ELEMENTARY NUMBER THEORY

Notes by

MOSHE JARDEN

School of Mathematics, Tel Aviv University

Ramat Aviv, Tel Aviv 69978, Israel

e-mail: jarden@post.tau.ac.il

web page: http://www.tau.ac.il/∼jarden/Courses

Forward

We start with the set of natural numbers, $\mathbb{N} = \{1, 2, 3, \ldots\}$ equipped with the familiar addition and multiplication and assume that it satisfies the induction axiom. It allows us to establish division with a residue and the Euclid's algorithm that computes the greatest commond divisor of two natural numbers. It also leads to a proof of the fundamental theorem of arithmetic: Every natural number is a product of prime numbers in a unique way up to the order of the factors. Euclid's theorem about the infinitude of the prime numbers is a consequence of that theorem.

Next we introduce congruences and the Euler's φ-function ($\varphi(n)$ is the number of the natural numbers between 1 and n that are relatively prime to n). Then we prove Euler's theorem: $a^{\varphi(n)} \equiv 1 \mod n$ for each natural number n and every integer a relatively prime to n. We also prove the Chinese remainder theorem and conclude the multiplicity of the Euler phi function: $\varphi(mn) = \varphi(m)\varphi(n)$ if m and n are relatively prime. This theorem is the main ingredient in the first and most applied public key in cryptography. Next we prove that each prime number p has $\varphi(p-1)$ primitive roots modulo prime numbers.

Most of this material enters into the proof of the quadratic reciprocity law: Every distinct odd prime numbers p, q satisfy $\left(\frac{q}{p}\right) = (-1)^{\frac{p-1}{2}\frac{q-1}{2}}\left(\frac{p}{q}\right)$, where $\left(\frac{q}{p}\right) = 1$ if there exists an integer x with $x^2 \equiv q \mod p$, and -1 otherwise.

The last part of these notes is devoted to the proof of Dirichlet's theorem about the Dirichlet density of the set prime numbers $p \equiv a \mod m$, where $\gcd(a, m) = 1$.

Mevasseret Zion, 4 April 2012

Table of Contents

1. Natural Numbers

AXIOMS
input, 8

Our starting point are the **natural numbers** $1, 2, 3, 4, \dots$. We denote the set of all natural numbers by $\mathbb{N}$ and take for granted that $\mathbb{N}$ is **well ordered**. That is there is a **relation** $<$ on $\mathbb{N}$ satisfying:

(1a) $n \not< n$.

(1b) If $a < b$ and $b < c$, then $a < c$.

(1c) If $a \neq b$, then either $a < b$ or $b < a$.

(1d) For every $a \in \mathbb{N}$ there is a $b \in \mathbb{N}$ such that $a < b$.

(1e) Every non-empty subset A of $\mathbb{N}$ has a smallest element.

(1f) 1 is the smallest element of $\mathbb{N}$.

We also assume that **addition** and **multiplication** have already been defined on $\mathbb{N}$ with the usual properties. In particular, each $b \in \mathbb{N}$ greater than 1 can be written as $b = a + 1$ for a unique $a \in \mathbb{N}$.

LEMMA 1.1 (Induction Axiom): *Suppose a subset A of $\mathbb{N}$ satisfies the following two conditions:* AXIa input, 38

(2a) $1 \in A$.

(2b) *If $n \in A$, then $n + 1 \in A$.*

Then $A = \mathbb{N}$.

Proof: Assume $B = \mathbb{N} \smallsetminus A \neq \emptyset$ and let b be the smallest elelment of B. Then $b \neq 1$, so $b = a + 1$ for some $a \in A$. It follows from (2b) that $b \in A$. This contradiction implies that $A = \mathbb{N}$. ∎

It is sometimes convenient to use the following form of the induction axiom:

LEMMA 1.2: *Let P be a property of natural numbers such that for each $n \in \mathbb{N}$ we have:* AXIb input, 62

(3) *if each $m < n$ has property P, then n has property P.*

Then every natural number has property P.

