Computational Number Theory

Lecture Notes

N.R.Aravind

Acknowledgements

The latex source file was built using the LegrandOrangeBook template (copyright 2022, Goro Akechi) available at BOOK-WEBSITE.COM and licensed under the Creative Commons Attribution-NonCommercial 4.0 License (the "License"). A copy of the License is available at https://creativecommons.org/licenses/by-nc-sa/4.0. Minor modifications were made in the use of the template.

Contents

II Quadratic Equations in Two Variables

Preface

The purpose of these notes is to present elementary algorithms in number theory from the point of view of solving polynomial equations - primarily over $\mathbb{Z}$ and over $\mathbb{Z}_p$ (the ring of integers modulo p with p prime). The simplest case, namely that of factorizing polynomials in one variable, already uses non-trivial ideas.

The two-variable case is non-trivial even when the degree is restricted to two. We will see (only) some of the classical examples and basic theory of binary quadratic forms. The fundamental problems of primality testing and integer factoring are included in this part as they concern the equation $xy = n$.

What about the multivariate cases? Linear equations are solvable efficiently, whether over $\mathbb{Z}$ or $\mathbb{Z}_n$. Factoring multivariate polynomials can also be done efficiently, i.e. in (randomized) polynomial time, over $\mathbb{Z}_p$. This requires some understanding of finite fields, so we shall study finite fields as well as one or two applications.

Over integers, the picture is very different. The problem of deciding if an arbitrary polynomial in any number of variables has an integer solution - mentioned by Hilbert among his 23 problems for the twentieth century, was famously shown to be undecidable by Matiyasevich in 1970 following a series of work by Julia Robinson, Martin Davis and Hilary Putnam.

Polynomials in One Variable

1. Two Equations from Ancient Times

1.1 The Cubic Equation

Solutions to linear and even quadratic equations have been known from a very long time. In 1800 BC, Egyptians solved quadratic equations by the "method of false position", i.e. by finding successively smaller intervals containing a root. From a clay tablet dated between 1800 BC to 1600 BC, we know that Babylonians of the period knew how to solve quadratic equations exactly.

A natural follow-up to the quadratic equation is: what about cubic equations? Solving the simplest cubic equation $x^3 = a$, boils down to finding cube-roots, and finding cube-roots numerically was also known for a long time; for example, Aryabhatta (around AD 500),gave a method for finding both square-roots and cube-roots.

What about general cubic equations? This was first studied by Omar Khayyam (AD 1100), where he gave a geometric solution.

1.1.1 Geometric solution

Consider the intersection of the parabola $y = \dfrac{x^2}{a}$ and the circle $(x-r)^2 + y^2 = r^2$. An intersection point satisfies $\dfrac{x^4}{a^2} + x^2 - 2rx = 0$, i.e. $x^3 + a^2x = 2ra^2$. Thus we get a solution to the cubic equation $x^3 + px = q$ when $p \geq 0$. In Omar Khayyam's time, negative numbers were avoided and thus the same equation above with a negative value for p would instead be written as $x^3 = px + q$ with p positive. For example, to solve the equation $x^3 + x = 12$, we intersect $y = x^2$ with $(x-6)^2 + y^2 = 36$. This is illustrated in Figure 1.1.1 shown below.

Omar Khayyam divided the cubic equation into various categories so that the coefficients would be positive, and gave different geometric solutions for them.

