8 on both sides we get $x \equiv 3 \pmod{11}$.

eg. Solve $12x \equiv 10 \pmod{14}$.

$(12, 14) = 2 | 10$, so we know there are two solutions over $\mathbb{Z}/14\mathbb{Z}$.

$x_0 = 2$ is a particular integer solution. So $\bar{2} = q_{14}(2 + 0 \cdot \frac{14}{2})$, $\bar{9} = q_{14}(2 + 1 \cdot \frac{14}{2})$ are the two solutions over $\mathbb{Z}/14\mathbb{Z}$.

The integer solution set is $S = \{2 + 14t | t \in \mathbb{Z}\} \cup \{9 + 14t | t \in \mathbb{Z}\}$.

Remark One should pay special attention to the difference between the cases where you mod out a prime number and the cases where you mod out a non-prime number. In the latter case, you should first check: (1) whether the equation has a solution; (2) if so, how many solutions there are over $\mathbb{Z}/m\mathbb{Z}$.

12.2 Chinese Remainder Theorem

The Chinese Remainder Theorem states:

**Theorem 7.** Suppose $m_1, ..., m_n$ are integers such that $(m_i, m_j) = 1, \forall i \neq j$. Given $b_1, ..., b_n \in \mathbb{Z}$, the system:

$$\begin{cases}
    x \equiv b_1 \pmod{m_1} \\
    x \equiv b_2 \pmod{m_2} \\
    \vdots \\
    x \equiv b_n \pmod{m_n}
\end{cases}$$

has a unique solution over $\mathbb{Z}/m\mathbb{Z}$, where $m = m_1 \cdot m_2 \cdot \ldots \cdot m_n$.

A little bit of a history. The following math puzzle appeared in an ancient Chinese math writing:
"Suppose there is bunch of objects whose quantity we don’t know. If you count them in threes, the remainder is 2; if you count them in fives, the remainder is 3; and if you count them in sevens, the remainder is 2. How many objects are there in all?"—Sun Zi Suanjing

This might be how the “Chinese Remainder Theorem” got its name. Formulated in modern language, the above problem is nothing but the following system of equations:

\[
\begin{align*}
x &\equiv 2 \pmod{3} \\
x &\equiv 3 \pmod{5} \\
x &\equiv 2 \pmod{7}
\end{align*}
\]

As we have discussed yesterday, it is somewhat saying a “function” over a disjoint union of some sets of points is determined by its value over every one of these sets separately.

First consider the case \(n = 2\). Intuitively, imagine a disjoint union of two sets, we show that one can always find a “function” that is 1 over one set, and zero over the other. If you have heard of characteristic function before, what we are doing here is similar. (Let \(A\) be a subset of some space \(X\). The characteristic function of \(A\) is defined to be the function that sends every point in \(A\) to 1, and every point not in \(A\) to 0.)

Recall a lemma from yesterday:

**Lemma 16.** Suppose \((m_1, m_2) = 1\). One can always find \(a \in \mathbb{Z}\) such that 
\[a \equiv 1 \pmod{m_1}\] and \(a \equiv 0 \pmod{m_2}\).

**Proof.** 
\((m_1, m_2) = 1\), so \(q_1, q_2\) such that \(q_1m_1 + q_2m_2 = 1\). Take \(a = 1 - q_1m_1 = q_2m_2\), it is easy to check \(a\) satisfies the conditions we want.

**Corollary 7.** The previous lemma implies the Chinese Remainder Theorem in the case \(n = 2\).

**Proof.** By the lemma, there exist \(f_1, f_2\), such that 
\[f_2 \equiv 1 \pmod{m_2}, f_1 \equiv 1 \pmod{m_1}, f_2 \equiv 0 \pmod{m_1}, f_1 \equiv 0 \pmod{m_2}\].

Take \(f = b_1f_1 + b_2f_2\). Since \(m_2|f_1\), \(f \equiv b_1f_1 \equiv b_1 \pmod{m_1}\).

Similarly, \(f \equiv b_2f_2 \equiv b_2 \pmod{m_2}\).

Then, \(q_m(f) \in \mathbb{Z}/m\mathbb{Z}\) is a solution to the system over \(\mathbb{Z}/m\mathbb{Z}\).

This solution over \(\mathbb{Z}/m\mathbb{Z}\) is unique: Suppose \(f, f'\) are both solutions to the original system. Then, \(f - f' = sm_1 = tm_2\), for some \(s, t\). Since \((m_1, m_2) = 1\), \(m_2|s, m_1|t\), i.e. \(m_1 \cdot m_2|f - f'\).

Eventually, we prove the general case:
Proof. (Proof of the Chinese Remainder Theorem.) Let \( m' = m_2 \cdot \ldots \cdot m_n \). Since \( m_i, m_j \) are coprime for all \( i \neq j \), \((m_1, m') = 1\). Then, we can apply the theorem for two factors, to get some \( f_1 \) such that

\[
 f_1 \equiv 1 \pmod{m_1}, \quad f_1 \equiv 0 \pmod{m'}.
\]

But this means \( f_1 \equiv 0 \pmod{m_i} \) for \( i = 2, 3, \ldots, n \).

By same argument, one can get \( f_i \) such that \( f_i \equiv 1 \pmod{m_i} \) and \( f_i \equiv 0 \pmod{m_j} \), for all \( j \neq i \).

Eventually, take \( f = b_1 f_1 + b_2 f_2 + \ldots + b_n f_n \), \( f \) is a solution to the original system of linear congruence equations.

To see uniqueness of solution over \( \mathbb{Z}/m\mathbb{Z} \), suppose \( f_1, f_2 \) are two solutions of the given system of equations. Then \( f_1 - f_2 \equiv 0 \pmod{m_i} \), for all \( i \). This implies

\[
 f_1 - f_2 = c_1 m_1 = c_2 m_2 = \ldots = c_n m_n, \quad \text{for some } c_1, \ldots, c_n \in \mathbb{Z}.
\]

Since \((m_i, m_j) = 1\) for all \( i \neq j \), this implies \( m_2 | c_1 \) and

\[
 f_1 - f_2 = c_{1,2} m_1 m_2 = c_3 m_3 = \ldots = c_n m_n, \quad \text{for some } c_{1,2} \in \mathbb{Z}.
\]

By induction, \( f_1 - f_2 = c (m_1 \cdot \ldots \cdot m_n) \), for some \( c \in \mathbb{Z} \), i.e. \( q_m(f_1) = q_m(f_2) \). \( \square \)

Let’s see some examples.

\textbf{eg.} Solve

\[
\begin{aligned}
 x &\equiv 2 \pmod{3} \\
 x &\equiv 3 \pmod{5} \\
 x &\equiv 2 \pmod{7}
\end{aligned}
\]

**Step 1:** Find an integer that is congruent to 0 modulo 5,7, and congruent to 1 modulo 3. This is the same as congruent to 0 modulo 35, and congruent to 1 modulo 3. Similarly, find another integer that is congruent to 0 modulo 21 and congruent to 1 modulo 5; and a third integer that is congruent to 0 modulo 15, and congruent to 1 modulo 7.

- Set 35n \( \equiv 1 \pmod{3} \), one gets 2n \( \equiv 1 \pmod{3} \), i.e. \( n \equiv 2 \). Take \( f_1 = 35 \cdot 2 = 70 \).
- Set 21m \( \equiv 1 \pmod{5} \), one gets m \( \equiv 1 \pmod{5} \). Take \( f_2 = 21 \cdot 1 = 21 \).
- Set 15s \( \equiv 1 \pmod{7} \), one gets s \( \equiv 1 \pmod{7} \). Take \( f_3 = 15 \cdot 1 = 15 \).

**Step 2:** \( x = 2f_1 + 3f_2 + 2f_3 = 2 \cdot 70 + 3 \cdot 21 + 2 \cdot 15 = 233 \) is one integer solution to the system.

**Step 3:** \( q_{105}(233) = \overline{23} \) is the unique solution over \( \mathbb{Z}/105\mathbb{Z} \). Here, 105 is the product of 3, 5, 7. One can also say the integer solution set is \( S = \{23 + 105t | t \in \mathbb{Z}\} \).

\textbf{eg.} Solve the following system:

\[
\begin{aligned}
 4x &\equiv 6 \pmod{10} \\
 5x &\equiv 4 \pmod{7}
\end{aligned}
\]

**Step 1:** First, solve \( 4x \equiv 6 \pmod{10} \). \( (4,10) = 2 | 6 \), so the equation has a
solution. \(3 \cdot 4 + (-1) \cdot 10 = 2\), so get a particular solution \(x_0 = 3 \cdot 3 = 9\). Since \((4, 10) = 2\), we are supposed to get two solutions over \(\mathbb{Z}/10\mathbb{Z}\). Namely, \(\overline{9} + q_{10}(0 - \frac{10}{2}) = \overline{9}, \overline{9} + q_{10}(1 - \frac{10}{2}) = \overline{4}\).

Second, solve \(5x \equiv 4 \pmod{7}\). Since 7 is prime, every non-zero element in \(\mathbb{Z}/7\mathbb{Z}\) has a multiplicative inverse and \(\overline{5}^{-1} = \overline{3}\). Hence, just multiply both sides by \(\overline{3}\), one gets \(x \equiv 5 \pmod{7}\).

**Step 2:** We end up having two systems.
First:
\[
\begin{aligned}
&x \equiv 9 \pmod{10} \\
&x \equiv 5 \pmod{7}
\end{aligned}
\]
Since \(5 \cdot 10 \equiv 1 \pmod{7}\), \(3 \cdot 7 \equiv 1 \pmod{10}\), we get \(x \equiv 5 \cdot 50 + 9 \cdot 21 \equiv 19 \pmod{70}\) is the unique solution over \(\mathbb{Z}/70\mathbb{Z}\) to this system.
Second,
\[
\begin{aligned}
&x \equiv 4 \pmod{10} \\
&x \equiv 5 \pmod{7}
\end{aligned}
\]
we get \(x \equiv 5 \cdot 50 + 4 \cdot 21 \equiv 54 \pmod{70}\) is the unique solution over \(\mathbb{Z}/70\mathbb{Z}\) to this system.

**Step 3:** We get two solutions over \(\mathbb{Z}/70\mathbb{Z}\): \(x \equiv 19, 54 \pmod{70}\). The integer solution set is
\[
S = \{19 + 70t | t \in \mathbb{Z}\} \cup \{54 + 70t | t \in \mathbb{Z}\}.
\]

**Remark** This does not violate the uniqueness part of the Chinese Remainder Theorem: the theorem only applies to system of linear congruence equations such that the coefficients in front of \(x\) are all 1. Indeed, we first reduce the more general system into one for which the Chinese Remainder Theorem does apply and then use the theorem.

**eg.** There is a different way to solve a system of linear congruence equations. We shall demonstrate it using the following example.

Solve the system of linear congruence equations
\[
\begin{aligned}
x &\equiv 1 \pmod{4} \\
x &\equiv 4 \pmod{7} \\
x &\equiv 18 \pmod{19}
\end{aligned}
\]

**Step 1:** Set \(4n + 1 \equiv 4 \pmod{7}\), get \(4n \equiv 3 \pmod{7}\), and hence \(n \equiv 6 \pmod{7}\). (We multiply 2 on both sides of the equation, since \(\overline{4} \cdot \overline{2} = \overline{1}\) in \(\mathbb{Z}/7\mathbb{Z}\).) Take \(N = 4 \cdot 6 + 1 = 25\).

**Step 2:** Set \(28w + 25 \equiv 18 \pmod{19}\), get \(9w \equiv 12 \pmod{19}\), and hence \(w \equiv 14 \pmod{19}\). (We multiply 2 on both sides of the equation and \(\overline{2} \cdot \overline{9} = -\overline{1}\) in \(\mathbb{Z}/19\mathbb{Z}\); so, \(-w \equiv 24 \pmod{19}\), i.e. \(w \equiv 14 \pmod{19}\).) Take \(x = 28 \cdot 14 + 25 = 417\).
417 is the unique solution over \( \mathbb{Z}/532\mathbb{Z} \).

**Remark** The idea behind this approach is to first find a number \( N \) satisfying the first two equations, and then add an appropriate multiple of \( m_1 \cdot m_2 \) to \( N \), so that the resulting integer satisfies all three equations.

### 12.3 Linear Algebra over \( \mathbb{Z}/m\mathbb{Z} \)

Another question we want to answer about linear congruence equations is the following: how do we solve a system of the form

\[
\begin{align*}
a_1x + a_2y &\equiv b_1 \pmod{m} \\
a_3x + a_4y &\equiv b_2 \pmod{m}
\end{align*}
\]

Such systems can be re-written in terms of matrices and vectors:

\[
AX \equiv \tilde{b} \pmod{m}.
\]

Similar to our discussion for univariate linear congruence equations, this is equivalent to a linear equation over \( \mathbb{Z}/m\mathbb{Z} \):

\[
\begin{bmatrix}
q_m(a_{1,1}) & \ldots & q_m(a_{1,n}) \\
q_m(a_{2,1}) & \ldots & q_m(a_{2,n}) \\
\vdots & \ddots & \vdots \\
q_m(a_{s,1}) & \ldots & q_m(a_{s,n})
\end{bmatrix}
\begin{bmatrix}
x_1 \\
x_2 \\
\vdots \\
x_n
\end{bmatrix}
= 
\begin{bmatrix}
q_m(b_1) \\
q_m(b_2) \\
\vdots \\
q_m(b_n)
\end{bmatrix}
\]

This is justified via the following terminology and proposition:

**Definition 22.** We say two \( s \times n \) matrices with integer entries, \( A = (a_{i,j}) \), \( B = (b_{i,j}) \) are congruent modulo \( m \), if \( a_{i,j} \equiv b_{i,j} \pmod{m} \) for all \( i, j \). We write \( A \equiv B \pmod{m} \).

**Proposition 37.** If \( A, B \) are \( s \times n \) matrices congruent modulo \( m \) and \( C, D \) are \( c \times s \) and \( n \times d \) matrices respectively. Then, \( CA \equiv CB \pmod{m} \) and \( AD \equiv BD \pmod{m} \).

**Proof.** This follows from the fact the map \( q_m \) (which sends integers to their remainders modulo \( m \)) preserves multiplication and addition. Yet matrix multiplication only involves multiplying and adding the entries. \( \square \)

So all we doing is linear algebra over \( \mathbb{Z}/m\mathbb{Z} \).

**Remark** The main focus here is to solve a linear system of congruence equations over \( \mathbb{Z}/m\mathbb{Z} \). It is quite different from the scenario of the Chinese Remainder Theorem, since in that case \( m_1, \ldots, m_n \) are all different, while in this context we fix a single \( m \).

The general theorem is the following:
Theorem 8. Let $A$ be an $n \times n$ invertible matrix. Denote $D = \det(A)$. If $(D,m) = 1$, then

$$A \tilde{X} \equiv \tilde{b} \pmod{m}$$

has a unique solution modulo $m$.

To proof this theorem, we recall a few notions/results from linear algebra:

Definition 23. Let $A$ be an $n \times n$ matrix. The adjoint of $A$, denoted as $\text{adj}(A)$, is defined to be the transpose of the co-factor matrix:

$$\begin{bmatrix}
(-1)^{1+1} \det(A_{1,1}) & \cdots & (-1)^{1+n} \det(A_{1,n}) \\
(-1)^{2+1} \det(A_{2,1}) & \cdots & (-1)^{2+n} \det(A_{2,n}) \\
\vdots & \ddots & \vdots \\
(-1)^{n+1} \det(A_{n,1}) & \cdots & (-1)^{n+n} \det(A_{n,n})
\end{bmatrix}$$

where $A_{i,j}$ is the matrix obtained from $A$ by removing $i$-th row and $j$-th column.

eg. $\text{adj}\left(\begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}\right) = \begin{pmatrix} 4 & -2 \\ -3 & 1 \end{pmatrix}$

Lemma 17. If $A$ is invertible, then $\frac{1}{\det(A)} \text{adj}(A)A = \text{Id}$.

(proof of Theorem 2). Notice that if $A$ is an $n \times n$ matrix with integer coefficients, so is $\text{adj}(A)$. Moreover, if $(\det(A), m) = 1$, then $q_m(\det(A))$ has a multiplicative inverse in $\mathbb{Z}/m\mathbb{Z}$. Therefore,

$$\tilde{x} = q_m(\det(A))^{-1} q_m(\text{adj}(A)) q_m(\tilde{b})$$

is the unique solution modulo $m$. Here, $q_m(A)$ denotes the matrix over $\mathbb{Z}/m\mathbb{Z}$ obtained from $A$ by applying $q_m$ to every entry of the matrix.

eg. Solve

$$\begin{bmatrix} 3 & 1 & 1 \\ 2 & 1 & 3 \\ 0 & 2 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} \equiv \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} \pmod{9}.$$

$\det(A) = -14 \equiv 4 \pmod{9}$. $q_9(\det(A))^{-1} = 7$.

Also, $\text{adj}(A) = \begin{bmatrix} (-1)^2 \cdot (-6) & (-1)^3 \cdot (-2) & (-1)^4 \cdot 2 \\ (-1)^3 \cdot 0 & (-1)^4 \cdot 0 & (-1)^5 \cdot 2 \\ (-1)^4 \cdot 4 & (-1)^5 \cdot (-3) & (-1)^6 \cdot 1 \end{bmatrix} \equiv \begin{bmatrix} 3 & 2 & 2 \\ 0 & 0 & 2 \\ 4 & 3 & 1 \end{bmatrix} \pmod{9}$.

Hence, $\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = 7 \cdot \begin{bmatrix} \frac{3}{2} \\ \frac{2}{2} \\ \frac{3}{2} \end{bmatrix} = \begin{bmatrix} \frac{7}{2} \\ \frac{7}{2} \\ \frac{7}{2} \end{bmatrix} = \begin{bmatrix} 4 \\ 5 \\ 7 \end{bmatrix}$ is the unique solution over $\mathbb{Z}/9\mathbb{Z}$. 
Chapter 13

Lecture 13

13.1 Application to Cryptography

As one important application of modular arithmetic, in this section we introduce the basic idea behind cryptography.

13.1.1 Caesar Cipher

A most basic idea behind ciphering is that one can identify Roman letters with numbers or remainders modulo certain integer. Take the 26 English letters as an example, one can identify them with the 26 elements in $\mathbb{Z}/26\mathbb{Z}$:

$$
A = 0, B = 1, C = 2, ..., Z = 25
$$

Of course, if we just use this identification, there is no secrecy at all. So, we want to apply some operation. The simplest case is to take a linear transformation:

$$
\mathbb{Z}/26\mathbb{Z} \rightarrow \mathbb{Z}/26\mathbb{Z} : \bar{i} \mapsto a \cdot \bar{i} + b.
$$

eg. Say we take the linear transformation $\bar{i} \mapsto 3 \cdot \bar{i} + 2$. Then, the 26 letters now correspond to the following elements:

$$
A \rightarrow \overline{2}, B \rightarrow \overline{5}, C \rightarrow \overline{8}, D \rightarrow \overline{11}, E \rightarrow \overline{14}, F \rightarrow \overline{17}, G \rightarrow \overline{20}, H \rightarrow \overline{23}, I \rightarrow \overline{0}, J \rightarrow \overline{3}, K \rightarrow \overline{6}, L \rightarrow \overline{9}, M \rightarrow \overline{12}, N \rightarrow \overline{15}, O \rightarrow \overline{18}, P \rightarrow \overline{21}, Q \rightarrow \overline{24}, R \rightarrow \overline{1}, S \rightarrow \overline{4}, T \rightarrow \overline{7}, U \rightarrow \overline{10}, V \rightarrow \overline{13}, W \rightarrow \overline{16}, X \rightarrow \overline{19}, Y \rightarrow \overline{22}, Z \rightarrow \overline{25}.
$$

This means, one writes “C” for “A”, “F” for “B”, etc.

What is the message: “HXOQOCHXOBAEBOCJJWPAIO”?

To decipher a secret message in this system, without knowing the actual linear transformation, one needs to guess/obtain at least partial information of the

\[1\] I think I wrote “The weather is really nice”.
13.1. APPLICATION TO CRYPTOGRAPHY

system. One way to do this is to look at the frequency a letter appears in the ciphered code. The most frequent letter may correspond to "E", since the latter is the most frequently used English letter according to some empirical data.

**eg.** Suppose we know/guess "O" stands for "E" and "H" stands for "T". We have obtained two equations:

\[
\begin{align*}
4a + b &= 14 \\
19a + b &= 7
\end{align*}
\]

Subtract the second equation for the first, we get

\[-15a = 7, \text{ i.e. } 11a = 7.\]

Easy to see \(7 \cdot 11 = 78 = 26 \cdot 3 - 1\). This is good enough, since we then get \(-a = 7 \cdot 7 = 23\).

**Remark** Since the system is based on certain linear transformation, deciphering boils down to solving linear congruence equations. However, this system is evidently not secure, since the possibilities are very limited.

### 13.1.2 Exponentiation Cipher

Now we introduce a slightly more complicated ciphering method.

Start with a large prime number \(p\), together with some positive integer \(e\) such that \((e, p - 1) = 1\). (Notice: \(p - 1 = \varphi(p)\).)

This time, we identify A,B,C,...,Z with two-digital numbers 00,01,02,...,25.

We first identify our text as a chain of letters:

**eg.** "What is geometry" becomes WHATISGEOMETRY.

Then, this chain is naturally identified with an even-digital number:

**eg.** WHATISGEOMETRY=2207001908180604...

We cut this number into "blocks". Each block is a \(2m\)-digital number, where \(m\) is the largest integer such that \(25^{\overbrace{25\ldots25}^{m\text{ copies of }25}} < p\). This means that the "largest" block possible, \(\underbrace{Z\ldots Z}_{m\text{ copies of }Z}\) should still correspond to an element in \(\mathbb{Z}/p\mathbb{Z}\).

Eventually, to cipher the text, we raise each block to its \(e\)-th power inside \(\mathbb{Z}/p\mathbb{Z}\).

**eg.** Take \(p = 2633\), \(e = 29\). Since \(2525 < p\), \(252525 > p\), \(m = 2\). So,

\[WHATISGEOMETRY = 2207001908180604131504191724\]

The ciphered text is 2207\(^{29}\) 0019\(^{29}\) 0818\(^{29}\) 0604\(^{29}\) 1315\(^{29}\) 0419\(^{29}\) 1724\(^{29}\).
Remark The exponentiation is taken inside $\mathbb{Z}/2633\mathbb{Z}$, and the result should be another 4-digital number smaller than 2633.

If the original text has odd many letters in all, and hence the number it corresponds to has $4n + 2$ many digits, one needs to append some dummy letter (say, “X”) in the end of your text to get a 4n-digital number, that one can cut into blocks of 4-digital numbers.

Eventually, we briefly mention the procedure of recovering the message, given $e$. It is another application of Fermat’s Little Theorem.

