4 From Euclid to Fermat to Euler to Gauss and to RSA algorithm

4.1 The fundamental theorem of arithmetic

Here I will give a detailed proof of the fundamental theorem of arithmetic noting that there is a very interesting discussion on why this theorem is not "obvious" in the Internet[1]. Indeed, it is quite naive to expect that the products like 1357×4183 and 1081×5251 are not the same only because all four numbers in these products are prime. Another point about this theorem is that in its proof it is very easy unconsciously to assume it itself thus falling in a circular trap. Finally, on many occasions various proofs of this theorem are based on the so-called *Euclid's lemma* (see, e.g., the textbook) but here I do not require this lemma in my proof and eventually prove Euclid's lemma using the fundamental theorem of arithmetic.

Let $m, n \in \mathbf{N}$ be two natural numbers. I say that m divides n, denoted $m \mid n$, if there is $k \in \mathbf{N}$ such that $n = km$. A number $p \in \mathbf{N}$ is called *prime* if $p \geq 2$ and it is divided by only 1 and itself. Here are a few first prime numbers: $2, 3, 5, 7, 11, 13, 17, 23, 29, \ldots$ If a number is not prime and different from 1 it is called composite. In other words $m \in \mathbf{N}$ is composite if and only if $m = ab$, $a, b \in \mathbf{N}$ and $2 \leq a \leq b < m$.

Theorem 4.1 (Fundamental theorem of arithmetic)**.** *Any natural number $n \geq 2$ can be uniquely written as the product of prime numbers:*

$$n = p_1^{\alpha_1} \cdot p_2^{\alpha_2} \cdot \ldots \cdot p_k^{\alpha_k}, \quad 2 \leq p_1 < p_2 < \ldots < p_k, \quad \alpha_1, \alpha_2, \ldots, \alpha_k \geq 1,$$

where p_j are prime numbers and α_j are natural.

Remark 4.2. Note that 1 is not a prime number, and one of the main reasons for it is to have the uniqueness of factoring of any natural number into product of primes. Sometime it is convenient to define that 1 is equal to the empty product of prime numbers, thus removing the condition $n \geq 2$ in the theorem.

I will split the proof of Theorem 4.1 into three steps.

Lemma 4.3 (Existence)**.** *Any natural $n \geq 2$ can be written as a product of primes.*

Proof. (By strong induction) The statement is true for the base step $n = 2$. Assume that it is true for any $2 \leq k \leq n - 1$ and consider $n \in \mathbf{N}$. If it is prime we are done. If it is not prime, it is composite, or a product $n = ab$ of two natural $2 \leq a \leq b < n$, for which the induction assumption is true, which finishes the proof. ∎

Remark 4.4. Note that I also proved a somewhat weaker statement that any natural number (other than 1) is either prime or divisible by prime.

The next step will be somewhat auxiliary, the main reason I separate it is the fact that for many students this statement seems so obvious that they do not realize that it still requires a proof.

Math 478/678: History of Mathematics by Artem Novozhilov
e-mail: artem.novozhilov@ndsu.edu. Spring 2024

[1]https://gowers.wordpress.com/2011/11/13/why-isnt-the-fundamental-theorem-of-arithmetic-obvious/

Lemma 4.5. *If $n = p_1p_2 \dots p_k$ (here I do not assume that all p_j are distinct, the only assumption is that $p_1 \leq p_2 \leq \dots \leq p_k$) is the unique factoring of n into product of primes and prime p divides n then $p = p_j$ for at least one index $1 \leq j \leq k$.*

Proof. (By contradiction) Assume that $p \mid n$ and $p \neq p_j$ for all j. Since $n = pa = pq_1 \dots q_l$ by Lemma 4.3 for some primes q_i I found a different factoring of n into product of primes, which contradicts the assumption that such factoring is unique. ■

Lemma 4.6 (Uniqueness of prime factorization)**.** *The prime factorization*

$$n = p_1p_2 \dots p_k, \quad 2 \leq p_1 \leq p_2 \leq \dots \leq p_k$$

is unique.

Proof. (By contradiction) It is clear that at least for the first several natural numbers this factorization is unique. Assume that there are natural numbers for which the prime factorization is not unique. Let $n \in \mathbf{N}$ be the smallest such number (which must exist by the well-ordering principle). That is,

$$n = p_1p_2 \dots p_k = q_1q_2 \dots q_l,$$

where (and below) all the factors are arranged in nondecreasing order. First I note that $p_i \neq q_j$ for all possible i and j, because otherwise I would have found, after canceling identical factors, a smaller natural number that would have non unique prime factorization. Hence I can assume that $p_1 < q_1$ (otherwise I can switch the notation). I have that $n \geq p_1^2$ (since there must be at least two factors in the product of primes and $p_2 \geq p_1$) and hence $n > q_1^2$. Together this implies that $n^2 > p_1^2q_1^2$ or

$$n > p_1q_1.$$

Consider now the natural number $n - p_1q_1$. This number is smaller than n (and hence has a unique prime factorization) and is divisible by construction by both p_1 and q_1. By Lemma 4.5 natural number $n - p_1q_1$ has the unique prime factorization p_1q_1a, where a is a product of some primes. This implies that

$$n = p_1p_2 \dots p_k = p_1q_1(1 + a),$$

and by canceling p_1 and recalling that $q_1 \neq p_i$ for any i I found two different prime factorizations for the natural number $p_2 \dots p_k = q_1(1 + a)$, which is smaller than n. Since n by assumption was the smallest such number I reached a contradiction thus finishing the proofs of both the lemma and Theorem 4.1. ■

Euclid himself did not prove Theorem 4.1 in his Elements. He went, however, very close to the same exactly statement. The key fact, which is proved in Elements, and can be used to show the uniqueness of prime factorization, is Proposition 30 in Book VII.