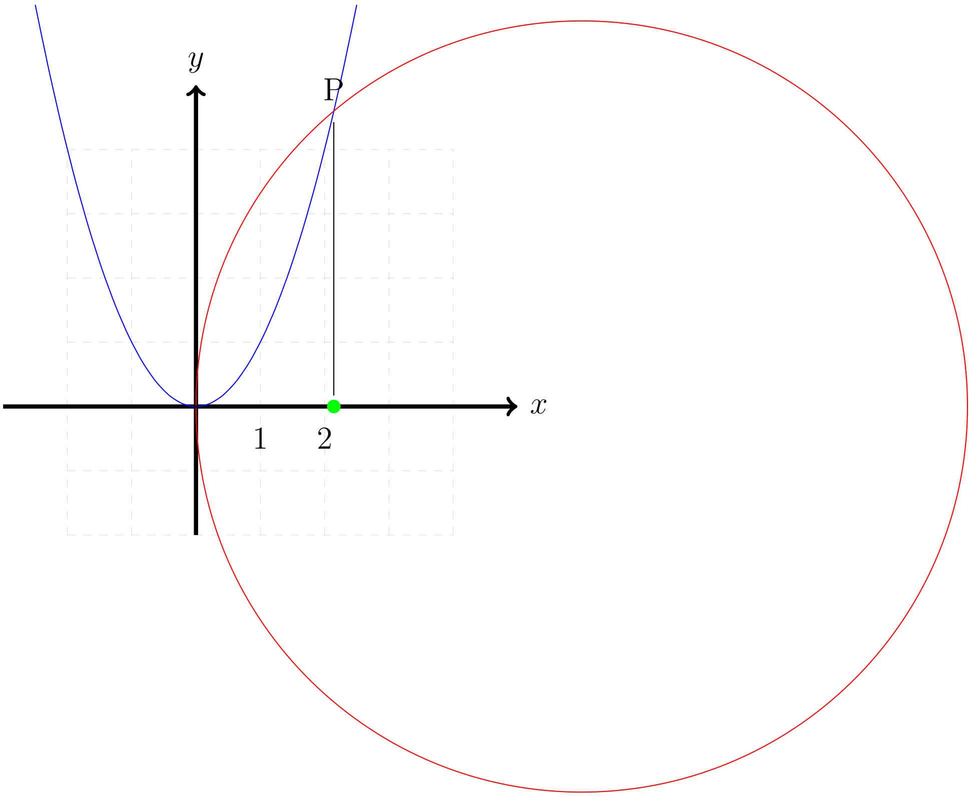

Figure 1.1: Illustration of the geometric solution of $x^3 + x = 12$

1.1.2 Algebraic solution and depressed cubics

While a geometric solution is useful, does the cubic equation have a "closed-form" expression for its solutions in terms of its coefficients, like the quadratic equation does? The answer is Yes and here is a solution. A solution to $x^3 + px + q = 0$ is given by:

$$x = \sqrt[3]{\frac{-q}{2} + \sqrt{D}} + \sqrt[3]{\frac{-q}{2} - \sqrt{D}}. \tag{1.1}$$

where $D = \frac{q^2}{4} - \frac{p^3}{27}$.

But how do we obtain this? And what about the more general cubic with a non-zero x^2 term? To answer the latter question first, it turns out that we can reduce the solution of any cubic polynomial to a solution of a cubic without the x^2 term: such a cubic polynomial is called a depressed cubic.

Consider the polynomial $f(x) = x^3 + ax^2 + bx + c$. Substitute $x = y + r$ to obtain

$$f(x) = y^3 + (3r + a)y^2 + (3r^2 + 2ar + b)y + (r^3 + ar^2 + br + c).$$

If we choose $r = -a/3$, then we obtain a depressed cubic in y, let's call it $g(y)$. Thus to solve $f(x) = 0$, we can solve $g(y) = 0$ and then find the corresponding roots of f.

History

Around AD 1500, Scipione del Ferro, professor at the University of Bologna, discovered a formula for depressed cubic equations (cubic equations with a missing x^2 term) and shortly before his death in 1526, communicated it to his disciple, Antonio del Fiore. After del Ferro's death, around 1535, Fiore issued a challenge to another mathematician Tartaglia, with a list of 30 problems all of whose solutions depended on knowing how to solve the cubic equations $x^3+px=q$ and $x^3=px+q$.

Interestingly, Tartaglia had five years earlier, independently figured out solutions to $x^3+ax^2=b$ and $x^3=ax^2+b$. After accepting the challenge he figured out shortly how to solve the other kind of cubic equation and thus solved all the 30 problems posed by Fiore; he himself posed both kinds of equations in the counter-challenge which Fiore could not solve, and hence won the duel. Here are two of the equations that Tartaglia solved in the duel: $x^3+x=12$ and $x^3+3x=15$.

After the duel, Tartaglia was approached by Gerolamo Cardano, to share his secret as Cardano was writing a book on arithmetic etc. [Incidentally, Cardano was also the first to write a book on probability although he made many mistakes in it.] Initially, Tartaglia refused, but later shared his formula by means of a poem on the promise that Cardano would keep it secret.

Cardano kept his secret for some time, but a few things changed his mind. Firstly, he himself figured out how to reduce the most general equation to a depressed cubic; secondly his student Ferrari figured out how to solve the biquadratic equation, i.e. an equation of degree four. Thirdly, he visited del Ferro's house and examined his manuscripts and was convinced that del Ferro was the original discoverer of the solution to the cubic. He published the solutions in his new book Ars Magna, which led to a fallout between Cardano and Tartaglia.