Proof: We denote the set of all natural numbers n such that each $m < n$ has property P by A. Since there are no natural numbers less than 1, Condition (3), implies that $1 \in A$.

Now suppose that $n \in A$. Then each $m < n$ has property P. By (3), n has property P, so all $m < n+1$ has property P. By definition, $n+1 \in A$. It follows from the induction axiom that $A = \mathbb{N}$. Since each $n \in \mathbb{N}$ satisfies $n < n+1$ and $n+1 \in A$, we get that n has property P. ∎

If $a < b$, then there exists a unique $x \in \mathbb{N}$ such that $a + x = b$. As usuall, we write $x = b - a$. If $b \le a$, then $b - a$ is not defined within $\mathbb{N}$. So, we consider also the number 0 and the negative number $-n$ for each $n \in \mathbb{N}$ and extend addition and multiplication to the set $\mathbb{Z} = \{0, \pm 1, \pm 2, \pm 3, \ldots\}$ of **integers**. Then $\mathbb{Z}$ is a **commutative ring** with 1, that is it satisfies the following conditions for all $x, y, z \in \mathbb{Z}$:

(4a) $x + y = y + x$ (commutative law for addition).

(4b) $(x + y) + z = x + (y + z)$ (associative law of addition).

(4c) $0 + x = x$.

(4d) For each $a \in \mathbb{Z}$ there exists a unique element $-a \in \mathbb{Z}$ such that $a + (-a) = 0$.

(5a) $xy = yx$ (commutative law of multiplication).

(5b) $(xy)z = x(yz)$ (associative law of multiplication).

(5c) $1 \cdot x = x$.

(6) $x(y + z) = xy + xz$ (distributive law).

It is also convenient to consider also the **field** $\mathbb{Q}$ of all quotients $\frac{a}{b}$ with $a, b \in \mathbb{Z}$ and $b \neq 0$. It satisfies Conditions (4), (5), and (6), for all $x, y, z \in \mathbb{Q}$ and in addtion also the following one:

(5d) For each $x \neq 0$ there exists a unique element $x^{-1} \in \mathbb{Q}$ such that $x \cdot x^{-1} = 1$.

Exercise 1.3: *Prove that* $\mathbb{N}$ *is an* ***infinite set****, that is there exist no* $n \in \mathbb{N}$ *and a bijective function* $f\colon \mathbb{N} \to \{1, \ldots, n\}$*. Hint: You are allowed to use the following set theoretic rule: Every injective map* f *from a finite set* A *onto itself is bijective. Then apply this rule to* $A = \{1, \ldots, n\}$ *and observe that* $n + 1 \notin \{1, \ldots, n\}$*.* AXIc input, 127

Exercise 1.4: *Use induction on* n *(or otherwise) to prove the following formulas:* AXId input, 139

$$(7a) \qquad \sum_{k=1}^{n} k = \frac{n(n+1)}{2}$$

$$\sum_{k=1}^{n} k^2 = \frac{n(n+1)(2n+1)}{6} \tag{7b}$$

$$\sum_{k=1}^{n} k^3 = \frac{n^2(n+1)^2}{4} \tag{7c}$$

EXERCISE 1.5: *Prove that every real number $x \neq 1$ satisfies $\sum_{i=0}^{n} x^i = \frac{x^{n+1}-1}{x-1}$.* AXIe input, 155

EXERCISE 1.6: *Define $F_0 = 0$, $F_1 = 1$, and assuming that F_n and F_{n+1} have beed defined, let $F_{n+2} = F_{n+1} + F_n$. One calls $F_0, F_1, F_2, \ldots$ the* **Fibonacci numbers**. *Prove by induction on n that $F_n < \left(\frac{7}{4}\right)^n$. Hint: Use that $44 < 49$.* AXIf input, 160

EXERCISE 1.7: *Let A be a nonempty subset of $\mathbb{Z}$. Suppose that A is bounded from above. That is, there exist $r \in \mathbb{Q}$ such that $a \leq r$ for all $a \in A$. Prove that A has a maximal element. Hint: Let b the largest integer k satisfying $b \leq r$ (one denotes that number by $[r]$). Then consider the set $A' = \{b - a \mid a \in A\}$.* AXIg input, 168

2. Division with a Residue

RESIDUE input, 6

Starting from division with a residue in $\mathbb{N}$, we establish the Euclidean algorithm that allows an efficient computation of the greatest common divisor of natural numbers.

