Number Theory-Lecture Notes

Eugen J. Ionascu

Contents

List of Figures

List of Tables

Preface

These lecture notes are written over a few years, beginning with the summer semester of 2007 for my students enrolled in a Number Theory course (R. Foley, M. Huckaby, S. Kwon, L. Storm, S. Meredith, S. Thrasher, and A. Markov) and continued in the summer of 2011 (students: E. Driver, M. Redmond, J. Patterson, Y. Robinson and R. Roop-Eckart). There are so many books on number theory and some are technically available to everyone in pdf format on the web. Each one of them, in a sense, is the author(s) perspective of the subject and preference of the topics of interest. We make no exception in these notes. It is sometimes a help for the student who likes to inform himself/herself, to have the possibility of reading a topic from a different perspective so that that particular material will have a better chance "to sink". We made a list of the books that we had the chance to consult and point out which one of these is freely accessible on the web.

There are lots of topics that usually go into an introductory course in number theory depending on the scope of the course and the background of the students. Our scope is to bring the students to a point where he/she may be interested in asking and solving open questions in this field. As a result, no matter what a particular topic goes in, we would like to pursue it, as much as possible, to show the connections that can be made and go toward other developments in the field.

There is also an opportunity in this course to go through lots of techniques of proofs in mathematics. We are also interested in using the power of the computer in the study and learning of number theory.

Another characteristic in our approach is the pursuit of the spectacular in the area called by Carl Friedrich Gauss "the queen of mathematics". One must give Gauss a considerable amount of credit here. Paranthetically, in 1796 at the age of nineteen, Gauss decided to dedicate his life to mathematics after he has shown how a regular polygon of seventeen sides can be constructed with a straight edge and a collapsible compass.

We end this preface with two examples of such striking facts in number theory whose statements are nevertheless easy enough to understand. The first is an

exercise that we learned of it from [19] (page 33):

$$2^{32}+1=(2^9+2^7+1)(2^{23}-2^{21}+2^{19}-2^{17}+2^{14}-2^9-2^7+1)$$

which, for convincing one of its validity, it requires just a little algebra. Interesting enough is the fact that it shows that $2^{2^n}+1$ is not a prime number for all $n\in\mathbb{N}$ (as Fermat predicted and allowed Euler's to show off with his calculational powers by giving this counterexample).

The second is a theorem of Hurwitz from 1891 for which we give as a reference [3], a book connecting number theory with ergodic theory:

Theorem 0.0.1. *For every irrational number x there exist infinitely many pairs of integers p and q, such that*

$$|x-\frac{p}{q}|\leq\frac{1}{\sqrt{5}q^2}.$$

The constant $\frac{1}{\sqrt{5}}$ is the best possible in the sense that if we replace it by something smaller, say $C>0$, then there are infinitely many irrationals x, for which only finitely many pairs of integers p and q satisfy

$$|x-\frac{p}{q}|\leq\frac{C}{q^2}.$$

Finally, number theory abounds in old and new conjectures but one can come up easily with his own. A good reference for lots and lots of interesting and dramatic conjectures or facts in this area is [10]. One of the million-dollar conjectures or millennium problems is at the intersection of number theory and complex analysis. It is known as the Riemann Hypothesis. A reformulation of it is to show that ([8] and [11])

$$|\int_1^x\frac{1}{\ln t}dt-\#\{p|p\ is\ prime\ such\ that\ p\leq x\}|\leq\sqrt{x}\ln x, for\ x\geq 3.$$

Let us conclude with the observation that the topics of number theory are basically at the heart and very good introductions to other, more abstract and technical, branches of mathematics.

Figure 1: Arnold 102

Figure 2: University Hall 025

Figure 3: University Hall 345

Chapter 1

Some Basics about the objects of study

Quotation: *" The development of mathematics toward greater precision has led, as is well known, to the formalization of large tracts of it, so that one can prove any theorem using nothing but a few mechanical rules... One might therefore conjecture that these axioms and rules of inference are sufficient to decide any mathematical question that can at all be formally expressed in these systems. It will be shown below that this is not the case, on the contrary there are in the two systems mentioned relatively simple problems in the theory of integers that cannot be decided based on the axioms." (Kurt Gödel)*

1.1 Natural numbers and rings

We are going to use the following classical notations for the various sets of real numbers $\mathbb{R}$: the natural numbers $\mathbb{N} := \{1, 2, 3, ...\}$, the integers $\mathbb{Z} := \{..., -3, -2, -1, 0, 1, 2, 3, ...\}$, and rationals $\mathbb{Q} := \{\frac{a}{b} : a \in \mathbb{Z}, b \in \mathbb{N}\}$. We assume the operations of addition and multiplication on $\mathbb{N}$ are well-defined and have the well-known properties (see the Peano's axiomatic model for the construction of $\mathbb{N}$): for every a, b, c in $\mathbb{N}$, we have

$$a + b = b + a, ab = ba \quad \diamondsuit \text{ commutative property,}$$

$$(a + b) + c = a + (b + c), (ab)c = a(bc) \quad \heartsuit\text{-associative property,}$$

$$a(b+c) = ab + ac \quad \clubsuit\text{-distributivity of multiplication with respect to addition,}$$

$$1(a) = a \quad \spadesuit\text{-1 is the multiplicative identity .}$$

The order on $\mathbb{N}$ is defined as usual: we say that $a < b$ if $b = a + c$ for some $c \in \mathbb{N}$. Also, we also assume the Principle of Mathematical Induction and the Well Ordering Principle of $\mathbb{N}$. These operations can be extended to $\mathbb{Z}$ and obtain, what is usually called a commutative ring with unity. All properties above are preserved (with the specific changes in terms of PMI and WOP-every set of integers bounded from below has a least element). The only two elements in $\mathbb{Z}$ that have an inverse with respect to multiplication are ± 1 (these will be called *units*).

Let us introduce another useful ring which it will help put some concepts in perspective, that is the ring of Gaussian integers: $\mathbb{Z}[i] := \{a + bi | a, b \in \mathbb{Z}\}$. The addition here is done on components, i.e. $(a+bi)+(c+di) = (a+c)+(b+d)i$, and the multiplication is as usual as complex numbers: $(a+bi)(c+di) = (ac-bd)+(ad-bd)i$. All properties above are preserved, so we get a commutative ring with unity, but we do not have a well defined order and we do not have a PMI. It is easy to see that the only units here are $\{\pm 1, \pm i\}$. The set of units in a ring with unity R, will be denoted by $\mathcal{U}_R$, so $\mathcal{U}_{\mathbb{Z}} = \{1. - 1\}$ and $\mathcal{U}_{\mathbb{Z}[i]} = \{1, -1, i, -i\}$.

Let us make the observation that while for $\mathbb{Z}$, the equation $|m| = |n|$ implies $m = \pm n$, for Gaussian integers it is not true that $|z| = |w|$ implies $z = uw$ for some $u \in \{1, -1, i, -i\}$. For instance, $|z| = |w| = \sqrt{65}$ if $z = 8 + i$ and $w = 4 + 7i$ but $zu \neq w$ for any $u \in \{1, -1, i, -i\}$.

Definition 1.1.1. *For a and b in $\mathbb{N}$, we say that a* divides *b (written $a|b$), or b is divisible by a (written $b \vdots a$), if there exists $c \in \mathbb{N}$ such that $b = ac$. In this case, a is called a* divisor *of b.*

Examples: Clearly, 1 is a divisor of every natural number and $n|n$ for every $n \in \mathbb{N}$. We have $63|2016$, since $2016 = 63(32)$. The set of all divisors of 2016 is

$$\{1, 2, 3, 4, 6, 7, 8, 9, 12, 14, 16, 18, 21, 24, 28, 32, 36, 42, 48, 56,$$

$$63, 72, 84, 96, 112, 126, 144, 168, 224, 252, 288, 336, 504, 672, 1008, 2016\}$$

The set of all divisors of $\{2017\}$ is $\{1, 2017\}$. A number $p > 1$, $p \in \mathbb{N}$ is called a *prime*, if the only divisors of p are 1 and p. So, 2017 is a prime. The greatest

known prime, $2^{74,207,281}-1$, was just recently discovered. We are going to show that the set of primes is infinite.

If in the Definition 1.1.1, we change $\mathbb{N}$ with $\mathbb{Z}$, we get the concept of divisibility in within the integers. As a result we can say, for instance that the set of integer divisors of 16 is $\{\pm 1, \pm 2, \pm 4, \pm 8, \pm 16\}$. If in the Definition 1.1.1, we change $\mathbb{N}$ with $\mathbb{Z}[i]$, we get the concept of divisibility in within the Gaussian Integers. Let us look at two examples here. We have $44+9i$ is a divisor of 2017, because $2017 = 1936+81 = 44^2+9^2 = (44+9i)(44-9i)$. Or, $2+3i$ is a divisor of $-1+5i$ since $-1+5i = (2+3i)(1+i)$.

One of the important functions in number theory, *the divisor function*, is denoted by d, $d:\mathbb{N}\to\mathbb{N}$, and it is defined by $d(n)$ is the number of positive divisors of $n\in\mathbb{N}$. For instance, we have seen that $d(16)=5$ and $d(2016)=36$. (Johann Peter Gustav Lejeune) Dirichlet (1805-1859) has shown that the average of the divisor function, i.e.

$$\frac{1}{n}\sum_{k=1}^{n} d(k) \approx \ln n + 2\gamma - 1,$$

where $\gamma = \lim_{n\to\infty}(-\ln\sum_{k=1}^{n}\frac{1}{k}) \approx 0.5772156649$ is the Euler-Mascheroni constant.

1.2 Division Algorithm

The order we defined earlier is in fact the usual order on the real numbers $\mathbb{R}$. As a result, due to the WOP on the integers, we can define the following function called *the integer part*.

Definition 1.2.1. *Given a real number x, by $\lfloor x\rfloor$, we understand the greatest integer k such that $k\le x$.*

Let us show that for every real number x, we have

$$\lfloor x\rfloor \le x < \lfloor x\rfloor + 1. \tag{1.1}$$

Let us say that $\lfloor x\rfloor = k$. We want to show that $k\le x<k+1$. By definition, we must have $k\le x$. By way of contradiction, let us assume that $x<k+1$ is not true. In other words $k+1\le x$. Since $k+1$ is an integer less than or equal to x, by definition of k as being the greatest with this property, we must have $k\ge k+1$. But this leads to the contradiction $0\ge 1$. ■

The function $x\to x-\lfloor x\rfloor$ is usually called the *fractional part* and it is denoted by $\{\cdot\}$, i.e. $\{x\} = x-\lfloor x\rfloor$ which is a number in $[0,1)$ for every $x\in\mathbb{R}$.

Theorem 1.2.2. [Division Algorithm for Integers] *Given a nonzero positive integer n ($n \in \mathbb{N}$) and an integer k, then there exists two (unique) integers q and r, $r \in \{0, 1, 2, ..., n-1\}$, such that $k = nq + r$.*

The number q is called **quotient** and the number r is called the **remainder**. Let us use a direct proof of this theorem.

(Existence) Assume that the hypothesis is true. In other words, we have an integer k and positive natural number n. To prove the conclusion we need to show the existence of q and r satisfying the required conditions. We define $q = \lfloor \frac{k}{n} \rfloor$ (greatest integer function defined earlier) and $r = k - nq$. We observe that $x = \frac{k}{n}$ is a well defined real number since n is not zero. Next, all we need to show is that $r \in \{0, 1, 2, ..., n-1\}$ or $0 \leq r < n$. Because we have proved in (1.1) that $\lfloor x \rfloor \leq x < \lfloor x \rfloor + 1$, we have $\lfloor \frac{k}{n} \rfloor \leq \frac{k}{n} < \lfloor \frac{k}{n} \rfloor + 1$. This can be written as $q \leq \frac{k}{n} < q + 1$ or $0 \leq r < n$, which is exactly what we wanted.

(Uniqueness) By way of contradiction, suppose that we have two distinct writings, $k = nq_1 + r_1$ and $k = nq_2 + r_2$ with $r_1, r_2 \in \{0, 1, 2, ..., n-1\}$. We may assume, without loss of generality, that $q_1 \neq q_2$. Indeed, if $q_1 = q_2$ then $r_1 = k - nq_1 = k - nq_2 = r_2$ and so we have the same writing, but we assumed these writings were distinct. Since $|q_1 - q_2|$ is a positive integer, it must by at least 1. Then $n(q_1 - q_2) = r_2 - r_1$ which implies $|r_2 - r_1| = n|q_1 - q_2| \geq n(1) = n$. Suppose, without loss of generality, that $r_1 \geq r_2$. Then $n \leq r_1 - r_2 \leq r_1 < n$, which leads to the contradiction $n < n$. It remains that the two writings must be the same and the uniqueness in Theorem 1.2.2 is shown. ■

The set $\{0, 1, 2, ..., n-1\}$ is usually called a *complete set of residues modulo* n (CRS). In general, a CRS is just a set R of n integers with the property that for every $i \in \{0, 1, 2, ..., n-1\}$, there exists a unique $r \in R$ such that $r - i$ is divisible by n. We observe that any set of the form

$$\{r | r = i + nk_i \text{ for some } \quad i = 0, 1, 2, ...n-1, \text{ and } k_i \in \mathbb{Z}\}$$

is a CRS modulo n. Another very common one is the following

$$\mathcal{R}S_n := \{r | r = i \quad \text{if} \quad 0 \leq i < n/2 \text{ or } \quad r = i - n \quad \text{if} \quad n/2 \leq i < n\}.$$

For a complex number $z = a + bi$, we review some of the customary notation: $\overline{z} = a - bi$, $|z| = \sqrt{z\overline{z}} = \sqrt{a^2 + b^2}$ and $\mathbf{Re}(z) = a$, $\mathbf{Im}(z) = b$. It is very common in the algebraic number theory to denote $N(z) = |z|^2 = a^2 + b^2$.

Theorem 1.2.3. [Division Algorithm for Gaussian Integers (Version 1)] *Given a nonzero z ($z \in \mathbb{Z}[i]$) and a Gaussian integer w, then there exists two*

(unique) Gaussian integers q and r, such that $w = qz + r$ where r satisfies

$$0 \leq \mathbf{Re}(r\overline{z}) < |z|^2,\ 0 \leq \mathbf{Im}(r\overline{z}) < |z|^2. \tag{1.2}$$

PROOF. **(Existence)** We represent w in the orthonormal base $\{z, iz\}$: $w = \alpha z + \beta(iz)$, for some real numbers α and β. This is basically equivalent to calculating $\frac{w}{z} = \alpha + \beta i$ as a complex number. We take $q = \lfloor\alpha\rfloor + \lfloor\beta\rfloor i$. Then, we simply define $r = w - zq$. Let us observe first that q and r are Gausssian integers. Since $\frac{w}{z} = \alpha + \beta i = q + s$ where $s = \{\alpha\} + \{\beta\}i$, by our definitions $r = (\{\alpha\} + \{\beta\}i)z$. So, wee need to show that the two equalities in (1.2) are valid. But $r\overline{z} = (\{\alpha\}+\{\beta\}i)|z|^2$, which imply that $0 \leq \mathbf{Re}(r\overline{z}) = \{\alpha\}|z|^2 < |z|^2$ and $0 \leq \mathbf{Im}(r\overline{z}) = \{\beta\}|z|^2 < |z|^2$. Hence (1.2) is valid.

(Uniqueness) By way of contradiction, if we have two different writings $w = q_1 z + r_1 = q_2 z + r_2$ with r_1 and r_2 satisfying (1.2), then

$$0 \leq \mathbf{Re}(r_1\overline{z}) < |z|^2,\ 0 \leq \mathbf{Im}(r_1\overline{z}) < |z|^2, \text{ and}$$
$$0 \leq \mathbf{Re}(r_2\overline{z}) < |z|^2,\ 0 \leq \mathbf{Im}(r_2\overline{z}) < |z|^2.$$

Clearly if $q_1 = q_2$, then $r_1 = r_2$ which contradicts the assumption that we start with two different writings. So, we may assume that $q_1 \neq q_2$.