1.1.3 Solving a depressed cubic

We consider the polynomial $f(x)=x^3+px+q$. Suppose we write $x=\sqrt[3]{u}+\sqrt[3]{v}$. Then

$$x^3=u+v+3\sqrt[3]{uv}x. \tag{1.2}$$

Now we notice that if $u+v=-q$ and $3\sqrt[3]{uv}=-p$, then $x=\sqrt[3]{u}+\sqrt[3]{v}$ satisfies $f(x)=0$. Clearly we can find such a pair u,v as roots of the quadratic polynomial $z^2+qz-\frac{p^3}{27}$. Thus we find $u=\frac{-q}{2}+\sqrt{D}, v=\frac{-q}{2}-\sqrt{D}$, where $D=\frac{q^2}{4}-\frac{p^3}{27}$. The corresponding root is: $x=\sqrt[3]{\frac{-q}{2}+\sqrt{D}}+\sqrt[3]{\frac{-q}{2}-\sqrt{D}}$.

Exercise 1.1 Find a cube-root of the polynomial x^3+x^2-10. ■

1.1.4 Quartic and higher powers: More history

Quartic (fourth-degree) polynomials also have a closed-form expression, found by Ferrari (as mentioned in the history), and naturally this led to the question of a formula for polynomials of degree five and higher. By formula, we mean an expression using the four standard arithmetic operations plus the operation of taking nth roots.

No such formula was however found despite many attempts and around 1800, Ruffini (and later Abel), proved that no general formula can exist. This does not however rule out individual polynomials having roots expressed in terms of radicals; it only rules out the absence of a common formula that works for all polynomials.

Nevertheless Galois in 1830 figured out the exact conditions under which a polynomial has a radical solution; in particular most polynomials of degree five and higher do not have closed form expressions for their roots. The simplest example of such a polynomial is x^5-x-1. Galois thus solved the problem completely and the theory he built (and further refined by subsequent mathematicians) is called Galois theory.

In a different direction, that every polynomial of degree n has exactly n complex roots (the fundamental theorem of algebra) was proved by d'Alembert, Gauss and Argand around 1800.

1.2 The Equation $ax+by=c$

In contrast to solving equations over the reals ($\mathbb{R}$) or over the complex numbers ($\mathbb{C}$), the focus in number theory is to consider equations over integers ($\mathbb{Z}$) or over rational numbers ($\mathbb{Q}$).

Such equations are called Diophantine equations in honor of Diophantus (ÃD 250). Diophantus wrote a book called Arithmetica in which he collected over 200 problems and explained their solutions. He was mainly interested in solutions in terms of positive rationals; for many problems he found general parametric solutions. His book famously inspired Fermat to write in its margins. He was also one of the first persons to use symbolic notation (combined with reasoning by words).

1.2.1 Bezout's Lemma

The simplest equations, over $\mathbb{Z}$, are, as in the case of reals, linear equations. Aryabhatta was the first to explain how to solve an equation of the form $ax+by=c$; an example that he used was 137x+10=60y.

First, let's look at an example without a solution. Consider the equation $4x+6y=5$. This has no solution in integers because 2 divides the LHS but not the RHS. In general, we see that if $d|a$ and $d|b$, then for $ax+by=c$ to have a solution, d must divide c. In particular, the greatest common divisor of (a,b) must divide c. This is a necessary condition but it also turns out to be sufficient.

Theorem 1.1 — Bezout's Lemma. Let a,b,c be natural numbers. The equation

$ax+by=c$ has a solution in integers if and only if $gcd(a,b)$ divides c.

Proof. It is both necessary and sufficient to prove that we can express $gcd(a,b)$ as $ax+by$ for some integers x,y. Let $d=gcd(a,b)$. For a given value of d, we prove by induction on $a+b$ (over pairs (a,b) with $(gcd(a,b)=d)$ that d can be written as $ax+by$. The base case is when $a+b=d$, i.e. $a=d$ and $b=0$. In this clearly $x=1,y=0$ is a solution. Now consider an arbitrary pair (a,b) with $gcd(a,b)=d$ and without loss of generality, let $a\geq b$. Then $gcd(a-b,b)=d$ and by the induction hypothesis, we have: $d=(a-b)x+by$. This implies that $d=ax+b(y-x)$ and thus d is an integer-linear combination of a,b as desired. This completes the proof. ■