Since $(e, p - 1) = 1$, $q_{p-1}(e) \in U_{p-1}$. Hence, there is some integer $d$ smaller than $p - 1$ such that $c \cdot d = s \cdot (p - 1) + 1$, for some $s$. Therefore, raising every block to the $d$-th power will recover the original block:

eg. $(29)^{-1} = 2269$ in $\mathbb{Z}/2632\mathbb{Z}$. So, $((2207)^{29})^{2269} = ((2207)^{2632})^q \cdot (2207)^1 = 2207$.

13.1.3 RSA Cryptosystem

One generalization of the exponentiation ciphering leads to the so-called RSA cryptosystem.

1. For the original exponentiation ciphering, we use one large prime number $p$ together with some $e$ such that $(e, \varphi(p)) = 1$.

2. For the RSA cryptosystem, $p$ is replaced by a product $n$ of two large prime numbers $p_1, p_2$; we still need some $e$ such that $(e, \varphi(n)) = 1$. Note that $\varphi(n) = (p_1 - 1)(p_2 - 1)$.

The procedure of ciphering remains the same: one breaks his message into blocks of even many letters, turn it into even-digital numbers, and then raise each number to its $e$-th power modulo $n$. In order to decipher the message, one needs to find the multiplicative inverse $d$ of $q_{\varphi(n)}(e)$ in $\mathbb{Z}/\varphi(n)\mathbb{Z}$ and raise each block of numbers to the $d$-th power to recover the information.

eg. Say the text is “FERMATS LITTLE THEOREM”. Take $n = 43 \cdot 59 = 2537$, $e = 13$. Then, the text becomes

0504^{13} 1712^{13} 0019^{13} 1811^{13} 0819^{13} 1911^{13} 0419^{13} 0704^{13} 1417^{13} 0412^{13}$ (mod 2537)

To decipher, one needs to raise each number to the 937-th power, since 937 is the multiplicative inverse of 13 in $\mathbb{Z}/2436\mathbb{Z}$.

What is the purpose of making such a change?

Notice that in the exponentiation ciphering case, knowing $p$ is the same thing as knowing $\varphi(p)$, which is $p - 1$. So, there is no secret in this part. However, knowing $n$ is not the same thing as knowing $\varphi(n)$, because knowing the latter requires factorizing $n$ as $p_1 \cdot p_2$. When $p_1, p_2$ are really large, this is not an easy thing even using large computers. This illustrates the basic idea of Public Key Cryptography:
A cryptosystem may have many users, labeled from 1 to \(N\). For each user, the pair \((e_i, n_i)\) is pre-chosen and is revealed to all the users. If the \(i\)-th user sends a message to the \(j\)-th user, everyone can actually see the ciphered text. However, deciphering is essentially impossible and only the authorized user, the \(j\)-th user is given the information of \(p_1, p_2\) and is able to decipher the message.

**Remark** The main point is to make brute force deciphering computationally infeasible.

We assume that everyone sees the ciphered text. This is because, when the \(i\)-th user input his ciphered message into the system, he is not just inputting integers. Rather, he tells the system that these are remainders modulo \(n\). So, the information of \(n\) is available on the network. Moreover, once made public, \((e_i, n_i)\) can also be thought of as a digital signature of the \(i\)-th user. This is useful in some other way as well.

### 13.1.4 The Knapsack Cipher

This cryptosystem originates from the following question about integers:

Given positive integers \(a_1, ..., a_n\) and \(S\), do some or all of \(a_i\)’s sum up to \(S\)?

This is the same as solving \(a_1x_1 + ... + a_nx_n = S\), for \(x_i = 0\) or \(1\).

It turns out that one can solve this question in an easy way, if \(\{a_1, ..., a_n\}\) has special properties.

**Definition 24.** A sequence \(a_1, ..., a_n\) is called super increasing if

\[
\sum_{i=1}^{j-1} a_i < a_j, \text{ for } j = 2, ..., n.
\]

The algorithm to solving the Knapsack Problem for a super increasing sequence: Suppose \(\{a_1, ..., a_n\}\) is super increasing. Then, take

\[
x_j = \begin{cases} 
1 & \text{if } S - \sum_{i=j+1}^{n} x_i a_i \geq a_j \\
0 & \text{otherwise}
\end{cases}
\]

Suppose \(S - \sum a_j x_j = 0\), the problem has a solution; otherwise, it doesn’t have a solution.

**eg.** Suppose the sequence is 2, 3, 6, 13, 25, 51 and \(S = 97\). Using the given algorithm, one gets

\[
x_6 = 1, x_5 = 1, x_4 = 1, x_3 = 1, x_2 = 0, x_1 = 1.
\]

We now show how one may turn this into a cryptosystem.
1. Each user of the cryptosystem chooses a super increasing sequence of the same length \(N\), together with some \(m > 2a_N\) and \(w\) such that \((m, w) = 1\).

2. Generate another sequence \(b_1, \ldots, b_N\): \(b_i \equiv wa_i (\text{mod } m)\).

3. When ciphering a message, turn the string of letters into a sequence of zeros and ones, using the binary equivalents of roman letters.

4. Cut the sequence into blocks \(x_1x_2\ldots x_N\) of length \(N\) and compute the quantity \(S = b_1x_1 + \ldots + b_nx_n\) for each block.

5. The ciphered text is just the \(S\)'s generated in each block.

To decipher the message, one needs to find the multiplicative inverse of \(q_m(w)\) in \(\mathbb{Z}/m\mathbb{Z}\) to recover the original super increasing sequence and solve for the string of zeros and ones from there.

eg. Let the text be “Fermat’s last theorem”. The binary equivalent is

\[
\begin{align*}
0010100100 & \quad 1000101100 & \quad 0000010011 & \quad 1001001011 & \quad 0000010010 \\
1001100100 & \quad 0110011001 & \quad 1001100100 & \quad 0110011001 & \quad 0010001100
\end{align*}
\]

Take \((a_1, \ldots, a_{10}) = (2, 11, 14, 29, 58, 119, 241, 480, 959, 1917)\), \(m = 3837\), \(w = 1001\).

Then, \((b_1, \ldots, b_{10}) = (2002, 3337, 2503, 2170, 503, 172, 3347, 855, 709, 417)\).

So, \(S_1 = 3839\), \(S_2 = 6707\), \(S_3 = 1298\), \(S_4 = 8645\), \(S_5 = 881\), \(S_6 = 5973\), \(S_7 = 6031\), \(S_8 = 8599\), \(S_9 = 6705\) is the ciphered text you sent out.

The multiplicative inverse of 1001 in \(\mathbb{Z}/3837\mathbb{Z}\) is 23. You use it to recover \(a_i\) from \(b_i\). Once you get \(\{a_i\}\), since it is super increasing, you can solve for the string of one’s and zeros and thus recover the original message.
13.1.5 Summary

The secretness for each of the cryptosystem mentioned above:

1. Caesar ciphering: lies in the knowledge of the linear coefficients $a, b$;
2. Exponential ciphering: lies in the knowledge of $e$ and $p$;
3. RSA: lies in the knowledge of the factorization $n = p_1p_2$;
4. Knapsack ciphering: lies in the difference between $\{b_i\}$ and $\{a_i\}$.
Chapter 14

Lecture 14

14.1 Arithmetic Progression: a Summary

So far, we have proved various cases of the statement that there are infinitely many prime numbers of the form $aq + r$, for specific $a, r$ such that $(a, r) = 1$, using two techniques:

1. If in $\mathbb{Z}/a\mathbb{Z}$, $\bar{i}$ is a unit such that the product it cannot be a product of the other units, we approach in the following way:

   (a) Find some $p_1$ of the correct form;
   (b) let $Q_n = a(p_1p_2...p_{n-1}) + r$ and argue that $Q_n$ has a prime factor of the desired form not equal to any of $p_i$.

   Typical cases are "$6q + 4, 4q + 3, 3q + 2$".

2. If $\mathbb{Z}/p\mathbb{Z}$ contains a unit of order $a$, $a|p-1$ and $p = a \cdot c + 1$.

   (a) Find some $p_1$;
   (b) let $Q_n = (p_1..p_{n-1})^{a-1} + (p_1...p_{n-1})^{a-2} + ... + (p_1...p_{n-1}) + 1$ and argue that $Q_n$ has a prime factor of the desired form not equal to any of $p_i$.

   Typical cases are "$qn + 1$", where $q$ is some fixed prime number.

Remark Sometimes, one needs to skip the smallest prime of the desired form to facilitate the induction argument; or, one needs to stuck in a factorial term in the construction of $Q_n$. We see this through examples:

eg. Show that there are infinitely many prime numbers of the form $4q + 3$.

(1) Start with $p_1 = 7$.
(2) In general, suppose $p_1,...,p_{n-1} > 3$ and are of the desired form. Let $Q_n = 4(p_1p_2...p_{n-1}) + 3$. One can see that $2, 3, p_i$ do not divide $Q_n$:
14.2. Euler’s Theorem, Fermat’s Little Theorem, Order of Units etc.

1. 2 does not divide $Q_n$, since 2 does not divide 3.
2. 3 does not divide $Q_n$, since 3 does not divide $4(p_1p_2...p_{n-1})$.
3. $p_i$ does not divide $Q_n$, since it does not divide 3.

$Q_n$ is an integer greater than one and must have a prime factor. Its prime factors cannot be all of the form $4q + 1$, since $\mathbb{Z}^N \neq \mathbb{F}$ in $\mathbb{Z}/4\mathbb{Z}$. Therefore, $Q_n$ must have a prime factor $> 3$ of the form $4n + 3$, not equal to any of $p_1, ..., p_{n-1}$: call it $p_n$.

By induction, we can conclude there are infinitely many prime numbers of the form $4q + 3$.

**Remark** If we start with $p_1 = 3$, we cannot argue that $p_n > 3$ in the inductive step.

**eg.** Show that there are infinitely many prime numbers of the form $qn + 1$, where $q$ is some fixed prime number.

**Lemma 18.** Any prime factor $p$ of $a^{n-1} + a^{n-2} + ... + a + 1 (a > 1)$ greater than $q$ is of the form $nq + 1$.

**Proof.** $q_p(a^{n-1}) = q_p(a-1)q_p(a^{n-1} + a^{n-2} + ... + a + 1) = \bar{0}$, so $ord_p(a_p) = 1,q$.

Since $p > q$, the order must be $q$; otherwise $q_p(a^{n-1} + a^{n-2} + ... + a + 1) = \bar{n} \neq \bar{0}$ and we get contradiction.

Now, take $Q_1 = (q!)^{q-1} + ... + (q!) + 1$. By the construction, none of $1, ..., q$ divides $Q_1$. So the prime factors of $Q_1$ are greater than $q$. By the lemma, they are of the form $qn + 1$. Take one of them to be $p_1$.

Define $Q_n = (q!p_1...p_{n-1})^{q-1} + ... + q!p_1...p_{n-1} + 1$. Then, none of $1, ..., q$ divides $Q_n$; nor does any $p_i$. By the lemma, $Q_n$ has a prime factor of the desired form, not equal to any of $p_i$. Call it $p_n$.

By induction, we are done.

**eg.** There are infinitely many primes of the form $(2 \cdot q)n + 1$, where $q$ is prime.

**Proof.** We have already shown that there are infinitely many primes of the form $q \cdot n + 1$, where $q$ is prime. Moreover, any prime $> 2$ is odd.

Suppose $q > 2$. If $p = qn + 1$ is an odd prime, $qn = p - 1$ is even, i.e. $2|qn$. Since $(2, q) = 1$, $2|n$. So, $p = (2q)n + 1$.

Suppose $q = 2$, then $2q = 4$. We have already shown in a previous lecture that there are infinitely many primes of the form $4n + 1$.

Hence, we are done.

**14.2 Euler’s Theorem, Fermat’s Little Theorem, Order of Units etc.**

**eg.** Find a positive integer $m$ smaller than 13 such that $25^{10001} \equiv m (\mod 13)$. $25 \equiv -1 (\mod 13)$, so $25^{10001} \equiv 12 (\mod 13)$. 

eg. Find a positive integer \( m \) smaller than 13 such that \( 7^{10001} \equiv m \pmod{13} \).

By Fermat’s little theorem, \( 7^{12} \equiv 1 \pmod{13} \). Since \( 10001 = 12 \cdot 833 + 5 \), get \( 7^{10001} \equiv 7^5 \pmod{13} \).

So, \( 7^{10001} \equiv 11 \pmod{13} \).

eg. Solve \( x^5 + y^5 \equiv z^5 \pmod{5} \).

Notice that Fermat’s Little Theorem implies \( x^5 \equiv x \pmod{5} \) for any \( x \). Hence, the equation reduces to \( x + y \equiv z \pmod{5} \), which is equivalent to understanding the table of addition of \( \mathbb{Z}/5\mathbb{Z} \).

As for the computation of Euler-\( \varphi \), we recall the following facts:

1. Suppose \( n = \prod p_i^{m_i} \).

\[
\varphi(n) = n \cdot \prod_{p \mid n} \left(1 - \frac{1}{p}\right) = \prod_{p \mid n} p_i^{m_i} \left(1 - \frac{1}{p_i}\right) = \prod_{p \mid n} p_i^{m_i-1} (p_i - 1).
\]

2. \( \sum_{d \mid n} \varphi(d) = n \).

3. \( \varphi(mn) = \varphi(m)\varphi(n) \) only when \( (m,n) = 1 \).

### 14.3 Computations

eg. Solve the following system of linear congruence equations:

\[
\begin{cases}
6x \equiv 4 \pmod{8} \\
3x \equiv 4 \pmod{7} \\
2x \equiv 3 \pmod{13}
\end{cases}
\]

From the first equation, one gets \( (6,8) = 2 \mid 4 \), so there are two solutions over \( \mathbb{Z}/8\mathbb{Z} \). \( 2, 2 + 4 = 6 \) are the solutions.

From the second equation, one gets \( x \equiv 6 \pmod{7} \). One sees this by multiplying 5 on both sides, since \( 5 \equiv 5^{-1} \).

From the third equation, one gets \( x \equiv 8 \pmod{13} \). One can see this by multiplying 6 on both sides: since \( 6 = -2^{-1} \), one gets \(-x \equiv 5 \pmod{13} \); in other words, \( x \equiv 8 \pmod{13} \).
14.4. GREATEST COMMON DIVISOR

So, one needs to solve two systems:

\[
\begin{align*}
&x \equiv 2 \pmod{8} & x \equiv 6 \pmod{8} \\
&x \equiv 6 \pmod{7} & x \equiv 6 \pmod{7} \\
&x \equiv 8 \pmod{13}, & x \equiv 8 \pmod{13}.
\end{align*}
\]

I’ll leave the rest to you.

**eg.** Solve the following system of linear congruence equations:

\[
\begin{align*}
3x - y + 7z & \equiv 2 \\
10x + 9y + z & \equiv 1 \pmod{7} \\
2x - 5y + 3z & \equiv 0.
\end{align*}
\]

Let \( A = \begin{bmatrix} 3 & -1 & 7 \\ 10 & 9 & 1 \\ 2 & -5 & 3 \end{bmatrix} \equiv \begin{bmatrix} 3 & -1 & 0 \\ 3 & 2 & 1 \\ 2 & 2 & 3 \end{bmatrix} \pmod{7}. \)

Then, \( \det(A) \equiv 3 \cdot (6 - 2) + (9 - 2) \equiv 5 \pmod{7}. \) So, \( q_7(\det(A)^{-1}) = 3. \)

\[
\text{adj}(A) \equiv \begin{bmatrix} (-1)^{1+1} \cdot 4 & (-1)^{1+2} \cdot (-3) & (-1)^{1+3} \cdot (-1) \\ (-1)^{2+1} \cdot 7 & (-1)^{2+2} \cdot 9 & (-1)^{2+3} \cdot 3 \\ (-1)^{3+1} \cdot 2 & (-1)^{3+2} \cdot 8 & (-1)^{3+3} \cdot 2 \end{bmatrix} \equiv \begin{bmatrix} 4 & 3 & -1 \\ 0 & 2 & 4 \\ 2 & -1 & -2 \end{bmatrix} \pmod{7}.
\]

So, \( \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 4 & 3 & -1 \\ 0 & 2 & 4 \\ 2 & -1 & -2 \end{bmatrix} \begin{bmatrix} 0 \\ 2 \\ 6 \end{bmatrix} = \begin{bmatrix} 5 \\ 6 \end{bmatrix}. \)

14.4 Greatest Common Divisor

**eg.** Prove: \((a_1, ..., a_i, (a_{i+1}, ..., a_n)) = (a_1, ..., a_n).\)

**Proof.**

\[
((a_1, ..., a_i, (a_{i+1}, ..., a_n)) = b_1 \cdot (a_1, ..., a_i) + b_2 \cdot (a_{i+1}, ..., a_n)
= b_1 \cdot (c_1a_1 + ... + c_ia_i) + b_2 \cdot (c_{i+1}a_{i+1} + ... + c_na_n)
= (b_1c_1)a_1 + ... + (b_1c_i)a_i + (b_2c_{i+1})a_{i+1} + ... + (b_2c_n)a_n
\]

So, it can be written as a linear combination of \(a_1, ..., a_n.\) Hence, it must be a positive multiple of \((a_1, ..., a_n),\) since \(\{m_1a_1 + ... + m_na_n | m_i \in \mathbb{Z} \} = ((a_1, ..., a_n)).\)

Thus,

\[
((a_1, ..., a_i, (a_{i+1}, ..., a_n)) \geq (a_1, ..., a_n).
\]

On the other hand,

\[
(a_1, ..., a_n) = d_1a_1 + ... + d_na_n
= (d_1a_1 + ... + d_ia_i) + (d_{i+1}a_{i+1} + ... + d_na_n)
= e_1 \cdot (a_1, ..., a_i) + e_2 \cdot (a_{i+1}, ..., a_n)
\]
So, it can be written as a linear combination of \((a_1,...,a_i)\) and \((a_{i+1},...,a_n)\). Hence, it must be a positive multiple of \(\((a_1,...,a_i),(a_{i+1},...,a_n)\)\). Thus,

\[
((a_1,...,a_i),(a_{i+1},...,a_n)) \leq (a_1,...,a_n).
\]

Thus, \(\((a_1,...,a_i),(a_{i+1},...,a_n)\) = (a_1,...,a_n)\) \(\square\)

Remark \((a_1,...,a_n)\) is the smallest positive integer in

\[
\{m_1a_1 + ... + m_na_n|m_i \in \mathbb{Z}\} = ((a_1,...,a_n)).
\]

14.5 Divisibility Rule for \(m = 29\)

eg. Show that \(29|a_1a_2...a_n\) if and only if the number \(29|a_1...a_{n-1} + 3 \cdot a_n\).

Proof. \(a_1a_2...a_n = 10 \cdot (a_1a_2...a_{n-1}) + a_n = 10 \cdot (a_1...a_{n-1} + 3 \cdot a_n) + (-29) \cdot a_n\).

So, \(q_{29}(a_1a_2...a_n) = q_{29}(10) \cdot q_{29}(a_1...a_{n-1} + 3 \cdot a_n)\).

Since 29 is a prime number, two non-zero elements in \(\mathbb{Z}/29\mathbb{Z}\) do not multiply to \(\overline{0}\). Since \(q_{29}(10)\) is non-zero in \(\mathbb{Z}/29\mathbb{Z}\), \(q_{29}(a_1a_2...a_n) = \overline{0}\) if and only if \(q_{29}(a_1...a_{n-1} + 3 \cdot a_n) = \overline{0}\) \(\square\)
Chapter 15

Lecture 15

If there is one thing in mathematics that fascinates me more than anything else (and doubtless always has), it is neither “number” nor “size”, but always form. And among the thousand-and-one faces whereby form chooses to reveal itself to us, the one that fascinates me more than any other and continues to fascinate me, is the structure hidden in mathematical things.

Alexander Grothendieck, Récoltes et Semailles

15.1 A Quick Summary

Our course can be divided into three parts:

- Basic theory of integers
- Basic theory of congruence, focusing on linear congruence
- High-degree congruence, with a focus on quadratic congruence equations

We are now in the last third of the course. We will first talk about univariate high-degree congruence equations in general. This is essentially the same as theory of polynomials over $\mathbb{Z}/p\mathbb{Z}$. We will then explore in a more concrete way this idea of going from “global” equations to “local” equations. Eventually, we will end our course with the law of quadratic reciprocity.
15.2 High Degree Congruence Equations

Previously, when discussing arithmetic progression, we have already implicitly utilized some high-degree congruence equations.

**eg.** When \( p > 2 \) is a prime number and \( p|a^2 + 1, p = 4n + 1 \) for some \( n \). Another way to put it, \( q_p(a) \) is a solution to the quadratic equation \( x^2 + 1 = 0 \) over \( \mathbb{Z}/p\mathbb{Z} \). So, we re-state the result as follows: if \( x^2 + 1 = 0 \) has a solution over \( \mathbb{Z}/p\mathbb{Z} \) and \( p > 2 \), then \( p \) is of the form \( 4n + 1 \).

**eg.** When \( p > 3 \) is a prime number and \( p|a^2 + a + 1, p = 3n + 1 \) for some \( n \). Another way to put it, \( q_p(a) \) is a solution to the quadratic equation \( x^2 + x + 1 = 0 \) over \( \mathbb{Z}/p\mathbb{Z} \). So, we re-state the result as follows: if \( x^2 + x + 1 = 0 \) has a solution over \( \mathbb{Z}/p\mathbb{Z} \) and \( p > 3 \), then \( p \) is of the form \( 3n + 1 \).

Some natural questions to ask:

1. Q1: when do these polynomial equations have solution(s)?
2. Q2: how is the existence of solution related to \( p \)?

These will form the central theme of the rest of our course.

A first point to make: these polynomial should no longer be regarded as integer coefficient polynomials. Rather, we should think of them as \( \mathbb{Z}/p\mathbb{Z} \)-coefficient polynomials. Our first task is then to develop some general terminologies to talk about them:

**Definition 25.** Denote \((\mathbb{Z}/b\mathbb{Z})[x]\) to be the set of all polynomials with \( \mathbb{Z}/b\mathbb{Z} \)-coefficients, i.e. its elements are of the form \( f(x) = a_n x^n + a_{n-1} x^{n-1} + \ldots + a_1 x + a_0 \), where all \( a_i \) are in \( \mathbb{Z}/b\mathbb{Z} \).

We call \( f(c) = a_n c^n + a_{n-1} c^{n-1} + \ldots + a_1 c + a_0 \) to be the value of \( f \) at \( x = c \).

If \( f(c) = 0 \), we call \( c \) a root of \( f \).

**Remark** \((\mathbb{Z}/b\mathbb{Z})[x]\) is also a ring: you can add and multiply these polynomials just like you add and multiply usual polynomials; the constants \( 1, 0 \) are the multiplicative unit and zero element in this ring. The only difference is that when you add and multiply coefficients, you need to use modular arithmetic instead of usual integer arithmetic. However, this already makes a big difference.

**eg.** Take \( b = 6 \). \( x^2 - 5x = (x - 5) \) is the factorization is not unique in \( \mathbb{Z}/6\mathbb{Z} \). Hence, \( 0, 2, 3, 5 \) are all roots of \( x^2 - 5x \). In this example, the degree of the polynomial is two but it has more than two roots.