This implies that $|\mathbf{Re}(r_1\overline{z}) - \mathbf{Re}(r_2\overline{z})| < |z|^2$ and $|\mathbf{Im}(r_1\overline{z}) - \mathbf{Im}(r_2\overline{z})| < |z|^2$. But $0 = (q_1 - q_2)z + (r_1 - r_2)$ which implies that $(r_1 - r_2)\overline{z} = (q_2 - q_1)|z|^2$. Therefore, $|\mathbf{Re}(q_2 - q_1)|z|^2| < |z|^2$ and $|\mathbf{Im}(q_2 - q_1)|z|^2| < |z|^2$. This implies that $|\mathbf{Re}(q_2 - q_1)| < 1$ and $|\mathbf{Im}(q_2 - q_1) < 1$. These two inequalities can be satisfied only if $q_1 = q_2$ and this a contradiction. ∎

Theorem 1.2.4. [Division Algorithm for $\mathbb{Z}[i]$ (Version 2)] *Given a nonzero z ($z \in \mathbb{Z}[i]$) and a Gaussian integer w, then there exists two (unique) Gaussian integers q and r, such that $w = qz + r$ where r satisfies $|r| < |z|$ and in addition, either*

$$0 \leq \mathbf{Re}(r\overline{z}) < |z|^2,\ 0 \leq \mathbf{Im}(r\overline{z}) < |z|^2 \quad \textit{or} \tag{1.3}$$

$$-|z|^2 \leq \mathbf{Re}(r\overline{z}) < 0,\ -|z|^2 \leq \mathbf{Im}(r\overline{z}) < 0 \textit{ with } |r + z(1+i)| \geq |z|. \tag{1.4}$$

PROOF. **(Existence)** We construct r first as in the Theorem 1.2.3. Let us denote the set of all $r's$ that satisfy (1.2) by $\mathcal{R}$. We have two disjoint possibilities for each $r \in \mathcal{R}$: either $|r| < |z|$ or $|z| \geq |z|$. So, we can write $\mathcal{R} = \mathcal{R}_1 + \mathcal{R}_2$ where

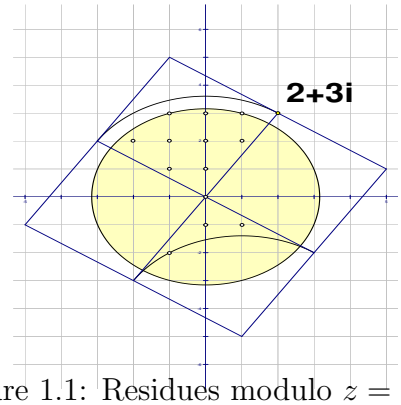


Figure 1.1: Residues modulo $z = 2 + 3i$

$\mathcal{R}_1 := \{r \in \mathcal{R} | |r| < |z|\}$ and $\mathcal{R}_2 := \{r \in \mathcal{R} | |r| \geq |z|\}$ (see Figure 1.1). We are going to replace every $\widetilde{r} \in \mathcal{R}_2$ by $r := \widetilde{r} - z(1+i)$. So, we define

$$\mathcal{R}_2' = \{r | r = \widetilde{r} - z(1+i) \text{ for some } \widetilde{r} \in \mathcal{R}_2\}.$$

For every $r \in \mathcal{R}_2'$ we need to change the corresponding q. Because we know that for some $\widetilde{q} \in \mathbb{Z}[i]$ we have

$$w = \widetilde{q}z + \widetilde{r} = \widetilde{q}z + r + (1+i)z = (\widetilde{q} + 1 + i)z + r,$$

we need to take $q := \widetilde{q} + 1 + i$. Let us observe that for $r \in \mathcal{R}_2'$ we have $\mathbf{Re}(r\overline{z}) = \mathbf{Re}(\widetilde{r}\overline{z}) - |z|^2$ which satisfies $-|z|^2 \leq \mathbf{Re}(r\overline{z}) < 0$ and similarly we have the same observation for $\mathbf{Im}$. Also, since $r + (1+i)z = \widetilde{r}$ and by the definition of $\mathcal{R}_2$ we must have $|\widetilde{r}| = |r + (1+i)z| \geq |z|$. This means that the values in $\mathcal{R}_2'$ satisfy (1.4). For those $r \in \mathcal{R}_2'$ we need to show that $|r| < |z|$. From the previous proof

$$r = \widetilde{r} - z(1+i) = [(\{\alpha\} - 1) + (\{\beta\} - 1)i]z.$$

Hence we have to show that $(\{\alpha\} - 1)^2 + (\{\beta\} - 1)^2 < 1$. Because $\widetilde{r} \in \mathcal{R}_2$ we have $|\widetilde{r}| \geq |z|$ or $\{\alpha\}^2 + \{\beta\}^2 \geq 1$. From here we see that $\{\alpha\} + \{\beta\} > \{\alpha\}^2 + \{\beta\}^2 \geq 1$ with the observation that the first inequality is a strict one because $\{\alpha\}, \{\beta\} \in [0, 1)$. Hence, $\{\alpha\} + \{\beta\} > 1$ and then

$$(\{\alpha\} - 1)^2 + (\{\beta - 1\})^2 \leq (1 - \{\alpha\}) + (1 - \{\beta\}) < 1.$$

This shows the existence of an q and r in $\mathbb{Z}[i]$ satisfying the requirements of the theorem.

(Uniqueness) To show that the writing is unique, we reduce the problem to the uniqueness already shown in Theorem 1.2.3. By way of contradiction if we have two writings $w = q_1 z + r_1 = q_2 z + r_2$ with r_1 and r_2 satisfying (1.3) or (1.4). We have

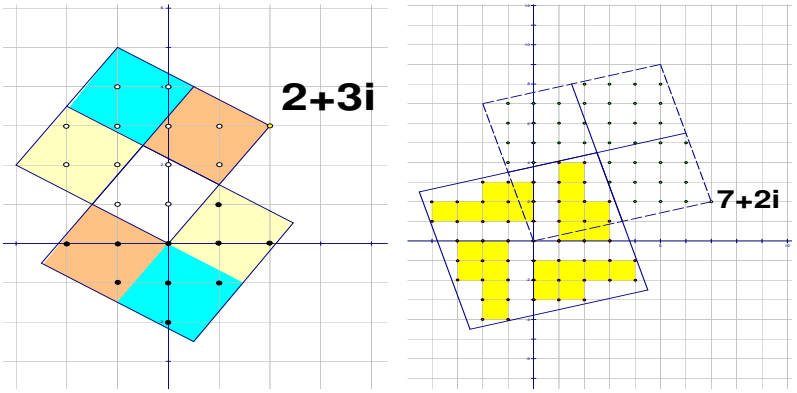


Figure 1.2: Residues modulo $z = 2+3i$ and $z = 7+2i$ with the property $|r| \leq \frac{1}{\sqrt{2}}|z|$

either r_1, r_2 both satisfying (1.3) in which case we know that $r_1 = r_2$ and $q_1 = q_2$. If r_1, r_2 both satisfy (1.4), the proof follows the same arguments as in Theorem 1.2.3. Let us observe that we cannot have, for instance r_1 satisfying (1.3) and r_2 satisfying (1.4). Indeed, assuming this is possible, this implies that $r := r_2 + (1+z)i$ satisfies (1.3) and by uniqueness shown in Theorem 1.2.3, we must have $r_1 = r_2 + (1+z)i$ and of course $q_1 = q_2 - (1+i)z$. But this is not possible since we get the contradiction $|z| > |r_1| = |r_2 + (1+z)i| \geq |z|$. ■

In Figure 1.1, we have shown a complete set of residues modulo $z = 2+3i$:

$$\mathcal{R}_z = \{-2+2i, -1+i, -1+2i, -1+3i, 0, i, 2i, 3i, 1+2i, 1+3i, -1-2i, -i, 1-i\}$$

constructed as in the proof of Theorem 1.2.4.

We will show that the cardinality of a complete set of residues modulo z is precisely $|z|^2$. Let us observe that Theorem 1.2.3 can be made stronger in the following way.

Theorem 1.2.5. [Division Algorithm for $\mathbb{Z}[i]$ (Version 3)] *Given a nonzero z ($z \in \mathbb{Z}[i]$) and a Gaussian integer w, then there exists two unique Gaussian integers q and r, such that $w = qz + r$ where r satisfies $|r| \leq \frac{1}{\sqrt{2}}|z|$ and*

$$-\frac{1}{2}|z|^2 \leq \mathbf{Re}(r\overline{z}) < \frac{1}{2}|z|^2, \ \frac{1}{2}|z|^2 \leq \mathbf{Im}(r\overline{z}) < \frac{1}{2}|z|^2 \tag{1.5}$$

The proof of this theorem is similar to what we have done earlier and it is left to the reader. The idea of proof is basically described in Figure 1.2

where the region $\mathcal{R}$ is divided into four smaller squares. Three of them are translated with $-z$, $-iz$ or $-(1+i)z$ respectively. The (CSR) is very symmetric in this case:

$$\mathcal{R}S_z = \{0\} \cup \bigcup_{u \in \mathcal{U}(\mathbb{Z}[i])} \{u, 2u, (1+i)u\}.$$

Let us denote by $\mathcal{R}S_z$ this (CSR) in general, which is the equivalent of $\mathcal{R}S_n$ in $\mathbb{Z}$.

Definition 1.2.6. *A Gaussian number z is a* prime *if it is not a unit, and $z = xy$ with $x, y \in \mathbb{Z}[i]$ implies either x or y is a unit.*

Comment: In general, this definition and the similar one we had for $\mathbb{N}$, is the definition that one has for so called *irreducible* elements. The definition of a prime instead, in general, is an element p with the property that $p|ab$ implies $p|a$ or $p|b$. We will show that this is a property that takes place in $\mathbb{Z}$ and $\mathbb{Z}[i]$.

Proposition 1.2.7. *The norm $N : \mathbb{Z}[i] \to \mathbb{N}$ is multiplicative, i.e., $N(zw) = N(z)N(w)$ for all $z, w \in \mathbb{Z}[i]$, where $N(z) = N(a+bi) = a^2+b^2$ for $z = a+bi \in \mathbb{Z}[i]$.*

Proof If $z = a + bi$ and $w = c + di$ we have $zw = (ac - bd) + (ad + bc)i$. Hence, the identity we want to show is $(ac - bd)^2 + (ad + bc)^2 = (a^2 + b^2)(c^2 + d^2)$. This identity is usually called the Lagrange's identity and it can be checked by usual algebra manipulations. ■

We can show that for instance $z = 1 + i$ is a prime. Indeed, if $z = xy$, then by Proposition 1.2.7, $N(z) = 2 = N(x)N(y)$. This implies $N(x)$ or $N(y) = 1$, so x or y is a unit. For the same reason, we can think of a lot of other Gaussian primes: $1 + 2i$, $2 + 3i$, etc. We will show that every prime p in $\mathbb{N}$, of the form $p = 4k + 3$, $k \in \mathbb{Z}$, is also a prime in $\mathbb{Z}[i]$.

Proposition 1.2.8. *We have the following properties of the divisibility:*

(1) for every $n \in \mathbb{Z}$, we have $1|n$ and $n|n$;

(2) for every $a, b, d, \alpha, \beta \in \mathbb{Z}$, if $d|a$ and $d|b$ then have $d|\alpha a + \beta b$;

(3) if d is a divisor of a, $n \in \mathbb{Z} \setminus \{0\}$, then $d \leq |a|$.

For two numbers $a, b \in \mathbb{N}$ we have 1 as a common divisor and any common divisor d must be at most as big as $min(a, b)$.

Definition 1.2.9. *Given two integers a and b not both equal to zero, the biggest number that divides both a and b is called the* ***greatest common divisor*** *of a and b and it is denoted by* $\gcd(a, b)$.

For instance $\gcd(141, 235) = 47$ since 47 divides both numbers, $141 = 47(3)$ and $235 = 47(5)$, and since $141(2) - 235 = 47$ we see that any number d dividing both 141 and 235 must be a divisor of $141(2) - 235$ so a divisor of 47. Hence 47 is the biggest divisor. This idea works in general, and the procedure is called the *Euclidean Algorithm*. In The Elements, Euclid included this algorithm to calculate the greatest common divisor of two numbers.

Let us start with an example. We let $a = 22099$ and $b = 4223$. One can check the following calculations based on the Division Algorithm:

$$\begin{array}{l} a = 5b + 984, \quad \text{let's define} \quad r_1 = 984 \\ b = 4(984) + 287, \quad \text{let's define} \quad r_2 = 287 \\ r_1 = 3(287) + 123, \quad \text{define} \quad r_3 = 123 \\ r_2 = 2(123) + 41, \quad \text{define} \quad r_4 = \boxed{41} \\ r_3 = 3(41) + 0, \quad r_5 = 0. \end{array}$$

We observe that the sequence of remainders, r_1, r_2, ... satisfies

$$b > r_1 > r_2 > r_3 > r_4 > r_5 = 0.$$

With a moment of thought, working our way backwards with these equalities, using the same argument as in showing that $\gcd(141, 235) = 47$, we conclude that $41 = \gcd(r_3, r_2) = \gcd(a, b) = 41$. Let us record this as a general statement and point out to the essential steps in its proof.

Theorem 1.2.10. *We let a and b be two integers with b different of zero. Then there exists a finite sequence of equalities $r_j = q_j r_{j+1} + r_{j+2}$ with $r_{-1} = a$, $r_0 = |b|$, $j = -1, 0, 1, ..., k-1$ such that*

(1) $q_i \in \mathbb{Z}$,

(2) $|b| > r_j > r_{j+1} \geq 0$ for $j = 1, ..., k$ and

(3) $r_{k+1} = 0$,

(4) $\gcd(a, b) = r_k$.

(**Bézout Lemma**) *In addition, there exists $x, y \in \mathbb{Z}$ such that $ax + by = r_k$.*

Proof. We obtain all these equalities from the Division Algorithm. First we divide a by $|b|$ and get $a = q_{-1}|b| + r_1$ with $r_1 \in \{0, 1, ..., |b| - 1\}$. Since the sequence r_j is strictly decreasing and is bounded below by 0, we need to have only finitely many r_j's until we reach zero. We let $r_{k+1} = 0$. Then $r_k > 0$ divides r_{k-1} and because we have $r_{k-2} = r_{k-1}q_{k-2} + r_k = r_k q_{k-1} q_{k-2} + r_k = r_k(q_{k-1}q_{k-2} + 1)$ we see that r_k divides r_{k-2} too. In fact, we have $r_k = \gcd(r_{k-1}, r_{k-2})$. This equality spreads all the way to $r_k = \gcd(a, b)$ using an induction on $j = k-1, k-2, ..., -1$. The second part of the theorem is obtained by eliminating the variables r_j all the way to a and b. ∎

For instance, for our example before $41 = 287 - 2(123) = 287 - 2[984 - 3(284)] = 287(7) - 2(984) = [b - 4(984)]7 - 2(984) = 7b - 30(984) = 7b - 30(a - 5b)$ which implies $41 = 157b - 30a$.

1.3 Euclidean Algorithm and The greatest common divisor in $\mathbb{Z}[i]$

If $a, b \in \mathbb{Z}[i]$ and $a|b$, then $N(a)|N(b)$, the the set of divisors of a non-zero Gaussian integer is finite. We observe that for integers, the greatest common divisor has the property that any other common divisor must divide it. We will take this property as the definition of the greatest common divisor for Gaussian integers.

Definition 1.3.1. *Given two Gaussian integers a and b not both equal to zero, a common divisor d that has the property that any other common divisor divides it, it is called a* ***greatest common divisor*** *of a and b and it is denoted by* $\gcd(a, b)$.

The Euclidean algorithm in the case of Gausssian integers is basically following the same pattern as in the case of integers. It is showing that the $\gcd(a, b)$ exists. Next, let us make the observation that this concept is defined up to a unit. Indeed, if d' is another greatest common divisor, the by definition $d'|d$ and $d|d'$. Hence, d/d is a Gausssian integer of norm 1, so it must be a unit. So, we are going to use the terminology *the greatest common divisor* knowingly it is unique up to the four units of $\mathbb{Z}[i]$.

The following Maple code is implementing the Division Algorithm in Theorem 1.2.5:

DivAlg:=proc(w,z)

local i,u,a,b,aa,bb,aaa,bbb,q,r;

u:=$w * conjugate(z)/abs(z)^2$**;**

a:=Re(u);b:=Im(u);aa:=floor(a);bb:=floor(b);aaa:=a-aa;bbb:=b-bb;

if $aaa < 1/2$ **and** $bbb < 1/2$ **then q:=aa+bb*I; r:=w-z*q;fi;**

if $aaa \geq 1/2$ **and** $bbb < 1/2$ **then q:=(aa+1)+bb*I; r:=w-z*q;fi;**

if $aaa < 1/2$ **and** $bbb \geq 1/2$ **then q:=aa+(bb+1)*I; r:=w-z*q;fi;**

if $aaa \geq 1/2$ **and** $bbb \geq 1/2$ **then q:=(aa+1)+(bb+1)*I; r:=w-z*q;fi;**

[q,r];

end:

Let us look at an example. The same method we used in the case of integers can be used here to determine the greatest common divisor of the following two Gaussian

integers: $a = 2016 + 2017i$ and $b = 118 + 81i$. We used the Division Algorithm in Theorem 1.2.5 and the Maple program above to help with the computations. One can check that

$$a = (20+4i)b-20-75i, b = (20+75i)(1-i)+23+26i, \text{ and } 20+75i = (23+26i)(2+i).$$

As we have seen before, we can say that $\gcd(a, b) = 23 + 26i$. Let us record at this point the result that is usually known as Bezout's Lemma:

Lemma 1.3.2. (**Bézout Lemma**) *Given two integers of Gaussian integers a and b then there exists x, y integers or Gaussian integers such that* $\gcd(a, b) = ax + by$.

We say that two integers a and b are *relatively prime* or *coprime* if $\gcd(a, b) = 1$. The same definition applies to Gaussian integers.

Proposition 1.3.3. (**Euclid's Lemma**) *If a and b are relatively prime, and $a|bc$ then $a|c$.*

Proof. By Bézout Lemma, we have $ax + by = 1$ for some integers x and y. Then, if we multiply this by c, we get $acx + bcy = c$. Because $a|bc$, we have $bc = a\alpha$ for some $\alpha \in \mathbb{Z}$. Hence,

$$c = acx + bcy = acx + a\alpha y = a(cx + \alpha y),$$

which shows that $a|c$. ■

Corollary 1.3.4. *(1) Given a and b are relatively prime, if $a|c$ and $b|c$, then $ab|c$.*

(2) If p is a prime and $p|ab$, then $p|a$ or $p|b$.

Proof. (1) Since $a|c$ and $b|c$ we can write $c = a\alpha = b\beta$ for some integers α and β. Hence, $a\alpha = b\beta$ implies that a divides $b\beta$. By Euclid's Lemma, a must divides β, and so $\beta = a\gamma$ for some integer γ. Therefore, $c = b\beta = ba\gamma$ which implies $ab|c$.

(2) If p is a prime and p is not a divisor of a, then $\gcd(p, a) = 1$. Hence, by Euclid's Lemma, $p|b$. ■

The similar statements and proofs for Gaussian integers are left as an exercise. Let us use these results to find a decomposition of $w = 7+6i$ as a product of primes in $\mathbb{Z}[i]$. Since $N(w) = (7+6i)(7-6i) = 7^2+6^2 = 49+36 = 85 = 5(17) = (2+i)(2-i)17$ we see that the prime $2+i$ divides $7+6i$ or $7-6i$. It is easy to check that $(2+i)|(7+6i)$ and $7 + 6i = (2 + i)(4 + i)$. Since $4 + i$ is also a Gaussian prime, we have found a decomposition of $w = 7 + 6i$ as a product of primes in $\mathbb{Z}[i]$. In a similar way we can find the following decomposition for $2016 + 2017i$:

$$2016 + 2017i = (-1)(1 + 2i)(4 + i)(4 + 15i)(6 + 19i).$$

Let us observe that we can define an order on the Gaussian primes, as suggested in these decompositions, which will allow to obtain a similar result as in the case of $\mathbb{N}$.