We make some remarks: firstly, the proof can be made algorithmic by finding successively smaller pairs (a,b) till we reach the pair $(d,0)$ and work backwards. A simple recursive algorithm is the following:

Algorithm 1 Recursive-Bezout

```
procedure SIMPLE-EUCLID($(a,b)$)
    if $a=0$ then return $(0,1)$
    end if
    if $b=0$ then return $(1,0)$
    end if
    $x \leftarrow$ SIMPLE-EUCLID$(a-b,b)(1)$
    $y \leftarrow$ SIMPLE-EUCLID$(a-b,b)(2)$
    if $a>b$ then return $(x,y-x)$
    else return $(x-y,y)$
    end if
end procedure
```

Now the second remark: as in the case of Euclid's algorithm, we can make it more efficient by considering not just $a-b$ but $a-qb$ for $q=\lfloor a/b \rfloor$. We first look at Euclid's algorithm to thus find the gcd of two numbers.

1.2.2 Euclid's Algorithm

Euclid ($\tilde{3}$50 BC) wrote his algorithm in his famous book The Elements, along with a few other statements in number theory.

Algorithm 2 Euclid's algorithm

```
procedure EUCLID$(a,b)$                          ▷ Returns $gcd(a,b)$
    $A \leftarrow max(a,b)$, $B \leftarrow min(a,b)$
    while $B \neq 0$ do
        $r \leftarrow A$ mod $B$                    ▷ $gcd(A,B)=gcd(B,r)$
        $A \leftarrow B$
        $B \leftarrow r$
    end while
    return $A$
end procedure
```

One of our concerns in this course will be the design of efficient algorithms, often algorithms running in time polynomial in the input size. How efficient is Euclid's algorithm in terms of its input size?

Theorem 1.2 Let $m = max(\log_2 a, \log_2 b)$. Then the number of iterations in Euclid's algorithm $2m$. Further, the time complexity of each iteration is at most $O(m \log m)$.

Proof. The last statement follows from the fact that the complexity of each iteration is essentially the cost of integer division, which has the same complexity as integer multiplication. In 2020, Harvey and van der Hoeven gave a $O(n \log n)$ time algorithm to multiply two n-bit integers; for comparison, the complexity of FFT-based integer multiplication is $O(n \log^2 n)$. Thus the cost of all the basic arithmetic operations performed on two n-bit integers is $\tilde{O}(n)$.

To prove the first statement, let (A_i, B_i) denote the value of the pair after i iterations, with $(A_0, B_0) = (max(a,b), min(a,b))$. Further, let $A_i = B_i q_i + r_i$ with $0 < r_i < B_i$. Note that $A_{i+2} = B_{i+1} = r_i < B_i$.

We have $A_i = B_i q_i + A_{i+1} \geq B_i + A_{i+2} > 2A_{i+2}$. The A_is reduce by a factor of at least 2 after every two iterations, so that the number of iterations is at most $2 \log_2 m$. ■

1.2.3 The Extended Euclidean algorithm

We now look at the extended Euclid's algorithm which, in addition to finding $gcd(a,b)$, finds two integers x, y such that $ax + by = gcd(a,b)$.

Algorithm 3 Extended Euclid's algorithm

```
procedure EXTENDED-EUCLID($(a,b)$)    ▷ Finds $x, y$ such that $ax + by = gcd(a,b)$
    $A \leftarrow max(a,b)$, $B \leftarrow min(a,b)$
    $x \leftarrow 1, y \leftarrow 0$                      ▷ $ax + by = A$ will be invariant.
    $u \leftarrow 0, v \leftarrow 1$                      ▷ $au + bv = B$ will be invariant.
    while $B \neq 0$ do
        $r \leftarrow A \bmod B$                          ▷ $gcd(A,B) = gcd(B,r)$
        $q \leftarrow \lfloor A/B \rfloor$
        $A \leftarrow B$
        $B \leftarrow r$
        $\begin{pmatrix} x & u \\ y & v \end{pmatrix} \leftarrow \begin{pmatrix} u & x - qu \\ v & y - qv \end{pmatrix}$
    end while
    return $(x, y, A)$
end procedure
```

How the algorithm works: The key observation is that the pair (A_{i+1}, B_{i+1}) is obtained from (A_i, B_i) by a linear transformation, namely:

$$(A_{i+1}, B_{i+1}) = (A_i, B_i) \begin{pmatrix} 0 & 1 \\ 1 & -q_i \end{pmatrix} \tag{1.3}$$

Let $M = \prod_i \begin{pmatrix} 0 & 1 \\ 1 & -q_i \end{pmatrix}$. Then we have: $(A_0, B_0)M = (gcd(a,b), 0)$. Thus the first column of M yields a solution to $ax+by = gcd(a,b)$. Also, the equality $A_0x + B_0y = A$ is maintained as an invariant after every iteration.