Corollary 4.7 (Euclid's lemma)**.** *If $p \in \mathbf{N}$ is prime and $p \mid ab$ for some $a, b \in \mathbf{N}$ then either $p \mid a$ or $p \mid b$.*

Proof. By the fundamental theorem of arithmetic, ab is uniquely factored into product of primes $p_1 \dots p_kq_1 \dots q_l$, where $a = p_1 \dots p_k$ and $b = q_1 \dots q_l$. By the same theorem (Lemma 4.5), p must coincide with either one of p_i or q_j, and hence p divides either a or b. ■

Exercise 1. Prove a more general variant of Euclid's lemma: If $a \mid bc$, and a and b are relatively prime then $a \mid c$.

Exercise 2. Give an example of a, b, c such that $c \mid ab$ and at the same time $c \nmid a$ and $c \nmid b$.

Exercise 3.
There was a young lady named Chris
Who, when asked her age, answered this
Two-thirds of its square
Is a cube I declare
Now what was the age of the miss?

Exercise 4. Assuming that Euclid's lemma is true, give a different proof of the uniqueness of prime factorization.

Exercise 5. Prove that $\sqrt{p}$ is irrational for any prime p.

To finish this short section I would like to mention one common misconception, which can be found in many number theory textbooks[2]. Euclid in his Elements gave a proof of the fact that there are infinitely many prime numbers. In many books it is claimed that he did this by contradiction, which is incorrect (to be fully honest he did use a small bit of contradiction inside his proof, but he never started his proof with the sentence like "Suppose that there are finitely many primes.")

Here are Euclid's arguments using modern notation.

Theorem 4.8. *There are infinitely many primes.*

Proof. Let a, b, c be prime numbers. Consider the number

$$abc + 1.$$

This number is either prime (which is different from a, b, c) or divisible by prime (Lemma 4.3). In the latter case this prime cannot be a, b, c otherwise it would mean that a, b, or c divide 1, which is absurd. In either case we found another prime number different from a, b, c, call it d. Now we can repeat the process starting with a, b, c, d. In other words, given k prime numbers it is always possible to find a $k + 1$-st prime number, which finishes the proof. ■

4.2 Congruences and divisibility rules

4.3 Fermat's little and Euler's theorems

Now that we have some experience working with congruences, we can prove Fermat's little theorem and its generalization Euler's theorem. Several times in the proofs below I will need a basic fact when one can cancel factors in congruences, so let me start with it.

Lemma 4.9. *If integers c and n are relatively prime then the congruence*

$$ac \equiv bc \pmod{n}$$

implies

$$a \equiv b \pmod{n}.$$

[2]See Hardy, M., & Woodgold, C. (2009). Prime simplicity. The Mathematical Intelligencer, 31, 44–52, for further details.

Proof. By the definition of congruences we have that

$$n \mid (a-b)c$$

and by the assumptions n and c are relatively prime. Hence by Euclid's lemma $n \mid (a-b)$ or

$$a \equiv b \pmod{n}$$

as required. ■

Theorem 4.10 (Fermat's little theorem)**.** *Let p be a prime number, and $a \in \mathbf{N}$ be relatively prime with p. Then*

$$a^{p-1} \equiv 1 \pmod{p}.$$

Remark 4.11. Often Fermat's little theorem is formulated as $a^p \equiv a \pmod{p}$ for any natural a. Think out why this does not add much to the statement I have given.

Proof of Theorem 4.10. Consider $p-1$ numbers $a, 2a, \ldots, (p-1)a$ modulo p:

$$\begin{aligned} a &\equiv r_1 \pmod{p}, \\ 2a &\equiv r_2 \pmod{p}, \\ &\vdots \\ (p-1)a &\equiv r_{p-1} \pmod{p}. \end{aligned}$$

Since a and p are relatively prime by the assumption, $1, \ldots, p-1$ are relatively prime with p because p is prime, then none of $r_j \neq 0$. Moreover, $r_i \neq r_j$ for any $1 \leq i, j \leq p-1$, $i \neq j$. Indeed, if it happened that $r_i = r_j = r$, it would mean $ia \equiv ja \pmod{p}$, or, by Lemma 4.9, $i \equiv j \pmod{p}$ or simply $i = j$, which is impossible. Therefore we conclude that $\{r_1, r_2, \ldots, r_{p-1}\} = \{1, 2, \ldots, p-1\}$ (possibly in a different order, but this is not important for us).