We start by building some simple results about roots of polynomial over \( \mathbb{Z}/b\mathbb{Z} \). The next few results show that much of your intuition about polynomials still works in this new setting.

**Lemma 19.** Let \( f(x) = a_n x^n + a_{n-1} x^{n-1} + \ldots + a_1 x + a_0 \) be a polynomial with coefficients in \( \mathbb{Z}/b\mathbb{Z} \). It has a root \( a \) if and only if \( f \) can be written as \( f(x) = (x - a) \cdot g(x) \).
Proof. If \( f(x) = (x-a) \cdot q(x) \), \( f(a) = \overline{0} \circ q(a) = \overline{0} \).

Conversely, re-write \( f(x) \) as:
\[
f(x) = a_n[(x-a) + a]^n + a_{n-1}[(x-a) + a]^{n-1} + ... + a_1[(x-a) + a] + a_0
\]

Now, expand the terms out, we see that \( f(x) = a_n(x-a)^n + \tilde{a}_{n-1}(x-a)^{n-1} + ... + \tilde{a}_1(x-a) + b \). If you plug in \( x = a \), all but last term becomes \( \overline{0} \). This implies \( f(\overline{a}) = b = \overline{0} \), i.e. \( b = 0 \).

Therefore, \( f(x) = a_n(x-a)^n + \tilde{a}_{n-1}(x-a)^{n-1} + ... + \tilde{a}_1(x-a) \). Now, pull out a common factor \( x-a \) from each term on the right hand side and we are done. \( \square \)

This immediately implies the following:

**Proposition 38.** (Lagrange) Suppose \( p \) is prime, \( f(x) \) is a polynomial of degree \( n \) over \( \mathbb{Z}/p\mathbb{Z} \). Then \( f(x) \) has at most \( n \) roots (not necessarily distinct) over \( \mathbb{Z}/p\mathbb{Z} \).

(Sketch of the proof). From the previous lemma, suppose \( f \) has \( n \) roots, \( f \) can be written as a product of \( n \) linear factors, and it has exhausted the degree of \( n \).

One might think the proof is already done. However, this is only so if one can show that \( f \) over \( \mathbb{Z}/p\mathbb{Z} \) factorizes uniquely. This turns out to be true when \( p \) is prime. (See \[15.3\]) \( \square \)

### 15.2.1 \( x^p - x \) over \( \mathbb{Z}/p\mathbb{Z} \).

A most important polynomial over \( \mathbb{Z}/p\mathbb{Z} \): Consider
\[
f(x) = x^p - x
\]
over \( \mathbb{Z}/p\mathbb{Z} \). Fermat’s little theorem says \( \forall \overline{i} \in \mathbb{Z}/p\mathbb{Z} \), \( f(\overline{i}) = \overline{0} \). But \( f(x) \) is not the zero polynomial! This also tells us that polynomials over \( \mathbb{Z}/p\mathbb{Z} \) are not determined by the values over all numbers in \( \mathbb{Z}/p\mathbb{Z} \).

**Theorem 9.** Let \( f(x) \) be a polynomial over \( \mathbb{Z}/p\mathbb{Z} \) of degree \( n \). It has \( n \) distinct roots if and only if it divides \( x^p - x \).

**Proof.** As mentioned above, Fermat’s little theorem tells us that \( \overline{0}, ..., \overline{p-1} \) are the \( p \) distinct roots of \( x^p - x \) over \( \mathbb{Z}/p\mathbb{Z} \). By Lemma \[19\]
\[
x^p - x = x(x-\overline{1}) \cdot ... \cdot (x-\overline{p-1}).
\]

If \( f(x) \) is a factor of \( x^p - x \), it is of the form \( b(x-a_1)...(x-a_n) \), where \( a_1, ..., a_n \) are distinct elements in \( \mathbb{Z}/p\mathbb{Z} \). So, it has \( n \) distinct roots.

Conversely, if \( f(x) \) has \( n \) distinct roots, by previous lemma, it can be written as \( b(x-a_1)...(x-a_n) \), where \( a_1, ..., a_n \in \mathbb{Z}/p\mathbb{Z} \) are all distinct. It is then clear that \( f/x^p - x \):
\[
x^p - x = f(x) \cdot (b^{-1} \prod_{c \neq a_i} (x-c)).
\]
\( \square \)
eg. Consider $\mathbb{Z}/7\mathbb{Z}$.

\[
x^7 - x = x(x^6 - 1) = x(x - 1)(x^5 + x^4 + x^3 + x^2 + x + 1)
= x(x - 1)(x^4(x + 1) + x^2(x + 1) + (x + 1))
= x(x - 1)(x - 6)(x^4 + x^2 + 1)
= x(x - 1)(x - 6)(x^4 - 9x^2 + 81)
= x(x - 1)(x - 6)(x - 2)(x^2 - 3)
= x(x - 1)(x - 6)(x - 2)(x - 3)(x + 3)
= x(x - 1)(x - 6)(x - 2)(x - 3)(x - 4).
\]

Corollary 8. (Wilson) If $p$ is prime, $(p - 1)! \equiv -1 \pmod{p}$.

Proof. We already know $x^{p-1} - 1 = (x - 1) \cdots (x - p + 1)$. Plug in $x = 0$, we get

\[-1 = (-1)^{p-1} q_p((p - 1)!).
\]

When $p - 1$ is even, this gives $(p - 1)! \equiv (-1)^{p-1} \pmod{p}$ directly;
when $p = 2$, $-1 \equiv 1 \pmod{2}$, so the result still holds. \hfill \Box

eg. Take $p = 7$. $(p - 1)! = 6! = 720 = 102 \cdot 7 + 6 \equiv -1 \pmod{7}$.

15.2.2 Greatest Common Divisor of Two polynomials

In this section, we demonstrate the idea that $x^p - x$ can be thought of as the “root detector” for polynomials over $\mathbb{Z}/p\mathbb{Z}$.

Definition 26. Let $f, g$ be two polynomials in $(\mathbb{Z}/p\mathbb{Z})[x]$. Define the greatest common divisor of $f, g$ to be a polynomial $d(x)$ satisfying the following properties:

1. $d(x)$ is a factor of $f$, and a factor of $g$.

2. $d(x)$ has the highest degree among all common divisors of $f$ and $g$.

3. The highest degree coefficient of $d$ is 1.

We denote the g.c.d. of $f, g$ as $(f, g)$. If $(f, g) = 1$, then we say $f, g$ are coprime.

Remark The last condition helps to kill ambiguity in the definition: notice that if $f(x)|g(x)$, then $\overline{7} \cdot f(x)|g(x)$, where $\overline{7}$ is a unit.

Proposition 39. Let $g(x)$ be any polynomial over $\mathbb{Z}/p\mathbb{Z}$. The number of distinct roots of $g(x)$ over $\mathbb{Z}/p\mathbb{Z}$ is equal to the degree of $(g(x), x^p - x)$.
15.3. FACTORIZATION OF POLYNOMIALS

Proof. Any factor of \(x^p - x\) is of the form \(b(x-a_1)\ldots(x-a_m)\), where all \(a_j \in \mathbb{Z}/p\mathbb{Z}\) are distinct. If \(b(x-a_1)\ldots(x-a_m)\) is the g.c.d. of \(g(x)\) and \(x^p - x\), \(a_1, \ldots, a_m\) are the roots of \(g(x)\). Moreover, any other element \(c\) in \(\mathbb{Z}/p\mathbb{Z}\) cannot be a root of \(g(x)\): otherwise, \(b(x-c)(x-a_1)\ldots(x-a_m)\) divides \(g(x)\) and \(x^p - x\), which violates the assumption that \(b(x-a_1)\ldots(x-a_m)\) is the g.c.d. of \(g(x)\) and \(x^p - x\). Thus, we have found all the distinct roots of \(g(x)\).

\textbf{eg.} Let \(g(x) = x^5 + 2x^3 - 4x + 1\). Compute the g.c.d of \(g\) and \(x^7 - x\) in \((\mathbb{Z}/7\mathbb{Z})[x]\).

\[
x^7 - x = (x^5 + 2x^3 - 4x + 1)(x^2) - 2x^5 + 4x^3 - x^2 - x
\]

\[
= (x^5 + 2x^3 - 4x + 1)(x^2 - 2) + x^3 - x^2 - 2x + 2
\]

\[
x^5 + 2x^3 - 4x + 1 = (x^3 - x^2 - 2x + 2)(x^2) + x^4 + 4x^3 - 2x^2 - 4x + 1
\]

\[
= (x^3 - x^2 - 2x + 2)(x^2 + x) + 5x^3 - 6x + 1
\]

\[
= (x^3 - x^2 - 2x + 2)(x^2 + x + 5) + 5x^2 + 4x - 2
\]

\[
x^3 - x^2 - 2x + 2 = (5x^2 + 4x - 2)(3x) + x^2 + 4x + 2
\]

\[
= (5x^2 + 4x - 2)(3x + 3) + 6x + 1
\]

\[
5x^2 + 4x - 2 = (6x + 1)(2x) + 2x - 2
\]

\[
= (6x + 1)(2x + 5)
\]

Now, take \(6x + 1\) and turn it into a polynomial with highest coefficient 1:

\[
\overline{6} \cdot (6x + 1) = x + \overline{6}.
\]

Hence, \((x^5 + 2x^3 - 4x + 1, x^7 - x) = x + 6 = x - 1\). From here, we see that \(\overline{1}\) is the only solution to the congruence equation \(x^5 + 2x^3 - 4x + 1 \equiv 0(\text{mod } 7)\).

15.3 Factorization of Polynomials

In chapter 2, we have seen that there is an Euclidean algorithm for \(\mathbb{Z}\), which follows from division with remainder. Here we show that there is also an Euclidean algorithm for polynomials.

\textbf{Proposition 40.} Let \(k\) be any field, say \(\mathbb{R}, \mathbb{C}, \mathbb{Z}/p\mathbb{Z}\), and \(f(x), g(x)\) are two \(k\)-coefficient polynomials. Then, we can find a pair of unique polynomials \(q(x), r(x)\) such that \(f(x) = q(x) \cdot g(x) + r(x)\), where \(r(x)\) is either 0, or \(\deg(r) < \deg(g)\).

\textbf{Proof.} (Existence) When \(\deg(g) > \deg(f)\), one can simply take \(q = 0, r = f\), and we are done. So, can assume \(\deg(f) \geq \deg(g)\).

In this case, suppose \(f(x) = a_nx^n + \ldots + a_1x + a_0, g(x) = b_m x^m + \ldots + b_1x + b_0\), take \(q_1(x) = a_nb_m^{-1}x^{n-m}\). We can compute directly that \(\deg(f - q_1 \cdot g) \leq n - 1\). If it is smaller than \(\deg(g)\), take \(r(x) = f(x) - q_1(x) \cdot g(x)\), and we are done. Otherwise, call \(f_1(x) = f(x) - q_1(x) \cdot g(x)\), and we can find some \(q_2(x)\) such that
\( \text{deg}(f_1 - q_2 \cdot g) \leq \text{deg}(f_1) - 1 \). Since \( \text{deg}(g) \) is a fixed number, and at each step the degree drops by at least one, we conclude that after finitely many steps, we get \( \text{deg}(f - q_1 \cdot g - \ldots - q_k \cdot g) < \text{deg}(g) \). Take \( q = q_1 + q_2 + \ldots + q_k \), and \( r = f - q_1 \cdot g - \ldots - q_k \cdot g \), then we are done.

(Uniqueness) Suppose \( q(x)g(x) + r(x) = t(x)g(x) + s(x) \), then \( (q(x) - t(x))g(x) = s(x) - r(x) \). Since the degree of \( s(x) - r(x) \) is supposed to be smaller than \( \text{deg}(g) \), we know that both sides must be zero, and hence uniqueness follows. \( \square \)

**Corollary 9.** Let \( k \) be a field. Let \( k[x] \) be the ring of \( k \)-coefficient polynomials. Then, every ideal of \( k[x] \) is of the form \( (f(x)) \).

**Proof.** Let \( I \) be an ideal, and take \( f \) to be some element of \( I \) with the smallest degree. Then \( I = (f(x)) \), and the rest of the proof uses division with remainder: if there is an element \( g \in I \) which is not of the form \( f(x)q(x) \), then by taking division with remainder, one can write \( g(x) = f(x)q(x) + r(x) \), where \( \text{deg}(r) < \text{deg}(f) \). By the definition of ideal, since \( f \in I \), \( f \cdot q \in I \), and hence \( r = q - f \cdot q \in I \).

So \( r(x) \) is a polynomial in \( I \) with degree even smaller than the degree of \( f \). This violates our choice of \( f \). \( \square \)

In this case, we can talk about the g.c.d of two polynomials \( f, g \), that is, some linear combination \( a(x)f(x) + b(x)g(x) \) with smallest degree possible, and we always assume the highest degree coefficient to be 1. We can say \( f, g \) are coprime, if \( (f,g) = 1 \).

**eg.** Let \( f(x) = x^2 + 4 \), \( g(x) = x + 1 \) be polynomials over \( \mathbb{Z}/5\mathbb{Z} \), then \( g|f \), because \( (x + 1)(x + 4) = x^2 + 4 \).

We are ready to state the unique factorization result for polynomials with coefficients in a field.

**Definition 27.** A polynomial is irreducible if it cannot be written as a product of two non-constant polynomials.

**Corollary 10.** Let \( k \) be any field, and \( f(x) \) be any \( k \)-coefficient polynomial. Then \( f \) can be written uniquely as a product of irreducible polynomials.

**Proof.** By the Euclidean algorithm, just as for integers, two irreducible polynomials in \( k[x] \) are either coprime, or one divides another. Then the same proof for the unique factorization of \( \mathbb{Z} \) applies here as well. \( \square \)

**Remark** Why does this result fail in the case \( \mathbb{Z}/b\mathbb{Z} \) when \( b \) is not a prime number? In \( (\mathbb{Z}/6\mathbb{Z})[x] \), \( x - 2 \) and \( x - 5 \) are both irreducible, since any product of two non-constant polynomials will have degree at least two. However, \( x - 2 \) and \( x - 5 \) are not coprime. In fact if a constant in \( \mathbb{Z}/6\mathbb{Z} \) is a linear combination of the two, it is always a multiple of 3, and 3 is not a unit in \( \mathbb{Z}/6\mathbb{Z} \). Thus, as a consequence \( x^2 - 5x = (x - 2)(x - 3) = x(x - 5) \) does not factorize uniquely in \( (\mathbb{Z}/6\mathbb{Z})[x] \).
eg. What polynomials are irreducible depends on the field we assume the coefficients are in. Take $f(x) = x^2 + 1$ as a real-coefficient polynomial, it is irreducible, but as complex-coefficient polynomial, $x^2 + 1 = (x + i)(x - i)$.

eg. Moreover, if you consider $f(x) = x^2 + 1$ over $\mathbb{Z}/2\mathbb{Z}$, $f(x) = (x + 1)^2$. 
Chapter 16

Lecture 16

16.1 Analogy between $\mathbb{Z}$ and $(\mathbb{Z}/p\mathbb{Z})[x]$

As we have mentioned before, solving univariate congruence equation modulo some prime number $p$ is equivalent to understanding the ring of polynomials with coefficients in $\mathbb{Z}/p\mathbb{Z}$. It turns out that $\mathbb{Z}$ and $(\mathbb{Z}/p\mathbb{Z})[x]$ have similar properties, which shall list here:

1. Division with remainder:
   
   - In $\mathbb{Z}$: $\forall a, b$, there exists unique $q, r$ such that $a = q \cdot b + r$, where $0 \leq r < |b|$.
   
   - In $(\mathbb{Z}/p\mathbb{Z})[x]$: $\forall a(x), b(x)$, there exists unique $q(x), r(x)$ such that $a(x) = q(x) \cdot b(x) + r(x)$, where $\text{deg}(r(x)) < \text{deg}(b(x))$.

   **Remark** Even the proof for uniqueness are very similar: in $\mathbb{Z}$, if the pair $(q, r)$ were not unique, the condition $r < |b|$ would be violated; in $(\mathbb{Z}/p\mathbb{Z})[x]$, if the pair $(q(x), r(x))$ were not unique, the condition $\text{deg}(r(x)) < \text{deg}(b(x))$ would be violated.

   For the detail of the proof, see [15.3]

2. Euclidean algorithm:
   
   - In $\mathbb{Z}$: apply division with remainder multiple times to find g.c.d.
   
   - In $(\mathbb{Z}/p\mathbb{Z})[x]$: do the same thing.

   **Remark** For the definition of g.c.d of two polynomials, see Lecture 15.

3. Ideals:
   
   - In $\mathbb{Z}$: every ideal is of the form $(n)$. This follows from division with remainder.
• In \((\mathbb{Z}/p\mathbb{Z})[x]\): every ideal is of the form \((f(x))\), i.e. \(\{q(x)f(x)q(x) \in (\mathbb{Z}/p\mathbb{Z})[x]\}\). This also follows from division with remainder.

We point out that existence of Euclidean Algorithm in both situations is the key to the above resemblance. Such rings are called Euclidean domain.

Remark For more detail, see 15.3

4. Prime ideals:
   • In \(\mathbb{Z}\): \((0),(p)\) are the only prime ideals, where \(p\) is prime.
   • In \((\mathbb{Z}/p\mathbb{Z})[x]\): the prime ideals are \((f(x))\) where \(f(x)\) is an irreducible polynomial.

Definition 28. A polynomial is irreducible if it cannot be written as a product of two non-constant polynomials.

Remark One cares about irreducible polynomials since they are more or less the “smallest” building blocks in \((\mathbb{Z}/p\mathbb{Z})[x]\). Just like you cannot write a prime \(p\) as a product other than \(p = p \cdot 1\), you cannot write an irreducible \(f(x)\) as a product other than \(f(x) = c \cdot (1/c)f(x)\), where \(c \in \mathbb{Z}/p\mathbb{Z}\) is a non-zero constant.

5. Unique factorization:
   • In \(\mathbb{Z}\): every integer greater than 1 can be written uniquely as a product of primes.
   • In \((\mathbb{Z}/p\mathbb{Z})[x]\): every polynomial can be written uniquely as a product of irreducible polynomials.

Remark See also 15.3

16.2 Irreducible Polynomials over \(\mathbb{Z}/p\mathbb{Z}\)

The next two results classify irreducible polynomials over \(\mathbb{Z}/p\mathbb{Z}\).

Proposition 41. Let \(p\) be a prime number. The irreducible polynomials of degree \(n\) over \(\mathbb{Z}/p\mathbb{Z}\) are precisely the degree \(n\) irreducible factors of the polynomial \(x^p - x\) over \(\mathbb{Z}/p\mathbb{Z}\).

Proposition 42. The irreducible factors of \(x^p - x\) are the irreducible polynomials in \((\mathbb{Z}/p\mathbb{Z})[x]\) whose degree divides \(n\).

The proofs involve knowledge of finite extension of finite fields, and hence are not required. You do need to know these results themselves.

eg. Over \(\mathbb{Z}/2\mathbb{Z}\), \(x^4 - x = x(x^3 - 1) = x(x - 1)(x^2 + x + 1)\). On one hand, we know \(x, x - 1\) are the only degree 1 irreducible polynomials over \(\mathbb{Z}/2\mathbb{Z}\), and
one can check directly that neither of them divides $x^2 + x + 1$. This tells us the only irreducible polynomial in $(\mathbb{Z}/2\mathbb{Z})[x]$ of degree 2 is $x^2 + x + 1$.

**eg.** Over $\mathbb{Z}/2\mathbb{Z}$, $x^8 - x = x(x-1)(x^7 + x^6 + x^5 + x^4 + x^3 + x^2 + x + 1)$, which shows that the only degree 3 irreducible polynomials over $\mathbb{Z}/2\mathbb{Z}$ are $x^3 + x^2 + 1, x^3 + x + 1$. This tells us the only irreducible polynomial in $(\mathbb{Z}/2\mathbb{Z})[x]$ of degree 2 is $x^2 + x + 1$.

**Remark** To verify a degree 2 or a degree 3 polynomial is irreducible, one can simply check it does not have roots. This is because, if such a polynomial factorizes, it must contain at least one degree 1 factor.

In summary, we point out that these two propositions tell us everything about irreducible polynomials over $\mathbb{Z}/p\mathbb{Z}$. Suppose we know all the irreducible polynomials up to degree $n$, to find the degree $n+1$ ones, one factorizes $x^{p^n+1} - x$ and look at the degree $n+1$ factors. (If a factor has degree larger than $n+1$, by Proposition 2, you know it must factorize into a product of polynomials with smaller degrees.) In doing so, you can simply test whether any of the irreducibles of degree $\leq n$ divides the factor you are looking at. Inductively, we know all irreducibles.

We can also say that we completely understand univariate congruence equations modulo a prime number $p$. Most importantly, to solve $f(x) \equiv 0 (mod \ p)$, one can either plug in $0, \ldots, p-1$ and see whether any of them is a root; or, one can compute the g.c.d. of $f(x)$ and $x^p - x$.

### 16.3 Application 1

As an important application of univariate congruence equation, we show that if $d > 0$ and $d | p-1$, then $\mathbb{Z}/p\mathbb{Z}$ always has a unit of order $d$. Notice this is not true in general!

**eg.** In $\mathbb{Z}/8\mathbb{Z}$, the units are $1, 3, 5, 7$. They have orders 1, 2, 2, 2 resp. But $\phi(8) = 4$, and there is no unit of order 4!

To prove the claimed result, we need a lemma:

**Lemma 20.** Let $p$ be a prime. If $d | p-1$, then $x^d - 1 \equiv 0 (mod \ p)$ has $d$ solutions.

**Proof.** Suppose $p-1 = q \cdot d$. $x(x^d - 1)(x^{(q-1)d} + x^{(q-2)d} + \ldots + 1) = x^p - x$. Hence, $(x^d - 1, x^p - x) = x^d - 1$. Hence, the polynomial has $d$ distinct roots over $\mathbb{Z}/p\mathbb{Z}$. And therefore the corresponding congruence equation has $d$ solutions.

Now we state and prove the following proposition:

**Proposition 43.** Let $p$ be a prime number. Suppose $d | p-1$, then $\mathbb{Z}/p\mathbb{Z}$ has exactly $\phi(d)$ many units of order $d$. 

Proof. Step 1 We have shown before that for any unit \( a \) of \( \mathbb{Z}/p\mathbb{Z} \), \( \text{ord}_p(a) \mid p - 1 \).

Step 2 Now, suppose \( a \) is a unit such that \( \text{ord}_p(a) = d \mid p - 1 \). We claim: for any \( i \) such that \( 1 \leq i \leq d \) and \((i, d) = 1\), \( a^i \) is also of order \( d \). This is because suppose \( \text{ord}_p(a^i) = m \), i.e. \( a^{im} = 1 \). By division with remainder, \( im = qd + r \), where \( 0 \leq r < d \). But \( r \) cannot be positive, since \( a^r = a^{im-qd} = 1 \) would otherwise violate the condition that \( d \) is the smallest positive integer \( N \) such that \( a^N = 1 \).

Hence, \( im = qd \), i.e. \( d \mid im \).