PROPOSITION 2.1 (Division with residue): *For all $m \in \mathbb{N}$ and $n \in \mathbb{Z}$ there exist unique $q, r \in \mathbb{Z}$ such that $n = qm + r$ and $0 \le r \le m - 1$.* RESa input, 13

Proof of existence: The set $A = \{a \in \mathbb{Z} \mid \ am \le n\}$ is not empty. Indeed, if $n < 0$, then $nm \le n$, so $n \in A$. If $n \ge 0$, then $0 \cdot m \le n$, so $0 \in A$.

By definition $a \le \frac{n}{m}$ for each $a \in A$, so A is bounded from above. Hence, there exists a maximal element q in A. In other words, $qm \le n$ but $n < (q+1)m$. The difference $r = n - qm$ satisfies the requirements of the lemma.

Proof of uniqueness: Suppose r' and q' are additional integers satisfying $n = q'm + r'$ and $0 \le r' \le m - 1$. Then $0 = (q' - q)m + (r' - r)$. Thus, $r - r' = (q' - q)m$, so $|r' - r| = |q' - q|m$. If $q' \ne q$, then $|q' - q| \ge 1$ (because q and q' are integers), hence

$$|r' - r| \ge m. \tag{1}$$

By symmetry, we may assume that $r' \ge r$, so $m > r' - r \ge 0$. Hence $|r' - r| < m$, a contradiction to (1). ∎

We exploit Proposition 2.1 to achieve radix represenentation of integers.

PROPOSITION 2.2: *For each integer $g \ge 2$ every $n \in \mathbb{N}$ can be uniquely represented as* RESb input, 54

$$n = a_k g^k + a_{k-1} g^{k-1} + \cdots + a_0$$

with $0 \le a_0, \ldots, a_{k-1}, a_k \le g - 1$ and $a_k \ne 0$.

Proof of existence: Using Proposition 2.1, we write $n = qg + a_0$ with $q \in \mathbb{N}$ and $0 \le a_0 \le g-1$. Since $g \ge 2$, we have $q < n$. By induction, there exists k and $a_1, \ldots, a_{k-1}, a_k$ such that $q = a_k g^{k-1} + a_{k-1} g^{k-2} + \cdots + a_1$, where $0 \le a_1, \ldots, a_{k-1}, a_k \le g - 1$ and $a_k \ne 0$. It follows that $n = qg + a_0 = a_k g^k + a_{k-1} g^{k-1} + \cdots + a_1 g + a_0$, as desired.

Proof of uniqueness: Assume we have two presentations

$$n = \sum_{i=0}^{k} a_i g^i, \quad 0 \le a_i \le g-1 \quad a_k \ne 0 \tag{2a}$$

$$n = \sum_{i=0}^{k} b_i g^i, \quad 0 \le b_i \le g-1 \quad b_k \ne 0 \tag{2b}$$

Assume without loss that $k \le l$ and let $a_{k+1} = \cdots = a_l = 0$. This allows us to rewrite (2a) in the form

$$n = \sum_{i=0}^{l} a_i g^i, \quad 0 \le a_i \le g-1 \quad a_k \ne 0 \tag{3}$$

Assume there exists $0 \le i \le l$ with $a_i \ne b_i$. Let r be the greatest integer between 0 and l with

$$a_r \ne b_r. \tag{4}$$

Substruction of (3) from (2b) yields $\sum_{i=0}^{r}(b_i - a_i)g^i = 0$, so

$$(b_r - a_r)g^r = \sum_{i=0}^{r-1}(a_i - b_i)g^i. \tag{5}$$

By (4), $|b_r - a_r| \ge 1$. Hence, by (5)

$$\begin{aligned} g^r \le |b_r - a_r|g^r &\le \sum_{i=0}^{r-1} |a_i - b_i|g^i \le \sum_{i=0}^{r-1}(g-1)g^i \\ &= (g-1)\sum_{i=0}^{r-1} g^i = (g-1)\frac{1-g^r}{1-g} = g^r - 1 < g^r, \end{aligned}$$

a contradiction. ∎

Examples 2.3:

RESc
input, 129

(7a) $2012 = 2 \cdot 10^3 + 0 \cdot 10^2 + 1 \cdot 10 + 2$ (decimal representatio).

(7b) $17 = 1 \cdot 2^4 + 0 \cdot 2^3 + 0 \cdot 2^2 + 1$ (binary representation).

(7c) $17 = 2 \cdot 8 + 1$ (base 8). ∎

Exercise 2.4: *The greengrocer Reuven has a scale and for each integer $n \geq 0$ he has one weight weighing 2^n kilos. Explain how can Reuven weigh each merchandise weighing a positive integral number of kilos.* RESd input, 139

Exercise 2.5: *Also the butcher Shimeon has a scale and for each $n \geq 0$ he has one weight weighing 3^n kilos. Explain how can Shimeon weigh each merachndise weighing a positive integral number of kilos.* RESe input, 146

3. Greatest Common Divisor

DIVISION
input, 9

We say that an integer a **divides** an integer b and write $a|b$, if there exists an integer a' such that $aa' = b$. The following properties of division hold for all $a, b, c \in \mathbb{Z}$.

(1a) $a|a$.

(1b) If $a|b$ and $b|c$, then $a|c$.

(1c) $a|b$ implies $a|bc$.

(1d) If $a|b$, then $\pm a| \pm b$.

(1e) If $a|b$ and $a|c$, then $a|(b \pm c)$.

(1f) $\pm 1|d$ for each $d \in \mathbb{Z}$.

(1g) $a|0$.

(1h) If $a \neq 0$, then $0 \nmid a$.

(1i) If $a|b$ and $b \neq 0$, then $|a| \leq |b|$.

Definition 3.1: Let $a, b, d \in \mathbb{N}$. We say that d is a **greatest common divisor** of a and b if

DIVa
input, 35

(2a) $d|a$ and $d|b$, and

(2b) if $c|a$ and $c|b$, then $c|d$.

For nonzero $a, b \in \mathbb{Z}$ we define the greatest common divisor of a and b as the greatest common divisor of $|a|$ and $|b|$. ∎

By (1i), $|c| \leq d \leq |a|, |b|$ for each common divisor c of a and b. This implies that a and b have at most one greatest common divisor. Note that if $a|b$, then a is a greatest common divisor of a and b. More generally, we have:

Proposition 3.2: *Every nonzero $a, b \in \mathbb{Z}$ have a greatest common divisor.*

DIVb
input, 57

Proof: It suffices to prove the proposition for $a, b \in \mathbb{N}$. We do it by induction on $\min(a, b)$. To this end we assume without loss that $a \leq b$ and divide b by a with a residue:

$$b = qa + r, \quad q \in \mathbb{N}, \text{ and } 0 \leq r \leq a - 1. \tag{3}$$

Then $\min(a, r) = r < a = \min(a, b)$. If $r = 0$, then $a|b$, so a is the greatest common divisor of a and b. If $r > 0$, then $\min(r, a) = r < a = \min(a, b)$. Hence, the induction hypothesis yields a common divisor d of r and a. By (3), d is also the greatest common divisor of a and b. ∎

Following Proposition 3.2, we write $\gcd(a,b)$ for the unique greatest common divisor of a and b. It turns out that $\gcd(a,b)$ is a linear combination of a and b with integral coefficients:

PROPOSITION 3.3: *For all $a, b \in \mathbb{N}$ there exist $x, y \in \mathbb{Z}$ such that* $\gcd(a,b) = ax + by$. DIVc input, 86

Proof: We proceed again by induction on $\min(a,b)$. Without loss we assume that $a \leq b$ and divide b by a as in (3). If $r = 0$, then $a = \gcd(a,b) = 1 \cdot a + 0 \cdot b$. Otherwise $d = \gcd(a,b) = \gcd(a,r)$ and $\min(a,r) = r < a = \min(a,b)$. The induction hypothesis gives $x', y' \in \mathbb{Z}$ such that $d = ax' + ry'$. Substitution r from (3) in the latter equality, we get

$$d = ax' + (b - qa)y' = ax' + by' - qay' = a(x' - qy') + by',$$

as desired. ∎

The usual algorithm to compute the greatest common divisor of natural numbers uses factorization of those numbers into products of prime numbers. If the numbers involved are big, the factorization is a lengthy operation. We propose a much quicker procedure:

Procedure 3.4: The Euclid algorithm for the computation of the greatest common divisor. DIVd input, 118 Given natural numbers $a \leq b$ we apply division with residue several times to compute a descending sequence of natural numbers $a = r_0 > r_1 > \cdots > r_n > 0$ and a sequence of natural numbers $q_1, \dots, q_n, q_{n+1}$ such that

$$
\begin{aligned}
b &= q_1 a + r_1 \qquad 0 < r_1 < a = r_0 \\
r_0 &= q_2 r_1 + r_2 \qquad 0 < r_2 < r_1 \\
r_1 &= q_3 r_2 + r_3 \qquad 0 < r_3 < r_2 \\
&\vdots \\
r_{n-3} &= q_{n-1} r_{n-2} + r_{n-1} \qquad 0 < r_{n-1} < r_{n-2} \\
r_{n-2} &= q_n r_{n-1} + r_n \qquad 0 < r_n < r_{n-1} \\
r_{n-1} &= q_{n+1} r_n
\end{aligned} \tag{4}
$$

Then $\gcd(b, a) = \gcd(a, r_1) = \gcd(r_1, r_2) = \cdots = \gcd(r_{n-1}, r_n) = r_n$. ∎

Example 3.5: We apply Euclid's algorith to compute $\gcd(4147, 10672)$. DIVe input, 152

$$
\begin{aligned}
10672 &= 4147 \cdot 2 + 2378 \\
4147 &= 2378 \cdot 1 + 1769 \\
2378 &= 1769 \cdot 1 + 609 \\
1769 &= 609 \cdot 2 + 551 \\
609 &= 551 \cdot 1 + 58 \\
551 &= 58 \cdot 9 + 29 \\
58 &= 29 \cdot 2
\end{aligned}
$$

It follows that $\gcd(4147, 10672) = 29$. ∎

Since $a = r_0 > r_1 > \cdots > r_{n-1} > r_n \geq 1$, the number of steps in the procedure (4) is at most a. The next result improves this estimate considerably.

PROPOSITION 3.6: *The number n of steps to compute $\gcd(a, b)$ for natural numbers $a \leq b$ by the Euclid algorithm (Remark 3.4) satisfies $n < 2\log_2 a = \frac{2}{\log 2}\log a$.* DIVf input, 179

Proof: We use the notation of Remark 3.4.

CLAIM: $r_i < \frac{1}{2} r_{i-2}$. By construction

(5a) $r_i < r_{i-1} < r_{i-2}$ and

(5b) $r_{i-2} = q_i r_{i-1} + r_i$.