Theorem 1.3.5. (**Fundamental Theorem of Arithmetic**) *Every natural number $n > 1$ can be written in a unique way as a product of primes $p_1p_2...p_k$ where $p_1 \leq p_2 \leq p_3 \leq \cdots \leq p_k$, with $k \in \mathbb{N}$.*

Proof. **Existence:** We proceed by Strong Induction on n. For $n = 2$, the statements is true since $p_1 = 2$ is a prime. Assume that we can find such decompositions for every $k \leq n$, with $n \geq 2$. Then $n+1$ is either a prime, in which case we have a decomposition, or it is composite. If it is composite $n+1 = ab$ with $a, b \in \mathbb{N}$, and $a > 1$, $b > 1$. This shows that $a \leq n$ and $b \leq n$ ($a \geq n+1$ implies $n+1 = ab > (n+1)(1)$, a contradiction). Therefore, by the induction hypothesis, a and b can be written as a product of primes. Putting all these primes together in a non-increasing list, gives the decomposition of $n+1$ as required. Therefore, by the PMI we have the statement true for all $n > 1$.

Uniqueness: Let us proceed by Strong Induction on n, again. For the basic step, $n = 2$ if $2 = p_1...p_k$ then since 2 is a prime $2|p_j$ form some j. This is possible only if $p_j = 2$. Hence, $p_1 = p_j$ and then after simplifying by 2 we get $1 = p_2...p_k$. Automatically, we observe that this is not possible if $k > 1$. It remains that $k = 1$ and $p_1 = 2$ in other words the writing is unique. Assume that the uniqueness is valid for all $2 \leq k \leq n$, with $n \geq 2$. Let us say we have two writings for $n+1$:

$$n+1 = p_1p_2...p_k = q_1q_2...q_s,\ p_1 \leq p_2 \leq \cdots \leq p_k \text{ and } q_1 \leq q_2 \leq \cdots \leq q_s.$$

As before, p_1 divides the product $q_1q_2...q_s$ so by Corollary 1.3.4, then p_1 divides at least one of the q_j. This is possible only if it is equal to one of them. Simplifying the equality by p_1 we get $(n+1)/p_1 = p_2...p_k = q_1p_2...q_s/p_1$. If $(n+1)/p_1 = 1$ then we proceed as the case $n = 2$ and arrive at a contradiction if $s > 1$. In this case then $n+1 = p_1 = q_1$ and so we have unique writing. Because $1 < (n+1)/p_1 \leq (n+1)/2 < n$, we can use the induction hypothesis, and conclude that these two writings must be the same. Therefore, the writings of $n+1$ must be the same. Thus, by the PMI we have the statement true for all $n > 1$. ∎

Let us observe that if $p = a + bi$ is a Gaussian prime, we can look at its associates ip, $-p$, and $-ip$, at least one of these is in the first quadrant. Also, $\overline{p} = a - bi$ must be also a Gaussian prime. It is also clear that $a \neq b$ if $a + bi$ is prime, unless $a = b = \pm 1$. An associate of $\overline{p} = a - bi$ is $i\overline{p} = b + ai$, which shows that we can always take an associate of a prime in the first quadrant $\{re^{it} | r \geq 0, t \in [0, \pi/2)\}$ (polar coordinates), of these primes. We will then list the primes in the decomposition of a Gaussian number in nondecreasing order of their norm $N(p)$,

and in the case of the same norm as for primes like $p = a + bi$, $q = c + di$ with $a^2 + b^2 = c^2 + d^2$ we just use the order on a and c. For instance, $p = 2 + i$ comes first and then $1 + 2i$ if we have something like this occurring. The decomposition of 2, according to this writing is then $2 = -i(1 + i)^2$.

Chapter 2

Some Diophantine equations

2.1 Pythagorean Triples

Quotation: *"The result of the mathematician's creative work is demonstrative reasoning, a proof, but the proof is discovered by plausible reasoning, by GUESSING" (George Polya, [21])*

Notions, concepts, definitions, and theorems: *Characterization, co-prime numbers, primitive solution, parametrization*

Perhaps one of the first encounters that a student may have had with a number theory problem comes from a geometry or a trigonometry course. It is about triples of numbers such as $(3, 4, 5)$ or $(5, 12, 13)$ called *Pythagorean triples* for the obvious reason that if a triple like this, say (a, b, c), we take a, b, c (assuming $0 < a \leq b \leq c$) to represent the sides of a triangle (measured with a certain but inessential unit) then the triangle is a right triangle. By the Pythagorean theorem we must have $a^2+b^2 = c^2$. This was a discovery believed to be known to the Babilonians (Plimpton 322, 1800 BC) but some recent interpretations of Plimpton 322 point out in other directions. A rather simple question here is whether or not there are other triples of positive integers like these. Of course a simple question since we can multiply any of the two examples above by a positive integer and obtain other Pythagorean triples such as $(6, 8, 10)$ or $(15, 36, 39)$. But let us impose the condition that such a triple cannot be simplified by any integer greater than 1. One such triple is called a *primitive* one.

One can use an exhaustive search and find that there are only sixteen such triples with $c < 100$. These are included in the Table 2.1. Investigating the Table 2.1 we can start making all sorts of conjectures about these triples. First of all, we guess there are infinitely many.

$(3,4,5)$	$(9,40,41)$	$\boxed{(16,63,65)}$	$(36,77,85)$
$(5,12,13)$	$(11,60,61)$	$(20,21,29)$	$(39,80,89)$
$(7,24,25)$	$(12,35,37)$	$(28,45,53)$	$(48,55,73)$
$(8,15,17)$	$(13,84,85)$	$\boxed{(33,56,65)}$	$(65,72,97)$

Table 2.1: Primitive Pythagorean Triples with the hypothenuse less than 100.

One of the first two numbers in a triple (called *legs*) is an odd number and the other is even (in fact divisible by 4). One of the legs is divisible by 3. One of the numbers in a triple is divisible by 5. There are only two triples with the same hypothenuse: $65^2 = 56^2 + 33^2 = 16^2 + 63^2$. There are only two triples in which one leg is one more than the other. The corresponding triangles are close to be isosceles and since there are no examples of isosceles ones we may conjecture that there are no Pythagorean triples with equal legs. There are six examples in which the hypothenuse is one unit more than a leg. This suggests that there are infinitely many examples such as these?

Let us pursue this question. If $c = b+1$ then the equation $c^2 = a^2+b^2$ becomes $b^2+2b+1 = b^2+a^2$ which gives $b = \frac{a^2-1}{2}$. If we are careful to take a an odd number then we find plenty (infinitely many) of such triples:

$$(a, \frac{a^2-1}{2}, \frac{a^2-1}{2}+1), a \in \mathbb{N}, a \ odd, a \geq 3. \tag{2.1}$$

For $a = 3,5,7,9,11,13$ we get all of the above ones. Increasing a past 13 will give a $c = \frac{a^2+1}{2} > 100$. The recipe (2.1) was known to the Pythagorean school. Let us observe that it produces primitive Pythagorean triples (two consecutive integers are coprime (*coprime* or *relatively prime* integers are two integers a and b with the property that their greatest common divisor is 1).

We are going to show next that every Pythagorean triple has the property that one of the legs is divisible by 3. Let us denote this triple as before by (a,b,c) with $0 < a \leq b \leq c$, $a,b,c \in \mathbb{N}$. In general, a natural number is of the form $3k$, $3k+1$ or $3k-1$ with $k \in \mathbb{Z}$. Let us assume by way of contradiction that a and b are not divisible by 3 so they must be of the form $3k \pm 1$. This implies

$$c^2 = a^2 + b^2 = (3k \pm 1)^2 + (3k' \pm 1)^2 = 3(3k^2 + 3k'^2 \pm 2k \pm 2k') + 2$$

is of the form $3l+2$. But no perfect square is of the form $3l+2$. This contradiction shows that we cannot have this case and so it remains that a or b must be divisible by 3.

Problem 1 (Homework) *Prove that in every Pythagorean triple (a, b, c) at least one of the numbers a, b, c is divisible by 5.*

All the conjectures formulated above are true for all primitive Pythagorean triples. The reader is invited to prove them all.

Before we find a *general parametrization* of all primitive Pythagorean triples let us look also at the case $b = a+1$. The equation becomes $a^2+a^2+2a+1 = c^2$ or $2c^2-(2a+1)^2 = 1$. We will see in the last section of this Chapter that this equation leads to the so called Pell's equation. We know that there are at least two particular solutions for this equation: $c = 5$, $a = 4$ and $c = 29$, $a = 20$. One can check that the following recurrence gives an infinite sequence of solutions:

$$a_{n+1} = 3a_n + 2c_n + 1, \text{ and } c_{n+1} = 4a_n + 3c_n + 2, \ c_0 = 1, \ a_0 = 0, \ n \in \mathbb{N}.$$

The fact that these formulae generate Pythagorean triples of the form $(a, a+1, c)$, reduces to a simple algebra calculation and an induction argument. The first ten such triples that are generated this way are: $(3, 4, 5)$, $(20, 21, 29)$, $(119, 120, 169)$, $(696, 697, 985)$, $(4059, 4060, 5741)$, $(23660, 23661, 33461)$, $(137903, 137904, 195025)$, $(803760, 803761, 1136689)$, $(4684659, 4684660, 6625109)$ and $(27304196, 27304197, 38613965)$. Is this method generating all of such triples? We will show that this is indeed the case.

Next, we are going to make another connection with geometry. If we introduce $x = \frac{a}{c}$, $y = \frac{b}{c}$ we see that $x^2 + y^2 = 1$ and $x, y \in \mathbb{Q}$. So, any solution of $a^2 + b^2 = c^2$ gives a point (x, y) on the unit circle with coordinates which are positive rational values (Figure 2.1). Then the slope of the line connecting this point and the point of coordinates $(-1, 0)$ is a rational positive number given by $m = \frac{y}{x+1}$. This gives $y = m(x+1)$ and so $m^2(x+1)^2 + x^2 = 1$ which gives

$$(x+1)[(m^2+1)x + m^2 - 1] = 0.$$

Since x is assumed positive we have only the solution $x = \frac{1-m^2}{m^2+1}$. Then $y = \frac{2m}{m^2+1}$. Notice that if $m = \frac{u}{v}$ with u, v positive integers such that the fraction m is in its reduced form we then have

$$x = \frac{a}{c} = \frac{v^2-u^2}{u^2+v^2} \ \text{ and } \ y = \frac{b}{c} = \frac{2uv}{u^2+v^2}.$$

The fraction $\frac{v^2-u^2}{u^2+v^2}$ could only be simplified by a power of 2. Indeed, if p is an odd prime that simplifies it, then it must divide their sum which is $2v^2$ and also it must divide their difference which is $2u^2$. Hence p divides u and v. But this is not possible

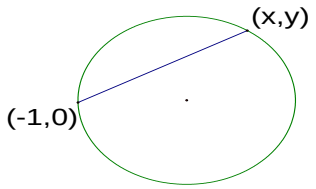


Figure 2.1: Chord idea

by our assumption on u and v. So the only factor that $\frac{v^2-u^2}{u^2+v^2}$ can be simplified by is 2. The same is true with the fraction $\frac{2uv}{u^2+v^2}$: it is either in the reduced form or it can be simplified only by a factor of 2.

Let us notice that both fractions can be simplified by 2 if and only if u and v are both odd numbers. In this case if we set $k = (u+v)/2$ and $l = (v-u)/2$ we see that $2(k^2+l^2) = u^2+v^2$, $k^2-l^2 = uv$ and $v^2-u^2 = (v-u)(v+u) = 4kl$. Then after the simplification the two fractions become something very similar to what we started:

$$x = \frac{2kl}{k^2+l^2}, \quad \textit{and } y = \frac{k^2-l^2}{k^2+l^2},$$

where basically the roles of x and y have interchanged if we identify the parameters the obvious way.

Because $v = k+l$ and $u = k-l$ we see that $gcd(k,l) = 1$ and k and l cannot be both odd. Hence the new fraction cannot be simplified any further by any positive integer. Hence we have $a = v^2-u^2$, $b = 2uv$ and $c = u^2+v^2$ in case u and v have different parity or $a = 2kl$, $b = k^2-l^2$ and $c = k^2+l^2$ in case u and v are odd which implies that k and l again having different parity. We see that we end up in either case with basically the same parametric formulae and the only difference is the parity of a and b. So, we have proved the following theorem:

Theorem 2.1.1. *Every primitive Pythagorean triple (a,b,c) with b even is given by the formula $(u^2-v^2, 2uv, u^2+v^2)$ with u and v relatively prime natural numbers of different parity and $u > v$.*

What is interesting here is that one can use this same method to show a symilar characterization of primitive Pythagorean triples in Gaussian integers (PPTGI):

triple of non-zero Gaussian integers (A, B, C), such that $A^2 + B^2 = C^2$ and

$$\gcd(A, B, C) = 1.$$

Let us remember that a Gaussian integer $z = a + bi$ is even if $N(z) = a^2 + b^2$ is even and otherwise is called odd. This concept is the same as saying the z is even iff $\alpha = 1 + i$ (a prime) divides z.

We observe that if A and B are both even then C must be even and that implies $\alpha | gcd(A, B, C) = 1$, a contradiction. Also if both A and B are odd, then the sum is even and then C is even. In this case, we can make a little change and observe that (A, iC, iB) is a primitive Pythagorean triple with the second component even. Therefore, in general we do not loose much of generality assuming that A and C are odd and B is even, just as Theorem 2.1.1.

Theorem 2.1.2. *Every primitive Pythagorean triple of non-zero Gaussian integers (A, B, C), with B even is given by the formula $(M^2 - N^2, 2MN, M^2 + N^2)$ with M and N non-zero relatively prime Gaussian integers of different parity.*

Problem 2: (Homework) *Show that similar properties happen in this situation:*

(a) one of the numbers in a PPTGI is a multiple of α^3

(b) one of the numbers in a PPTGI is a multiple of $1+2i$ and one is a multiple of $1 - 2i$

(c) there are examples in which none of the numbers in a PPTGI is divisible by 3 (a prime here two)

(d) the smallest PPTGI, in the sense that $max(N(A), N(B), N(C))$ is the smallest non-zero positive number, is essentially $(1 - 2i, 2 + 2i, 1 + 2i)$

(e) the smallest Gaussian integer which is the sum of two non-zero squares, in two different ways (not necessarily relatively prime) is

$$-6 + 8i = 2(1 + 2i)^2 = (1 + 4i)^2 + 3^2.$$

Let us use this characterization to find two Pythagorean triples with the same hypothenuse as in the case of $\mathbb{N}$. We observe that in the case of $\mathbb{N}$, in Theorem 2.1.1 the values of (u, v) are given by $(8, 1)$ and $(7, 4)$. Also, $65 = 5(13)$ is the product of the smallest two numbers which can be written as sums of two non-zero squares in two different ways and with different numbers ($50 = 7^2 + 1^2 = 5^2 + 5^2$ does not satisfy this condition). For Gaussian integers we have $1 + 2i = (1 + i)^2 + 1^2$ and $3 - 2i = (1 + i)^2 + (2 - i)^2$ are the smallest in $\mathbb{Z}[i]$ (in terms of their norm). Using Lagrange's Identity, $(a^2+b^2)(c^2+d^2) = (ac-bd)^2+(ad+bc)^2 = (ac+bd)^2+(ad-bc)^2$ we obtain that

$$7 + 4i = (2 + i)^2 + 2^2 = (2 - 3i)^2 + (4 + 2i)^2.$$

This gives the two PPTGI with the desired property: $(-1+4i, 8+4i, 7+4i)$ and $(-17-28i, 28-16i, 7+4i)$.

Problem 3: (Homework) *Show that every $z \in \mathbb{Z}[i]$, $z = a+2bi$ is a sum of two non-zero squares for all a and b, except if $a \equiv 2$ (mod 4) and $b \equiv 1$ (mod 2).*

Problem 4: (Homework)([17]) *Show that for $z \in \mathbb{Z}[i]$, $z = a+2bi$, we have $a \equiv 2$ (mod 4) and $b \equiv 1$ (mod 2) iff $(1+i)^2|z$ and $(1+i)^3 \not| z$.*

Problem 5: (Homework) Find three different writings of $z = 2017 + 2016i$ as a sum of two squares.

Problem 6: (Homework) *Find a similar theorem that characterizes all the primitive solutions of the Diophantine equation $a^2 + 2b^2 = c^2$.*

A good question here is this: *how general is this method?* Maybe another example will be helpful to understand some subtleties that may appear. We consider the case of a triangle whose sides are a, b and c and one of the angles measures $60°$. Suppose this angle is opposite to side c. Then by the Law of Cosines $c^2 = a^2 + b^2 - 2ab\cos 60°$ or

$$c^2 = a^2 - ab + b^2 \tag{2.2}$$

We can use this equation as our starting point and apply the same arguments with the scope of finding a characterization of all primitive solutions of (2.2), i.e. $a, b, c \in \mathbb{N}$ and $gcd(a, b, c) = 1$. This excludes that solutions of the form (a, a, a) with $a > 1$. Since (2.2) is symmetric in a and b, and $a = b$ is excluded, we may assume that $a < b$. We observe that this condition is equivalent to $gcd(b, c) = 1$ or $gcd(a, c) = 1$. Introducing $x = \frac{a}{c}$ and $y = \frac{b}{c}$, we get the equation of an ellipse this time (Figure 2.2): $x^2 - xy + y^2 = 1$. We have a point of rational coordinates on this curve that we can use: $A(-1, 0)$.