Since \( a^{id} = (a^d)^i = 1 \), for \( m \) to be the order of \( a^i \), one must have \( d = m \).

Thus we’ve concluded that if \( \mathbb{Z}/p\mathbb{Z} \) has a unit of order \( d \) it is has at least \( \varphi(d) \) many units of order \( d \).

Step 3 Next, we show that there can’t be more than \( \varphi(d) \) many units of order \( d \).

Step 3.1 Any unit of order \( d \) must be a solution to the equation \( x^d - 1 = 0 \) over \( \mathbb{Z}/p\mathbb{Z} \). Let \( a \) be a unit or order \( d \). Then, \( a, a^2, a^3, \ldots, a^d \) are \( d \) distinct elements which all satisfy the equation \( x^d - 1 = 0 \). By the previous lemma, they are all the solutions. Hence, to look for unit of order \( d \), it suffices to look at \( a, a^2, a^3, \ldots, a^d \).

Step 3.2 If \( i > 1 \) and \( id \), it is easy to see that \((a^i)^{d/i} = 1\), hence \( \text{ord}_p(a^i) \neq d \).

Consequently, if \( c = (i, d) > 1 \), then \((a^i)^{d/c} = 1\) and its order cannot be \( d \). Therefore, only the \( i \)'s such that \((i, d) = 1\) can be a unit of order \( d \). There are \( \varphi(d) \) many such \( i \)'s.

Step 3.3 Combine Step 3.2 with Step 2, if there is one unit of order \( d \) in \( \mathbb{Z}/p\mathbb{Z} \), then there are exactly \( \varphi(d) \) many units of order \( d \).

Step 4 Since the order of a unit is always one of the \( d \)'s such that \( d \mid p - 1 \), the number of units in \( \mathbb{Z}/p\mathbb{Z} \) must be less or equal than

\[
\sum_{d \mid p-1} \varphi(d) \leq \sum_{d \mid p} \varphi(d).
\]

The inequality follows from Step 1.

But we also know that \( \mathbb{Z}/p\mathbb{Z} \) has \( p - 1 \) units (only \( \bar{0} \) is not a unit). And we have a proposition saying that \( \sum_{d \mid p-1} \varphi(d) = p - 1 \). This shows that every of the factors of \( p - 1 \) must be realized as the order of some unit of \( \mathbb{Z}/p\mathbb{Z} \). (Similar to the Pigeon-hole principle.)

This concludes the proof.

\( \square \)

eg. In your HW, you will conclude that a prime number \( p \) is of the form \( 4n + 1 \) if and only if \( \mathbb{Z}/p\mathbb{Z} \) has a unit of order \( 4 \).

eg. You have shown the fact before that \( \overline{p-1} \) is of order \( 2 \) in \( \mathbb{Z}/p\mathbb{Z} \) \((p > 2)\). In fact, that’s the only order \( 2 \) unit, because \( \varphi(2) = 1 \).
98

CHAPTER 16. LECTURE 16

eg. Also, now you can say a prime $p$ is of the form $5n + 1$ if and only if $\mathbb{Z}/p\mathbb{Z}$ has a unit of order 5. In fact, there are $\varphi(5) = 4$ many of them.

16.4 Application 2: Solving all Univariate Congruence Equations

In the recent lectures, our main focus was to solve univariate congruence equations modulo a prime number $p$. From there, I claim we can solve any univariate congruence equation. The general procedure is as follows:

**Step 1**: Given $f(x) \equiv 0 \pmod{n}$, use Fundamental Theorem of Arithmetic, $n = p_1^{m_1}...p_r^{m_r}$, where $p_1,...,p_r$ are distinct prime numbers.

**Step 2**: $f(x) \equiv 0 \pmod{n}$ is equivalent to the system:

$$
\begin{align*}
  f(x) &\equiv 0 \pmod{p_1^{m_1}} \\
  f(x) &\equiv 0 \pmod{p_2^{m_2}} \\
  & \quad \vdots \\
  f(x) &\equiv 0 \pmod{p_r^{m_r}}
\end{align*}
$$

so it is enough to solve congruence equations modulo powers of prime numbers.

**Step 3** There is a procedure which we shall see tomorrow that reduces to $f(x) \equiv 0 \pmod{p_i^{m_i}}$ to congruence equations modulo $p_i$. In this procedure, we only need to solve linear congruence equations.

**Step 4** Once you have solutions over $p_i^{m_i}$ for every $i$, you get various systems (possibly more than one) of the form:

$$
\begin{align*}
  x &\equiv a_{1,i_1} \pmod{p_1^{m_1}} \\
  x &\equiv a_{2,i_2} \pmod{p_2^{m_2}} \\
  & \quad \vdots \\
  x &\equiv a_{r,i_r} \pmod{p_r^{m_r}}
\end{align*}
$$

Then, Chinese Remainder Theorem will give you the solutions to the original equation over $\mathbb{Z}/n\mathbb{Z}$.

**Remark** Notice that in this procedure, we use Fundamental Theorem of Arithmetic, method of solving congruence equations modulo a prime, knowledge of linear congruence equations and Chinese Remainder Theorem. It puts together many of the topics we have covered so far!
Chapter 17

Lecture 17

17.1 Warm-up

e.g. Classify all units of $\mathbb{Z}/13\mathbb{Z}$ by their orders. $\varphi(13) = 12$ and the positive factors of 12 are 1, 2, 3, 4, 6, 12.

<table>
<thead>
<tr>
<th>Orders</th>
<th>Number of Units with this Order</th>
<th>List of them</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>$\varphi(1) = 1$</td>
<td>$\bar{1}$</td>
</tr>
<tr>
<td>2</td>
<td>$\varphi(2) = 1$</td>
<td>$-\bar{1} = \bar{12}$</td>
</tr>
<tr>
<td>3</td>
<td>$\varphi(3) = 2$</td>
<td>$\bar{3}, \bar{9}$</td>
</tr>
<tr>
<td>4</td>
<td>$\varphi(4) = 2$</td>
<td>$\bar{5}, \bar{8}$</td>
</tr>
<tr>
<td>6</td>
<td>$\varphi(6) = 2$</td>
<td>$\bar{10}, \bar{4}$</td>
</tr>
<tr>
<td>12</td>
<td>$\varphi(12) = 4$</td>
<td>$\bar{2}, \bar{11}, \bar{6}, \bar{7}$</td>
</tr>
</tbody>
</table>

You have to consider $\bar{2}, \bar{3}, ..., \bar{6}$ and $\bar{7} = -\bar{6}, ..., \bar{12} = -\bar{1}$.

$2^6 = 64 = 5 \cdot 13 - 1$, so $ord_{13}(\bar{2}) \neq 1, 2, 3, 6$. And $2^4 = 13 + 3$, so $ord_{13}(\bar{2}) \neq 4$; conclude $ord_{13}(\bar{2}) = 12$. Also conclude that $\bar{1}^{10} = -\bar{1}, \bar{1}^{11}, \bar{1}^{12} \neq \bar{1}$ and $\bar{1}^{14} = \bar{2}^4 \neq \bar{1}$; so, $ord_{13}(\bar{1}) = 12$.

$3^3 = 27 = 2 \cdot 13 + 1$, so $ord_{13}(\bar{3}) = 3$ and you get $(-\bar{3})^3 = -\bar{1}$, so that $\bar{10}^6 = \bar{1}$. Since $\bar{10}^2 \neq \bar{1}$, conclude that $ord_{13}(\bar{10}) = 6$.

$4^6 = \bar{2}^{12} = \bar{1}$. Its order can only be 6, since otherwise the order of 2 would not be 12. Also get $\bar{5}^6 = \bar{1}$. But notice that $\bar{5}^3 = \bar{1}$, its order is 3 not 6.

$5^2 = -\bar{1}$, so $ord_{13}(\bar{5}) = 4$. Also conclude $\bar{8}^2 = -\bar{1}$ and hence $ord_{13}(\bar{8}) = 4$.

Since at this stage we have already filled in the first five rows, conclude that $ord_{13}(\bar{6}) = ord_{13}(\bar{7}) = 12$. 

99
17.2 Reduction to Modulo $p$

Following yesterday’s lecture, we now describe the general procedure of solving the equation $f(x) \equiv 0 \pmod{p^n}$ starting from solving $f(x) \equiv 0 \pmod{p}$.

In Calculus, we have seen the usage of Taylor polynomials in approximating a certain function. In particular, the Taylor polynomial of degree $n$ at $x = x_0$ is given as follows:

$$f(x) \sim f(x_0) + \frac{f'(x_0)}{1!}(x-x_0) + \frac{f''(x_0)}{2!}(x-x_0)^2 + \ldots + \frac{f^{(n)}(x_0)}{n!}(x-x_0)^n$$

**Remark** We use the symbol $\sim$ since in general the Taylor polynomial is not equal to the function you started with.

In the case where $f(x)$ is a polynomial, we have an even better result. The degree $n$ Taylor polynomial of a degree $n$ polynomial $f(x)$ is equal to $f$ itself. Using this fact, we have the following lemma:

**Lemma 21.** Let $f(x)$ be an integer-coefficient, degree $n$ polynomial. For any integer $x_0$, the following identity holds:

$$f(x_0 + tp^N) = f(x_0) + \frac{f'(x_0)}{1!}(tp^N) + \ldots + \frac{f^{(n)}(x_0)}{n!}(tp^N)^n, \quad (*)$$

**Proof.** By the definition of Taylor polynomial, and the fact that $f$ is a degree $n$ polynomial to start with, one can check directly that

$$f(x_0 + tp^N) = f(x_0) + \frac{f'(x_0)}{1!}(tp^N) + \ldots + \frac{f^{(n)}(x_0)}{n!}(tp^N)^n,$$

by plugging $x = x_0 + tp^N$ into the general formula. \hfill $\square$

This is the essential lemma with which one reduces $f(x) \equiv 0 \pmod{p^n}$ to $f(x) \equiv 0 \pmod{p}$. We take an inductive procedure to achieve this.

**(Base Case)** Using knowledge from previous lectures, solve $f(x) \equiv 0 \pmod{p}$.

**(Induction Step)** Suppose you have solved the equation

$$f(x) \equiv 0 \pmod{p^n},$$

and you want to solve the equation $f(x) \equiv 0 \pmod{p^{n+1}}$.

Clearly, if $x$ is a solution over $\mathbb{Z}/p^{n+1}\mathbb{Z}$, it is also a solution over $\mathbb{Z}/p^n\mathbb{Z}$.

Moreover, if $x_0$ is a solution over $\mathbb{Z}/p^n\mathbb{Z}$, any number of the form $x_0 + tp^N$ is also a solution over $\mathbb{Z}/p^{n+1}\mathbb{Z}$. (But $x_0$ and $x_0 + tp^N$ may be different over $\mathbb{Z}/p^{n+1}\mathbb{Z}$!) Therefore, one starts from solutions over $\mathbb{Z}/p^n\mathbb{Z}$, and tries to “enhance” it to a solution over $\mathbb{Z}/p^{n+1}\mathbb{Z}$. We summarize the main steps here:

Step 1: Pick a $x_0$ which is a solution of $f(x) \equiv 0 \pmod{p^n}$.
Step 2: Try to solve for some \( t \) such that \( f(x_0 + tp^N) \equiv 0 \pmod{p^{N+1}} \).

By the lemma above, \( f(x_0 + tp^N) = f(x_0) + f'(x_0)(tp^N) + \ldots + \frac{f^{(n)}(x_0)}{n!}(tp^N)^n \).

It is congruent to 0 modulo \( p^{N+1} \) if and only if \( f(x_0) + f'(x_0)(tp^N) \) is congruent to 0 modulo \( p^{N+1} \), since all the other terms on the right-hand side are congruent to 0 modulo \( p^{N+1} \). Notice also that \( p^N | f(x_0) \), since \( x_0 \) is a solution to the equation over \( \mathbb{Z}/p^N \mathbb{Z} \).

This is further equivalent to solving \( f'(x_0)t = \frac{\ell(x_0)}{p^n} \pmod{p} \). Notice that since we’ve fixed an \( x_0, f'(x_0) \) and \( \frac{\ell(x_0)}{p^n} \) are fixed numbers. So we are solving a concrete linear equation over \( \mathbb{Z}/p \mathbb{Z} \).

Step 3: We now apply our knowledge of linear congruence equations to discuss all possible situations:

1. Case 1: \( (f'(x_0), p) = 1 \). In this case, there is a unique \( t \) such that the linear equation holds.

2. Case 2: \( p | f'(x_0) \) and \( p | \frac{\ell(x_0)}{p^n} \). Then, the equation (1) becomes \( 0 \equiv 0 \).

   Any \( t \) is a solution. Hence, \( x_0 + tp^N \) is a solution over \( \mathbb{Z}/p^{N+1} \mathbb{Z} \) for all \( t = 0, \ldots, p-1 \).

3. Case 3: \( p | f'(x_0) \) but \( p \) does not divide \( \frac{\ell(x_0)}{p^n} \). Then, \( f(x) \equiv 0 \pmod{p^{N+1}} \) does not have a solution of the form \( x_0 + tp^N \).

Step 4: Try step 1-3 for every solution of \( f(x) \equiv 0 \pmod{p^N} \) and we’re done with the induction step.

We now look at some examples:

**eg.** Suppose \( f(x) = x^4 + 2x + 13 \). Solve \( f(x) \equiv 0 \pmod{25} \).

First solve \( f(x) \equiv 0 \pmod{5} \). The equation reduces to \( x^4 + 2x + 3 = 0 \) over \( \mathbb{Z}/5 \mathbb{Z} \). Since 0 is not a solution, one can conclude that the solution must be a unit. Hence, \( x^4 = 1 \). One can directly solve \( 2x + 4 = 0 \) over \( \mathbb{Z}/5 \mathbb{Z} \), which gives \( x = 3 \).

Next, compute \( f'(x) = 4x^3 + 2 \), and \( f'(3) = 110, f(3) = 100 \). Hence, \( \frac{f(3)}{3} = 20 \). We need to solve \( f'(3)t \equiv \frac{f(3)}{3} \pmod{5} \), which is the same as \( 0 \cdot t = 0 \) over \( \mathbb{Z}/5 \mathbb{Z} \). This means that \( t \) can take any of 0, 1, 2, 3, 4. Therefore, \( 3 + 0 \cdot 5 = 3, 3 + 1 \cdot 5 = 8, 3 + 2 \cdot 5 = 13, 3 + 3 \cdot 5 = 18, 3 + 4 \cdot 5 = 23 \) are the solutions of \( f(x) = 0 \) over \( \mathbb{Z}/25 \mathbb{Z} \).

**eg.** Suppose \( f(x) = x^3 + 2x + 2 \). Solve \( f(x) \equiv 0 \pmod{7^3} \).
First solve $f(x) \equiv 0 \pmod{7}$. Since there are only 7 elements in $\mathbb{Z}/7\mathbb{Z}$, try them one by one, and we get 2, 3 are the solutions over $\mathbb{Z}/7\mathbb{Z}$.

Alternatively, let’s compute $(x^7 - x, x^3 + 2x + 2)$.

\[
x^7 - x = (x^3 + 2x + 2)(x^4 - 2x^2 - 2x) + 4x^3 - x^2 + 3x = (x^3 + 2x + 2)(x^4 - 2x^2 - 2x + 4) + x^2 - 5x - 1
\]
\[
x^3 + 2x + 2 = (x^2 - 5x - 1)(x) + 5x^2 + 3x + 2 = (x^2 - 5x - 1)(x + 5)
\]

Hence, $(x^7 - x, x^3 + 2x + 2) = x^2 - 5x - 1 = (x - 2)(x - 3)$ and $2, 3$ are solutions over $\mathbb{Z}/7\mathbb{Z}$.

Second, $f'(x) = 3x^2 + 2$. So, $f'(2) = 3 \cdot 2^2 + 2 = 14$, $f(2) = 14$. Solve

\[
14t \equiv -2 \pmod{7},
\]

one gets there are no such $t$’s.

**Remark** Notice: if you take remainder before you divide $f(2)$ by 7, you get 0, and arrive at a wrong conclusion!

On the other hand, $f''(3) = 29, f(3) = 35$. Solve $29t \equiv -5 \pmod{7}$, one gets $t \equiv 2 \pmod{7}$. Conclude that $3 + 2 \cdot 7 = 17$ is a unique solution over $\mathbb{Z}/7^2\mathbb{Z}$.

Eventually, $f'(17) = 3 \cdot 17^2 + 2 = 869, f(17) = 4949$. Now solve

\[
\frac{f(17)}{7^2} + f'(17) \cdot t \equiv 0 \pmod{7}.
\]

**Remark** Reminder: at the $N$-th step, should divide $f(x_0)$ by $p^{N-1}$ (not just $p$), where $x_0$ is some solution to $f(x) \equiv 0 \pmod{p^{N-1}}$. This is important!

Solve $q_7(869)t = -q_7(101)$ over $\mathbb{Z}/7\mathbb{Z}$, one gets $t = 4$. Further conclude that $17 + 4 \cdot 49 = 213$ is the unique solution to the equation $f(x) \equiv 0 \pmod{7^3}$.

**eg.** Suppose $f(x) = x^3 + 8$. Solve $f(x) \equiv 0 \pmod{27}$.

We can get 1 is the only solution to $f(x) \equiv 0 \pmod{3}$.

$f'(x) = 3x^2$, and $f'(1) = 3, f(1) = 9$. We need to solve $f'(1)t \equiv -\frac{f(1)}{3} \pmod{3}$.

This leads to the equation $0 \cdot t \equiv 0 \pmod{3}$, so that $t = 0, 1, 2$. Thus, 1, 4, 7 are the solutions of $f(x) \equiv 0 \pmod{9}$.

$f'(1) = 3, f(1) = 9$; $f'(4) = 48, f(4) = 72$; $f'(7) = 147, f(7) = 351$.

$f'(1)t \equiv -\frac{f(1)}{9} \pmod{3}$ implies $0 \cdot t \equiv 1 \pmod{3}$, which has no solution.

$f'(4)t \equiv -\frac{f(4)}{9} \pmod{3}$ implies $0 \cdot t \equiv 2 \pmod{3}$, and has no solution as well.
17.2. REDUCTION TO MODULO $P$

$f'(7)t \equiv -\frac{f(7)}{9} \pmod{3}$ implies $0 \cdot t \equiv 0 \pmod{3}$, and $t = 0, 1, 2$.

Therefore, 7, 16, 25 are the solutions of the original congruence equation.
Chapter 18

Lecture 18

18.1 Warm-up

eg. Solve $x^2 + 1 \equiv 0 \pmod{100}$.
First of all, take the prime factorization of 100: $100 = 2^2 \cdot 5^2$.
This is equivalent to solving the system
\[
\begin{cases}
  x^2 + 1 \equiv 0 \pmod{2^2} \\
  x^2 + 1 \equiv 0 \pmod{5^2}.
\end{cases}
\]

The first equation does not have a solution since $\mathbb{Z}/4\mathbb{Z}$ does not have a unit of order 4. ($x^2 + 1 = 0$ implies $x^4 - 1 = 0$ and the root cannot have order 1 or 2.) Therefore, the original equation does not have a solution.

eg. Solve $x^2 - 1 \equiv 0 \pmod{100}$.
First of all, take the prime factorization of 100: $100 = 2^2 \cdot 5^2$.
This is equivalent to solving the system
\[
\begin{cases}
  x^2 - 1 \equiv 0 \pmod{2^2} \\
  x^2 - 1 \equiv 0 \pmod{5^2}.
\end{cases}
\]

Let’s start by solving $x^2 - 1 = 0$ over $\mathbb{Z}/4\mathbb{Z}$.
First, over $\mathbb{Z}/2\mathbb{Z}$, one gets $x_0 = \overline{1}$ is the only root.
Next, compute $f'(x) = 2x$ and $f(1) = 0, f'(1) = 2$.
Set $2 \cdot t = -\frac{1}{2}$ over $\mathbb{Z}/2\mathbb{Z}$, and one gets $t = \overline{0}, \overline{1}$. Therefore, roots over $\mathbb{Z}/4\mathbb{Z}$ are $\overline{1} + 0 \cdot \overline{2} = \overline{1}$ and $\overline{1} + 1 \cdot \overline{2} = \overline{3}$.

Then, solve $x^2 - 1 = 0$ over $\mathbb{Z}/25\mathbb{Z}$.
First, over $\mathbb{Z}/5\mathbb{Z}$, one gets $x_0 = \overline{1}, \overline{4}$.
$f(1) = 0, f'(1) = 2$. Set $2 \cdot t = -\frac{1}{2}$ over $\mathbb{Z}/5\mathbb{Z}$, one gets $t = \overline{0}$ and $\overline{1} + 0 \cdot \overline{5} = \overline{1}$ is a root over $\mathbb{Z}/25\mathbb{Z}$.
18.2. “GLOBAL VS LOCAL” REVISITED

\( f(4) = 15, f'(4) = 8. \) Set \( 8 \cdot t = -\frac{15}{5} \) over \( \mathbb{Z}/5\mathbb{Z} \), one gets \( t = 4 \) and \( 4 + 4 \cdot 5 = 24 \). is another root over \( \mathbb{Z}/25\mathbb{Z} \).

All in all, we get four systems:

\[
\begin{align*}
\begin{cases}
x \equiv 1 \pmod{4} \\
x \equiv 1 \pmod{25}
\end{cases},
\begin{cases}
x \equiv 1 \pmod{4} \\
x \equiv 24 \pmod{25}
\end{cases},
\begin{cases}
x \equiv 3 \pmod{4} \\
x \equiv 1 \pmod{25}
\end{cases},
\begin{cases}
x \equiv 3 \pmod{4} \\
x \equiv 24 \pmod{25}
\end{cases}.
\end{align*}
\]

By applying Chinese Remainder Theorem four times, we get \( x = 1, 49, 51, 99 \) are the only roots over \( \mathbb{Z}/100\mathbb{Z} \).

18.2 “Global vs Local” Revisited

Suppose \( f(x, y, z) \) is an integer coefficient polynomial, and we are looking for integer solutions to the equation \( f(x, y, z) = 0 \). It is clear that

\[ f(x, y, z) = 0 \Rightarrow f(x) \equiv 0 \pmod{n}, \forall n > 0. \]

Therefore, we can analyze congruence equations to prove non-solvability of certain diophantine equation.

For a reason which will become self-evident through the following examples, often we seek to solve homogeneous equations, namely every monomial is of the same degree.

Moreover, when we talk about integer solutions, we mean non-trivial ones, namely non-zero solutions.

\textbf{eg.} In the case of Fermat’s equation \( x^n + y^n = z^n \), notice that \((a, 0, a)\) is always a solution. This does not tell you much about integers. In fact, Fermat’s Last Theorem is concerned with the non-existence of non-trivial solutions.

\textbf{eg.} Consider the equation \( x^2 + y^2 = 3z^2 \). By an earlier HW problem,

\[ x^2 + y^2 \equiv 0, 1, 2 \pmod{4}. \]

(In \( \mathbb{Z}/4\mathbb{Z} \), \( 0^2 = 0, 1^2 = 1, 2^2 = 0, 3^2 = 1 \).)