If $r_{i-1} \leq \frac{1}{2} r_{i-2}$, then $r_i < \frac{1}{2} r_{i-2}$ (by (5a). If $r_{i-1} > \frac{1}{2} r_{i-2}$, then $q_i r_{i-1} > \frac{1}{2} r_{i-1}$, so by (5b),

$$r_i = r_{i-2} - q_i r_{i-1} < r_{i-2} - \frac{1}{2} r_{i-2} = \frac{1}{2} r_{i-2}.$$

Thus, in both cases,

$$r_{i-2} > 2r_i \tag{6}$$

Now we consider the inequalities $a > r_1 > r_2 > \cdots > r_n$ and distinguish between two cases:

Case A: $n = 2m$. Then, $a > r_2 > r_4 > r_6 > \cdots > r_{2m} \geq 1$. Hence,

$$a > 2r_2 > 2^2 r_{2\cdot 2} > 2^3 r_{2\cdot 3} > 2^4 r_{2\cdot 4} > \cdots > 2^m r_{2\cdot m} \geq 2^m,$$

Hence, $2^m < a$, so $\frac{n}{2} = m < \log_2 a$.

Case B: $n = 2m + 1$. Then $a > r_2 > r_4 > \cdots > r_{2m} > r_{2m+1} = r_n \geq 1$. Hence, $r_{2m} \geq 2$ and $a > 2^m r_{2m} \geq 2^{m+1}$. Therefore, $m + 1 < \log_2 a$, so $n = 2m + 1 < 2m + 2 < 2\log_2 a$. ▮

Example 3.7: In example 3.5 we have computed $\gcd(4147, 10672)$ in 6 steps. The estimate that Proposition 3.6 is $\frac{2\log 4147}{\log 2} \approx 24$, so even bigger. ▮ DIVg input, 237

Procedure 3.8: Computation of $\gcd(a, b)$ *as a linear combination of* a *and* b. Given $a, b \in \mathbb{N}$, Proposition 3.3 guarantees the existence of $x, y \in \mathbb{Z}$ such that $d = \gcd(a, b) = ax + by$. Procedure 3.4 allows us to compute x and y in $2\log_2 \min(a, b)$ steps. DIVi input, 247

Indeed, in the notation of that procedure, $d = r_n$. By the equation berfore the last of (4)

$$d = r_{n-2} - q_n r_{n-1}. \tag{7}$$

By the second before the last equation of (4), $r_{n-1} = r_{n-3} - q_{n-1} r_{n-2}$. Substituting this value of r_{n-1} in (7), we get

$$d = (1 + q_{n-1}) r_{n-2} - q_n r_{n-3}. \tag{8}$$

Next we write r_{n-2} as a linear combination of r_{n-3} and r_{n-4}, substitute that in (8), and get an expression of d as a linear combination of r_{n-3} and r_{n-4}. By Proposition 3.6, after $2\log_2 \min(a,b)$ steps like this, we express d as a linear combination of a and b. ∎

Example 3.9: We use the data of Example 3.5 and apply Procedrue 3.8 to express $\gcd(4147, 10672)$ as a linear combination of 4147 and 10672 with integral coefficients. DIVk input, 277

$$\begin{aligned} 29 &= 551 - 58 \cdot 9 \\ &= 551 - (609 - 551) \cdot 9 = 551 \cdot 10 - 609 \cdot 9 \\ &= (1769 - 609 \cdot 2) \cdot 10 - 609 \cdot 9 = 1769 \cdot 10 - 609 \cdot 29 \\ &= 1769 \cdot 10 - (2378 - 1769) \cdot 29 = 1769 \cdot 39 - 2378 \cdot 29 \\ &= (4147 - 2378) \cdot 39 - 2378 \cdot 29 = 4147 \cdot 39 - 2378 \cdot 68 \\ &= 4147 \cdot 39 - (10672 - 4147 \cdot 2) \cdot 68 = 4147 \cdot 175 - 10672 \cdot 68. \end{aligned}$$ ∎

Exercise 3.10: Write a computer program to compute $\gcd(a,b)$ and apply it to compute the following examples: DIVh input, 303

$$\begin{array}{rl} \gcd(46368, 987) & \gcd(196418, 39088169) \\ \gcd(196418, 3524578) & \gcd(121393, 3524578) \\ \gcd(10610209857723, 4807526976) & \gcd(211485077978050, 259695496911122585) \end{array}$$