Introducing the slope of the line connecting A and $P(x, y)$, $x, y > 0$ given by a primitive solution of (2.2), we get $m = \frac{y}{x+1}$. Using the equation of the curve we see that

$$x = \frac{1-m^2}{1-m+m^2},\ \textit{and}\ y = \frac{2m-m^2}{1-m+m^2}.$$

From the Figure 2.2, we can see that $0 < m < 1$ (or $x > 0$). Since $m \in \mathbb{Q}$ we may write $m = \frac{u}{v}$ with $gcd(u, v) = 1$, $u < v$, $u, v \in \mathbb{N}$. Substituting in terms of u and v, we obtain

$$x = \frac{v^2-u^2}{u^2-uv+v^2},\ \textit{and}\ y = \frac{2uv-u^2}{u^2-uv+v^2}.$$

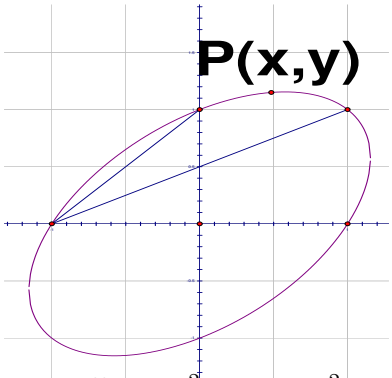


Figure 2.2: Ellipse $x^2 - xy + y^2 = 1$, $x < y$

Let p be a prime that divides $v^2 - u^2$ and $u^2 - uv + v^2$. It must divide $2v^2 - uv = v(2v - u)$. We can exclude that p divides v since this attracts p divides u contradicting $gcd(u, v) = 1$. It must be true that p divides $2v - u$. Because $4v^2 - u^2 = (2v - u)(2v + u)$ we get that p divides $4v^2 - u^2$. Hence p must divide $3v^2 = (4v^2 - u^2) - (v^2 - u^2)$. This says that p is equal to 3. A similar argument can be used to show that the second fraction is either in the reduced form or it could be simplified only by a factor of 3 (a priory a power of 3). Actually we can see that both fraction can be simplified by 3 if and only if $u = 3k \pm 1$ and $v = 3l \mp 1$. So, if $u + v$ is not a multiple of three the two fractions are in the reduced form and therefore

$$a = v^2 - u^2,\ b = 2uv - u^2 \ and\ c = u^2 - uv + v^2.$$

Let us see what happens if, for instance, $u = 3k + 1$ and $v = 3l - 1$. We notice that $\frac{u+v}{3} = s \in \mathbb{N}$ and $\frac{2v-u}{3} = t \in \mathbb{N}$. This gives $v = s + t$ and $u = 2s - t$ with $gcd(s, t) = 1$. Substituting into the fractions for y and x we obtain

$$x = \frac{2st - s^2}{s^2 - st + t^2},\ and\ y = \frac{2st - t^2}{s^2 - st + t^2}.$$

We observe that these fractions cannot be simplified by a 3 this time, because that will imply that $s \equiv \pm 1 \pmod 3$ and $t \equiv \mp 1 \pmod 3$ which in turn gives u and v divisible by 3. So in this case

$$a = 2st - s^2,\ b = 2st - t^2 \ and\ c = s^2 - st + t^2.$$

This proves the following characterization.

Theorem 2.1.3. *Every primitive solution (a, b, c) with $0 < a < b$ of the Diophantine equation*

$$a^2 - ab + b^2 = c^2$$

$(1,1,1)$	$(3,8,7)$	$(5,8,7)$	$(7,15,13)$	$(8,15,13)$	$(16,21,19)$
$(5,21,19)$	$(11,35,31)$	$(24,35,31)$	$(7,40,37)$	$(33,40,37)$	$(13,48,43)$
$(35,48,43)$	$(16,55,49)$	$(39,55,49)$	$(9,65,61)$	$(56,65,61)$	$(32,77,67)$
$(45,77,67)$	$(17,80,73)$	$(63,80,73)$	$(40,91,79)$	$(51,91,79)$	$(11,96,91)$
$(85,96,91)$	$(80,99,91)$	$(19,99,91)$	$(55,112,97)$	$(57,112,97)$	

Table 2.2: Primitive solutions of $a^2 - ab + b^2 = c^2$, with $a < b$ and c less than 100.

is in one of the two forms:

$$\begin{gathered} a = v^2 - u^2,\ b = 2uv - u^2 \text{ and } c = u^2 - uv + v^2,\ \text{with } v > u,\ \text{or} \\ a = 2uv - v^2,\ b = 2uv - u^2 \text{ and } c = u^2 - uv + v^2,\ \text{with } 2v > u > v/2, \end{gathered} \tag{2.3}$$

where $u, v \in \mathbb{N}$, such that $gcd(u, v) = 1$, $u + v \not\equiv 0$ (mod 3). Conversely, every triple given by (2.3) is a primitive solution of (2.2).

The primitive solutions of (2.2) with $c < 100$ are listed in Table 2.2.

We include one more example like this which was proved in [13]:

Theorem 2.1.4. *For every positive integers l and k such that, $\gcd(k, l) = 1$ and k is odd, then a, b and c given by*

$$c = 2l^2 + k^2 \text{ and } \begin{cases} b = |2l^2 + 2kl - k^2|,\ a = |k^2 + 4kl - 2l^2|,\ if\ k \not\equiv l \pmod 3 \\ a = |2l^2 - 2kl - k^2|,\ b = |k^2 - 4kl - 2l^2|,\ if\ k \not\equiv -l \pmod 3 \end{cases}$$

is a primitive solution for $a^2 + 2b^2 = 3c^2$. Conversely, with the exception of the trivial solution $a = b = c = 1$, every primitive solution for $a^2 + 2b^2 = 3c^2$ appears in the way described above for some l and k.

There are exceptional situations when this method cannot be applied. As an example, let us see what happens with the Diophantine equation $2a^2 + 3b^2 = c^2$. We cannot use the same technique as before since $(-1, 0)$ is not on the curve $2x^2 + 3y^2 = 1$. As a matter of fact there is no point of rational coordinates on this curve that we can use instead. The idea of proof is to use the trick that we have seen already about the classification of integers modulo 3. Without loss of generality, suppose we

have a solution (a, b, c) of $2a^2 + 3b^2 = c^2$ which is also primitive. Notice that if we have a nonzero solution then we have a primitive one. Then if a is a multiple of 3 we see that 3 divides c^2. This implies 3 divides c. Hence $a = 3a'$ and $c = 3c'$ which attracts $6a'^2 + b^2 = 3c'^2$. Hence, 3 must divide $3c'^2 - 6a'^2 = b^2$ which leads us to a contradiction: 3 divides a, b and c. It remains that a is not a multiple of three. Thus, $a = 3k \pm 1$ and so $a^2 = 3l + 1$. Then $2a^2 + 3b^2 = 3s + 2$ but c^2 cannot be of the form $3s + 2$.

Problem 7: (Homework) *Prove that the equation* $3a^2+5b^2 = c^2$ *has no integer solution other than the trivial one:* $a = b = c = 0$.

As a classical notation for the greatest common divisor of two integers u and v (or even more then two but finitely many) not both (all) zero, we will use $gcd(a, b)$. For instance using the prime factorization one can check that $gcd(2012, 3521) = 503$.

A book dealing only with the topic of Pythagorean triples is [25]. There are ramifications of this topic that go deep into the theory of abstract algebra and analysis. Let us just state the following three facts, without proof, that one can read more about in [16]:

$$\lim_{x\to\infty} \frac{\#\{(a,b,c) \in \mathbb{Z}^3 | 0 < a < b < c \leq x, a^2 + b^2 = c^2, gcd(a,b,c) = 1\}}{x} = \frac{1}{2\pi}, \tag{2.4}$$

$$\begin{aligned} &\lim_{x\to\infty} \frac{1}{x} \#\{(a,b,c) \in \mathbb{Z}^3 | 0 < a < b < c, a^2 + b^2 = c^2, \\ &a + b + c \leq x, gcd(a,b,c) = 1\} = \frac{\ln 2}{\pi^2}, \end{aligned} \tag{2.5}$$

and

$$\begin{aligned} &\lim_{x\to\infty} \frac{1}{x} \#\{(a,b,c) \in \mathbb{Z}^3 | 0 < a < b < c, a^2 + b^2 = c^2, \\ &ab \leq 2x, gcd(a,b,c) = 1\} = \frac{\Gamma(1/4)}{\pi^2\sqrt{2\pi}}, \end{aligned} \tag{2.6}$$

where $\Gamma(x) = \int_0^\infty e^{-t}t^{x-1}dt$ defined for $x > 0$.

A nice undergraduate project could be to find the equivalent of these statements for the primitive solutions of $a^2 - ab + b^2 = c^2$ as characterized in Theorem 2.1.3. Let us see what is the relevance of (2.4). We counted all the pythagorean

triples which are primitive disregarding the order of legs with the hypothenuse less than or equal to 100 and we got 16. Then $\frac{16}{100}$ must be close to $\frac{1}{2\pi}$. Actually this is a very good approximation since $|\frac{16}{100} - \frac{1}{2\pi}| \approx 0.00084$.

Finally another interesting question here is to find Pythagorean triples that have the same hypothenuse. From Table 2.1 we see for instance that $65^2 = 16^2 + 63^2 = 33^2 + 56^2$. So, we may ask if one can find such examples with an arbitrary number of representations. In view of Theorem 2.1.1, this question leads us into a big topic in number theory: representation of numbers as sums of two squares which we will study later.

2.2 Linear Diophantine Equations

Quotation: *"I'm not a religious man, but it's almost like being in touch with God when you're thinking about mathematics. Learning mathematics is always extraordinarily hard work – reading it, listening to lectures. I enjoy a kind of mathematical "gossip", when people sit in easy chairs with their feet up and tell me their mathematics; then I can learn." – Paul R. Halmos, Want To Be A Mathematician: An Automathography, Springer-Verlag, 1985.)*

Notions, concepts, definitions, and theorems: *Characterization, various facts, some problems and some curiosities*

One other classical equation that has a definite answer in number theory is the linear one:

$$ax + by = c, \text{where } a,\ b,\ \text{and } c \text{ are integers not all zero.} \tag{2.7}$$

Let us observe that if there exist a solution (x_0, y_0) then there are infinitely many solutions:

$$x = x_0 + bt \text{ and } y = y_0 - at \text{ with } t \in \mathbb{Z}.$$

If we look at the equation $2x+4y = 3$, say, we see that $3 = 2(x+2y)$. Because 3 is not divisible by 2 this equation is impossible in integers. So, if a number $d > 1$ divides a and b and it does not divide c then there is no solution of this equation. Therefore, it is necessary for the existence of a solution that $gcd(a, b)$ divides c. Next we need a preliminary result about the greatest common divisor of two numbers which is well defined only for two integers not both zero.

Proposition 2.2.1. *Suppose a, b, c and d are integers such that $ad - bc = 1$. Then $gcd(u, v) = gcd(au + bv, cu + dv)$ for all integers u and v (not both zero).*

PROOF. Let us denote by $d_1 = gcd(u, v)$ and $d_2 = gcd(au+bv, cu+dv)$. Clearly d_1 divides $au+bv$ and $cu+dv$ and so $d_1 \leq d_2$ by the definition of the greatest common divisor. Similarly since the system

$$\begin{cases} au + bv = U \\ cu + dv = V \end{cases}$$

can be solved for u and v and get $u = dU - bV$ and $v = aV - cU$ we see that d_2 divides u and v. Hence $d_2 \leq d_1$. By the trichotomy property of integers we get $d_1 = d_2$. ■

Corollary 2.2.2. *Let $x \in \mathbb{Z}$ be arbitrary. For integers u and v not both zero, we have $gcd(u, v) = gcd(u, v + xu)$.*

PROOF. We let $a = 1$, $b = 0$, $c = x$ and $d = 1$ in the Theorem 2.2.1. ■

Let us use the method of strong mathematical induction to regain the **Bézout Lemma** in the case of co-prime numbers:

Lemma 2.2.3. *Two integers a and b are relatively prime (or coprime) if and only if there exists two integers x and y such that $ax + by = 1$.*

PROOF. For the sufficiency part of this theorem, certainly if $ax + by = 1$ holds then any positive common divisor of a and b divides $ax + by$ and so it must be equal to 1. Hence $gcd(a, b) = 1$. For necessity first let us observe that we can assume that a and b are coprime natural numbers. Then we use strong induction on $k \geq max(a, b)$.

Basis step: $[k = 1]$ This forces $a = b = 1$ and the statement is true if $x = 1$, $y = 0$.

Inductive step: $k \geq 1$ Suppose the statement is true for every coprime natural numbers a and b with $a, b \leq k$. Consider two natural numbers A and B such that $max(A, B) = k + 1$. Clearly A and B cannot be equal and we can say there is an order one them: $A < B = k+1$. Hence $A \leq k$. Then if we set $C = B-A$ we see that we see that $C \leq k$. By Corollary 2.2.2 we see that $1 = gcd(A, B) = gcd(A, B - A)$ $(x = -1)$. So, we can use the induction hypothesis on A and C: there exits $x', y' \in \mathbb{Z}$ such that $Ax'+Cy' = 1$. But this implies $Ax'+(B-A)y' = 1$ or $A(x'-y')+By' = 1$. Therefore, the conclusion follows by taking $x = x' - y'$ and $y = y'$. ■

Problem 8.(Homework) *Show that for two integers a and b not both zero we have*

$$gcd(a, b) = min\{ax + by | ax + by > 0 \text{ and } x, y \in \mathbb{Z}\}.$$

Problem 9. *As a corollary of this fact show that if $m \in \mathbb{N}$ and $a, b \in \mathbb{Z}$ are not both zero then $gcd(ma, mb) = mgcd(a, b)$.*

Corollary 2.2.4. [Inverse modulo m]: *If $gcd(a, m) = 1$ then there exists x such that $ax \equiv 1$ (mod m).*

At this point we can show:

Theorem 2.2.5. *Assume a and b are positive integers. The equation (2.4) has solutions if and only if $gcd(a, b)$ divides c. In case this last condition happens every solution of (2.4) is given by*

$$x = x_0 + \frac{b}{gcd(a,b)}t,\ y = y_0 - \frac{a}{gcd(a,b)}t, t \in \mathbb{Z}, \tag{2.8}$$

where (x_0, y_0) is a particular solution. The formula (2.8) gives a family of infinitely many solutions of (2.4) called the general solution.

Proof. It is clear that if there is a solution then $d = gcd(a, b)$ must divide c. If on the other hand $d|c$ then $c = dk$ for some $k \in \mathbb{Z}$. If $k = 0$ then a particular solution of (2.4) is $x_0 = y_0 = 0$. Dividing by d we get the equation $a'x + b'y = 0$ where $a = da'$, $b = db'$ with $gcd(a', b') = 1$. So, if (x, y) is an arbitrary solution then a' divides $-b'y$. This implies a' divides y. Similarly b' divides x. Then $x = b't$ and $y = a's$. Hence $a'b'(t + s) = 0$ which gives $s = -t$. Therefore (2.8) is satisfied for any solution of (2.4).

If c is not equal to zero, then if $c = dc'$ the equation becomes $a'x + b'y = c'$ with $gcd(a', b') = 1$. We know that there exists $x', y' \in \mathbb{Z}$ such that $a'x' + b'y' = 1$. Then a particular solution can be obtained: $x_0 = x'c'$ and $y_0 = y'c'$. The rest of the theorem follows the same way as above by observing that if (x, y) is a general solution then $(x - x_0, y - y_0)$ is a solution of (2.4) in which $c = 0$. ■

Example: Say we want to solve the Diophantine equation $94x + 17y = 5$. We use the Euclidean Algorithm. Since $94 = 17(5) + 9$, $17 = 9 + 8$ and $9 = 8 + 1$. Substitute the remainder of the second to last into the last equation we get $9 = (17 - 9) + 1$ or $9(2) = 17 + 1$. Substitute the remainder of the first equation into the last one we just got we obtain $[94 - 17(5)](2) = 17 + 1$ or $94(2) - 17(11) = 1$. So, a particular solution of the given equation is $(10, -55)$. The general solution can be expressed by $x = 10 - 17t$ and $y = 94t - 55$, $t \in \mathbb{Z}$.

The following theorem is well known under the name of Fermat's Little Theorem and it can be proved in various ways. One can try to prove this statement by induction using the binomial formula.

Theorem 2.2.6. (Fermat's Little Theorem) *For $a \in \mathbb{N}$ and p a prime number, then*

$$a^p \equiv a \pmod{p}.$$

Here is yet another proof that we learned from Arthur Engel's book [5].

Proof. This is a combinatorial proof. Let us count the number of necklaces that have p stones and each stone may be of a different colors. We can think of building them: so, the first stone can be of a different colors, the second can be also of a different colors, and so on. This gives a^p choices but lots of them give the same necklaces. As an example let's say $p = 3$ and $a = 2$. We build a necklaces with black (B) and white stones (W). So if we build the necklaces say WBB, the order BBW and BWB will basically give the same necklace. The only exceptions to this cycling procedure is if we start we BBB or WWW (also see Problem 10 below). So, in general if we take away the number of choices that are build with stones of the same color, i.e. a of them, what is left needs to be divisible by p which is the number of distinct circular cycling that will give the same necklace. Hence $a^p - a$ must be divisible by p. ■

Problem 10 (Homework) *How does the hypothesis of p, being prime, factors out into this proof?*

A corollary of this theorem, which is commonly called the same way and it is clearly the nontrivial part of the theorem, is stated in the following way.

Corollary 2.2.7. *For $a \in \mathbb{N}$ and p a prime number such that $\gcd(a, p) = 1$, then $a^{p-1} \equiv 1 \pmod{p}$.*

The idea of looking for some sort of converse of Corollary 2.2.7 had given a series of rich developments in number theory. For example, if an odd number p has the property that $2^{p-1} \equiv 1$ (mod p), should p be a prime? The first number n for which $2^{n-1} \equiv 1$ (mod n) and n is not a prime is $n = 341 = (11)(31)$. To see that 341 divides $2^{340} - 1$ we see that

$$2^{340} - 1 = (2^{10} - 1)(2^{330} + ... + 1) \text{ and } 2^{10} - 1 = 1023 = 341(3).$$

Other numbers with this property, in order, are 561, 645, 1105, 1387, 1729, 1905, etc. This defines a sequence of integers which is A001567 (pseudoprimes base 2) in the The On-Line Encyclopedia of Integer Sequences ([20]).

If we require the property that for every a, such that $\gcd(p, a) = 1$, to have $a^{n-1} \equiv 1$ (mod p), and if p is composite, it turns out that this is still possible and we get a sequence called the Carmichael pseudoprimes numbers (A002997, [20]): 561, 1105,

1729, 2465, 2821, 6601, 8911, etc. Pseudoprimes are of primary importance in public-key cryptography.