If \( z \) is odd, \( 3z^2 \equiv 0 \pmod{4} \). Then, there the equality cannot hold over \( \mathbb{Z}/4\mathbb{Z} \); and hence there is no integer solution with \( z \) being odd.

So we suppose \( z \) is even, and hence \( z^2 \equiv 0 \pmod{4} \). Correspondingly, \( x^2 + y^2 \equiv 0 \pmod{4} \). This is only possible when \( x, y \) are even. Hence, \( 2|(x, y, z) \).

The identity then looks like \((2n)^2 + (2m)^2 = 3 \cdot (2k)^2 \). But by dividing 4 on both sides of the equation, we arrive at a “smaller” solution.

We can then conclude that the equation does not have a non-trivial integer solution to start with. Otherwise, one can divide the solution by 2 infinitely many times and still get an integer solution. This is impossible!
Remark This leads to the important idea of infinite descent in number theory, which dates back to Euler. He used this idea to prove Fermat’s Last Theorem for $n = 3$. See Extra Credit HW5.

Notice it is important that we start with a homogeneous equation; otherwise, infinite descent fails.

Another example of degree 3: $x^3 + 6y^3 + 63z^3 = 0$.

Notice that over $\mathbb{Z}/3\mathbb{Z}$, this gives $x^3 \equiv 0 \pmod{3}$. In particular $x \equiv 0 \pmod{3}$.

Next look at the equation over $\mathbb{Z}/9\mathbb{Z}$, since

$$x^3 + 63z^3 = (3m)^3 + 9 \cdot (7z^3) \equiv 0 \pmod{9},$$

we get $6y^3 \equiv 0 \pmod{9}$. Hence, $y \equiv 0 \pmod{3}$.

Eventually, look at the equation over $\mathbb{Z}/27\mathbb{Z}$, since $x^3 + 6y^3 \equiv 0 \pmod{27}$, conclude that $63z^3 \equiv 0 \pmod{27}$, and therefore $z \equiv 0 \pmod{3}$. But then $3\gcd(x, y, z)$, and again by infinite descent this is impossible. Thus the original equation does not have a non-trivial integer solution.

### 18.3 Multiple Roots

Motivation: how do we distinguish between $(x - 1)^2g(x)$ and $(x - 1)g(x)$? This is mainly done through derivative tests.

In order to do so, we need to introduce the formal derivative of a polynomial:

**Definition 29.** Let $f(x) = a_nx^n + \ldots + a_1x + a_0$ be a polynomial with coefficient in $\mathbb{Z}/k\mathbb{Z}$ for some arbitrary $k$. Define its derivative $f'(x) = n a_n x^{n-1} + \ldots + 2a_2x + a_1$.

Before we pass to polynomials over $\mathbb{Z}/k\mathbb{Z}$, we state the derivative test for multiple roots over $\mathbb{Z}$:

**Proposition 44.** Suppose $f(x)$ is an integer-coefficient polynomial. $x_0$ is a root of multiplicity $n$ if and only if $f(x_0) = f'(x_0) = \ldots = f^{(n-1)}(x_0) = 0$, and $f^{(n)}(x_0) \neq 0$.

**Proof.** By definition, $x_0$ is a root of multiplicity $n$ if $f(x) = (x - x_0)^ng(x)$, where $x - x_0$ does not divide $g(x)$.

Using the product rule for derivatives, it is clear that the stated condition for $f$ and its derivatives holds.

Conversely, one can get $f(x)$ is of the desired form by direct calculation:

$$f(x_0) = 0 \Rightarrow f(x) = (x - x_0)g_1(x) \Rightarrow f'(x) = g_1(x) + (x - x_0)g'(x) \Rightarrow f'(x_0) = g_1(x_0)$$

$$f'(x_0) = 0 \Rightarrow g_1(x) = (x - x_0)g_2(x) \Rightarrow f(x) = (x - x_0)^2g_2(x)$$

$$\ldots$$

$$f^{(n-1)}(x_0) = 0 \Rightarrow f(x) = (x - x_0)^ng_n(x)$$

$$f^{(n)}(x_0) \neq 0 \Rightarrow g_n(x_0) \neq 0$$

And thus we know the root $x_0$ has multiplicity $n$. □
Remark. Here, we are working over \( \mathbb{Z} \), and not modding out by anything.

The situation becomes subtler when working over \( \mathbb{Z}/p\mathbb{Z} \), even if \( p \) is prime. In that case, if a root \( x_0 \) is of multiplicity \( n \), then by the same argument as above \( f(x_0) = f'(x_0) = \ldots = f^{(n-1)}(x_0) = 0 \) still holds. However, funny things may happen.

\[
\text{eg. Consider } f(x) = x^{21} - 1 = 0 \text{ over } \mathbb{Z}/7\mathbb{Z}. \quad f(1) = 0, \text{ and } f'(x) = 21x^{20} \equiv 0, \text{ hence all derivatives are zero!}
\]

But \( f \) does not have a root of multiplicity \( 21 \), because one can check that

\[
x^{21} - 1 = (x - 1)^7(x - 2)^7(x - 4)^7.
\]

Although not perfectly satisfactory, one can still say the following:

**Lemma 22.** Let \( f(x) \in (\mathbb{Z}/p\mathbb{Z})[x] \). \( a \) is a repeated root iff \( (x-a)|(f(x), f'(x)) \).

**Proof.** If \( x = a \) is a repeated root, then \( f(x) = (x - a)^n g(x) \), where \( n > 1 \).

\[
f'(x) = n(x - a)^{n-1} g(x) + (x - a)^n g'(x),
\]

and hence \( x - a|(f(x), f'(x)) \). (Notice that even if \( n \) is a multiple of \( p \), it does not matter, since it merely kills the first term in the derivative.)

Conversely, if \( (x-a)|(f(x), f'(x)) \), then \( f(x) = (x-a)g(x) \), and hence

\[
f'(x) = g(x) + (x-a)g'(x).
\]

Since \( (x-a)|f'(x) \), we get \( x-a|g(x) \), i.e. \( g(x) = (x-a)g_2(x) \) and therefore

\[
f(x) = (x-a)^2g_2(x).
\]

This implies \( x = a \) is a repeated root. \( \Box \)

**Proposition 45.** Let \( f(x) \in (\mathbb{Z}/p\mathbb{Z})[x] \) with \( \deg(f(x)) \leq p \). Let \( a \) be a root of \( f(x) \). Then,

1. The multiple of \( a \) is \( n < p \) iff \( f(a) = f'(a) = \ldots = f^{(n-1)}(a) = 0, f^{(n)}(a) \neq 0 \).

2. The multiple of \( a \) is \( p \) iff \( \deg(f(x)) = p \) and \( f'(x) \) is constantly zero. In that case, \( f(x) = (x-a)^p = x^p - a \).

**Proof.** (1) (“Only if” part)

\[
f(a) = f'(a) = \ldots = f^{(n-1)}(a) = 0 \text{ follows from the fact that } f(x) = (x-a)^n g(x) \text{ and applying product rule.}
\]

\[
f^{(n)}(a) \neq 0 \text{ follows from the fact that } f^{(n)}(x) = n!g(x) + \ldots + g^{(n)}(x)(x-a)^n = n!g(x) + (x-a)q(x), \text{ and hence } f^{(n)}(a) = n!g(a). \text{ Here, } n < p, \text{ so } q_p(n!) \neq 0.
\]

(This is important!)

(“If” part) This computation is exactly the same as in the case over \( \mathbb{Z} \) and is hence omitted.

(2) By definition, a degree \( p \) polynomial has a root of multiple \( p \) if and only if

\[
f(x) = (x-a)^p = x^p - a.
\]

The second equality follows from binomial expansion and that \( p \equiv 0(\text{mod } p) \). \( \Box \)
eg. Over $\mathbb{Z}/2\mathbb{Z}$, $x^2 - 1 = (x - 1)^2$.

## 18.4 Quadratic Residues

Our last topic for this summer session will be quadratic congruence equations. As you may have learned in high school, a quadratic equation is the next simplest polynomial equation other than the linear equation. So, it is natural to look at quadratic congruence equations since we have already understood linear congruence equations.

So, let’s look at congruence equations of the form

$$ax^2 + bx + c \equiv 0 \pmod{m}.$$ 

As we’ve discussed before, this problem reduces to solving equations of the form

$$ax^2 + bx + c \equiv 0 \pmod{p^N},$$

where $p$ is a prime number, and further reduces to solving

$$ax^2 + bx + c \equiv 0 \pmod{p}.$$

However, since $\mathbb{Z}/p\mathbb{Z}$ is a field, by multiplying $ax^2 + bx + c$ by $a^{-1}$, one can further assume the equation is monic, i.e. the highest degree coefficient equals 1. In other words, it suffices to study equations of the form $x^2 + ax + b \equiv 0 \pmod{p}$.

By completing the squares, it further suffices to discuss the quadratic equation of the form $x^2 \equiv a \pmod{p}$.

Notice first when $p = 2$, the situation is trivial. $x^2 \equiv 0 \pmod{2}$ has unique solution 0, and $x^2 \equiv 1 \pmod{2}$ has unique solution 1. Therefore, hereafter we assume $p > 2$. In particular, $p - 1$ is even and $\frac{p-1}{2}$ is an integer.

**Definition 30.** Suppose $p > 2$ is prime, $a \in \{1, 2, ..., p - 1\}$. If $x^2 \equiv a \pmod{p}$ has a solution, then $a$ is called a quadratic residue modulo $p$. Otherwise, it is called a quadratic non-residue modulo $p$.

**eg.** Consider $p = 7$. $1^2 \equiv 1 \pmod{7}$, $2^2 \equiv 4 \pmod{7}$, $3^2 \equiv 2 \pmod{7}$, $4^2 \equiv 2 \pmod{7}$, $5^2 \equiv 4 \pmod{7}$, $6^2 \equiv 1 \pmod{7}$. Therefore, 1, 2, 4 are the quadratic residues modulo 7, while 3, 5, 6 are quadratic non-residues modulo 7.

**Remark** Since we are solving an equation over $\mathbb{Z}/p\mathbb{Z}$, it suffices to look at $a \in \{1, 2, ..., p - 1\}$. Any other integer, not divisible by $p$, is just congruent to one of these numbers. However, we can certainly extend the definition to any integer $a$.

Mathematicians have defined the Legendre symbol to act as an indicator as for when the equation $x^2 \equiv a \pmod{p}$ is solvable, which we shall now introduce:

**Definition 31.** Let $p$ be a prime number and $a$ an integer. The Legendre symbol is defined as follows:
\[
(a/p) = \begin{cases} 
1 & \text{if } a \text{ is a quadratic residue of } p \\
-1 & \text{if } a \text{ is a quadratic non-residue of } p \\
0 & \text{if } p | a 
\end{cases} 
\]

In particular, this defines a new symbol, which is **has nothing to do with fractions!!!**
Chapter 19

Lecture 19

19.1 Warm-up

eg. Show that $x^n + p^iy^n + p^jz^n = 0$ does not have a non-zero integer solution, where $p$ is prime and $1 < i < j < n$.

Proof. Suppose $(x_0, y_0, z_0)$ is a non-zero solution. Then, $x_0^n + p^iy_0^n + p^jz_0^n = 0$.
This implies $x_0^n + p^iy_0^n + p^jz_0^n = 0 \equiv 0 \pmod{p}$, so $x_0^n \equiv 0 \pmod{p}$.
Therefore, $p|x_0^n$. Since $p$ is prime, $p|x_0$. Thus, $p^n|x_0^n$.
Now, $x_0^n + p^iy_0^n + p^jz_0^n = 0 \equiv 0 \pmod{p^i}$, so $p^iy_0^n \equiv 0 \pmod{p^i}$.
Therefore, $p^i|p^iy_0^n$; hence, $p|y_0$ and $p^n|y_0^n$.
Eventually, $x_0^n + p^iy_0^n + p^jz_0^n = 0 \equiv 0 \pmod{p^n}$, so $p^jz_0^n \equiv 0 \pmod{p^n}$.
Therefore, $p^n|p^jz_0^n$; hence, $p|z_0$.
Thus, $p|(x_0, y_0, z_0)$. By infinite descent, this is impossible. □

19.2 Quadratic Residues (Continued)

Yesterday, we introduced the Legendre Symbol:

Definition 32. Let $p$ be a prime number, and $a$ an integer. The Legendre symbol is defined as follows:

$$\left(\frac{a}{p}\right) = \begin{cases} 1 & \text{if } a \text{ is a quadratic residue of } p \\ -1 & \text{if } a \text{ is a quadratic non-residue of } p \\ 0 & \text{if } p|a \end{cases}$$

We shall develop three main theorems related to quadratic residues:

1. Euler’s Theorem;
2. Gauss’s Lemma;
3. Quadratic Reciprocity Law.

Example. Let \( p = 7 \). We mentioned yesterday that 1, 2, 4 are the quadratic residues modulo 7 and 3, 5, 6 are the quadratic non-residues modulo 7. One can check directly that 1, 2, 4 are roots to the equation \( x^3 - 1 = 0 \), and 3, 5, 6 are the roots to the equation \( x^3 + 1 = 0 \).

In general, \( x^{p-1} - 1 = (x^{\frac{p-1}{2}} - 1)(x^{\frac{p-1}{2}} + 1) \). It turns out the first factor keeps track of all the quadratic residues modulo \( p \), while the second factor remembers the quadratic non-residues.

In particular, this tells us that exactly half of the units are quadratic residues, and the other half of them are quadratic non-residues. This result is due to Euler.

**Theorem 10.** (Euler) Let \( p \) be an odd prime. Then, \( a^{\frac{p-1}{2}} \equiv (a/p) (\text{mod } p) \).

**Proof.** First, notice that \( x^{p-1} - 1 = (x^{\frac{p-1}{2}} - 1)(x^{\frac{p-1}{2}} + 1) \). Therefore, \( x^{\frac{p-1}{2}} - 1 \) is a factor of \( x^p - x = x(x^{p-1} - 1) \), and has exactly \( \frac{p-1}{2} \) distinct roots. By exactly the same argument, \( x^{\frac{p-1}{2}} + 1 \) has also \( \frac{p-1}{2} \) distinct roots. Clearly, the set of roots of \( x^{\frac{p-1}{2}} + 1 \) and the set of roots of \( x^{\frac{p-1}{2}} - 1 \) have no intersection, and they together run through all \( p - 1 \) units of \( \mathbb{Z}/p\mathbb{Z} \).

Next, if \( (a/p) = 1 \), that means there is some \( b \in \mathbb{Z}/p\mathbb{Z} \) such that \( b^2 = a \). Fermat’s Little Theorem tells us \( b^{p-1} = a^{\frac{p-1}{2}} \equiv 1 (\text{mod } p) \), i.e. \( a \) is also one of the roots of \( x^{\frac{p-1}{2}} - 1 \).

On the other hand,

\[
x^{p-1} - 1 = (x^{p-1} - a^{\frac{p-1}{2}}) + (a^{\frac{p-1}{2}} - 1) = ((x^2)^{\frac{p-1}{2}} - a^{\frac{p-1}{2}}) + (a^{\frac{p-1}{2}} - 1)^2
\]

We conclude that, \( (x^2)^{\frac{p-1}{2}} - a^{\frac{p-1}{2}} = (x^2 - a)((x^2)^{\frac{p-1}{2}} - 1) + (x^2)^{\frac{p-1}{2}} - 2a + \cdots + a^{\frac{p-1}{2}} - 1 \), so \( x^2 - a \) \((x^2)^{\frac{p-1}{2}} - a^{\frac{p-1}{2}} \). Thus, if \( a^{\frac{p-1}{2}} - 1 \equiv 0 (\text{mod } p) \), \( x^2 - a \) is a factor of \( x^{p-1} - 1 \), regarded as polynomials with coefficients in \( \mathbb{Z}/p\mathbb{Z} \). Hence, \( x^2 - a \) has two roots, and \( (a/p) = 1 \).

By the previous two paragraphs, we have shown \( (a/p) = 1 \) if and only if \( a^{\frac{p-1}{2}} \equiv 1 (\text{mod } p) \). As said before, these cover \( \frac{p-1}{2} \) many units of \( \mathbb{Z}/p\mathbb{Z} \). The left-over ones are quadratic non-residues, and have to be the roots of \( x^{\frac{p-1}{2}} + 1 \). This further shows that \( (a/p) = -1 \) if and only if \( a^{\frac{p-1}{2}} \equiv -1 (\text{mod } p) \).

Thus, \( a^{\frac{p-1}{2}} \equiv (a/p) (\text{mod } p) \). \( \square \)

In particular, we know:

**Corollary 11.** Let \( p \) be an odd prime. Among the numbers \( 1, 2, \ldots, p - 1 \), exactly half of them are quadratic residues modulo \( p \).

The most useful consequence is the following:

\[ a - B = (A - B)(A^{n-1} + A^{n-2}B + \cdots + B^{n-1}). \]

\[ \]
Corollary 12. $(ab/p) = (a/p)(b/p)$.

Proof. By Euler’s Theorem, $(a/p)(b/p) \equiv a^{\frac{p-1}{2}} \cdot b^{\frac{p-1}{2}} \equiv (ab)^{\frac{p-1}{2}} \equiv (ab/p)(\text{mod } p)$.

Corollary 13. $(-1/p) = (p - 1/p) = (-1)^{\frac{p-1}{2}}$.

Notice this recovers our statement that $\mathbb{Z}/p\mathbb{Z}$ has a unit of order 4 if and only if $p = 4k + 1$ for some $k$:

Proof. $(-1/p) = 1$ iff $x^2 = -1$ has a root $x = b$ in $\mathbb{Z}/p\mathbb{Z}$.

$b^2 = -1$ in $\mathbb{Z}/p\mathbb{Z}$ iff $\text{ord}_p(b) = 4$ iff $p$ is of the form $4k + 1$.

eg. Find the quadratic residues and quadratic non-residues modulo 11.

We can look at the roots of $x^5 - 1 = 0$ over $\mathbb{Z}/11\mathbb{Z}$. They are units of order 1 and 5, which turn out to be 1, 3, 4, 5, 9. Correspondingly, 2, 4, 6, 7, 10 are the quadratic non-residues.

Next, we shall present Gauss’s Lemma. In order to discuss this result, a different complete residue system is introduced. For $p$ an odd prime, denote

$$N_p = \{-\frac{p-1}{2}, ..., -1, 0, 1, ..., \frac{p-1}{2}\}.$$ 

This is nothing but a different way to present remainders modulo $p$.

Lemma 23. Suppose $(a, p) = 1$. Compute the set $Q = \{a, 2a, ..., \frac{p-1}{2}a\}$. Let $n$ be the number of elements in $Q$ whose remainder modulo $p$ is presented by a negative number in $N_p$. Then, $(a/p) = (-1)^n$.

eg. Consider $p = 7$, $a = 2$. We know $(2/7) = 1$, since $4^2 = 2 \cdot 7 + 2$.

Gauss’s Lemma tells you to compute the set $\{2, 2 \cdot 2, 3 \cdot 2\} = \{2, 4, 6\}$. Presenting these remainders by numbers in $N_7$, we get $Q = \{2, -3, -1\}$, so there are two negative numbers in $Q$.

So, $(2/7) = (-1)^2 = 1$, which agrees with what we know.

Proof. For $i = 1, 2, ..., \frac{p-1}{2}$, suppose $i \cdot a \equiv b_i(\text{mod } p)$, where $b_i \in N_p$.

We claim that $|b_1|, ..., |b_{\frac{p-1}{2}}|$ are distinct elements in $N_p$.

To see the claim, first we show $b_1, ..., b_{\frac{p-1}{2}}$ are distinct. Since $q_p(a)$ is a unit in $\mathbb{Z}/p\mathbb{Z}$, and all $i = 1, 2, ..., \frac{p-1}{2}$ are units as well, $\forall i \neq j$, $b_i \neq b_j$.

$(q_p(a) \cdot i = q_p(a) \cdot j$ implies $q_p(a)^{-1} \cdot q_p(a) \cdot i = q_p(a)^{-1} \cdot q_p(a) \cdot j$, i.e. $i = j$.)

Next, we show $b_i \neq -b_j$, $\forall i \neq j$. We make the following observation: $b_i + b_j = \ldots$
19.3 Quadratic Reciprocity

$q_p(i + j) \cdot q_p(a)$, where $i, j \leq \frac{p-1}{2}$; since $1 < i + j < p - 1$, $q_p(i + j)$ is another unit. So, it is clear that $b_1 + b_j \neq 0$. Therefore, not only $b_1, ..., b_{\frac{p-1}{2}}$ are all distinct, but $|b_1|, ..., |b_{\frac{p-1}{2}}|$ are all distinct as well. But there are precisely $\frac{p-1}{2}$ many positive numbers in $N_p$, so $|b_1|, ..., |b_{\frac{p-1}{2}}|$ must run through all of them. This means

$$\prod b_i = (-1)^n \prod |b_i| = (-1)^n (\frac{p-1}{2})!,$$

Eventually, compute $\prod_{i=1}^{\frac{p-1}{2}} (i \cdot a) = a^{\frac{p-1}{2}} (\frac{p-1}{2})!$, and hence $q_p(a^{\frac{p-1}{2}} (\frac{p-1}{2})!) = \prod b_i$.

Thus, $q_p((\frac{p-1}{2})! a^{\frac{p-1}{2}}) = q_p(b_1 b_2 \ldots b_{\frac{p-1}{2}}) = q_p((-1)^n |b_1| \ldots |b_{\frac{p-1}{2}}|) = (-1)^n q_p((\frac{p-1}{2})!)$.

Since $q_p((\frac{p-1}{2})!)$ is nothing but another unit, we can cancel them out on both sides, and get $q_p(a^{\frac{p-1}{2}}) = (-1)^n$.

By Euler’s theorem, $(a/p) \equiv a^{\frac{p-1}{2}} \equiv (-1)^n$ (mod $p$).

**Remark** One can also consider $Q' = \{ a \cdot i | i = 1, 2, ..., p-1 \}$, and let $m$ be the number of elements in $Q'$ which has remainder modulo $p$ positive (represented by numbers in $N_p$, of course). Then, $m = n$, and you get the same result for $(a/p)$.

19.3 Quadratic Reciprocity

Recall that the study of number theory is to find (non-obvious) structures governing the behavior of integers. The main theorem regarding quadratic residues is the Quadratic Reciprocity Law, which is easy to state and presents an extremely good example of some deep structure among prime numbers.

**Theorem 11.** Let $p, q$ be distinct odd primes, then $(p/q)(q/p) = (-1)^{(\frac{p-1}{2})(\frac{q-1}{2})}$.

Putted differently, one can also state it as follows:
Theorem 12. Let $p, q$ be two distinct odd prime numbers.
(1) When one of $p, q$ is of the form $4k + 1$, $p$ is a quadratic residue of $q$ if and only if $q$ is a quadratic residue of $p$.
(2) When both of $p, q$ are of the form $4k + 3$, then $p$ is a quadratic residue of $q$ if and only if $q$ is not a quadratic residue of $p$. 
Chapter 20

Lecture 20

20.1 Warm-up

eg. Find \((\frac{6}{7})\).

\[(\frac{6}{7}) = \left(\frac{1}{7}\right) = (-1)^{-\frac{6-1}{2}} = -1.\] The last step uses Euler’s Theorem.

eg. Find \((\frac{181}{7})\).