Exercise 3.11: Prove that a necessary and sufficient condition on $a, b, c \in \mathbb{N}$ to satisfy $\gcd(a,b,c) = 1$ is that there exist $x, y \in \mathbb{Z}$ with $\gcd(ax + by, c) = 1$. ∎ DIVj input, 315

Exercise 3.12: Prove that if $\gcd(a,b) = 1$, then $\gcd(a+b, a-b)$ is 1 or 2. ∎ DIVl input, 321

4. The Prime Numbers

PRIME input, 8

The fundamental theorem of Number Theory says that every natural number is a product of prime numbers in a unique way up to the order of the factors.

Definition 4.1: A natural number p is said to be **prime** if $p > 1$ and if 1 and p are the only positive divisors of p. PRIa input, 13

A natural number n is **composite** if $n > 1$ and n is not a prime number. Thus, n has at least one divisor d with $1 < d < n$. $\blacksquare$

The first prime numbers are $2, 3, 5, 7, 11, 13, 17, 19, 23, \ldots$ and the first composite numbers are $4, 6, 8, 9, 10, 12, 14, 15, 16, \ldots$.

The following lemma helps to determine whether a natural number is prime.

LEMMA 4.2: *Let $p \geq 2$ be a natural number that has no divisor $d > 1$ with $d \leq \sqrt{p}$. Then p is prime.* PRIb input, 30

Proof: Assume p is composite. Then $p = ab$ with $1 < a < p$, so also $1 < b < p$. By assumption, $a, b > \sqrt{p}$. Hence, $p = ab > p$, a contradiction. $\blacksquare$

Example 4.3: In order to find out whether 83 is prime we need to check only the natural numbers up to $\sqrt{83} \approx 9$. Indeed, $2 \nmid 83$, $3 \nmid 83$, $4 \nmid 83$, $5 \nmid 83$, $6 \nmid 83$, $7 \nmid 83$, $8 \nmid 83$, $9 \nmid 83$, so 83 is indeed a prime number. $\blacksquare$ PRIc input, 43

Here is an effective way to compute all prime numbers less that a given natural number n.

Algorithm 4.4: The sieve of Eratosthenes. We write down the sequence of natural numbers from 2 to n. We circle the first number 2 in the sequence and then cross out (but not erase) each second number. Having done so, we find that the first non-circled and non-crossed number, namely 3, circle it and cross out every third number (taking into account also numbers that have been crossed out before). Then we continue with the next non-circled and non-crossed number, namely 5, and so on. When we go over $\sqrt{n}$, we circle all of the remaining numbers. By Lemma 4.2, the circled numbers are all of the prime numbers up to n. $\blacksquare$ PRId input, 61

Definition 4.5: We say that nonzero integral elements a and b are **relatively prime** if $\gcd(a,b) = 1$. In view of Proposition 3.3, this condition is equivalent to the existence of $x, y \in \mathbb{Z}$ such that $ax + by = 1$. PRIe input, 77

In particular, a prime number p is relatively prime to b if and only if $p \nmid b$. Two primes p and q are relatively prime if and only if $p \neq q$. ∎

LEMMA 4.6: *Every natural number n can be represented as a product of prime numbers.* PRIf input, 90

Proof by induction on n: If $n = 1$, then n is a product of an empty set of prime numbers. If n is prime, then the product contains only one factor, namely n. If n is composite, then $n = ab$ with $1 < a, b < n$. By assumption, $a = p_1 \cdots p_r$ and $b = q_1 \cdots q_s$, where the p_i's and the q_j's are prime numbers. Thus, $n = p_1 \cdots p_r q_1 \cdots q_s$ is the required representation of n. ∎

LEMMA 4.7: PRIg input, 107

(a) *If $a|mb$ and a is relatively prime to b, then $a|m$. In particular, if p is a prime number and $p|ab$, then $p|a$ or $p|b$.*

(b) *If $p, q_1, \ldots, q_r$ are prime numbers and $p|q_1 \cdots q_m$, then $p = q_i$ for some $1 \leq i \leq r$.*

(c) *If a and b are relatively prime and both divide m, then $ab|m$.*

Proof of (a): Our assumption gives $x, y \in \mathbb{Z}$ with $ax + by = 1$. Hence, $max + mby = m$. It follows that $a|m$.