2.3 Representations as sums of two squares

The first result that we need is the fact that we can solve a quadratic congruency modulo a prime only in the trivial way:

Theorem 2.3.1. *Let p be a prime number. Then the quadratic equation $x^2 \equiv 1$ (mod p) has only "two" solutions: $x \equiv \pm 1$ (mod p).*

Proof. The equation is equivalent to saying that p divides $x^2 - 1 = (x - 1)(x + 1)$. Hence by Euclid's Lemma 1.3.3 p must divide $x - 1$ or $x + 1$. Hence $x \equiv \pm 1$ (mod p). Conversely if $x \equiv \pm 1$ (mod p) we square this congruency and obtain $x^2 \equiv 1$ (mod p). ■

In other words, this theorem is saying that between the numbers $\{0, 1, 2, \ldots, p-1\}$ there are only two that satisfy the equation $x^2 \equiv 1$ (mod p): 1 and $p - 1$. This idea gives the following theorem:

Theorem 2.3.2. (Wilson's Theorem): *If p is a prime, we have $(p-1)! \equiv -1$ (mod p).*

Proof. We look at the numbers $A := \{1, 2, \ldots, p-1\}$ and for each $a \in A$, by Corollary 2.2.4 there exits x such that $ax \equiv 1$ (mod p). Dividing x by p we may take its remainder instead of x. This remainder cannot be zero. In other words, we can say that x can be chosen in A. Let us denote this number in A by $\overline{a}$. For some numbers a this association may turn out to be the same as a: those for which $aa \equiv 1$. By Theorem 2.3.1 there are only two numbers with this property: 1 and $p - 1$. So, if we take the product of all numbers in A with the exception of these two numbers, they can be grouped together, in the following way:

$$1(2)(3)\ldots(p-1) \equiv 1(p-1)(a_1\overline{a_1})\ldots(a_k\overline{a_k}) \equiv p - 1 \equiv -1 \ (mod\ p),$$

or $(p-1)! \equiv -1$ (mod p). ■

An important consequence of this theorem is the following fact about quadratic congruences $x^2 \equiv -1$ (mod p) with p prime.

Theorem 2.3.3. *Consider a prime number p. The equation $x^2 \equiv -1$ (mod p) has solutions if and only if $p = 2$ or $p \equiv 1$ (mod 4).*

Let us investigate this numerically. If $p = 5$, we get $x = \pm 2$. Or, if $p = 61$, then we need to search a little to find $x = \pm 11$. We will see a simple method by which x can be calculated.

PROOF. Let us assume first that $p = 4k + 3$ and by way of contradiction let x be a solution of $x^2 \equiv -1$ (mod p). Since $p - 1 = 4k + 2 = 2(2k + 1)$ implies that

$$x^{p-1} \equiv (x^2)^{(2k+1)} \equiv (-1)^{2k+1} \equiv -1 \ (mod\ p).$$

Let us observe that we must have $gcd(x, p) = 1$ and so, by Fermat's Little Theorem 2.2.6, $x^{p-1} \equiv 1$ (mod p). So, $-1 \equiv 1$ (mod p). This is possible only if $p = 2$ but 2 is not of the form $4k + 3$. Hence if $p \equiv 3$ (mod 4) the equation given has no solution.

On the other hand if $p = 4k + 1$ then Wilson's Theorem 2.3.2 implies that

$$-1 \equiv (p - 1)! \equiv 1(2) \dots (2k)(2k + 1) \dots (4k) \ (mod\ p).$$

So, if we set $x = 1(2)(3) \dots (2k)$ since $j \equiv -(4k - j + 1)$ (mod p) for $j = 1, 2, \dots, 2k$, we have $x^2 \equiv -1$ (mod p). ■

We are ready to show another theorem that was proven first by Fermat but known since 1632 (375 years ago Albert Girard stated it on the basis of numerical evidence but the first proof was given by Fermat, see [19], page 54)

Theorem 2.3.4. (Fermat) *Let p be a prime such that $p \equiv 1$ (mod 4) or $p = 2$. Then there is a representation of p as sum of two perfect squares: $p = x^2 + y^2$ with $x, y \in \mathbb{N}$ and $x < y$. This representation is unique.*

There are various proofs of this important theorem. One of the simplest proofs is described next and we learned it from [19].

PROOF. For $p = 2$ we clearly have $p = 1^2 + 1^2$. By Theorem 2.3.3 we can find an x such that $x^2 \equiv -1$ (mod p). We look at the function $f(a, b) = a + bx \ (mod\ p)$ defined on $A \times A$ with values in $B := \{0, 1, \cdots p - 1\}$, where $A = \{0, 1, ..., k\}$ and k is the positive integer such that $k^2 < p < (k + 1)^2$. This gives a function from a set with $(k + 1)^2$ values into a set with p values. Since $(k + 1)^2 > p$, by the Pigeonhole principle we must have at least one of the values of the function f attained for two different inputs: (a, b), (a', b'). Hence $a + bx \equiv a' + b'x$ (mod p). If we set $a - a' = u$ and $b' - b = v$ we get $u \equiv xv$ (mod p) and $u, v \in [-k, k]$. This implies $u^2 + v^2 \equiv v^2(x^2 + 1) \equiv 0$ (mod p). But $0 < u^2 + v^2 \leq 2k^2 < 2p$. There is only one integer divisible by p in the interval $(0, 2p)$. Therefore $u^2 + v^2 = p$.

For uniqueness we follow the steps recommended in [22], page 132, Exercises 11-13. This is in fact, going back to Euler's idea to factor a number which can

p	5	13	17	29	37	41	53	61	73	89
(x,y)	(1,2)	(2,3)	(1,4)	(2,5)	(1,6)	(4,5)	(2,7)	(5,6)	(3,8)	(5,8)

97	101	109	113	137	149	157	173	181	197
(4,9)	(1,10)	(3,10)	(7,8)	(4,11)	(7,10)	(6,11)	(2,13)	(9,10)	(1,14)

Table 2.3: Primes of the form $4k+1$ and their representation $p = x^2 + y^2$, $x < y$

be written as a sum of two squares in two different ways. Let us assume that $p = a^2 + b^2 = c^2 + d^2$ are two different representations of an odd prime p (the representation of $p = 2$ is clearly unique). We must have a and b of different parity and so we may assume a, c are odd and b, d are even. This implies that $u = gcd(a-c, b-d)$ is even and well defined because the representations are different. We set $a-c = ru$ and $d-b = us$ for some $r, s \in \mathbb{Z}$ not zero and so $gcd(r, s) = 1$. We have $a^2-c^2 = d^2-b^2$ or $(a-c)(a+c) = (d-b)(d+b)$ which implies $r(a+c) = s(d+b)$. By Lemma 1.3.3 we know that s must divide $a+c$. Then we let $a+c = sv$ for some $v \in \mathbb{Z}$. This says that $d+b = rv$ and $v = gcd(a+c, b+d)$ and v is an even number because $a+c$ and $b+d$ are even. Then one can check that $p = [(\frac{u}{2})^2 + (\frac{v}{2})^2](r^2 + s^2)$. This is a contradiction since p is a prime number. ■

The first twenty primes of the form $4k+1$ together with their unique representation as sum of two squares is ($p = x^2 + y^2$, $0 < x < y$) is included in Table 2.3.

Corollary 2.3.5. *Every Gaussian prime is in one of the following forms, or their associates:*

(a) $1 + i$ *(notice that* $1 - i = (-i)(1 + i)$*)*

(b) for every prime (in $\mathbb{N}$*) of the form* $p = 4k+1$, $p = a^2+b^2$ *(as in Theorem 2.3.4), we have two different primes* $z_p = a + bi$ *and* $z'_p = b + ai$

(c) for every prime (in $\mathbb{N}$*) of the form* $p = 4k + 3$*, this is also a prime in* $\mathbb{Z}[i]$*.*

Problem 11: *Prove this corollary.*

We are going to state Fermat's Little Theorem for the case of Gaussian integers and use a different idea for the proof.

Theorem 2.3.6. *For* $a \in Z[i]$ *and* q *a prime in* $\mathbb{Z}[i]$ *such that* $\gcd(a, p) = 1$*, then* $a^{N(q)-1} \equiv 1$ *(mod* q*).*

Proof. Let $\mathcal{R}C_q$ a complete set of residues modulo q (as in Theorem 1.2.5, for example). Let us take away the zero residue: $\mathcal{R}C_q^\star := \mathcal{R}C_q \setminus \{0\}$. We define a map

$S_a : \mathcal{R}C_q^\star \to \mathcal{R}C_q^\star$, by $S_a(x) \equiv ax \pmod{q}$, for all $x \in \mathcal{R}C_q^\star$. We observe that $ax \equiv 0 \pmod{q}$ implies q divides a or x and both of these options are excluded by our hypothesis and for $x \in \mathcal{R}C_q^\star$. Hence, S_a is well-defined. Also, we said earlier that we will show that the cardinality of $\mathcal{R}C_q$ is $N(q)$ and so the number of elements in $\mathcal{R}C_q^\star$ is $N(q) - 1$. The function S_a is one-to-one, because if $S_a(x) = S_a(y)$ then $ax \equiv ay \pmod{q}$ which implies $a(x - y) \equiv 0 \pmod{q}$. Therefore, $x \equiv y \pmod{q}$. But for x and y in $\mathcal{R}C_q$, this attracts $x = y$. This implies that S_a is actually a bijection, or in other words, it permutes the elements of $\mathcal{R}C_q^\star$. Therefore, $S_a(\mathcal{R}C_q^\star) = \mathcal{R}C_q^\star$ and so taking the product P of all of the elements in $\mathcal{R}C_q^\star$ gives

$$P = \prod_{x \in \mathcal{R}C_q^\star} S_a(x) \equiv \prod_{x \in \mathcal{R}C_q^\star} (ax) \equiv a^{N(q)-1} P \pmod{q}.$$

Obviously, since $\gcd(q, P) = 1$ we can simplify by P and obtain the desired conclusion. ■

Conventionally we have two types of unique factorizations: one in which the primes are written in nondecreasing order all to power 1, and one representation in which the primes appear only one time with a unique positive integer exponent. In the later case, the primes do not repeat. For instance, the first representation of 441000 is $441000 = (2)(2)(2)(3)(3)(5)(5)(5)(7)(7)$ as opposed to the second unique writing $441000 = (2^3)(3^2)(5^3)(7^2)$. The last one will be called here the **canonical** representation.

Theorem 2.3.7. [**Fermat**] *If the canonical representation of $n > 1$ is given by*

$$n = 2^\alpha \left(\prod_{p \equiv 1 \pmod 4} p^\beta \right) \left(\prod_{p \equiv 3 \pmod 4} p^\gamma \right)$$

then n can be represented as a sum of two squares if and only if all the γ's are even exponents.

Proof. Suppose n is of the form stated. Then each prime of the form $4k + 1$ can be written as a sum of two squares and the other factors are simply squares or (e.g. the factor 2^α) a sum of two perfect squares. Using the identity

$$(a^2 + b^2)(c^2 + d^2) = (ac + bd)^2 + (ad - bc)^2 \tag{2.9}$$

we see that the product of two numbers which are sums of squares is also a sum of two squares.

On the other hand, by way of contradiction let us say there is an odd γ corresponding to q prime of the form $4k+3$ and $n = x^2 + y^2$. Without loss of generality we may assume that this writing is simplified by any even power of q or in other words q does not divide x or y. Since γ is assumed odd we cannot make q disappear as a factor of n after such simplifications and then $x^2 + y^2 \equiv 0 \pmod{q}$. Since q does not divide x we can find a multiplicative inverse of x modulo q, say $\overline{x}$, and then $1 + t^2 \equiv 0 \pmod{q}$ with $t = y\overline{x}$. But this is impossible according to Theorem 2.3.3. ■

We are going to address the number of representations, r_2', of a number as a sum of two squares next counting only representations of the form $a^2 + b^2$, with $0 < a \leq b$. For the canonical representation

$$n = 2^{\alpha} \left(\prod_{p_i \equiv 1 \ (mod\ 4),\ p_i | n} p^{\beta_i} \right) \left(\prod_{p \equiv 3 \ (mod\ 4),\ p | n} p^{\gamma} \right)$$

let us define

$$d(n) = \prod_i (\beta_i + 1).$$

Theorem 2.3.8. [Euler (1738)] *The number of representations of n, as the sum of two squares of natural numbers, ignoring order, is*

$$r_2'(n) = \begin{cases} \frac{d(n)}{2} & \text{if } d(n) \text{ is even} \\ \\ \frac{d(n)+(-1)^{\alpha+1}}{2} & \text{if } d(n) \text{ is odd.} \end{cases} \tag{2.10}$$

We refer the reader to [19] or [22] for the ideas of a proof.

Examples: For instance, if $n = 5^2$ we have $n = 3^2 + 4^2$ and $2(5^2) = 1^2 + 7^2 = 5^2 + 5^2$ so $r_2'(25) = 1$ and $r_2'(50) = 2$. For $n = 325 = 5^2(13)$, $d(n) = 6$ and so $r_2'(n) = 3$: $325 = 1^2 + 18^2 = 6^2 + 17^2 = 10^2 + 15^2$.

One interesting thing related to these representations is that there are infinitely many primes even of the form $x^2 + y^4$ (see [7]). There are other results similar to the one in Theorem 2.3.4. We include here a series of such results which one can read more about in D. Cox book ([4]).

Theorem 2.3.9. [14] *For an odd prime p we have*

(i) $p = x^2 + 2y^2$ for some integers x, y if and only if $p \equiv 1$ or $3 \pmod{8}$;

(ii) $p = x^2 + 3y^2$ for some integers x, y if and only if $p = 3$ or $p \equiv 1 \pmod{3}$;

(iii) $p = x^2 + 4y^2$ for some integers x, y if and only if $p \equiv 1 \pmod{4}$;

(iv) $p = x^2 + 5y^2$ *for some integers* x, y *if and only if* $p = 5$ *or* $p \equiv 1$ *or* 3^2 *(mod 20);*

(v) $p = x^2 + 6y^2$ *for some integers* x, y *if and only if* $p \equiv 1$ *(mod 6);*

(vi) $p = x^2 + 7y^2$ *for some integers* x, y *if and only if* $p = 7$ *or* $p \equiv 1$, 3^2 *or* 5^2 *(mod 14);*

(vii) $p = x^2 + 8y^2$ *for some integers* x, y *if and only if* $p \equiv 1$ *(mod 8);*

(viii) $p = x^2 + 9y^2$ *for some integers* x, y *if and only if* $p \equiv 1$ *(mod 12);*

(ix) $p = x^2 + 10y^2$ *for some integers* x, y *if and only if* $p \equiv 1$ *or* 3^2 *(mod 10);*

(x) $p = x^2 + 11y^2$ *for some integers* x, y *if and only if* $p \equiv 1$, 3^2, 5^2, 7^2 *or* 9^2 *(mod 22);*

(xi) $p = x^2 + 14y^2$ *for some integers* x, y *if and only if the equations* $x^2 \equiv -14$ *and* $(x^2 + 1)^2 \equiv 8$ *(mod* p*) have solutions;*

(xii) $p = x^2 + 27y^2$ *for some integers* x, y *if and only if* $p \equiv 1$ *(mod 3) and the equation* $x^3 \equiv 2$ *(mod* p*) has a solution.*

Now, how can one obtain a result like the one in Theorem 2.3.9 (i)? Are our methods good enough for this task? We need first a result as in Theorem 2.3.3, whose proof is based on Gauss's and Euler's ideas.

Theorem 2.3.10. *Let us consider* p *and odd prime. The equation* $x^2 \equiv -2$ *(mod* p*) has solutions if and only if* $p \equiv 1$ *or* 3 *(mod 8).*

Proof. We want to show the necessity first. So, we assume that the equation $x^2 \equiv -2$ (mod p) has at least one solution, say x_0. It is clear that $\gcd(x_0, p) = 1$. Then by Fermat's Little Theorem we have

$$(-2)^{\frac{p-1}{2}} \equiv x_0^{p-1} \equiv 1 \ (mod\ p). \tag{2.11}$$

Let us introduce a few sets of residues: $A := \{1, 2, 3, ..., p-1\}$, $B := \{2, 4, 6, ..., p-1\}$ the set of all even residues in A and $C := \{1, 3, 5, ...\}$ all odd residues. We obviously have $A = B \cup C$. We consider now the map: $g : A \to A$, $g(x) = p - x$ for all $x \in A$. Clearly $g(B) = C$ and $g(C) = B$ (because p is an odd number). Also, let us split $B = B_1 \cup B_2$ and $C = C_1 \cup C_2$, where $B_1 = \{x \in B | x \leq \frac{p-1}{2}\}$, $B_2 = \{x \in B | x \geq \frac{p+1}{2}\}$, $C_1 = \{x \in C | x \leq \frac{p-1}{2}\}$, and $C_2 = \{x \in C | x \geq \frac{p+1}{2}\}$. We also observe that $g(B_2) = C_1$. Next, let us define r to be the number of elements in B_2 and let s be the number of elements in B_1.

It is easy to check that

$$\begin{cases} r = 2k, s = 2k \ if \ p = 8k+1 \\ r = 2k+1, \ s = 2k \ if \ p = 8k+3 \\ r = 2k+1, \ s = 2k+1 \ if \ p = 8k+5 \\ r = 2k+2, \ s = 2k+1 \ if \ p = 8k+7. \end{cases} \tag{2.12}$$

We observe that we always have $r + s = \frac{p-1}{2}$, the number of elements in B. We are going to take the product of all elements in B:

$$R = \prod_{2i \in B} (2i) = \prod_{2i \in B_1} (2i) \prod_{2i \in B_2} (2i) = \prod_{2i \in B_1} (2i) \prod_{2i \in B_2} g^2(2i) \Rightarrow$$

$$R \equiv (-1)^r \prod_{2i \in B_1} (2i) \prod_{2i \in B_2} (p - 2i) = (-1)^r \prod_{j \in A, \ j \leq \frac{p-1}{2}} j.$$

On the other hand $R = 2^{\frac{p-1}{2}} \prod_{j \in A, \ j \leq \frac{p-1}{2}} j$ which by (2.11) becomes

$$R \equiv (-1)^{\frac{p-1}{2}} \prod_{j \in A, \ j \leq \frac{p-1}{2}} j.$$

Comparing the two congruencies involving R, we see that $(-1)^r \equiv (-1)^{\frac{p-1}{2}} \pmod{p}$. This implies $p = 8k+1$ or $p = 8k+3$.