Notice that \(181 \equiv 6 \pmod{7}\), so \((\frac{181}{7}) = (\frac{6}{7}) = -1.

eg. Find \((\frac{7}{181})\).

Since \(181 = 4 \cdot 45 + 1\), by Quadratic Reciprocity Law, \((\frac{7}{181}) = (\frac{181}{7}) = -1.

Remark Two guiding principles when computing Legendre symbols:

1. When \(a > p\), replace with a smaller integer that is congruent to \(a\) modulo \(p\);
2. when \(a \ll p\), may want to consider quadratic reciprocity law.

20.2 Quadratic Reciprocity

Last week, we stated the quadratic reciprocity law in two different forms:

**Theorem 13.** Let \(p, q\) be distinct odd primes, then \((p/q)(q/p) = (-1)^{(p-1)(q-1)/4}\).

Putted differently, one can also state it as follows:

**Theorem 14.** Let \(p, q\) be two distinct odd prime numbers.

1. When one of \(p, q\) is of the form \(4k + 1\), \(p\) is a quadratic residue of \(q\) if and
only if \( q \) is a quadratic residue of \( p \).

(2) When both of \( p, q \) are of the form \( 4k + 3 \), then \( p \) is a quadratic residue of \( q \) if and only if \( q \) is not a quadratic residue of \( p \).

These two are indeed the same, because:

\[
\frac{(p - 1)(q - 1)}{4} = \begin{cases} 
1 & \text{when one of } p, q \text{ is of the form } 4n + 1 \\
-1 & \text{when both } p, q \text{ are of the form } 4n + 3.
\end{cases}
\]

When \( (p/q)(q/p) = (-1)^{\frac{(p-1)(q-1)}{4}} = 1 \), this means either \((p/q) = (q/p) = 1\), or \((p/q) = (q/p) = -1\), i.e. either \( x^2 = p \) over \( \mathbb{Z}/q\mathbb{Z} \) and \( x^2 = q \) over \( \mathbb{Z}/p\mathbb{Z} \) both have solutions, or neither of them has a solutions.

When \( (p/q)(q/p) = (-1)^{\frac{(p-1)(q-1)}{4}} = -1 \), this means either \((p/q) = 1\), \((q/p) = -1\), or \((p/q) = -1\), \((q/p) = 1\), i.e. exactly one of \( x^2 = p \) over \( \mathbb{Z}/q\mathbb{Z} \) and \( x^2 = q \) over \( \mathbb{Z}/p\mathbb{Z} \) has solutions.

### 20.3 Proof of the Quadratic Reciprocity Law

The main idea behind this proof is that one needs to find some set of objects which relates units in \( \mathbb{Z}/p\mathbb{Z} \) to units in \( \mathbb{Z}/q\mathbb{Z} \). Indeed, we are interested in the relation between the following two questions: (1) whether any of the units in \( \mathbb{Z}/p\mathbb{Z} \) solves \( x^2 = q \); (2) whether any units in \( \mathbb{Z}/q\mathbb{Z} \) solve \( x^2 = p \).

The set of objects we find is \( S = \{ a \in \mathbb{Z} | (a, pq) = 1, 1 \leq a \leq \frac{pq-1}{2} \} \). Notice that by the **Chinese Remainder Theorem**, the units in \( \mathbb{Z}/pq\mathbb{Z} \) I-1 correspond to pairs \((a, b) \in U_p \times U_q\).

Notice also we are only using half of the units in \( \mathbb{Z}/pq\mathbb{Z} \). The other half of the units are nothing but additive inverses of these half units. Just as we have seen in Gauss’s Lemma, this turns out to be the right idea to gather information on quadratic congruence. (On the other hand, using all the units will cause redundancy which prevents us from gathering the correct information.) By the formula for Euler-\( \varphi \) function, there are \( \frac{(p-1)(q-1)}{2} \) many elements in \( S \).

**eg.** Let \( p = 3 \), \( q = 5 \). Then \( \mathbb{Z}/15\mathbb{Z} \) contains \( \varphi(15) = \varphi(3)\varphi(5) = 8 \) many units. They can be written as \( \pm 1, \pm 2, \pm 4, \pm 7 \).

**Proof.** First, notice that integers in \( S \) correspond to integers under \( \frac{pq-1}{2} \) that are coprime to both \( p \) and \( q \). Among integers \( \leq \frac{pq-1}{2} \), the ones which are coprime to \( p \) can be listed in a table:

<table>
<thead>
<tr>
<th>( \frac{q+1}{2} )-th row</th>
<th>1 ( p+1 )</th>
<th>2 ( p+2 )</th>
<th>...</th>
<th>...</th>
<th>...</th>
<th>( p-2 )</th>
<th>( p-1 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>2nd row</td>
<td>1 ( p+1 )</td>
<td>2 ( p+2 )</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>( p-2 )</td>
<td>( p-1 )</td>
</tr>
<tr>
<td>1st row</td>
<td>1</td>
<td>2</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>( p-2 )</td>
<td>( p-1 )</td>
</tr>
<tr>
<td>(( \frac{q+1}{2} )-th row</td>
<td>(( \frac{q+1}{2} ))p + 1</td>
<td>(( \frac{q+1}{2} ))p + 2</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>( (\frac{q+1}{2})p + \frac{p-1}{2} )</td>
<td>...</td>
</tr>
</tbody>
</table>
Among these numbers, the ones not coprime to \( q \) are \( q, 2q, \ldots, \frac{p-1}{2} \). (Notice that these integers are indeed coprime to \( p \)).

Similar to the proof of Gauss’s Lemma, define \( A = \prod_{a \in S} a \). Compute \( q_p(A), q_q(A) \).

The reciprocity law will follow from two different presentations of these two remainders.

By the above table, easy to see:

\[
A = \frac{(p-1)![ (p+1) \ldots (2p-1) ] \ldots [ (\frac{q-3}{2})p+1 ] \ldots [ (\frac{q-1}{2})p+1 ] \ldots [ (\frac{p-1}{2})p + \frac{p-1}{2} ]}{\prod_{i=1}^{\frac{q-1}{2}} (q \cdot i)}
\]

\[
= \frac{p^{-3}}{\prod_{i=1}^{\frac{q-1}{2}} (i \cdot p + 1) \cdot \ldots \cdot (i \cdot p + p - 1)} 
\frac{[ (\frac{q-1}{2})p+1 ] \ldots [ (\frac{q-1}{2})p + \frac{p-1}{2} ]}{q^{\frac{p-1}{2}} \cdot (\frac{q-1}{2})!}
\]

By Wilson’s theorem, and Euler’s Theorem, we get

\[
q_p(A) = q_p((-1)^{\frac{p-1}{2}} \cdot (\frac{q-1}{2})! / q^{\frac{p-1}{2}} \cdot (\frac{p-1}{2})!) \equiv (-1)^{\frac{p-1}{2}} \cdot (\frac{q}{p}).
\]

Similarly, by switching \( p \) and \( q \) we get \( q_q(A) = (-1)^{\frac{q-1}{2}} \cdot (\frac{p}{q}) \).

These two results can be presented as

\[
\begin{cases}
A \equiv (-1)^{\frac{p-1}{2}} \cdot (\frac{q}{p}) \pmod{p} \\
A \equiv (-1)^{\frac{q-1}{2}} \cdot (\frac{p}{q}) \pmod{q}.
\end{cases}
\]

Next, we shall present \( q_p(A), q_q(A) \) in a different way. We shall use the following result: any element in \( U_p \times U_q \) can be written as \( (a, b) \) or \( -(a, b) \) where \( 1 \leq a \leq p - 1, 1 \leq b \leq \frac{q-1}{2} \). (The claim is given as Lemma 24 after this proof.)

Notice there is a natural map \( F : S \to U_p \times U_q \), where \( c \) is sent to \( (q_p(c), q_q(c)) \).

By Chinese Remainder Theorem, this map is injective: for any integer \( a \), its integer modulo \( p \) and its integer modulo \( q \) uniquely determines its remainder modulo \( pq \) and vice versa. Then, the image set \( F(S) \) in \( U_p \times U_q \) is a set of \( \frac{(p-1)(q-1)}{2} \) many elements, which can be described as follows: for any \( 1 \leq a \leq p - 1, 1 \leq b \leq \frac{q-1}{2} \), exactly one of \( (a, b) \) and \(- (a, b) \) is in \( F(S) \). This also follows from the Chinese Remainder Theorem:

\[
(q_p(c), q_q(c)) = (0, 0) \text{ if and only if } pq \mid c. \text{ Since any element in } S \text{ is between } 1 \text{ and } \frac{p-1}{2}, \text{ the sum of any two of them cannot be a multiple of } pq. \text{ Therefore, } \]

\[
(q_p(c_1 + c_2), q_q(c_1 + c_2)) = (q_p(c_1), q_q(c_1)) + (q_p(c_2), q_q(c_2)) \neq (0, 0). \text{ So, the sum of any two pairs in } F(S) \text{ cannot be } (0, 0), \text{ i.e. } (a, b) \text{ and } -(a, b) \text{ cannot be both in } F(S). \text{ On the other hand, } F(S) \text{ has } \frac{(p-1)(q-1)}{2} \text{ many elements, so by Lemma 24 } F(S) \text{ must contain one of } (a, b) \text{ and } -(a, b), \text{ for any } 1 \leq a \leq p - 1, 1 \leq b \leq \frac{q-1}{2} \text{. Thus,}
\]

\[
F(S) = \{ (-1)^j (a_i, b_i) \mid 1 \leq a_i \leq p - 1, 1 \leq b_i \leq \frac{q-1}{2} \},
\]
where $\epsilon_i$ could be 0 or 1.

By the above observation, $q_p(A) = (-1)^N q_p((p-1)! \frac{q-1}{2})$, $q_q(A) = (-1)^N q_q((q-1)! \frac{q-1}{2})$.

The $(-1)^N$ term comes from the fact that some $\epsilon_i = 1$. (We do not care how many $-1$'s there are in total.)

By Wilson’s Theorem, $q_p(A) = (-1)^N(-1) \frac{q-1}{2}$.

Still by Wilson’s Theorem, $-\frac{1}{2} = q_q(1 \cdot 2 \cdot \ldots \cdot q-1) = q_q((-1) \frac{q-1}{2}((q-1)! \frac{q-1}{2}))$.

Hence,

$$q_q(A) = (-1)^N q_q([\frac{q-1}{2}]^{p-1}) = (-1)^N q_q([\frac{q-1}{2}]^{p-1})$$

By Wilson’s Theorem, 

$$\frac{q-1}{2} = q_q(q-1) = q_q((-1) \frac{q-1}{2}((-1) \frac{q-1}{2}))$$

Thus, by multiplying the two equations, we get

$$q_q(A) = (-1)^N q_q[\frac{q-1}{2}]^{p-1}(-1) \frac{q-1}{2}((-1) \frac{q-1}{2})$$

Therefore, we also have

$$q_q(A) = (-1)^N(-1) \frac{q-1}{2} \mod p$$

$$q_q(A) = (-1)^N(-1) \frac{q-1}{2} \mod q.$$

Eventually, combine systems (1) and (2), we get

$$A \equiv (-1)^N(-1) \frac{q-1}{2} \mod p$$

$$A \equiv (-1)^N(-1) \frac{q-1}{2} \mod q.$$ 

This further implies

$$\left(\frac{q}{p}\right) \equiv (-1)^N \mod p$$

$$\left(\frac{q}{q}\right) \equiv (-1)^N(-1) \frac{q-1}{2} \mod q.$$ 

Since $\left(\frac{q}{p}\right), \left(\frac{q}{q}\right)$ can only be 1 or -1, the numbers on the right-hand side are also 1 or -1 and $p, q > 2$, it follows that $\left(\frac{q}{p}\right) = (-1)^N, \left(\frac{q}{q}\right) = (-1)^N(-1) \frac{q-1}{2}$. 

Thus, by multiplying the two equations, we get

$$\left(\frac{q}{p}\right) \left(\frac{q}{q}\right) = (-1)^N \frac{q-1}{2}.$$

\qed

**Lemma 24.** Let $U_p, U_q$ be the groups of units in $\mathbb{Z}/p\mathbb{Z}, \mathbb{Z}/q\mathbb{Z}$ resp. Then,

$$U_p \times U_q = \{ \pm(a, b) | 1 \leq a \leq p-1, 1 \leq b \leq \frac{q-1}{2} \}.$$ 

**Proof.** Using the standard complete residue system, any element in $U_p \times U_q$ can be written as $(\overline{a}, \overline{b})$, where $1 \leq a \leq p-1, 1 \leq b \leq q-1$. If $b \leq \frac{q-1}{2}$, the pair is already of the desired form, and we have nothing to show. If $b > \frac{q-1}{2}$, one can re-write $(\overline{a}, \overline{b}) = (-(p-a), -(q-b)) = (\overline{p-a}, \overline{q-b})$, where $1 \leq q-b \leq \frac{q-1}{2}$.

\qed
**eg.** A concrete example of the lemma: take $p = 3, q = 5$. Then, $(2, 4) \in U_3 \times U_5$ can be written as

$$ (2, 4) = (-1, -1) = -(1, 1). $$

**Remark** This is just a matter of re-writing things.
Chapter 21

Lecture 21

21.1 Warm-up

We briefly go over some major topics throughout the course:

1. Chinese Remainder Theorem

   (a) Be able to solve a single system of linear congruence equations, where the dividends are mutually coprime;
   (b) be able to solve \( f(x) \equiv 0 \pmod{n} \).

   \begin{itemize}
   \item Factorize \( n = p_1^{m_1} \cdots p_r^{m_r} \);
   \item Solve \( f(x) \equiv 0 \pmod{p_i^{m_i}} \) separately;
   \end{itemize}

Step 0: Solve \( f(x) \equiv 0 \pmod{p} \), either by trying out \( 0, 1, ..., p - 1 \) or computing \( (f, x^p - x) \).

Now suppose you have solved \( f(x) \equiv 0 \pmod{p^N} \).

Step 1: Pick a \( x_0 \) which is a solution to \( f(x) \equiv 0 \pmod{p^N} \).

Step 2: Solve for some \( t \) such that \( f(x_0 + tp^N) \equiv 0 \pmod{p^{N+1}} \).

This further reduces to solving \( f'(x_0)t = -\frac{f(x_0)}{p^N} \pmod{p} \) (1).

Step 3: We can now use our knowledge of solving linear congruence equation to discuss all possible situations:

i. Case 1: \( (f'(x_0), p) = 1 \). In this case, there is a unique \( t \) such that the linear equation holds.

ii. Case 2: \( p | f'(x_0) \) and \( p | \frac{f(x_0)}{p^N} \). Then, the equation (1) becomes \( 0 \equiv 0 \). Any \( t \) is a solution. So one takes \( t = 0, 1, ..., p - 1 \).
21.1. WARM-UP

iii. Case 3: $p | f'(x_0)$ but $p$ does not divide $\frac{f(x_0)}{p}$. Then, $f(x) \equiv 0 \pmod{p^{N+1}}$ does not have a solution of the form $x_0 + tp^N$.

Step 4: Try step 1-3 for every solution of $f(x) \equiv 0 \pmod{p^N}$.
- Apply Chinese Remainder Theorem to every system obtained by combining solutions of $f(x) \equiv 0 \pmod{p_i^{m_i}}$. For example, if $n$ has two prime factors and $f(x) \equiv 0 \pmod{p_1^{m_1}}$ has 3 solutions, $f(x) \equiv 0 \pmod{p_2^{m_2}}$ has 2 solutions, then there are 6 systems in all.

2. Euler-phi function
- $\varphi(mn) = \varphi(m) \varphi(n)$, when $(m, n) = 1$.
- $\varphi(n) = n \cdot \prod_{p|n} \left(1 - \frac{1}{p}\right)$.
- $\varphi(p) = p - 1$, when $p$ is prime; $\varphi(1) = 1$.
- $\varphi(mn) = \frac{d \varphi(m) \varphi(n)}{\varphi(d)}$, where $d = (m, n)$.

3. Units and their orders
- $\text{ord}_n(a) | \varphi(n)$, $\text{ord}_{p}(a) | p - 1$.
- $a^{\varphi(n)} \equiv 1 \pmod{n}$, if $(a, n) = 1$; $a^{p-1} \equiv 1 \pmod{p}$, if $p(a, p) = 1$.

4. Euclidean algorithm

5. Fundamental Theorem of Arithmetic and Divisibility, in particular, topics related to least common multiple and greatest common divisor.

6. Polynomials over $\mathbb{Z}/p\mathbb{Z}$
   - (a) Understand $\mathbb{Z}/p\mathbb{Z}$ as a root detector;
   - (b) understand the two theorems on irreducible polynomials over $\mathbb{Z}/p\mathbb{Z};$
   - (c) Euclidean algorithm for polynomials;
   - (d) repeated roots: $(f, f') \neq 1$;
   - (e) Wilson’s Theorem.

7. Quadratic Reciprocity
- Know its precise statement;
- use it for computing Legendre symbols;
- theoretical applications: see the last example.

8. Euler’s Theorem
- Structural facts: quadratic residues are roots of $x^{\frac{p-1}{2}} - 1$; quadratic non-residues are roots of $x^{\frac{p+1}{2}} + 1$;
- Notice that $a^{p-1} \equiv 1$ if $(a, p) = 1$; $a^{\frac{p-1}{2}}$ is either $\overline{1}$ or $-\overline{1}$; when $a$ is a quadratic residue, it looks like $b^2$ and thus $a^{\frac{p-1}{2}} = (b^2)^{\frac{p-1}{4}} = b^{p-1} = \overline{1}$.

9. Gauss’s Lemma: see how we use it to find $(\frac{2}{p})$.
21.2 More Examples on Quadratic Congruences

Here we summarize what we know about equations of the form $x^2 \equiv a \pmod{p}$.

**Proposition 46.** $x^2 \equiv -1 \pmod{p}$ has a solution iff $p = 4k + 1$ for some $k$.

Stated in the notation of Legendre symbol, $(-1/p) = (-1)^{\frac{p-1}{2}}$.

**Proof.** $b^2 = -1$ in $\mathbb{Z}/p\mathbb{Z}$ iff $\text{ord}_p(b) = 4$, iff $p$ is of the form $4k + 1$. □

eg. Trivially, $(0/p) = 0$, $(1/p) = 1$.

**Proposition 47.** $(2/p) = \begin{cases} 0 & (p = 2) \\ 1 & (p = 8k + 1, 8k + 7) \\ -1 & (p = 8k + 3, 8k + 5) \end{cases}$

**Proof.** We apply Gauss’s lemma. In $Q = \{2, 4, 6, ..., p-1\}$, we count the number of elements having negative remainders, when applying the complete residue system $N_p = \{ -\frac{p-1}{2}, ..., -1, 0, 1, ..., \frac{p-1}{2} \}$. There are four cases. Notice that $Q$ has $\frac{p-1}{2}$ many elements.

When $p = 8k + 1$, $\frac{p-1}{2}$ is a multiple of 4, and precisely half of the elements have negative remainders in $N_p$. They are

$$2 \cdot \frac{p+3}{4}, 2 \cdot \frac{p+7}{4}, ..., 2 \cdot \frac{p-1}{2}.$$ 

Therefore, $(2/p) = (-1)^{\frac{p-1}{2}} = 1$.

When $p = 8k + 3$, $\frac{p-1}{2}$ is an odd integer, and there are $\frac{p+1}{4}$ many elements in $Q$ having negative remainders in $N_p$. They are

$$2 \cdot \frac{p+1}{4}, 2 \cdot \frac{p+5}{4}, ..., 2 \cdot \frac{p-1}{2}.$$ 

So, $(2/p) = (-1)^{\frac{p+1}{2}} = -1$.

When $p = 8k + 5$, $\frac{p-1}{2}$ is even, and there are $\frac{p-1}{4}$ many elements in $Q$ having negative remainders in $N_p$. They are

$$2 \cdot \frac{p+3}{4}, 2 \cdot \frac{p+7}{4}, ..., 2 \cdot \frac{p-1}{2}.$$ 

So, $(2/p) = (-1)^{\frac{p+1}{2}} = -1$.

When $p = 8k + 7$, $\frac{p-1}{2}$ and is an odd integer, and there are $\frac{p+1}{4}$ many elements in $Q$ having negative remainders in $N_p$. They are

$$2 \cdot \frac{p+1}{4}, 2 \cdot \frac{p+5}{4}, ..., 2 \cdot \frac{p-1}{2}.$$ 

So, $(2/p) = (-1)^{\frac{p+1}{2}} = 1$.

This concludes the proof. □
In the Legendre symbol notation, \((2/p) = (-1)^{(p^2 - 1)/2}\).

Eventually, in the general situation, where \(p, q\) are two different odd prime numbers, our main tool is the reciprocity law.

The quadratic reciprocity law induces the following computational result:

**Proposition 48.** Fix an odd prime \(q\). Suppose \(p\) is another odd prime. \((q/p) = 1\) iff \((r/q) = 1\), where \(r\) is the standard remainder of \(p\) modulo \(4q\) if \(p\) is of the form \(4t + 1\); and is the negative of the remainder, if \(p\) is of the form \(4t + 3\).

**Proof.** Suppose \((r/q) = 1\). When \(p = 4t + 1\), \(p \equiv r(\mod 4q)\) implies that \(p \equiv r(\mod q)\). Therefore, \((p/q) = (r/q) = 1\). By the quadratic reciprocity law, \((q/p) = 1\). Vice versa.

When \(p\) is of the form \(4t + 3\), we get \((p/q) = (-r/q) = (-1/q)(r/q)\). We know that \((-1/q) = 1\) iff \(q\) is of the form \(4t + 1\). Thus, when \(q\) is of the form \(4t + 1\), we get \((p/q) = (-r/q) = (-1/q)(r/q) = (r/q) = 1\); when \(q\) is of the form \(4t + 3\), \((p/q) = -(r/q) = -1\). By the reciprocity law, \((q/p) = 1\). Vice versa. \(\Box\)

**eg.** Compute \((3/p)\) for of all odd primes \(p > 3\).

In the spirit of the previous proposition, we should look at the remainder of \(p\) modulo 12. Since \(p > 3\) is an odd prime, the possible remainders are 1,5,7,11.

When \(p = 12t + 1\), \(p\) is of the form \(4t + 1\), so take \(r = 1\). \((1/3) = 1\), so \((3/p) = 1\).

When \(p = 12t + 5\), \(p\) is of the form \(4t + 1\), so take \(r = 5\). \((5/3) = (2/3) = -1\), so \((3/p) = -1\).

When \(p = 12t + 7\), \(p\) is of the form \(4t + 3\), so take \(r = -7\). \((-7/3) = (2/3) = -1\), so \((3/p) = -1\).

When \(p = 12t + 11\), \(p\) is of the form \(4t + 3\), so take \(r = -11\). \((-11/3) = (1/3) = 1\), so \((3/p) = 1\).