Multiplying all the lines above yields

$$(p-1)!a^{p-1} \equiv (p-1)! \pmod{p},$$

or, invoking Lemma 4.9 again,

$$a^{p-1} \equiv 1 \pmod{p},$$

because $(p-1)!$ and p are relatively prime. The theorem has been proven. ■

Euler's theorem replaces p in Theorem 4.10 with an arbitrary natural n. For the exact statement I will need a definition of Euler's φ-function, which is the number of integers $1 \leq j < n$ which are relatively prime with n. Convince yourself that, e.g., $\varphi(6) = 2$, $\varphi(9) = 6$, and $\varphi(p) = p-1$ for any prime p.

Theorem 4.12 (Euler's theorem)**.** *For relatively prime a and n*

$$a^{\varphi(n)} \equiv 1 \pmod{n}.$$

Proof. Let $\{\alpha_1 = 1, \alpha_2, \dots, \alpha_{\varphi(n)}\}$ be $\varphi(n)$ numbers that are relatively prime with n. Consider

$$\begin{aligned}
\alpha_1 a &\equiv r_1 \pmod{n},\\
\alpha_2 a &\equiv r_2 \pmod{n},\\
&\vdots\\
\alpha_{\varphi(n)} a &\equiv r_{\varphi(n)} \pmod{n}.
\end{aligned}$$

For exactly the same reasons as in the proof of Fermat's little theorem, $r_j \neq 0$ and $r_i \neq r_j$ if $i \neq j$. Moreover, I claim that r_j and n must be relatively prime. Looking for a contradiction assume not, i.e., assume n and r_j have a common factor for some j. Since $\alpha_j a \equiv r_j \pmod{n}$ means that $\alpha_j a - r = kn$ for some integer k, this yields $\alpha_j a = kn + r$. If n and r have a common factor then $kn + r$ is divisible by this factor, and hence, by the fundamental theorem of arithmetic, $\alpha_j a$ also must be divisible by this factor, which is impossible since both a and α_j are relatively prime with n. Therefore we conclude that $\{r_1, r_2, \dots, r_{\varphi(n)}\} = \{\alpha_1, \alpha_2, \dots, \alpha_{\varphi(n)}\}$.

Multiplying all the lines above yields

$$\alpha_1 \dots \alpha_{\varphi(n)} a^{\varphi(n)} \equiv \alpha_1 \dots \alpha_{\varphi(n)} \pmod{n},$$

or, invoking Lemma 4.9,

$$a^{\varphi(n)} \equiv 1 \pmod{n},$$

because all α_j and n are relatively prime. The theorem has been proven. ■

4.4 RSA algorithm

Now we will see a little bit of magic of the discussed number theory (there is much more to the story, see the literature review at the end of this lecture).

Assume that Bob needs to send Alice some secret information I. To do it securely, this information must be encrypted so that no one could read this information other than Alice. This can be done, for instance, if both Alice and Bob exchanged some (hopefully strong) cypher earlier, and only they have access to this cypher (this is called a symmetric cryptosystem). The weakness of course is that if the cypher is broken (or stolen) all the future correspondence will be available to eavesdropper. In the seventies of the twentieth century it was realized that there is another dramatically different approach to safe information transmission. Namely, it was suggested that an asymmetric cryptosystem with a public key would be used. Pretty much it means that the person, who is about to receive the message I, shares some public information, which includes the public key α, which is used to encipher the information I, but cannot be used to decipher the message (this is why it is *assymetric*).

The first actual implementation of this idea was done by Ron Rivest, Adi Shamir, and Leonard Adleman in 1977[3]. Here is how it works.

Alice takes two prime numbers p, q and computes $n = pq$. The algorithm is based on the fact that if p and q are sufficiently large, knowing n will not allow to determine p, q in a reasonable time. Yet Alice knows both p, q and hence knows that $\varphi(n) = (p-1)(q-1) = N$.

Exercise 6. Prove that for primes p and q

$$\varphi(pq) = (p-1)(q-1).$$

[3]See Gardner, M. (1977). Mathematical games. Scientific American, August, 120–124.

Further, she chooses natural α, which should be relatively prime with N, and shares publicly α and n.

At this point anyone can use this public information to send secret messages to Alice, including Bob. He computes

$$I^{\alpha} \equiv J \pmod{n}$$

and sends the encrypted message J to Alice.

In addition to α Alice computes β (the private key), that must satisfy

$$\alpha\beta \equiv 1 \pmod{N},$$

i.e., $\alpha\beta = kN + 1$. Note that this is just a Diophantine's equation for the unknowns β and k, which can always be solved for relatively prime α and N by the extended Euclid's algorithm.

Finally, using Euler's theorem 4.12, Alice computes

$$J^{\beta} \equiv (I^{\alpha})^{\beta} \equiv I^{kN+1} \equiv (I^{N})^{k} I \equiv (I^{\varphi(n)})^{k} I \equiv [\text{by Theorem } 4.12] \equiv 1 \cdot I \equiv I \pmod{n},$$

and hence recovers the original message I!

4.5 Literature review