For the other implication let us use an argument by way of contradiction. So, we assume there is no solution for the equation $x^2 = -2 \pmod{p}$. We employ the same idea about pairing the elements of the set $A := \{1, 2, 3, ..., p-1\}$ in a similar way but this time taking into account a different map: $f : A \to A$, $f(x) \equiv -2\overline{x} \pmod{p}$ for all $x \in A$. This map is well defined (p is not equal to 2). Also, let us observe that $f(f(x)) = x$ for all $x \in A$. Indeed, using some obvious properties of the inverse element modulo p, we have:

$$-2\overline{(-2)\overline{x}} \equiv 2\overline{2}x \equiv x \ (mod \ p), for \ all \ x \in A.$$

A map with this property ($f \circ f = id$) is called an *involution*. The map f has no fixed points; in other words, there is no $x \in A$ such that $f(x) = x$, because this implies $x^2 \equiv -2 \pmod{p}$. So, the elements of A can be paired as $(a, f(a))$,

with $a < f(a)$, and we denote the set of the smaller residues in a pair by T. From Wilson's Theorem, we get that

$$1(2)(3)...(p-1) \equiv -1 \ (mod\ p) \ \ or \prod_{a \in T}(af(a)) \equiv -1 \ (mod\ p)$$

or

$$(-2)^{\frac{p-1}{2}} \equiv -1 \ (mod\ p). \tag{2.13}$$

So, the equality (2.11) is replaced now by (2.13). So, in the previous analysis we have $(-1)^r = -(-1)^{\frac{p-1}{2}}$ which is true only if $p = 8k+5$ or $p = 8k+7$. This is in contradiction to what we assumed for the converse part and so, it remains that we must have a solution of the given congruency. ■

Now we can prove, in a similar way as before, more than part (i) of Theorem 2.3.9.

Theorem 2.3.11. *An odd prime p can be written as $p = a^2 + 2b^2$ with a and b natural numbers with a odd, in a unique way, if and only if $p \equiv 1$ or 3 (mod 8).*

Proof. If p is not of the form mentioned in the statement, then the equality $p = a^2 + 2b^2$ implies that $\gcd(b, p) = 1$ which attracts that $\overline{b}$ exists. So, $a^2 + 2b^2 \equiv 0 \pmod{p}$ or $(a\overline{b})^2 + 2 \equiv 0 \pmod{p}$. This is impossible according to the previous theorem.

We consider as before $h : E \times E \to A$ where $E := \{0, 1, 2, ...k\}$ and $A := \{0, 1, 2, ..., p-1\}$, where $k^2 < p < (k+1)^2$, and $h(a, b) \equiv a + bx_0$ for $(a, b) \in E \times E$, with $x_0^2 = -2 \pmod{p}$. Because $|E \times E| = (k+1)^2 > p = |A|$, the Pigeon Whole Principle insures that $h(a_1, b_1) = h(a_2, b_2)$ for two different pairs (a_1, b_1), (a_2, b_2). Then, we must have $a_1 + b_1x_0 \equiv a_2 + b_2x_0 \pmod{p}$. Then we set $u = a_1 - a_2$ and $v = b_2 - b_1$. The earlier congruency implies $u \equiv vx_0$. So, let us observe that

$$0 < u^2 + 2v^2 \equiv v^2x_0^2 + 2v^2 = v^2(x_0^2 + 2) \equiv 0 \ (mod\ p).$$

On the other hand, $u^2 + 2v^2 \leq k^2 + 2k^2 < 3p$. This implies that $u^2 + 2v^2 = p$ or $u^2 + 2v^2 = 2p$. In the second case we see that $u = 2u'$ and so $p = v^2 + 2u'^2$. Therefore, in either of the two cases, $p = a^2 + 2b^2$, with a and b in $\mathbb{N}$ (p is odd).

For the uniqueness, we use the same idea of Euler's. By way of contradiction, let us assume that $p = a^2 + 2b^2 = c^2 + 2d^2$ with a and c different odd natural numbers. Automatically, we need to have $b \neq d$.

Then $(a-c)(a+c) = 2(d-b)(d+b)$. We know that $a-c$ and $a+c$ are even; so, the left hand side contains a factor of 2^2. This implies that d and b must have the same parity also, otherwise the right hand side will contain only a factor of 2 in its prime factorization.

So, as before, we set $\gcd(a-c, d-b) = 2r$, $r \geq 1$. Then, $a-c = 2ru$ and $d-b = 2rv$ with $gcd(u,v) = 1$, $|u|, |v| \geq 1$. The previous equality becomes $u(a+c) = 2v(d+b)$. Since $gcd(v,u) = 1$ we should have $a+c = vs$ ($|s| \geq 1$) and then $2(d+b) = us$. This implies in particular that 2 must divide either u or s. We observe that if s is odd, then u must be even and so v must be odd. This in contradiction with $\gcd(a+c, b+d) = \gcd(vs, s\frac{u}{2}) = s\gcd(v, \frac{u}{2}) = s$ since $a+c$ and $b+d$ are both even. Then, we should have $s = 2t$, $d+b = ut$ and $a+c = 2vt$. Now, one can check that

$$(t^2+2r^2)(u^2+2v^2) = (b+d)^2 + \frac{(a+c)^2}{2} + \frac{(a-c)^2}{2} + (d-b)^2 \Rightarrow$$

$$(t^2+2r^2)(u^2+2v^2) = a^2+c^2+2d^2+2b^2 = 2p \Rightarrow$$

p is composite, because the factors in the left hand side are more than 2. In a similar way we can deal the situation in which u is even. In this case v must be odd. In either case, p is composite and this contradiction shows that the decomposition must be unique. ■

Problem 9 *Prove in a similar way that if p is a prime number, then $p = a^2 + 3b^2$ with a, b in $\mathbb{N}$, if and only if $p \equiv 1$ (mod 3).*

2.4 Linear Diophantine Equations with positive solutions

Let us consider the Diophantine linear equation

$$ax + by = n, \tag{2.14}$$

with $a, b \in \mathbb{Z}$, $a, b > 0$, $gcd(a,b) = 1$ and solutions $x, y \in \mathbb{Z}$, $x, y \geq 0$,

Let us observe that in view of the general solution of this equation we can consider u and v the particular solution of (2.14) in which $n = 1$ ($au + bv = 1$) and such that

$0 \leq u < b$. Then v is a negative integer or zero, so we change the signs so we will assume that $au - bv = 1$ and $u, v \geq 0$. Then the formulae for the general solution of (2.14) is given by:

$$\begin{cases} x = un - bt \\ y = at - vn, \quad t \in \mathbb{Z}. \end{cases} \tag{2.15}$$

We see that in order for $x \geq 0$ and $y \geq 0$ we need to have

$$\frac{un}{b} \geq t \geq \frac{vn}{a}.$$

So, in order for such an integer t to exist, it is enough to have $\frac{un}{b} - \frac{vn}{a} \geq 1$. The reason is that every interval of real numbers of length at least one contains an integer. This inequality can be solved in terms of n: $n\frac{au-bv}{ab} \geq 1$ or $n \geq ab$. So, we proved that (2.14) has positive solution for every $n \geq ab$. For example, if we have bills in denominations of \$3 and \$5 we can express every integer amount of money greater or equal to \$15 in terms of such bills (e.g. $15 = 3(5)$, $16 = 3(2)+5(2)$, $17 = 3(4)+5$, $18 = 3(6)$, $19 = 3(3)+5(2)$, and so on). But in this particular case we can go down a few steps: $14 = 3(3)+5$, $13 = 3(1)+5(2)$, $12 = 3(4)$, $11 = 2(3)+5$, $10 = 3(0)+5(2)$, $9 = 3(3)$ and $8 = 3+5$. But \$7 cannot be paid in these denominations.

There is a general question here, known as the *money-changing problem*, the *coin problem* or under a more technical term as the *Frobenius problem*. What is the smallest number $g(a, b)$ such that (2.14) has non-negative solutions for all $n \geq g(a, b)$. In our previous example we have $g(3, 5) = 8$. We have the following general result due to J.J.Sylvester [26].

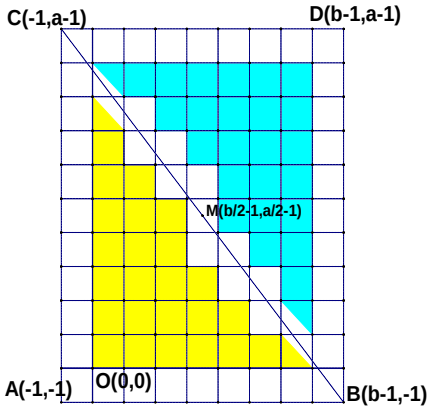
Theorem 2.4.1. [Sylvester, 1884] *The equation (2.14) has non-negative solutions for all $n \geq (a-1)(b-1)$ and no such solution if $n = ab - a - b$. (Hence, $g(a, b) = (a-1)(b-1)$.)*

Proof. Let us begin by showing that there is no solution for $n = ab - a - b$. By way of contradiction, suppose we have $ax + by = ab - a - b$ for some $x, y \geq 0$. This is equivalent to $a(x+1)+b(y+1) = ab$. Therefore a divides $y+1$ and b divides $x+1$. So if we substitute, $y + 1 = ay'$ $(x' > 0)$ and $x + 1 = bx'$ $(y' > 0)$ we get $x' + y' = 1$ which clearly has no positive integer solutions.

Let us write $n = ab - a - b + z$ where $z \geq 1$. The equation becomes

$$a(x + 1) + b(y + 1) = ab + z.$$

As discussed before, its general solution is



,

Figure 2.3: Theorem 2.4.2

$$\begin{cases} x' = x + 1 = u(ab + z) - bt \\ y' = y + 1 = at - v(ab + z).\ t \in \mathbb{Z} \end{cases} \tag{2.16}$$

Since $x = x' - 1$ and $y = y' - 1$, it is enough to show that $x' > 0$ and $y' > 0$. We know from the discussion before the theorem that a solution with $x \geq 0'$ and $y' \geq 0$ exists. We need to treat the case when $x' = 0$ or $y' = 0$. Let us assume $x' = 0$. Then $u(ab + z) = bt$. Since u is relatively prime with b it must divide t. So, $t = us$ and then $ab + z = bs$. This attracts z divisible by b. If $z = bz'$ then we have a writing for n using nonnegative coefficients: $a(b - 1) + b(z' - 1) = n$. Similar argument can be used in the case $y' = 0$.

Theorem 2.4.2. *If $gcd(a, b) = 1$, there are exactly $\frac{(a-1)(b-1)}{2}$ nonnegative integers $n < ab - a - b$ such that the equation (2.14) has a nonnegative solution. As a result, there are $\frac{(a-1)(b-1)}{2}$ nonnegative values of n that cannot be represented as $ax + by$ with $x, y \in \mathbb{Z}$, $x, y \geq 0$.*

Proof. In what follows we will refer to Figure 2.3. Consider the rectangle with vertices $A(-1, -1)$, $B(b - 1, -1)$, $C(-1, a - 1)$ and $D(b - 1, a - 1)$. Then the

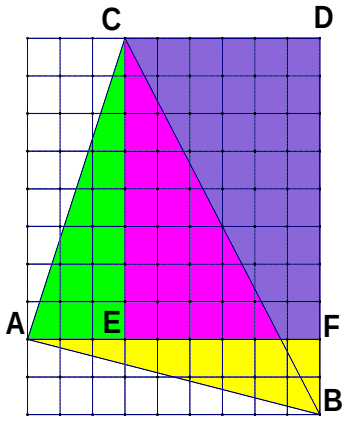


Figure 2.4: Theorem 2.4.2 and Theorem 2.4.3

diagonal $\overline{BC}$ has slope $-\frac{a}{b}$ and its equation is $aX+bY = ab-a-b$. Since, we know that there are no nonnegative solutions of this equation by the previous theorem, we conclude that there is no point of integer coordinates inside the rectangle that are on this line. So all the points of integer coordinates inside the rectangle are either above or below this diagonal.

Because of the symmetry half of them will be below and half will be above. Indeed we can use a reflection into the center of the rectangle: $O(b/2-1, a/2-1)$. This means that if (m,k) is such a point in one side of BC, then the point $(b-m-2, a-k-2)$ is going to be on the other side of BC. For each point (m,k) below BC, we can find a number $n = am+bk$ for which a nonnegative solution of (2.14) exists, namely $x = m$ and $y = k$. The other way around, if for some $n < ab-a-b$ there exists a nonnegative solution of (2.14), say (x,y), then (x,y) is unique and it is going to be in the interior of this rectangle and below BC its diagonal.

Hence the number of values of $n < ab-a-b$ for which the equation (2.14) has a nonnegative solution is given by the number of points below the diagonal BC that are in the interior of the rectangle. The total number of point inside the rectangle is $(a+1)(b+1) - 2(a+1) - 2(b-1) = (a-1)(b-1)/2$. The number of nonnegative values n for which (2.14) has no nonnegative solution is then $ab-a-b-(a-1)(b-1)/2+1 = (a-1)(b-1)/2$. ■

Problem 10 Homework: *The post office in a small town is left with stamps of only two values. They discover that there are exactly* 33 *postage amounts that*

cannot be made up using these stamps, including 46 *cents. What are the values of the remaining stamps?*

As a nice connection of this last theorem's proof with geometry is the following result known as Pick's Theorem.

Theorem 2.4.3. *Let $\mathcal{P}$ a polygon whose vertices have integer coordinates. Denote by $\#i$ the number of points of integer coordinates in the interior of this polygon and by $\#\partial$ the number of points on the sides of the polygon $\mathcal{P}$. Then the area of $\mathcal{P}$ in terms of the usual unit square area is*

$$Area(\mathcal{P}) = \frac{\#\partial}{2} + \#i - 1 \tag{2.17}$$

Proof. This is just a sketch of a proof. Let us observe that in our Figure **??** the area of the triangle ABC is equal to $\frac{ab}{2}$ and $\#\partial = a + b + 1$. Hence the formula above gives $\frac{\#\partial}{2} + \#i - 1 = \frac{a+b+1}{2} + (a-1)(b-1)/2 - 1 = ab/2 = Area(ABC)$. The formula (2.17) can be checked also to work for rectangles having sides parallel to the coordinates.

Next step is to show that the formula works in the case of a triangle in which all sides do not pass through additional points of integer coordinates. As before, we want to check that the formula (2.17) works in this case.

For instance, in the Figure **??** we have

$$Area(ABC) = Area(ABF) + Area(AEC) + Area(EFDC) - Area(CDB)$$

Using the formula already established for the areas in the right hand side of this equality we see that each point of integer coordinates that is on a common boundary and inside the triangle ABC counts twice and the formula (having a weight of $\frac{1}{2}$) is going to be account to exactly as every other interior point of ABC. The only problem is with the vertices A,B,C and the last -1 that appears in the formula. Vertex A is accounted twice and it should be accounted only one time with a weight of $\frac{1}{2}$, C is added twice and subtracted one time so there is no problem, B is added one time and subtracted in another. This is compensated by the accounting of A. The difference in a -1 is balanced out by what happens with E and F.

Every polygon has a diagonal that is contained in its interior. The formula (2.17) is invariant under gluing disjoint polygons along a side. This will help with using then an argument by induction on the number of vertices of the polygon. Hence the problem is reduced to a triangle. Another induction argument can be used, on the number of points on the sides of a triangle, to reduce the problem to the case discussed above. ■

2.5 Pell's Equation

Chapter 3

Arithmetic Functions

3.1 Euler's Totient Function

> **Quotation:** *"If people do not believe that mathematics is simple, it is only because they do not realize how complicated life is." John Louis von Neumann*

For every $n \in \mathbb{N}$ we define $\varphi(n)$ to be the number of all $k \in \{1, ..., n\}$ such that $gcd(k, n) = 1$. A more general theorem than Fermat's Little Theorem is included here:

Theorem 3.1.1. [Euler] *Let $a, n \in \mathbb{N}$, $n \geq 2$, such that $gcd(a, n) = 1$. Then $a^{\varphi(n)} \equiv 1$ (mod n).*

Proof. Let us denote by A the set of numbers less than n which are relatively prime with n. For instance, if $n = 10$ then $A = \{1, 3, 7, 9\}$. We take two numbers a and b in A. Then we multiply these numbers and then take the remainder r modulo n ($r \in \{0, 1, ..., n-1\}$). Let us show that $r \in A$.

Indeed if $r \notin A$ then $gcd(n, r) > 1$. So let us choose a prime q which divides n and r. Since $ab - r$ is divisible n we get that q divides $ab - r$. But then it must divide $ab = ab - r + r$. By Euclid's lemma q must divide a or b. Either way that contradicts the fact $a, b \in A$.

We obtain, in this way, a special operation in A which we are going to denote by $\star$. In our example, $n = 10$, we get the following table which gives this operation:

Let us prove next the following simplification rule that this operation has: $a \star b = a \star c$ implies $b = c$. Indeed, if $a \star b = a \star c$ we get $ab \equiv ac$ (mod n) or

Table 3.1: The multiplication modulo 10

$\star$	1	3	7	9
1	1	3	7	9
3	3	9	1	7
7	7	1	9	3
9	9	7	3	1

$n|a(b-c)$. Since $gcd(a,n)=1$ this implies $n|b-c$ which finally gives $b=c$ since $-(n-2) \leq b-c \leq n-2$).