Therefore,

\[
(3/p) = \begin{cases} 
1 & p \equiv 1,11(\mod 12) \\
-1 & p \equiv 5,7(\mod 12)
\end{cases}
\]

**Remark** You will be asked to do a similar computation in HW9.

### 21.3 More Examples

The following problem is from an earlier HW.

**eg.** Show that \(x^3 - 26y^3 = 24z^3\) does not have non-trivial solution.
Modulo 13, one gets $x^3 \equiv -2z^3 \pmod{13}$. We have two cases to discuss.

Case 1: $q_{13}(x) = q_{13}(z) = 0$. In this case, $13|\gcd(x, y, z)$ always holds. But this is a homogeneous equation, if there is any non-trivial solution, one can always assume the solution $(x, y, z)$ satisfies $\gcd(x, y, z) = 1$. Therefore, one can ignore Case 1.

Case 2: $q_{13}(x), q_{13}(z) \neq 0$. This implies $[q_{13}(x) \cdot (q_{13}(z))^{-1}]^3 = -\overline{2}$, which further implies $x^3 = -2$ has a solution over $\mathbb{Z}/13\mathbb{Z}$. Yet this is impossible. If $a$ is another unit such that $a^3 = -\overline{2}$, by Fermat’s Little Theorem, $a^{12} = (-2)^4 = 1$. But $(-2)^4 \equiv 3 \neq 1$, so we get a contradiction.

**eg.** When you compute Problem 8 on HW9, you may have encountered the following number: $\left(\frac{499^2 - 1}{997}\right)$.

Notice that $499^2 - 1 = (499 + 1)(499 - 1) = 500 \cdot 498$. So, $\left(\frac{499^2 - 1}{997}\right) = (\frac{500}{997}) \cdot (\frac{498}{997})$.

The point I want to make is that $500 = 5 \cdot 10^2$ and $(\frac{10^2}{997}) = (\frac{10}{997})^2 = 1$. So you know $(\frac{500}{997}) = (\frac{5}{997})$.

More generally,

$$\left(\frac{a^2}{p}\right) = \begin{cases} 0 & \text{if } p|a \\ 1 & \text{otherwise.} \end{cases}$$

**eg.** Let $p, q$ be two distinct primes, both of the form $4n + 3$. Then, $x^2 - py^2 = qz^2$ does not have non-trivial solution.

Modulo $p$, one gets $x^2 \equiv qy^2 \pmod{p}$. Since the original equation is homogeneous, by infinite descent one can ignore the case $x = y = 0$; otherwise $q|(x, y, z)$. Then, this congruence equation is the same as $(y^{-1}x)^2 \equiv q \pmod{p}$. Similarly, modulo $q$, one gets $(xz^{-1})^2 \equiv p \pmod{q}$. If the original equation has a non-trivial solution, then both these congruence equations have solutions. However, we know from the reciprocity law, this is impossible. Hence, the original equation does not have non-trivial solution.
Chapter 22

Lecture 22

22.1 More Examples

eg. Solve \(3x^3 - 4x + 5 \equiv 0 \pmod{5^2 \cdot 7^2}\).

First, solve \(3x^3 - 4x + 5 \equiv 3x^3 - 4x = x(3x^2 - 4) \equiv 0 \pmod{5}\). Note that \(3x^2 - 4 \equiv 0 \pmod{5}\) is equivalent to \(x^2 \equiv 3 \pmod{5}\) \((3^{-1} = 2)\) and 3 is a quadratic non-residue mod 5. So, \(x = 0\) is the only root.

Let \(f(x) = 3x^3 - 4x + 5\), \(f'(x) = 9x^2 - 4\). \(f(0) = 5\), \(f'(0) = -4\).

Solve \(-4t \equiv -\frac{5}{5} \pmod{5}\), we get \(t \equiv -1 \pmod{5}\). So, \(x = 20\) is the solution to

\[3x^3 - 4x + 5 \equiv 0 \pmod{5^2}\].

Then, solve \(3x^3 - 4x + 5 \equiv 0 \pmod{7}\). Since 1, 2, 4 are the quadratic residues mod 7, when \(x = 1, 2, 4, x^3 \equiv 1 \pmod{7}\); when \(x = 3, 5, 6, x^3 \equiv -1 \pmod{7}\); these follow from Euler’s Theorem. Hence, among 1, 2, 4, only \(x = 2\) is a solution over \(\mathbb{Z}/7\mathbb{Z}\); among 3, 5, 6, there is no solution.

\(f(2) = 21\), \(f'(2) = 32\).

Solve \(32t \equiv 4t \equiv -\frac{21}{7} \equiv -3 \pmod{7}\), we get \(t \equiv 1 \pmod{5}\).

Hence, \(x = 5\) is the solution to

\[3x^3 - 4x + 5 \equiv 0 \pmod{7^2}\].

125
Eventually, by Chinese Remainder Theorem,

\[
\begin{align*}
  \begin{cases}
    x \equiv 20 \pmod{25} \\
    x \equiv 9 \pmod{49}
  \end{cases}
\end{align*}
\]

leads to solving \( 9 + 49n \equiv 20 \pmod{25} \), i.e. \(-n \equiv 11 \pmod{25} \), i.e. \( n \equiv -11 \pmod{25} \).

Thus, \( x = 9 + 49 \cdot 14 = 695 \) is the only solution over \( \mathbb{Z}/5^2 \cdot 7^2 \mathbb{Z} \).

e.g. Does \( x^4 - 3y^4 = 31z^4 \) have an integer solution other than \((0, 0, 0)\)?

Assume \((x, y, z)\) is a non-trivial solution. Mod out by 3, one gets \( x^4 \equiv 31z^4 \pmod{3} \). By infinite descent argument, one can ignore the case where \( z = x = 0 \); otherwise \( 3 \mid (x, y, z) \) and one can always divide the original solution by 3 to obtain a new solution. So, assume \( z \) is a unit. Then, one gets a quadratic congruence equation:

\[
((xz^{-1})^2)^2 \equiv 31 \pmod{3}
\]

Similarly, mod out by 31 and using the same argument, one gets:

\[
((xy^{-1})^2)^2 \equiv 3 \pmod{31}
\]

Since 3, 31 are both of the form \( 4n + 3 \), quadratic reciprocity law tells you that these two congruence equations cannot both have solutions. Hence, the original equation does not have non-trivial integer solution.

e.g. Compute \((\frac{2}{p})\).

We know that

\[
(2/p) = \begin{cases}
0 & (p = 2) \\
1 & (p = 8k + 1, 8k + 7) \\
-1 & (p = 8k + 3, 8k + 5)
\end{cases}, \quad (3/p) = \begin{cases}
0 & p = 3 \\
1 & p = 12k + 1, 12k + 11 \\
-1 & p = 12k + 5, 12k + 7
\end{cases}
\]

Trivially, \((\frac{2}{p}) = 0\) for \( p = 2, 3 \). Now assume \( p > 6 \).

\((\frac{2}{p}) = 1\) if and only if \((\frac{2}{p}) = (\frac{2}{p}) = (\frac{2}{p}) = (\frac{2}{p}) = -1\).

Solve

\[
\begin{cases}
p \equiv 1 \pmod{8} \\
p \equiv 1 \pmod{12}
\end{cases} \iff \begin{cases}
p \equiv 1 \pmod{8} \\
p \equiv 1 \pmod{3}
\end{cases} \iff p = 24k + 1.
\]

Solve

\[
\begin{cases}
p \equiv 1 \pmod{8} \\
p \equiv 11 \pmod{12}
\end{cases} \iff \begin{cases}
p \equiv 1 \pmod{8} \\
p \equiv 3 \pmod{4}
\end{cases} \iff \text{no solution}.
\]

\[
\begin{cases}
p \equiv 1 \pmod{8} \\
p \equiv 2 \pmod{3}
\end{cases}
\]
Solve
\[
\begin{align*}
p \equiv 7 \pmod{8} & \quad \Leftrightarrow \quad p \equiv 7 \pmod{8} \\
p \equiv 1 \pmod{12} & \quad \Leftrightarrow \quad p \equiv 1 \pmod{4}\end{align*}
\]
no solution.

Solve
\[
\begin{align*}
p \equiv 7 \pmod{8} & \quad \Leftrightarrow \quad p \equiv 7 \pmod{8} \\
p \equiv 11 \pmod{12} & \quad \Leftrightarrow \quad p \equiv 3 \pmod{4}\end{align*}
\]
p = 24k + 23.

Solve
\[
\begin{align*}
p \equiv 3 \pmod{8} & \quad \Leftrightarrow \quad p \equiv 3 \pmod{8} \\
p \equiv 5 \pmod{12} & \quad \Leftrightarrow \quad p \equiv 1 \pmod{4}\end{align*}
\]
no solution.

Solve
\[
\begin{align*}
p \equiv 3 \pmod{8} & \quad \Leftrightarrow \quad p \equiv 3 \pmod{8} \\
p \equiv 7 \pmod{12} & \quad \Leftrightarrow \quad p \equiv 1 \pmod{3}\end{align*}
\]
p = 24k + 19.

Solve
\[
\begin{align*}
p \equiv 5 \pmod{8} & \quad \Leftrightarrow \quad p \equiv 5 \pmod{8} \\
p \equiv 5 \pmod{12} & \quad \Leftrightarrow \quad p \equiv 5 \pmod{3}\end{align*}
\]
p = 24k + 5.

Solve
\[
\begin{align*}
p \equiv 5 \pmod{8} & \quad \Leftrightarrow \quad p \equiv 5 \pmod{8} \\
p \equiv 7 \pmod{12} & \quad \Leftrightarrow \quad p \equiv 3 \pmod{4}\end{align*}
\]
no solution.

In summary,
\[
\left(\frac{6}{p}\right) = \begin{cases} 
0 & \text{if } p = 2, 3 \\
1 & \text{if } p = 24k \pm 1, 24k \pm 5 \\
-1 & \text{if } p = 24k \pm 7, 24k \pm 11.
\end{cases}
\]

**eg.** Which integers \(d\) can be the orders of units in \(\mathbb{Z}/p\mathbb{Z}\). Which one(s) has highest frequency?

We know if \(d|p-1\) then \(\mathbb{Z}/p\mathbb{Z}\) has \(\varphi(d)\) many units of order \(d\). And if \(d\) is an order \(d|p-1\). So the integers that can be the orders of units in \(\mathbb{Z}/p\mathbb{Z}\) are precisely positive factors of \(p-1\).

In your HW, you showed \(m|n\) implies \(\varphi(m)|\varphi(n)\). Thus, we know \(p-1\) has the highest frequency among all such \(d\).
Moreover, \( \varphi(q \cdot n) \geq (q - 1)\varphi(n) \) for all prime numbers \( q \):

1. If \( (q, n) = 1 \), then \( \varphi(q \cdot n) = \varphi(q)\varphi(n) = (q - 1)\varphi(n) \);
2. if \( (q, n) = q \), then \( \varphi(q \cdot n) = (qn)\prod_{p|n}(1 - \frac{1}{p}) = q\varphi(n) \).

Hence, \( d < p - 1 \) and \( \varphi(d) = \varphi(p - 1) \) if and only if \( d = \frac{p - 1}{2} \) and \( d \) is odd.

In summary, let \( d^* \) be the order with highest frequency,

\[
d^* = \begin{cases} 
    p - 1 & \text{if } p = 2, 4k + 1 \\
    \frac{p - 1}{2}, p - 1 & \text{if } p = 4k + 3.
\end{cases}
\]

**eg.** Does \( x^2 + x + 3 \equiv 0 \pmod{13} \) have a solution?

\[
x^2 + x + 3 \equiv (x + 7)^2 + 3 - 7^2 \equiv (x + 7)^2 - 46 \equiv 0 \pmod{13}.
\]

This is the same as solving \( y^2 = 7 \) over \( \mathbb{Z}/7\mathbb{Z} \). But \( \left( \frac{7}{13} \right) = \left( \frac{13}{7} \right) = \left( \frac{6}{7} \right) = -1 \). So the original equation does not have a solution.

**eg.** Let \( p \) be an odd prime and \( \left( \frac{a}{p} \right) = 1 \). Show that \( x^{2p} - a \) has two roots both of multiplicity \( p \).

\[
x^{2p} - a = (x^2)^p - a^p = (x^2 - a)^p.
\]

The first equality follow from the fact that \( a^p = a \) for all \( a \in \mathbb{Z}/p\mathbb{Z} \), which further follows from Fermat’s Little Theorem. The second equality from binomial expansion and the fact that \( p|\binom{p}{i} \) for all \( i \) except for \( i = 0, p \).

Since we are given \( \left( \frac{a}{p} \right) = 1 \) and \( p \) is odd, we know \( x^2 = a \) has two distinct roots over \( \mathbb{Z}/p\mathbb{Z} \). (You know it cannot have repeated roots, for example by seeing \( (x^2 - a, 2x) = 1 \).)

In particular, \( (x^2 - a)^p = ((x - a_1)(x - a_2))^p = (x - a_1)^p(x - a_2)^p \). This proves the claim.
Appendix A

Problem Sets

A.1 Problem Set 1

1. The Goldbach conjecture states: Every even integer greater than 2 can be written as a sum of two prime numbers. For example, 44 = 13 + 31.

(1) Verify the Goldbach conjecture for the following even numbers: 102, 114, 126, 138, 144.

(2) (Triple of primes.) Find all positive integer p such that p, p + 2, p + 4 are prime numbers.

(3) (Twin primes.) Suppose p, p + 2 are both prime numbers, and p > 7. Show that p cannot be written as a sum of two primes.

2. Find the greatest common divisor for the following pairs of integers using Euclidean Algorithm.

(1) 1875 and 105.

(2) 135 and 2001.

(3) −1960 and 555.

3. Let x, y be two non-zero integers. Show that if $x^2 + y^2$ is divisible by 4, then $x + y$ is divisible by 2.

4. (Numbers divisible by 11.)

(1) Show that a number of the form $a_1a_2a_3\ldots a_na_n$ is always divisible by 11. (Note: $a_1a_2a_3\ldots a_na_n$ does NOT mean multiplying a bunch of numbers. It simply means writing down all the digits of an integer.)

(2) More generally, prove that an integer $a_1a_2\ldots a_n$ is divisible by 11 if and only
if the following number is divisible by 11:
\[ N = (\text{Sum of its even-indexed digits}) - (\text{Sum of its odd-indexed digits}). \]

(For example, take 1234, its even-indexed digits (from left to right) are 2, 4, and its odd-indexed digits are 1, 3. Then, \( N = (2 + 4) - (1 + 3) = 2 \), which is not divisible by 11. In fact, 1234 is not divisible by 11. 1234 = 11 \cdot 112 + 2.)

5. Show that if \( a, b, c \) are integers such that \((a, b) = 1\) and \( c | (a + b) \), then \((c, a) = (c, b) = 1\).

### A.2 Problem Set 2

1. Let \( I \) be the set of all integers which can be written as a product of the form \( 2^m \cdot 3^n \cdot 5^\ell \), where \( m, n, \ell \geq 0 \). Compute the sum \( \sum_{n=1}^{\infty} \frac{1}{n^2} \).

2. Show that if \( a \) and \( b \) are positive integers and \( a^3 | b^2 \), then \( a | b \).

3. Show that \( \sqrt[3]{6} \) is irrational using divisibility.

4. Find a pair of positive integers \( a, b \) such that \((a, b) = 18\) and \([a, b] = 540\). How many pairs can you find in all? Why?

5. Suppose all the prime numbers are ordered from small to large:

\[ p_1 = 2, \, p_2 = 3, \, p_3 = 5, \, p_4 = 7, ... \]

In class, we discussed the numbers \( Q_n = p_1 p_2 ... p_n + 1 \), i.e. the product of the first \( n \) primes plus 1. We showed that \( Q_6 \) is not prime. Find the next \( n \) such that \( Q_n \) is not prime, and write down a factorization of that number. (You may find the following web-site helpful: [http://www.mathsisfun.com/prime-factorization-tool.php](http://www.mathsisfun.com/prime-factorization-tool.php). You do not need to justify your answer.)

6. Let \( W_n = n! + 1 \) where \( n \) is a positive integer. Prove that \( W_n \) has a prime factor greater than \( n \). Use this fact to give another proof of the fact that there are infinitely many prime numbers.

7. Write down in detail a proof of the identity \( \sum_{i=1}^{n} i^2 = \frac{n(n + 1)(2n + 1)}{6} \) using mathematical induction.

8. Suppose the smallest prime factor, \( p \), of a positive integer \( n \) satisfies the condition \( p > \sqrt[3]{n} \). Show that \( \frac{n}{p} \) is prime or 1.

9. For the following problems, if the number you claim to be prime is \( \geq 100 \), you must check that it is indeed a prime number.
A.3. **PROBLEM SET 3**

(1) Use Fermat factorization method to factorize 73920 into product of primes. (Hint: first factorize out all the 2’s.)

(2) Use any method you like to factorize 123456 into product of primes.

10. Show that \( n^2 - n + 41 \) is prime for all \( n : 0 \leq n \leq 40 \).

### A.3 Problem Set 3

1. Determine whether the following equations have integer solutions. If so, find the solution set to the following linear Diophantine Equations:

   (1) \( 21x - 14y = 49 \).

   (2) \( 19x + 20y = 1909 \).

   (3) \( 2072x + 1813y = 2849 \).

3. Suppose \((a, b) = 1, a > 0, b > 0\), and consider the equation \( ax + by = c \). Suppose \((x_0, y_0)\) is some integer solution to this equation. Can you give a criterion for the equation to have **positive** integer solution, namely a solution \((x, y)\) such that \( x > 0, y > 0 \)?

4. Suppose an orange costs 1 dollar and a pineapple costs 1.5 dollars. Tom has 130 dollars in all. What are all the possible combinations of orange and pineapples Tom can buy using exactly 130 dollars. (Note: one cannot buy negative amount of oranges or pineapples!)

5. Find the set of all integer solutions to the equation \( x + 2y + 3z = 5 \).

A.4 **Problem Set 4**

1. Write down the arithmetic tables of \( \mathbb{Z}/b\mathbb{Z} \) for (1) \( b = 6 \); (2) \( b = 7 \); (3) \( b = 12 \).

2. (1) Check that the multiplication of remainders modulo \( b \) (as defined in class) is indeed associative.

   (2) Check that the addition and multiplication of remainders modulo \( b \) (as defined in class) satisfy the two distributive laws.

3. Consider \( x^2 - y^2 = 2 \) as an equation over \( \mathbb{Z}/4\mathbb{Z} \). Show that it has no solution over \( \mathbb{Z}/4\mathbb{Z} \). Conclude that if \( n \) is an integer such that \( n \equiv 2 \) (mod 4), then \( n \) cannot be written in the form \( n = a^2 - b^2 \).
4. Show that $a_1^p + ... + a_n^p \equiv (a_1 + ... + a_n)^p \pmod{p}$ (Hint: Use the Multinomial Theorem.)

5. How many elements in $\mathbb{Z}/10\mathbb{Z}$ have multiplicative inverses? Find all of them.

6. Denote $\varphi(n)$ to be the Euler-\(\varphi\) function. Is $\varphi(n)$ an increasing function with respect to $n$? Why or why not? (Hint: try out for small $n$’s.)

7. In class, we defined a space whose points correspond to either $(0)$ or $(p)$, where $p$ is a prime number. Consider the following numbers as “functions” over this space, and determine their values at the points $(3)$, $(5)$, $(7)$; write “not defined” where necessary: (1) 1024; (2) $2^4 \cdot 121$; (3) $124 \cdot 35$; (4) $10 \cdot 2013$.

8. In class, we mentioned that if one knows the remainder of $n$ modulo 2, and the remainder of $n$ modulo 3, then one knows the remainder of $n$ modulo 6. Prove the following WITHOUT using Chinese Remainder Theorem: if $a, b > 0$ and $(a, b) = 1$, the remainder of $n$ modulo $a$ together with the remainder of $n$ modulo $b$ determine the remainder of $n$ modulo $a \cdot b$. We break the problem into the following two steps:

(1) Show that given any integer $r_1$ between 0 and $a - 1$, and any integer $r_2$ between 0 and $b - 1$, there is always some integer $n$ such that $n \equiv r_1 \pmod{a}$, and $n \equiv r_2 \pmod{b}$. You should only use results we have talked about.

(2) Show that if $n, n'$ both satisfy the conditions in part (1), then they differ by a multiple of $a \cdot b$.

A.5 Problem Set 5

1. We have discussed in class that $\varphi(mn) = \varphi(n)\varphi(m)$ holds when $(n, m) = 1$. Please prove a more general formula: if $(m, n) = d$, then $\varphi(mn) = \frac{\varphi(m)\varphi(n)}{\varphi(d)}$. (Hint: Write $m = \prod p_i^{\alpha_i}, n = \prod p_i^{\beta_i}$, and write $(m, n)$ correspondingly. Eventually, apply the formula for $\varphi$, and do some arithmetic.)

2. Prove that if $m|n$, then $\varphi(m)|\varphi(n)$.

3. (1) Let $p > 0$ be a prime number and $n$ be any positive integer. Show that $p^n - p^{n-1} \geq \sqrt{p^n}$ holds either when $p \geq 3$ or when $n > 1$.

(2) Conclude from part (1) that $\varphi(n) \geq \sqrt{n}$ when $n > 6$, and hence $\varphi(n) \to \infty$ when $n \to \infty$.

4. Show that there are infinitely many prime numbers of the form $5n + 1$.

5. Show that there are infinitely many prime numbers of the form $6n + 5$.

6. Some computation.
A.6. PROBLEM SET 6

(1) Compute \( \varphi(1023) \).
(2) Compute \( \varphi(7!) \).
(3) Write \( \varphi(2n) \) for arbitrary \( n > 0 \) as an expression involving \( \varphi(n) \); the result depends on the parity of \( n \).

A.6 Problem Set 6

1. Solve the following linear congruence equations.

(1) \( 4x \equiv 26 \pmod{33} \).

(2) \( 6x \equiv 12 \pmod{42} \).

(3) \( 4x \equiv 10 \pmod{13} \).

(4) \( 5x \equiv 48 \pmod{14} \).

2. Solve the following system of linear congruence equations.

\[
\begin{align*}
(1) & \quad \begin{cases} 
  x \equiv 3 \pmod{4} \\
  x \equiv 5 \pmod{21} \\
  x \equiv 7 \pmod{25} \\
  x \equiv 2 \pmod{5} 
\end{cases} \\
(2) & \quad \begin{cases} 
  x \equiv 7 \pmod{11} \\
  x \equiv 3 \pmod{24} 
\end{cases}
\end{align*}
\]

3. Consider

\[
\begin{align*}
\begin{cases} 
  x \equiv 1 \pmod{6} \\
  x \equiv 2 \pmod{10} 
\end{cases}
\end{align*}
\]

Show that this system of equations does not have a solution. Does this violate the Chinese remainder theorem? Why or why not?