An illustration of the algorithm for $a = 3, b = 8$:

A	B	x	y	u	v	q
8	3	1	0	0	1	2
3	2	0	1	1	-2	1
2	1	1	-2	-1	3	2
1	0	-1	3	3	-8	

Thus, we find that the gcd is 1 and $(-1,3)$ is a solution to $8x+3y = 1$. We may check that $8x+3y = A$ is valid in each iteration.

Exercise 1.2 Find all integers x, y such that $86x + 197y = 1$.

Exercise 1.3 Find three integers x, y, z such that $6x + 10y + 15z = 1$.

Exercise 1.4 Find all integers x, y, z such that $2x + 3y + 5z = 0$.

1.3 Classroom Exercises

Exercise 1.5 Find an integer solution to 32x+57y=1.

Solution: We iterate through the extended Euclidean algorithm.

A	B	x	y	u	v	q
57	32	1	0	0	1	1
32	25	0	1	1	-1	1
25	7	1	-1	-1	2	3
7	4	-1	2	4	-7	1
4	3	4	-7	-5	9	1
3	1	-5	9	9	-16	3
1	0	9	-16	-32	57	

Thus, we find an integral solution $x = 9, y = -16$.

Exercise 1.6 Find all integer solutions to 4x+7y=1.

Solution: We may find one solution by inspection or using the Euclidean algorithm. For example, $x = 2, y = -1$ is a solution. We may now use this to find all solutions: if (x,y) is a solution, then we have $4(x-2)+7(y+1) = 0$. We observe that 7 divides $4(x-2)$ and hence $7|(x-2)$. Thus, we get $7|(x-2)$; let $x = 7k+2$, this implies that $y = -4k-1$. Thus, the general solution is given by $\{(7k+2, -4k-1) : k \in \mathbb{Z}\}$.

Exercise 1.7 Prove that if $d|a$ and $d|b$, then $d|gcd(a,b)$.

Solution: By Bezout's lemma, there exist integers x, y such that $ax+by = gcd(a,b)$.

Since d divides the LHS, it must divide the RHS which proves the claim.

2. The Fundamental Theorem of Arithmetic

2.1 Prime numbers

A natural number n is defined to be a prime number if it has exactly 2 divisors (namely 1 and n). The sequence of prime numbers begins $2,3,5,7,\ldots$. The first interesting fact about prime number is that there are infinitely many of them.

Theorem 2.1 There are infinitely many prime numbers.

2.2 Euclid's lemma

Lemma 1 If p is prime and $p|ab$, then $p|a$ or $p|b$.

Proof. Suppose for contradiction that p does not divide a and p does not divide b. Then $gcd(p,a)=1$ and by Bezout's lemma, there exist integers x,y such that $px+ay=1$. Similarly, $gcd(p,b)=1$ and there exist integers u,v such that $pu+bv=1$. Multiplying the two relations, we get: $(px+ay)(pu+bv)=1$. Expanding the LHS, we get a contradiction because p divides each term on the LHS, but the RHS is equal to 1. This proves the lemma. ■

2.3 The fundamental theorem

Theorem 2.2 Every natural number $n>1$ can be uniquely factored into a product of prime numbers, i.e. we can write $n=p_1p_2\ldots p_k$, where the p_is are prime (not necessarily distinct), and if $n=q_1q_2\ldots q_l$ with each q_i prime, then $k=l$ and $\{q_1,q_2,\ldots,q_l\}$ is equal (as a multiset) to $\{p_1,p_2,\ldots,p_k\}$.

Proof. There are two statements to prove, (a) that every natural number larger than 1 can be factored into primes and (b) that the factorization is unique (up to ordering).

We first prove (a) by induction. The first few base cases are $2,3,4$ which we see have a prime factorization. The induction step: We consider an arbitrary natural number $n > 1$ and inductively assume that $1,2,\ldots,n-1$ have a prime factorization. If n is prime, then we are done, otherwise let $n = ab$ with $1 < a,b < n$. By the induction hypothesis, a,b have a prime factorization. The two factorizations may be combined to give a factorization for n. This proves (a).