Fix an element $a \in A$ and let us say $A = \{a_1, a_2, \ldots, a_{\varphi(n)}\}$. Then multiply all elements in A by a: aa_1, aa_2,...,$aa_{\varphi(n)}$. By the simplification rule, we get $\varphi(n)$ different elements in this list. That means we get all elements of A but maybe in a different order. In our example, this sequences are included as the numbers in the columns (or the rows) of the above Table 3.1. Let us say $A = \{a_1, a_2, \ldots, a_{\varphi(n)}\}$. Then

$$(aa_1)(aa_2)\cdots(aa_{\varphi(p)}) \equiv a_1a_2\cdots a_{\varphi(n)} \ (mod\ n).$$

After simplifying by $a_1a_2\cdots a_{\varphi(n)}$ we get

$$a^{\varphi(n)} \equiv 1 \ (mod\ n). \tag{3.1}$$

■

In particular, if $n=p$ with p a prime, $\varphi(p)=p-1$. Every $a \in \mathbb{N}$ is congruent with an $a' \in A = \{1,2,3,\ldots,p-1\}$ or a is divisible by p. Either way we have $a^p \equiv a$ (mod p) for every a as a result of Euler's Theorem. This shows that Fermat's Little Theorem is just a corollary of Theorem 3.1.1.

Let us find more information about this function φ around which there exists a huge literature and lots of open questions.

Proposition 3.1.2. *For every prime p and $n \in \mathbb{N}$ we have $\varphi(p^n) = p^{n-1}(p-1)$.*

Proof. The numbers a between 1 and p^n-1 which are not relatively prime with p^n are: p, $2p$,...,p^n-p. Hence there are $p^n-1-(p^{n-1}-1) = p^{n-1}(p-1)$ left that are coprime with p^n. ■

In general a function having the next property of φ is simply called *multiplicative*.

Theorem 3.1.3. *Let m and n two coprime positive integers. Then*

$$\varphi(mn) = \varphi(m)\varphi(n). \tag{3.2}$$

Proof. Let us introduce the sets

$$A := \{k \in \mathbb{N} | 1 \leq k \leq n-1, gcd(k,n) = 1\},$$
$$B := \{k \in \mathbb{N} | 1 \leq k \leq m-1, gcd(k,m) = 1\}, \text{ and}$$
$$C := \{k \in \mathbb{N} | 1 \leq k \leq mn-1, gcd(k,mn) = 1\}.$$

Define the function $f : A \times B \to \{0, 1, ..., mn-1\}$ by

$$f(k,l) = km + ln \ (mod\ mn) \ for\ all\ (k,l) \in A \times B.$$

Let us first show that our function takes values in C. Indeed, by Corollary 2.2.2 we may suppose $gcd(f(k,l), mn) = gcd(km + ln, mn) > 1$. Then for some prime q, $q|mn$ and $q|km + ln$. By Euclid's lemma $p|m$ or $p|n$. Either way we obtain that $q \leq gcd(k,n)$ or $q \leq gcd(l,m)$ contradicting that $k \in A$ or $l \in B$. Hence we actually have $f : A \times B \to C$.

Next we show that f is one-to-one. Suppose $f(k,l) = f(u,v)$ for some $(k,l),(u,v) \in A \times B$. Hence $km + ln - (um + vn) \equiv 0$ (mod mn). This is the same as $mn|(k-u)m + (l-v)n$. In particular $m|(k-u)m + (l-v)n$ which implies $m|(l-v)n$ and by Corollary **??** we have $m|l-v$. Since $m-2 \leq l-v \leq m-2$ we get $l = v$. Similarly we have $k = u$.

Finally let us show that f is onto on C. Let $c \in C$. This implies $gcd(c,m) = 1$ and $gcd(c,n) = 1$. By Corollary 2.2.4 there exists $\overline{m}$ such that $m\overline{m} \equiv 1$ (mod n) and also, there exits $\overline{n}$ such that $n\overline{n} \equiv 1$ (mod m). Define k to be the remainder of $\overline{m}c$ divided by n and l to be the remainder of $\overline{n}c$ divided by m. Let us observe that $k \in A$ and $l \in B$.

Then we just need to check that $f(k,l) = c$. For this it suffices to check that $km + ln \equiv c$ (mod mn). Let us observe that $km + ln \equiv (\overline{m}c)m + ln \equiv c$ (mod n) and $km + ln \equiv km + (\overline{n}c)n \equiv c$ (mod n). Hence, m and n divide $km + ln - c$. Since $gcd(m,n) = 1$ we have $mn|km + ln - c$.

The existence of the bijection f shows that $A \times B$ and C have the same number of elements: there are $\varphi(n)\varphi(m)$ elements in $A \times B$ and $\varphi(mn)$ in C. ∎

These two last results give the practical way to compute $\varphi(m)$ where $m = p_1^{\alpha_1} p_2^{\alpha_2} \cdots p_k^{\alpha_k}$ is one canonical factorizations:

$$\varphi(m) = m(1 - \frac{1}{p_1})(1 - \frac{1}{p_2}) \cdots (1 - \frac{1}{p_k}). \tag{3.3}$$

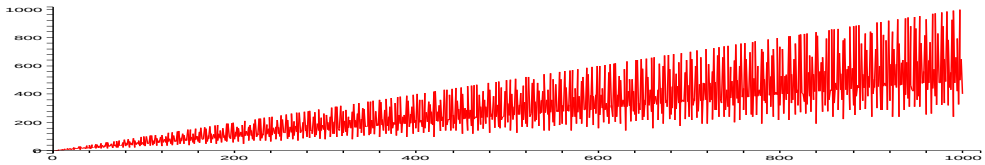


Figure 3.1: Graph of φ on $[1, 1000]$

For instance $\varphi(63) = \varphi(7(9)) = \varphi(7)\varphi(9) = 6(3)(2) = 36$. The graph of φ on the interval $[1, 1000]$ is included in Figure **??**.

From the graph we see that the behavior of this function is pretty interesting. There are basically two functions that give the upper and lower bounds for φ. Here are some surprising facts about these bounds:

$$\liminf_{n\to\infty} \frac{\varphi(n)\ln\ln n}{n} = e^{-\gamma}$$

where γ is Euler-Mascheroni constant ($\gamma = \lim_{n\to\infty}(1 + \frac{1}{2} + \cdots + \frac{1}{n} - \ln n)$) and on average, $\varphi(n)$ is closer to n as it is to this lower bound:

$$\lim_{n\to\infty} \frac{\varphi(1) + \varphi(2) + \cdots + \varphi(n)}{n^2} = \frac{3}{\pi^2}.$$

The values of $\varphi(n)$ are even for all $n \geq 3$. One open problem here is the following conjecture:

there are no $n \in \mathbb{N}$ such that the equation $\varphi(x) = n$ has exactly one solution.

Homework: Show that the equation $\varphi(x) = 14$ has no solutions.

One useful lower bound for $\varphi(n)$ is given by

Proposition 3.1.4. *For $n \notin \{2, 6\}$, $\varphi(n) \geq \sqrt{n}$.*

Proof. Let us use the formula (3.3). The inequality we need to prove is then equivalent to

$$\sqrt{n}(1 - \frac{1}{p_1})(1 - \frac{1}{p_2}) \dots (1 - \frac{1}{p_k}) \geq 1.$$

For every prime $p \geq 3$ we have $E(p) := p^{1/2}(1 - \frac{1}{p}) > 1$. Indeed the inequality can be written as $p - 1 > p^{1/2}$ or $p > \geq \frac{3+\sqrt{5}}{2}$ which is certainly true.

If $p = 2$ then the expression $E(p) = \frac{1}{\sqrt{2}}$. So our claim is not going to be true for $n = 2$ or $n = 6$ since $E(2)E(3) = \frac{\sqrt{2}}{\sqrt{3}} < 1$. But for every other n even number $n = 2k$, with $k \geq 4$, $E(2)$

The idea in the proof of Theorem 3.1.3 leads us to the solution of another classical problem in number theory: finding a solution of a simultaneous system of linear congruences such as

$$\begin{cases} x \equiv a \ (mod \ m) \\ x \equiv b \ (mod \ n) \end{cases}$$

Assuming $gcd(m, n) = 1$, we may look for a solution as before which is of the form $x = mk + nl$ with $k = \overline{m}b$ and $l = \overline{n}a$ where $\overline{m}$ is the inverse of m modulo n and $\overline{n}$ is the inverse of n modulo m.

This problem is known as the Chinese Remainder Theorem which can be easily generalized:

Theorem 3.1.5. *Let m_1, m_2, ..., m_k be mutually pairwise coprime numbers and $a_1, a_2, ..., a_k \in \mathbb{Z}$. Then the system of congruences*

$$\begin{cases} x \equiv a_1 \ (mod \ m_1) \\ x \equiv a_2 \ (mod \ m_2) \\ \dots \\ \dots \\ x \equiv a_k \ (mod \ m_k), \end{cases} \tag{3.4}$$

has a solution given by $x_0 = \sum_{j=1}^{k} M_j \overline{M_j} a_j$ *where* $M_j = (m_1 m_2 \cdots m_k)/m_j$ *and* $\overline{M_j}$ *is the inverse modulo* m_j *of* M_j*. Moreover, every solution of this system is of the form* $x = x_0 + k m_1 m_2 \cdots m_k$, $k \in \mathbb{Z}$.

PROOF. The fact that x_0 is s solution of (3.4) is essentially based on the observation that $M_j \equiv 0 \pmod{m_j}$ for all $k \neq j$, and $M_j \overline{M_j} a_j \equiv a_j \pmod{m_j}$ by hypothesis. Given another solution x of (3.4) implies that $x - x_0$ is divisible by m_j for all $j = 1, 2, \ldots, k$. Because m_1, m_2, $\ldots$, m_k are mutually pairwise coprime numbers we need to have $m_1 m_2 \cdots m_k | x - x_0$ and so the last statement of the theorem follows. ■

For an application of this theorem see the exercise after Theorem 3.3.1.

3.2 Construction of multiplicative functions

Two other functions which are classical fixtures in number theory are τ, and σ:

$$\tau(n) \textit{ number of positive divisors of } n,$$

$$\sigma(n) \textit{ sum of allpositive divisorsof } n.$$

Theorem 3.2.1. *If* $n = p_1^{a_1} p_2^{a_2} \cdots p_k^{a_k}$ *is one of the canonical prime factorizations of* n*, then* $\tau(n) = (a_1 + 1)(a_2 + 1) \cdots (a_k + 1)$ *and*

$$\sigma(n) = \frac{p_1^{a_1+1} - 1}{p_1 - 1} \frac{p_2^{a_2+1} - 1}{p_2 - 1} \cdots \frac{p_k^{a_k+1} - 1}{p_k - 1}.$$

Both functions, τ *and* σ*, are multiplicative.*

PROOF. Let us observe that the positive integer divisors of n are all of the form $p_1^{a_1'} p_2^{a_2'} \cdots p_k^{a_k'}$ with $0 \leq a_i' \leq a_i$, $i = 1, 2, \ldots, k$. All statements follow from this observation and the formula:

$$1 + r + \cdots + r^s = \frac{r^{s+1} - 1}{r - 1}.$$

There is a standard way to construct another multiplicative function out of a given one.

Theorem 3.2.2. *If f is multiplicative the function*

$$F(n) = \sum_{d|n} f(d),$$

is also multiplicative.

PROOF. If m and n are in $\mathbb{N}$ such that $gcd(m,n) = 1$, then

$$F(m)F(n) = \textstyle\sum_{d|m} f(d) \sum_{d|n} f(d) = \sum_{d_1|m,d_2|n} f(d_1)f(d_2)$$

$$= \textstyle\sum_{d_1|m,d_2|n} f(d_1d_2) = \sum_{d|mn} f(d) = f(mn).$$

The last before last equality is based on the fact that $\{d : d|mn\} = \{d_1d_2 : d_1|m \ and d_2|n\}$ which can checked by double inclusion and using the hypothesis that $gcd(m,n) = 1$. ■

Homework: If $gcd(m,n) = 1$ and $d|mn$, $d_1 = gcd(m,d)$, $d_2 = gcd(n,d)$. Show that $d = d_1d_2$.

A *perfect number* is a number n for which $\sigma(n) = 2n$. It is not known if there are any odd perfect numbers. This conjecture has been tested for all odd numbers less than 10^{400}. It was in the Elements of Euclid this proposition:

Theorem 3.2.3. *If $2^p - 1$ is a prime number (called Mersenne prime) then $n = 2^{p-1}(2^p - 1)$ is a perfect number. Every even perfect number n is of this form.*

PROOF. The first part is just an application of the Theorem 3.2.1:

$$\sigma(n) = \frac{2^p - 1}{2 - 1}\frac{q^2 - 1}{q - 1} = (2^p - 1)(q + 1) = 2^p(2^p - 1) = 2n$$

where $q = 2^p - 1$.

If n is perfect and even, let $n = 2^k t$ be its prime decomposition in which all odd primes are put together in the factor t. We have $k \geq 1$. Since σ is multiplicative we have

$$2n = 2^{k+1}t = \sigma(n) = (2^{k+1} - 1)\sigma(t). \tag{3.5}$$

Since $gcd(2^{k+1}, 2^{k+1} - 1) = 1$, this implies $2^{k+1}|\sigma(t)$ which in turn attracts $\sigma(t) = 2^{k+1}s$. Then, after substitution in (3.5), we get $t = (2^{k+1} - 1)s$ with $s \in \mathbb{N}$. By way of contradiction, we assume that $s > 1$. Then there are at least three distinct factors of t: 1, s and t (the hypothesis that $k \geq 1$ implies $t \neq s$).

Hence $2^{k+1}s = \sigma(t) \geq 1+s+t = 1+2^{k+1}s$ which is a contradiction. Therefore, it remains that $s = 1$ and so $\sigma(t) = t + 1$. This implies that t is a prime and so $n = 2^k(2^{k+1} - 1)$ which proves the claim of the theorem if we set $p = k + 1$. ■

There is a well known identity for sum of the powers of the consecutive positive integers:

$$\sum_{k=1}^{n} k^3 = \left(\sum_{k=1}^{n} k\right)^2, n \in \mathbb{N}.$$

A very similar identity takes place for the function τ:

$$\sum_{k|n} \tau(k)^3 = \left(\sum_{k|n} \tau(k)\right)^2, n \in \mathbb{N}. \tag{3.6}$$

Homework: Prove (3.6).

3.3 Möbius Inversion Formula

Suppose we have a formula $F(n) = \sum_{d|n} f(d)$ with f a multiplicative function. If we give values for n we get

$$F(1) = f(1),\ F(2) = f(1) + f(2),\ F(3) = f(1) + f(3),\ F(4) = f(1) + f(2) + f(4),$$

$$F(5) = f(1) + f(5),\ F(6) = f(1) + f(2) + f(3) + f(6),\ F(7) = f(1) + f(7),$$

$$F(8) = f(1) + f(2) + f(4) + f(8),\ F(9) = f(1) + f(3) + f(9),\ \ldots$$

These formulae can be solved for $f(n)$ in terms $F(m)$ to get

$$f(1) = F(1),\ f(2) = F(2) - F(1),\ f(3) = F(3) - F(1),$$

$$f(4) = F(4) - (F(2) - F(1)) - F(1) = F(4) - F(2),\ f(5) = F(5) - F(1),$$

$$f(6) = F(6) - F(1) - (F(2) - F(1)) - (F(3) - F(1)) = F(6) - F(2) - F(3) + F(1),$$

$$f(7) = F(7) - F(1),\ f(8) = F(8) - F(1) - (F(2) - F(1)) - (F(4) - F(2)) = F(8) - F(4),$$

$$f(9) = F(9) - F(1) - (F(3) - F(1)) = F(9) - F(3),\ \ldots$$

and this can be continued until a pattern is discovered. One is led to the introduction of the following function discovered by August Ferdinand Möbius (1790-1868) (well known for his twisted band):

$$\mu(n) = \begin{cases} 1 \ if \ n = 1; \\ (-1)^k \ if \ n = p_1 p_2 ... p_k, \ where \ p_i \ are \ distinct \ primes; \\ 0 \ \ otherwise. \end{cases}$$

This definition is clever enough to provide a general formula for $f(n)$ in terms of $F(m)$:

Theorem 3.3.1. *If* $F(n) = \sum_{d|n} f(d)$ *then*

$$f(n) = \sum_{d|n} \mu(d) F(\frac{n}{d}), \ n \in \mathbb{N}. \tag{3.7}$$

Proof. One can show that μ is multiplicative and

$$\sum_{d|n} \mu(d) = \begin{cases} 1 \ if \ n = 1; \\ 0 \ if \ n > 1. \end{cases}$$

To show (3.7) we get

$$\sum_{d|n} \mu(d) F(\frac{n}{d}) = \sum_{d|n} \mu(d) \sum_{c|\frac{n}{d}} f(c) = \sum_{d|n \ and \ c|\frac{n}{d}} \mu(d) f(c).$$

We notice that $d|n$ implies $n = dm$ and $c|\frac{n}{d} = m$ gives $m = ck$ with $m, k \in \mathbb{N}$. Hence $n = dck$. This implies $c|n$ and $d|\frac{n}{c}$. Conversely, if $c|n$ and $d|\frac{n}{c}$ then $d|n$ and $c|\frac{n}{d}$. Hence, in the above equality we can switch the order of summation but after all over the same set of pairs (d, c):

$$\sum_{d|n} \mu(d) F(\frac{n}{d}) = \sum_{c|n \ and \ d|\frac{n}{c}} \mu(d) f(c) = \sum_{c|n} f(c) \sum_{d|\frac{n}{c}} \mu(d).$$

But the last sum is equal to zero unless $\frac{n}{c} = 1$ or $c = n$. Hence it reduces to $f(n)$. ■

Let us observe that

$$\tau(n) = \sum_{d|n} 1,$$

so we can apply the inversion formula and obtain

$$1 = \sum_{d|n} \mu(d)\tau(\frac{n}{d}).$$

Similarly, we notice that

$$\sigma(n) = \sum_{d|n} d,$$

which gives

$$n = \sum_{d|n} \mu(d)\sigma(\frac{n}{d}).$$

Exercise: Show that there are infinitely many n such that $\mu(n)+\mu(n+1) = 0$.

Because $\mu(k) = 0$ for every k divisible by a perfect square we can look for n satisfying the system

$$\begin{cases} n \equiv 0 \ (mod\ 4) \\ n+1 \equiv 0 \ (mod\ 9) \end{cases}$$

This leads to the Chinese Remainder Theorem and so $x = 36k + 8$ with $k \in \mathbb{Z}$ will work.