4. Solve the system of linear congruence equations:

\[
\begin{align*}
\begin{cases} 
  2x \equiv 1 \pmod{5} \\
  4x \equiv 7 \pmod{19} \\
  2x \equiv 12 \pmod{22} 
\end{cases}
\end{align*}
\]

5. Solve the system of linear congruence equations:

\[
\begin{align*}
\begin{cases} 
  x + 2y \equiv 1 \pmod{7} \\
  4x - 3y \equiv 5 \pmod{7} 
\end{cases}
\end{align*}
\]
6. Find all units of \( \mathbb{Z}/k\mathbb{Z} \) and their orders, for the following \( k \)'s:
   (1) \( k = 7 \); (2) \( k = 8 \); (3) \( k = 9 \); (4) \( k = 12 \).

A.7 Problem Set 7

1. Prove that for \( 1 < k < p - 1 \), \((p - k)!(k - 1)! \equiv (-1)^k \pmod{p}\).

2. Let \( p, q \) be two distinct primes. Show that \( p^q - 1 + q^{p - 1} \equiv 1 \pmod{p \cdot q} \).

3. Factorize \( x^9 - x = x(x - 1) \ldots \) into a product of irreducible polynomials over \( \mathbb{Z}/3\mathbb{Z} \). Find all degree 2 irreducible polynomials over \( \mathbb{Z}/3\mathbb{Z} \). Please check that they are indeed irreducible.

4. Use Euclidean algorithm for \( (\mathbb{Z}/7\mathbb{Z})[x] \) to find the remainder of \( x^4 + 50x^2 + 30x - 10 \) modulo \( x^2 + 1 \).

5. Find the g.c.d. of \( x^3 + x^2 + x + 1 \) and \( x^7 - x \) over \( \mathbb{Z}/7\mathbb{Z} \). Is \( x^3 + x^2 + x + 1 \) an irreducible polynomial over \( \mathbb{Z}/7\mathbb{Z} \)? Justify.

6. Show that \( \mathbb{Z}/p\mathbb{Z} \) has a unit of order 4 if and only if \( p \) is of the form \( 4n + 1 \). (Notice that this time you can show both directions.)

7. Let \( p = 4n + 1 \) be a prime number. From Problem 6, We know that \( \mathbb{Z}/p\mathbb{Z} \) has units of order 4. Find all units of order 4 of \( \mathbb{Z}/p\mathbb{Z} \) by solving \( x^2 + 1 = 0 \) over \( \mathbb{Z}/p\mathbb{Z} \). (Hint: Use Wilson’s Theorem.)

8. Let \( p \) be a prime number. Show that \( x^p - x - 1 \) does not have a root over \( \mathbb{Z}/p\mathbb{Z} \).

A.8 Problem Set 8

1. Prove the following result: If \( f(x) \) is an integer coefficient polynomial, not constant, then there are always infinitely many prime numbers \( p \) such that \( f(x) \equiv 0 \pmod{p} \) has a solution.

   (Hint: Suppose \( f(x) = a_0 + a_1x + \ldots + a_nx^n \). If \( a_0 = 0 \), argue that there is nothing to prove; if \( a_0 \neq 0 \), then \( f(a_0x) = a_0 + a_1a_0x + \ldots + a_n(a_0x)^n = a_0(1 + x(a_1 + a_2a_0x + \ldots + a_n(a_0^{-1}x^{n-1}))). \)

   **Remark** This shows that the solvability of a congruence equation over \( \mathbb{Z}/p\mathbb{Z} \) is a much weaker condition than the solvability of the corresponding diophantine equation.

2.(1) Does the equation \( x^2 - 18xy + 81y^2 = 6z^2 \) have integer solution other than
(x, y, z) = (0, 0, 0)? Why or why not?

(2) How about the equation \( x^3 - 26y^3 = 24z^3 \)?

(3) And the equation \( x^2 - 3y^2 = 7z^2 \)?

3. Solve the following congruence equations, or show there are no solutions.
(1) \( x^3 - 2x + 3 \equiv 0 \pmod{27} \).

(2) \( x^3 - 2x + 4 \equiv 0 \pmod{125} \).

(3) \( x^5 + 7 \equiv 0 \pmod{25} \).

4. Show that \( x = 7 \) is a repeated root to \( x^4 - x^3 - 2x^2 + 2x - 5 = 0 \) over \( \mathbb{Z}/11\mathbb{Z} \). You do NOT need to compute the multiplicity of this root.

A.9 Problem Set 9

1. Evaluate the following values of Legendre symbols:
(1) \( \left( \frac{3}{53} \right) \); (2) \( \left( \frac{7}{79} \right) \); (3) \( \left( \frac{29}{31} \right) \); (4) \( \left( \frac{31}{641} \right) \).

2. Evaluate the Legendre symbols \( \left( \frac{503}{773} \right) \) and \( \left( \frac{501}{773} \right) \).

3. Find a criterion for the primes \( p \) such that \( \left( \frac{5}{p} \right) = 1 \).

4. Let \( p > 3 \) be a prime number. Show that \( x^2 \equiv -3 \pmod{p} \) is solvable iff \( p \equiv 1 \pmod{6} \).

5. Prove the following: if \( p = 4k + 1 \) and \( d|k \), then \( (d/p) = 1 \).

6. Let \( p \) be a prime number. We call a unit \( a \) in \( \mathbb{Z}/p\mathbb{Z} \) a primitive root, if \( \text{ord}_p(a) = p - 1 \), i.e. any unit in \( \mathbb{Z}/p\mathbb{Z} \) can be written as some power of \( a \). If \( p \) is of the form \( 2^n + 1 \), prove that the primitive roots in \( \mathbb{Z}/p\mathbb{Z} \) are precisely the quadratic non-residues modulo \( p \). If \( n > 1 \), prove \( 3 \) is always a primitive root.

7. Let \( p \) be an odd prime, and \( (a,p) = 1 \). Show that if \( x^2 \equiv a \pmod{p} \) has solutions, then \( x^2 \equiv a \pmod{p^N} \) always has solutions, for any \( N > 1 \).

8. Does \( x^2 + x + 1 \equiv 0 \pmod{997} \) have solutions? Why or why not?
Appendix B

Extra Credit Problem Sets

B.1 Problem Set 1

1. In this assignment, try to solve Pell’s equation for \( n = 2 \): \( x^2 - 2y^2 = 1 \).

(1) First, find a “smallest” non-trivial solution in the following way: start from \( y = 1 \), find the smallest positive integer \( y_0 \) such that \( 2y_0^2 + 1 \) is a complete square. Then, easy to see that \((\sqrt{2y_0^2 + 1}, y_0)\) is a solution. (You can think of \((1,0), (-1,0)\) as trivial solutions.)

(2.1) Define an operation “\( \circ \)”: \( (x_1, y_1) \circ (x_2, y_2) = (x_1x_2 + 2y_1y_2, x_1y_2 + x_2y_1) \). Explain why this is a natural definition. (Hint: what is the product of \( x_1 + \sqrt{2}y_1 \) and \( x_2 + \sqrt{2}y_2 \), under usual multiplication of real numbers?)

Hereafter, we denote \((x, y)^m = (x, y) \circ \ldots \circ (x, y)\) for \( m > 0 \). When \( m < 0 \), and \((x, y)\) is an integer solution to the equation, we denote \((x, y)^m = (x, y)^{-1} \circ \ldots \circ (x, y)^{-1}\), for \( m \) copies, where \((x, y)^{-1} = (x, -y)\). Define \((x, y)^0 = (1, 0)\).

(2.2) Explain why \((x, y)^{-1} = (x, -y)\) is a natural definition. (Hint: If \((x, y)\) is a solution, what is \(\frac{1}{x + \sqrt{2}y}\)?)

(3) Show that if \((x, y)\) is an integer solution to the equation, so are \(\pm (x, y)^m\), for any \( m \), where \(- (a, b)\) is just the pair \((-a, -b)\). (Hint: if \((x + \sqrt{2}y)(x - \sqrt{2}y) = 1\), then \((x + \sqrt{2}y)^m(x - \sqrt{2}y)^m = 1\). Similarly, if \((x, y)\) and \((x', y')\) are solutions, so is \((x, y) \circ (x', y')\).
Now, let \((x_0, y_0)\) be the solution you found in part (1), show that all solutions are of the form \(\pm (x_0, y_0)^m\) in the following way:

(4.1) Show that if there is a solution not of the given form, then there is a pair of positive integers \((x', y')\) which is a solution, such that \((x_0 + \sqrt{2}y_0)^M < x' + \sqrt{2}y' < (x_0 + \sqrt{2}y_0)^{M+1}\), for some \(M \geq 0\).

(4.2) Show that if \((x, y)\) and \((x', y')\) are both solutions, and \(x, x', y, y'\) are all positive, then \(y \neq y'\). In particular, either \(x > x', y > y'\), or \(x < x', y < y'\).

(4.3) Eventually, denote \((x'', y'') = (x', y') \circ (x_0, y_0)^{-M}\). Show that \(x'' + \sqrt{2}y'' < x_0 + \sqrt{2}y_0\). Argue that this contradicts part (1).

5. Eventually, consider Pell’s equation in general: \(x^2 - ny^2 = 1\) (\(n\) is not a perfect square). Guess a general pattern for solutions of the equation.

B.2 Problem Set 2

1. In this assignment, you are asked to give a proof of the fact that \(n^2 - n + 41\) is a prime number for \(n = 0, \ldots, 40\). Throughout this assignment, we only allow variables of polynomials to take INTEGER values.

1.1 Let \(f(x, y)\) be a polynomial of the form \(ax^2 + bxy + cy^2\). Define the discriminant to be \(\Delta(f) := b^2 - 4ac\). Let \(g(x, y) = x^2 + xy + 41y^2\) and find \(\Delta(g)\).

1.2 Suppose for some \(n : 0 \leq n \leq 40, n^2 - n + 41\) is not prime. Show that \(n^2 - n + 41\) has a prime factor \(q < 41\).

1.3 Suppose \(n^2 - n + 41 = q \cdot q', \) where \(q < 41\) is a prime number. Define \(h(x, y) = qx^2 + (2n - 1)xy + q'y^2\).

Show that \(\Delta(h) = \Delta(g)\).

1.4 (Key Observation) Show that any quadratic polynomial \(f(x, y)\) of the form \(ax^2 + bxy + cy^2\) such that \(a > 0\) and \(\Delta(f) = b^2 - 4ac < 0\) can be transformed into some \(s(u, v) = a'u^2 + b'uv + c'v^2\) such that \(0 \leq b' \leq a' \leq c'\) using an invertible integral linear change of variables.

Definition 33. An integral linear change of variables from \((x, y)\) to \((u, v)\) is given by

\[
\begin{align*}
x &= a_1u + a_2v \\
y &= a_3u + a_4v
\end{align*}
\]

\((a_1, a_2, a_3, a_4 \in \mathbb{Z})\).
It is called invertible if \( u, v \) can also be written in terms of integer linear combinations of \( x, y \):

\[
\begin{align*}
  u &= b_1 x + b_2 y \\
  v &= b_3 x + b_4 y \\
( & b_1, b_2, b_3, b_4 \in \mathbb{Z})
\end{align*}
\]

We break the proof into multiple steps.

1.4.1 First, by using the change of variables \( \begin{align*} x &= u \\
  y &= -v \end{align*} \) if necessary, show that one can always assume \( b \geq 0 \) in \( ax^2 + bxy + cy^2 \).

1.4.2 Secondly, by using the change of variables \( \begin{align*} x &= v \\
  y &= u \end{align*} \) if necessary, show that one can always assume \( a \leq c \).

1.4.3 Thirdly, by using the change of variables \( \begin{align*} x &= u - kv \\
  y &= v \end{align*} \) for some appropriate choice of \( k \) if necessary, show that one can always transform \( ax^2 + bxy + cy^2 \) into \( a'u^2 + b'uv + c'v^2 \), with \( |b'| \leq a' \).

1.4.4 Show that by applying changes of variables of the above types (possibly more than once), one can transform \( ax^2 + bxy + cy^2 \) such that \( a > 0 \) and \( b^2 - 4ac < 0 \), into \( a'u^2 + b'uv + c'v^2 \) such that \( 0 \leq b' \leq a' \leq c' \).

1.4.5 Show that all these changes of variables are indeed invertible integral linear change of variables, as defined above.

1.4.6 Show that these changes of variables all preserve discriminants.

1.5 Suppose \( a'u^2 + b'uv + c'v^2 \) is obtained from \( ax^2 + bxy + cy^2 \) using any invertible integral linear change of variables. Prove that they have the same range.

1.6 Let \( p(x, y) \) be any polynomial of the form \( ax^2 + bxy + cy^2 \). Show that if \( \Delta(p) = \Delta(g) \) (as defined in 1.1) and \( 0 \leq b \leq a \leq c \), then \( a = b = 1, c = 41 \).

1.7 Suppose \( n^2 - n + 41 = q \cdot q' \), where \( q : 1 < q < 41 \). Show that \( h(x, y) \) (as defined in 1.3) must have the same range as \( g(u, v) \) (as defined in 1.1).

1.8 Argue that the result in 1.7 leads to a contradiction by showing that \( q \) is in the range of \( h \) but not in the range of \( g \).
B.3 Problem Set 3

In this assignment, you are asked to show a few more results related to polynomials over \(\mathbb{Z}/p\mathbb{Z}\), where \(p\) is prime.

1. Show that \(x^p - x - 1\) is an irreducible polynomial over \(\mathbb{Z}/p\mathbb{Z}\).
   (1) Show that for any \(f,g,h \in (\mathbb{Z}/p\mathbb{Z})[x]\),
   \[(f(x) + g(x) + h(x))^p = f^p(x) + g^p(x) + h^p(x).\]
   (2) Show that \(x^{p^p} - x = (x^p - x - 1)^{p^{p-1}} + (x^p - x - 1)^{p^{p-2}} + \ldots + (x^p - x - 1)\) and conclude \(x^p - x - 1 \mid x^{p^p} - x\).
   (3) Conclude from (2) that \(x^p - x - 1\) is irreducible. You also need the fact that \(x^p - x - 1\) has no roots over \(\mathbb{Z}/p\mathbb{Z}\).

2. Let \(f,g\) be polynomials with coefficient in \(\mathbb{Z}/p\mathbb{Z}\), where \(p\) is prime. Write out carefully a proof of the fact that the g.c.d \((f, g)\) can be written as \(a(x)f(x) + b(x)g(x)\), where \(a(x), b(x)\) are some polynomials with coefficient in \(\mathbb{Z}/p\mathbb{Z}\).

3. Show that if \(d|n\), then \(x^d - x \mid x^n - x\). Conclude from here that all degree \(d\) irreducible polynomials over \(\mathbb{Z}/p\mathbb{Z}\) such that \(d|n\) are factors of \(x^n - x\).

4. (Up to a constant multiple) how many degree 2 irreducible polynomials are there over \(\mathbb{Z}/p\mathbb{Z}\)? Degree 3?

5. For which \(n\) among \(1, \ldots, p\) does \(x^n + x^{n-1} + \ldots + x + 1\) have a solution over \(\mathbb{Z}/p\mathbb{Z}\), where \(p\) is prime?

B.4 Problem Set 4

In this assignment, you are asked to prove the statement an odd prime number \(p\) can be written as a sum of two perfect squares \(a^2 + b^2\) if and only if \(p = 4n + 1\) for some \(n\). Along the way, you are also asked to explore the Gaussian integers.

1. Show that if \(p = a^2 + b^2\) and is an odd prime number, then \(p \equiv 1(\text{mod } 4)\).

2. To get the other direction, we utilize factorization in \(\mathbb{Z}[i] = \{a + bi | a, b \in \mathbb{Z}\}\), where \(i = \sqrt{-1}\). We define addition and multiplication on \(\mathbb{Z}[i]\) as follows (this definition agrees with usual addition and multiplication of complex numbers, but we won’t need this anywhere):

\[
\begin{align*}
(a + bi) + (c + di) &= (a + c) + (b + d)i \\
(a + bi) \cdot (c + di) &= (ac - bd) + (ad + bc)i
\end{align*}
\]
We take it as given that $\mathbb{Z}[i]$ is a commutative ring. In particular, it is closed under multiplication and addition.

2.1 For any number $a + bi \in \mathbb{Z}[i]$, define its norm to be $N(a + bi) := a^2 + b^2$. Show that $N(a + bi) \cdot N(c + di) = N((a + bi)(c + di))$. (Hint: $N(a + bi) = (a + bi)(a - bi)$.)

2.2 Show that $N(a + bi) = 1$ if and only if $a + bi$ is one of the following: $\pm 1, \pm i$. Show that these are the only units of $\mathbb{Z}[\omega]$. (Hint: how many choices of $a, b$ do you have to make $a^2 + b^2 = 1$?)

**Definition 34.** Suppose $a + bi \in \mathbb{Z}[i]$. It is called prime if its only factors are the units or a unit times itself.

2.3 Show that any element in $\mathbb{Z}[i]$ which is not a unit has a prime factorization. (Hint: you should mimic a similar result for integers and carefully apply the norm function.)

2.4 Show that the prime factorization is unique. (Hint: you should mimic the proof of Fundamental Theorem of Arithmetic and carefully apply the norm function.)

2.5 Show that if $a + bi$ is prime in $\mathbb{Z}[i]$ and $a + bi|(c + di)(c + fi)$, then $a + bi$ must divide one of $c + di, c + fi$.

2.6 Show that if $p = 4n + 1 \in \mathbb{Z}$ is not a prime in $\mathbb{Z}[i]$, then $p$ can be written as a sum of two perfect squares.

2.7 Given a prime integer $p = 4n + 1$, show that $p|((2n)!)^2 + 1$. (Hint: Wilson’s Theorem)

2.8 Show that $p$ does not divide $(2n)! + i$ or $(2n)! - i$. Hence, $p$ is not prime in $\mathbb{Z}[i]$.

**B.5 Problem Set 5**

Fermat’s Last Theorem states that there is no non-trivial solution to the equation $X^n + Y^n = Z^n$ when $n \geq 3$. We call a solution $(x, y, z)$ non-trivial, if $xyz \neq 0$.

In this assignment, we shall see that even a special case of Fermat’s Last Theorem may invoke interesting thoughts in number theory.

1. Denote $\omega = \frac{1}{2}(-1 + \sqrt{-3})$. Define $\mathbb{Z}[\omega] = \{a + b\omega | a, b \in \mathbb{Z}\}$. We define addition and multiplication as follows (this definition agrees with usual addition.
and multiplication of complex numbers, but we won’t need this anywhere):

\[(a + b\omega) + (c + d\omega) = (a + c) + (b + d)\omega\]
\[(a + b\omega) \cdot (c + d\omega) = (ac - bd) + (ad + bc - bd)\omega\]

We take it as given that \(\mathbb{Z}[\omega]\) is a commutative ring. In particular, it is closed under multiplication and addition.

We develop the theory of prime factorization in \(\mathbb{Z}[\omega]\) in the following steps:

1.1 For any number \(a+b\omega \in \mathbb{Z}[\omega]\), define its norm to be \(N(a+b\omega) := a^2 - ab + b^2\).
Show that \(N(a+b\omega) \cdot N(c+d\omega) = N((a+b\omega)(c+d\omega))\).
(Hint: denote \(\omega = \frac{1}{2}(-1 - \sqrt{-3})\), then \(N(a+b\omega) = (a+b\omega)(a+b\overline{\omega})\).

1.2 Check that \(\omega = \omega^2\), so it is also an element in \(\mathbb{Z}[\omega]\).

1.3 Show that \(N(a+b\omega) = 1\) if and only if \(a+b\omega\) is one of the following: \(\pm 1, \pm \omega, \pm \overline{\omega}\). Show that these are the only units of \(\mathbb{Z}[\omega]\).
(Hint: how many choices of \(a, b\) do you have to make \(a^2 - ab + b^2 = 1\)?)

**Definition 35.** Suppose \(a+b\omega \in \mathbb{Z}[\omega]\). It is called **prime** if its only factors are the units or a unit times itself.

1.4 Show that any element in \(\mathbb{Z}[\omega]\) which is not a unit has a prime factorization.
(Hint: you should mimic a similar result for integers and carefully apply the norm function.)

1.5 Show that the prime factorization is unique.
(Hint: you should mimic the proof of Fundamental Theorem of Arithmetic and carefully apply the norm function.)

1.6 Show that if \(a+b\omega\) is in one of the following two cases:

1. \(a+b\omega = p\) is a prime integer of the form \(3n + 2\),
2. \(N(a+b\omega)\) is 3 or a prime of the form \(3n + 1\);
then, \(a+b\omega\) is prime in \(\mathbb{Z}[\omega]\) in the sense of Definition 35.

1.7 Show that all primes of \(\mathbb{Z}[\omega]\) are of the form \((a+b\omega) \cdot (c+d\omega)\), where \(a+b\omega\) satisfies one of the conditions in 1.7 and \((c+d\omega)\) is a unit.

We now prove by contradiction that the equation does not admit non-trivial solutions.
2. Suppose $X^3 + Y^3 = Z^3$. Show that without loss of generality one can assume $X, Y, Z$ are pair-wise co-prime.

3. Show that without loss of generality one can assume $X, Y$ are odd and $Z$ is even.

4. Using 2, 3, show that one can find $a, b$ such that $(a, b) = 1$ and $X = a - b, Y = a + b$. Then, re-write the original equation as

$$2a(a^2 + 3b^2) = Z^3 \quad \text{(B.1)}$$

5. Right-hand side of (1) factorizes in $\mathbb{Z}$ as $2^{3n}p_1^{3m_1}...p_r^{3m_r}$, where $p_1, ..., p_r$ are distinct odd primes. Now factorize the left-hand side of (1) as $2a(a + \sqrt{-3}b)(a - \sqrt{-3}b)$ and study the following two cases:

5.1 Suppose $a + \sqrt{-3}b$ is a perfect cube in $\mathbb{Z}[\omega]$. Show that $2a, a - \sqrt{-3}b$ are also perfect cubes in $\mathbb{Z}[\omega]$. Conclude that in this case $a^2 + 3b^2$ and $2a$ are perfect cubes in $\mathbb{Z}$. Produce a “smaller” solution $(X', Y', Z')$ using these facts. (You may interpret “small” as $|X'Y'Z'| < |XYZ|.$)

5.2 Suppose $a + \sqrt{-3}b$ is not a perfect cube in $\mathbb{Z}[\omega]$. Show that $a^2 + 3b^2$ cannot be a perfect cube in $\mathbb{Z}$. Moreover, show that $(2a, a^2 + 3b^2) = (2a, a^2 + b^2, c^2)$ for some $c \in \mathbb{Z}$ such that $b^2 + 3c^2$ is a perfect cube in $\mathbb{Z}$. Then, by a same argument as in 5.1, one can produce a “smaller” solution.

6. Show that in either case one can keep on producing smaller solutions to the original equation. Argue carefully that this is impossible and hence Fermat's Last Theorem holds for $n = 3$.

Remark This gives you another example demonstrating the idea of enlarging one's scope from $\mathbb{Z}$ to some larger set of numbers. Notice the importance of the fact that elements in $\mathbb{Z}[\omega]$ factorizes in a unique way into primes.