We now prove (b), also by induction. As before, we may verify that (b) holds for the base cases $2,3,4$. For the induction step, we consider an arbitrary $n > 1$, assuming that (b) holds for all numbers lesser than n. Suppose that $n = p_1p_2\ldots p_k = q_1q_2\ldots q_k$. We apply Euclid's lemma to the product $q_1q_2\ldots q_k$ to deduce that p_1 divides some q_i. Since q_i is prime, we have $p_1 = q_i$. Now we apply the induction hypothesis to $n/p_1 = p_2\ldots p_k = \prod_{j\neq i} q_j$ and conclude that $\{p_2,\ldots,p_k\} = \{q_j | j \neq i\}$ (and $k-1 = l-1$). Together with $p_q = q_i$, this gives $\{p_1,\ldots,p_k\} = \{q_1,\ldots,q_l\}$, completing the proof of (b). ■

3. Congruences

3.1 Definition and properties

Let $a, b \in \mathbb{Z}$ and $n > 1$ be a natural number. We say that $a \equiv b \pmod{n}$ (read as a is *congruent to* b modulo n) if $a - b$ is divisible by n.

Examples: $17 \equiv 3 \pmod{7}$, $-20 \equiv -8 \pmod{3}$, $360 \equiv 0 \pmod{60}$.

Properties satisfied by congruences:

1. If $a \equiv b \pmod{n}$ and $c \equiv d \pmod{n}$, then $a + c \equiv b + d \pmod{n}$.
2. If $a \equiv b \pmod{n}$ and $c \equiv d \pmod{n}$, then $ac = bd \pmod{n}$.
3. Note that division doesn't work the same way as reals: in general, $ac \equiv bc \pmod{n}$ does not imply that $a \equiv b \pmod{n}$ OR $c \equiv 0 \pmod{n}$. A counter-example is $n = 6, a = 4, b = 2, c = 3$.
   To understand what we can deduce from $ac \equiv bc \pmod{n}$, we rewrite it as $(a-b)c \equiv 0 \pmod{n}$. Now we see that in some cases, we can draw some conclusions. For example, if $gcd(c,n) = 1$, then we can conclude that $a \equiv b \pmod{n}$. Also, if n is prime, then we can conclude that n divides one of $(a-b), c$, i.e. $a \equiv b \pmod{n}$ or $c \equiv 0 \pmod{n}$.
4. The congruence relation is an *equivalence relation*, i.e. it satisfies the following properties:
   (i) Reflexivity: $a \equiv a \pmod{n}$;
   (ii) Symmetry: $a \equiv b \pmod{n}$ implies that $b \equiv a \pmod{n}$;
   (iii) Transitivity: $a \equiv b \pmod{n}$ and $b \equiv c \pmod{n}$ implies that $a \equiv c \pmod{n}$.
   An equivalence relation partitions the set into *equivalence classes*, in this case *congruence classes*. The congruence classes for a given n are $\{0, \pm n, \pm - 2n, \ldots,\}, \{1, 1 \pm n, 1 \pm 2n, \ldots,\}, \ldots, \{n-1, n-1 \pm n, n-1 \pm 2n, \ldots\}$.
   We consider the numbers $\{0, 1, \ldots, n-1\}$ to be the canonical representatives of these congruence classes, as these numbers are the remainders when divided by n.

3.2 Arithmetic in $\mathbb{Z}_n$

The basic arithmetic operations of addition and subtraction have time complexity $O(\log n)$; the complexity of multiplication is $O(\log n \log\log n)$; the complexity of finding $a^b \pmod{n}$ (by repeated squaring) is $O(\log b \log n \log\log n) = O(\log^2 n \log\log n)$.

3.3 Linear congruences and the Chinese Remainder Theorem

3.4 Fermat's little theorem

3.5 Polynomials and Lagrange's theorem

3.6 Order and primitive roots

3.7 Euler's theorem

3.8 Application: The RSA Algorithm

4. The Quadratic Equation in $\mathbb{Z}_p$

5. Finite Fields

6. Polynomial Factorization over $\mathbb{Z}_p$

7. Polynomial Factorization over $\mathbb{Z}$

II

Quadratic Equations in Two Variables

8. Primality Testing: Before 2002

9. The Integer Factoring Problem

10. Primality Testing: The AKS algorithm

11. Binary Quadratic Forms