One important property of the totient function φ related with these type of sums is given in the next theorem.

Theorem 3.3.2. *For every $n \in \mathbb{N}$ we have*

$$\sum_{d|n} \varphi(d) = n. \tag{3.8}$$

Proof. First let us partition the set $A := \{1, 2, \ldots, n\}$ is classes

$$C_d := \{k \in A : gcd(n, k) = d\}$$

for each $d|n$. Next, we show that $\#C_d$ is $\varphi(n/d)$. Indeed, $k \in C_d$ implies $k = dl$, $n = dm$ with $gcd(m, l) = 1$. So, every $k \in C_d$ defines uniquely an element l in $\{1, 2, \ldots, m\}$ such that $gcd(l, m) = 1$. This correspondence is a bijection and so $\#C_d = \varphi(m) = \varphi(n/d)$. Since C_d forms a partition of A (each element in A is in one of these classes and every two different classes are disjoint) we have

$$\#A = \sum_{d|n} \#C_d = \sum_{d|n} \varphi(n/d) = \sum_{d|n} \varphi(d)$$

which gives (3.8). ■

Problem: Prove that $\dfrac{\varphi(n)}{n} = \displaystyle\sum_{d|n} \frac{\mu(d)}{d}$.

With a little thinking, one can see that this identity above is nothing else but (3.3).

Chapter 4

The Law of Quadratic Reciprocity

4.1 Euler's Criterion

> **Quotation:** *"It is a matter for considerable regret that Fermat, who cultivated the theory of numbers with so much success, did not leave us with the proofs of the theorems he discovered. In truth, Euler and Lagrange, who have not disdained this kind of research, have proved most of these theorems, and have even substituted extensive theories for the isolated propositions of Fermat. But there are several proofs which have resisted their efforts. Recherches d'Analyse Indéterminée,* Hist Acad Roy des Sciences (1785/1788) 513. *" Adrien-Marie Legendre (1752-1833).*

The law of quadratic reciprocity is referring to a relationship between two different prime numbers, say p and q, in terms of the existence or non-existence of solutions for the equations

$$x^2 \equiv p \ (mod\ q),$$

$$y^2 \equiv q \ (mod\ p).$$

In order to give the statement of this important theorem in one of its modern formulations we need to introduce yet another function, with notation $\left(\frac{\cdot}{p}\right)$, defined for every odd prime p and every a coprime with p known as the Legendre symbol:

$$\left(\frac{a}{p}\right) = \begin{cases} 1 \ if\ the\ equation\ x^2 \equiv a \ (mod\ p)\ has\ a\ solution, \\ \\ -1 \ if\ the\ equation\ x^2 \equiv a \ (mod\ p)\ has\ no\ solution \end{cases} \tag{4.1}$$

This relationship can be simply formulated in terms of the Legendre symbol:

Theorem 4.1.1. [**Law of Quadratic Reciprocity**] *For every p and q odd prime numbers we have*

$$\left(\frac{p}{q}\right)\left(\frac{q}{p}\right) = (-1)^{\frac{p-1}{2}\frac{q-1}{2}}. \tag{4.2}$$

For instance if $p = 17449$, which is the 2007th prime and $q = 7$. The right hand side of (**??**) is equal to 1. Also, $p \equiv 5 \pmod 7$ and so the equation $x^2 \equiv p \pmod 7$ has no solution. Theorem 4.1.1 says that the equation $y^2 \equiv 7 \pmod{17449}$ has no solution which is a fact a lot harder to check computationally.

In order to prove this theorem we need some preliminary results. First, we have almost stumbled over the following result when we looked at representations of numbers as sums of two squares.

Theorem 4.1.2. [**Cauchy's Criterion**] *Let p be an odd prime and a such that $gcd(a, p) = 1$. Then*

$$\left(\frac{a}{p}\right) \equiv a^{\frac{p-1}{2}} \ (mod\ p).$$

Proof. Suppose first that $\left(\frac{a}{p}\right) = 1$. Then there is a solution to the equation $x^2 \equiv a \pmod{\text{p}}$. Clearly we must have $gcd(x, p) = 1$. Hence by Fermat's Little Theorem we have $x^{p-1} \equiv 1 \pmod{\text{p}}$. This gives

$$a^{\frac{p-1}{2}} \equiv (x^2)^{(p-1)/2} \equiv x^{p-1} \equiv 1 = \left(\frac{a}{p}\right),\ (mod\ p).$$

Let us assume then that $\left(\frac{a}{p}\right) = -1$. Then the equation $x^2 \equiv a \pmod{\text{p}}$ has no solution. This makes the function $g(x) = a\overline{x} \pmod{\text{p}}$ defined on $X := \{1, 2, \ldots, p-1\}$, be a function that has no fixed point. Here we used the notation as before $\overline{x}$ for the inverse (in X) modulo p of x. Indeed, a fixed point x of g gives $a\overline{x} \equiv x \pmod{\text{p}}$, or after multiplying by x both sides of the congruence we get $a \equiv x^2 \pmod{\text{p}}$. By hypothesis this equation has no solution. Hence, g has no fixed point. Then we notice that $g(g(x)) = a(\overline{a\overline{x}}) \equiv a\overline{a}x \equiv x \pmod{\text{p}}$. [We used the following two properties of the inverse function: $\overline{uv} = \overline{u}\ \overline{v}$ and $\overline{\overline{u}} = u$ for all u and v coprime with p.]

This implies that $g(g(x)) = x$ for all $x \in X$. A map with this property is called and *involution*. The idea of using an involution to prove things in number theory

has been really successful (see [12] and [27]). Hence we can group the elements of X in pairs $(x, g(x))$ and notice that $xg(x) \equiv a \pmod{p}$ for every $x \in X$. Since each pair has exactly 2 elements ($g(x) \neq x$) there are precisely $\frac{p-1}{2}$ pairs. Let us choose from each pair the smallest of the two numbers and put it in a set T. Hence we have

$$1(2)(3)\cdots(p-1) = \prod_{x\in T} xg(x) \equiv a^{\frac{p-1}{2}} \ (mod\ p)$$

and so by Willson's Theorem $a^{\frac{p-1}{2}} \equiv -1 \pmod{p}$. ■

The next ingredient we need is the so called Gauss's Lemma:

Lemma 4.1.3. [Gauss] *Let p be an odd prime and a such that $gcd(a,p) = 1$. Consider the set*

$$U := \{u | \frac{p}{2} < u \leq p-1, u \equiv ka \ (mod\ p) \ for\ some\ k = 1, 2, ..., \frac{p-1}{2}\}$$

and $r = \#U$. Then $\left(\frac{a}{p}\right) = (-1)^r$.

Proof. Let us consider the similar set

$$V := \{u | 1 \leq u < \frac{p}{2}, u \equiv ka \ (mod\ p) \ for\ some\ k = 1, 2, ..., \frac{p-1}{2}\}$$

and define $s = \#B$. The residues in X of ka modulo p are distinct for all $k = 1, 2, \ldots, \frac{p-1}{2}$. Then we have $s + r = \frac{p-1}{2}$. Let us define

$$h : U \to W := \{1, 2, \ldots, (p-1)/2\}$$

defined by $f(x) = p - x$, $x \in U$. Let us show that $Range(h) = W \setminus V$. For this it suffices to show that $h(x) \notin V$ for all $x \in U$. Indeed, if $h(x) = p - x = u$ for some $x \in U$ and $u \in V$. This implies $p|x + u$ or $0 \equiv x + u \equiv (i + j)a \pmod{p}$ for some $i, j \in \{1, 2, \ldots, \frac{p-1}{2}\}$ $(i \neq j)$. Because $gcd(a, p) = 1$ we see that $p|(i + j)$. This is impossible since $i + j \in \{2, 3, \ldots, p - 1\}$.

Then modulo p we have

$$1(2)\ldots(\frac{p-1}{2}) = \prod_{x\in U}(p-x)\prod_{x\in V} x \equiv (-1)^r \prod_{x\in U} x \prod_{x\in V} x \equiv (-1)^r \prod_{j=1,2,\ldots,\frac{p-1}{2}} (ja)$$

Simplifying by $1(2)\ldots(\frac{p-1}{2})$ both sides we get

$$1 \equiv (-1)^r a^{\frac{p-1}{2}} \ (mod\ p)$$

which after using Euler's Criterion gives the conclusion of the Lemma. ■

The next fact we need is due to a simplification due to Eisenstein (Ferdinand Gotthold Max 1823-1852) to the third proof for the Law of Quadratic Reciprocity given by Gauss. To state this lemma we need to introduce the function $\lfloor x \rfloor$ which is called the greatest integer part and it is, what its name says it is, the greatest integer k such that $k \leq x$.

Lemma 4.1.4. *Let p be an odd prime and a and odd integer such that $gcd(a,p) = 1$. Then*

$$\left(\frac{a}{p}\right) = (-1)^{T(a,p)}$$

where

$$T(a,p) = \sum_{j=1}^{\frac{p-1}{2}} \lfloor \frac{ja}{p} \rfloor.$$

Proof. For every $j = 1, 2, \ldots, \frac{p-1}{2}$ we have $ja = p\lfloor \frac{ja}{p} \rfloor + r_j$ where $r_j \in U \cup V$ with U and V defined as in the proof of the Gauss's Lemma. Adding up these equalities we have

$$a(\sum_{j=1}^{\frac{p-1}{2}} j) = pT(a,p) + \sum_{x \in U} x + \sum_{x \in V} x. \tag{4.3}$$

By the facts shown in the previous Lemma, we have

$$\sum_{x \in U} (p - x) + \sum_{x \in V} x = \sum_{j=1}^{\frac{p-1}{2}} j,$$

or

$$pr - \sum_{x \in U} x + \sum_{x \in V} x = \sum_{j=1}^{\frac{p-1}{2}} j. \tag{4.4}$$

From (4.3) and (4.4) we obtain

$$(a-1)\sum_{j=1}^{\frac{p-1}{2}} j = p(T(a,p)-r) + 2\sum_{x\in U} x.$$

Because a and p are odd we can take the above equality modulo 2 and obtain that $T(a,p) \equiv r \pmod 2$ which in conjunction with Gauss's Lemma proves what we want. ■

Proof of the Theorem 4.1.1.

We need to prove the identity

$$T(q,p) + T(p,q) = \frac{p-1}{2}\frac{q-1}{2}, \tag{4.5}$$

where $T(p,q)$ is defined as in the proof of the previous Lemma. We observe that (4.5 actually is enough to obtain (**??**). Let us assume that $p < q$.

The idea to show (4.5 uses a counting of the lattice points inside and on two sides (avoiding axes) of the rectangle $\mathcal{R}$ of vertices $(0,0)$, $(0,(q-1)/2)$, $((p-1)/2,0)$ and $((p-1)/2,(q-1)/2)$. The Figure **??** shows the lattice points in the case $p = 17$, $q = 23$.

We consider the line of equation $y = qx/p$. We observe that this line does not pass through any of these lattice points. If there is a point $(a,b) \in \mathbb{Z}$ on this line then $bp = qa$ which attracts $b = qk$, $a = pk$ with $k \in \mathbb{Z}$. These points avoid the specified area in the rectangle $\mathcal{R}$.

The number of these lattice points is equal to $\frac{p-1}{2}\frac{q-1}{2}$. We count the points above the line $y = qx/p$ and then the ones below it in $\mathcal{R}$.

Those points below this line are points $(x,y) \in \mathbb{Z}\times\mathbb{Z}$ that satisfy $y < qx/p$, $1 \le x < (p+1)/2$, $1 \le y \le (q-1)/2$. Notice we can count these points on columns and for each $1 \le i \le (p-1)/2$ there are $\lfloor \frac{qi}{p} \rfloor$ of y's satisfying the inequality $y < qi/p$ since $\frac{q(p-1)}{2p} < \frac{(q-1)}{2}$ $(p < q)$. Hence there are $\sum_{i=1,2,\ldots,(p-1)/2} \lfloor qi/p \rfloor = T(q,p)$ points below the line $y = qx/p$.

For those above this line we see that the points are characterized by $(x,y) \in \mathbb{Z}\times\mathbb{Z}$ that satisfy $y > qx/p$, $1 \le x \le (p-1)/2$, $1 \le y \le (q-1)/2$. The inequality $y > qx/p$ is equivalent to $x < py/q$. Again we notice we can count these points on rows and for each $1 \le j \le (q-1)/2$ there are $\lfloor \frac{qi}{p} \rfloor$ of x's satisfying the inequality $x < pj/q$ since $\frac{p(q-1)}{2q} < \frac{(p+1)}{2}$. Hence there are $\sum_{j=1,2,\ldots,(q-1)/2} \lfloor pj/q \rfloor = T(p,q)$ points above the line $y = qx/p$. ■

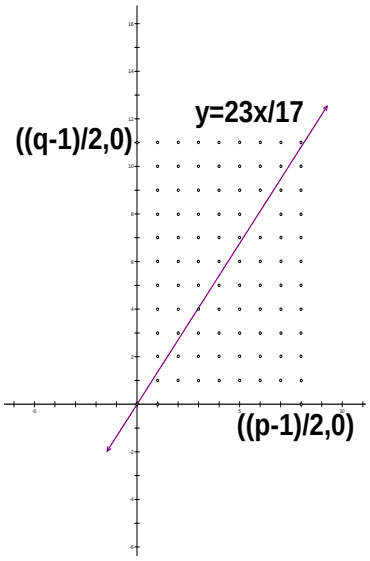


Figure 4.1: Lattice points in the case $p = 17$ and $q = 23$.

Other properties of the Legendre symbol are recorded in the next proposition.

Theorem 4.1.5. *Let p an odd prime and a, b such that $gcd(a, p) = gcd(b, p) = 1$. We have*

(i) if $a \equiv b \pmod p$, then $(\frac{a}{p}) = (\frac{b}{p})$;

(ii) $(\frac{a}{p})(\frac{b}{p}) = (\frac{ab}{p})$;

(iii) $(\frac{a^2}{p}) = 1$;

(iv) $(\frac{2}{p}) = (-1)^{\frac{p^2-1}{8}}$;

(v) $(\frac{-1}{p}) = (-1)^{\frac{p-1}{2}}$.

(vi) $(\frac{3}{p}) = 1$ iff $p \equiv 1$ or $11 \pmod{12}$.

These properties including the Law of Quadratic reciprocity can be used to calculate relatively easy the Legendre symbol.

Example: Suppose we want to compute

$$\left(\frac{2007}{29}\right) = \left(\frac{6}{29}\right) = \left(\frac{2}{29}\right)\left(\frac{3}{29}\right) = (-1)(-1) = 1.$$

Bibliography

[1] T. Andreescu and R. Gelca, *Mathematical Olympiad Challenges*, Birkhauser, **2000**.

[2] D. M. Burton, *Elementary Number Theory*, Sixth Edition, **2007**, McGraw Hill.

[3] K. Dajani and C. Kraaikamp, *Ergodic Theory of Numbers*, **2002**, The Carus Mathemmatical Monographs.

[4] D. A. Cox , *Primes of the Form* $x^2 + ny^2$, Wiley-Interscience, **1989**.

[5] A. Engel, *Problem-Solving Startegies*, Springer, **1997**.

[6] G. Everest and T. Ward, *An Introduction to Number Theory*, Springer, 2000.

[7] J. Friedlander and H. Iwaniec, *The polynomial* x^2+y^4 *captures its primes*, Ann. of Math. (2) 148 (1998), no. 3, 945–1040.

[8] A. Granville and G. Martin, *Prime number races*, American Mathematical Monthly, vol 113, No 1, **2006**.

[9] G. Greaves, *Sieves in Number Theory*, Springer, **2000**.

[10] R. Guy, Unsolved Problems in Number Theory, Third Edition, **2004**, Springer.

[11] V. Klee and S. Wagon, *Old and New Unsolved Problems in Plane Geometry and Number Theory*, **1991**, MAA The dolciani mathenatical expositions.

[12] T. Jackson, *A short proof that every prime* $p \equiv 3$ *(mod 8) is of the form* x^2+2y^2, *Amer. Math. Monthly* **107** (2000) 447.

[13] E. J. Ionascu, *Counting all regular tetrahedra in* $\{0, 1, 2, ..., n\}^3$,

[14] E. J. Ionascu and J. Patterson, *Primes of the form* $\pm a^2 \pm qb^2$, Stud. Univ. Babe¸s-Bolyai Math. **58** (2013), No. 4, pp. 421-430

[15] M.*Křížek*, F. Luca and L. Somer, *17 Lectures on Fermat Numbers*, CMS Books in Mathematics, Springer 2001.

[16] J. Lambek and L. Moser, *On the distribution of pythagorean triangles*, Pacific J. Math., 5(**1955**), pp. 73-83

[17] L. J. Mordell, *The Representation of a Gaussian Integer as a Sum of Two Squares*,, Mathematics Magazine, Vol. 40, No. 4 (**1967**), p. 209

[18] M. B. Nathanson, *Elementary methods in number theory*, Springer, **2000**.

[19] I. Niven, H. S. Zuckerman, H. L. Montgomery, *An Introduction to the Theory of Numbers*, Fifth Edition, **1991**, John Wiley & Sons, Inc.

[20] The On-Line Encyclopedia of Integer Sequences, http://oeis.org/Seis.html

[21] G. Polya, *Mathemeatics and plausible reasoning*, vol II, Princeton, 1968.

[22] K. H. Rosen, *Elementary Number Theory and its Applications*, Fifth Edition, **2005**, Addison Wesley.

[23] David Santos, *Number Theory for Mathematical Contests*, http://www.fmf.uni-lj.si/ lavric

[24] V. Shoup, *A Computational Introduction to Number Theory*, "Second Edition", http://www.shoup.net/ntb/

[25] W. Sierpinski, *Pythagorean Triangles*, Dover (**2003**).

[26] J. J. Sylvester, *Math. Quest. Educ. Times*, **37**(1884) 26.

[27] D. Zagier, *A one-sentence proof that every prime $p \equiv 1$ (mod 4) is a sum of two squares*, *Amer. Math. Monthly* **97** (1990) 144.

Index