Proof of (b): If $p = q_1$, we are done. Otherwise p is relatively prime to q_1. Hence, by (a), $p|q_2 \cdots q_r$. Now apply an induction hypothesis on r to conclude the existence of $2 \leq i \leq r$ with $p = q_i$.

Proof of (c): By assumption there exists a' such that $aa' = m$, so $b|aa'$. Since $\gcd(a,b) = 1$, it follows from (a) that $b|a'$. Thus, there exists b' with $bb' = a'$. It follows that $abb' = aa' = m$, so $ab|m$. ∎

THEOREM 4.8 (The fundamental theorem of arithmetic): *Every natural number n can be represented as a product of prime numbers in a unique way up to the order of the factors.* PRIh input, 141

Proof: In view of Lemma 4.6 we have to prove only the uniqueness of the representation.

Indeed, let $n = p_1 \cdots p_r$ and $n = q_1 \cdots q_s$ be two representations of n as a product of prime numbers. Assume without loss that $n \geq 2$, so that $r, s \geq 1$. By assumption, $p_1 | q_1 \cdots q_s$. By Lemma 4.7(b), $p_1 = q_j$ for some $1 \leq j \leq s$. Renumbering $q_1, \ldots, q_s$, we may assume that $j = 1$, so $p_1 = q_1$. Thus, $p_1 p_2 \cdots p_r = p_1 q_2 \cdots q_s$. It follows that $p_2 \cdots p_r = q_2 \cdots q_s$. Applying an induction hypothesis on r, we conclude that $r - 1 = s - 1$, hence $r = s$, and after a renumerations of $q_2, \ldots, q_s$, we have $p_2 = q_2, \ldots, p_r = q_r$, as claimed. ∎

Exercise 4.9: Let $f(X)$ be a non-constant polynomial with integral coefficients. Prove that $f(x)$ is composite for infinitely many $x \in \mathbb{N}$. Hint: Replace X by $Y = X + c$ for some $c \in \mathbb{Z}$ such that the zeroth term of $g(Y) = f(Y - c)$ is different from ± 1. ∎

DIVm
input, 165

5. The p-adic Valuation

PADIC
input, 8

We may gather all the prime factors of a natural number n to powers of distinct primes and represent n in the form $n = \prod_{i=1}^{m} p_i^{\alpha_i}$, where $p_1, \ldots, p_n$ are distinct primes and $\alpha_1, \ldots, \alpha_m \in \mathbb{N}$. For each prime number p we may define $v_p(n)$ as α_i if $p = p_i$ and 0 if $p \notin \{p_1, \ldots, p_n\}$. This yields the representation

$$n = \prod_p p^{v_p(n)}, \tag{1}$$

where p ranges over all prime numbers, $v_p(n) \geq 0$ for each p, and $v_p(n) = 0$ for all but finitely many p's (we may also say **for almost all** p). The uniqueness part of the fundamental theorem of Number Theory gurantees that $v_p(n)$ is well defined. It is the exponent of the highest power of p that divides n. For example, $60 = 2^2 \cdot 3 \cdot 5$, so $v_2(60) = 2$, $v_3(60) = 1$, $v_5(60) = 1$, and $v_p(60) = 0$ for all $p \notin \{2, 3, 5\}$.

For a fixed p the function v_p has the following properties:

(2a) $v_p(n) \geq 0$ for each $n \in \mathbb{N}$.

(2b) $v_p(1) = 0$.

(2c) $v_p(mn) = v_p(m) + v_p(n)$.

(2d) $v_p(m + n) \geq \min(v_p(m), v_p(n))$. Indeed, if $v_p(m) \leq v_p(n)$, then $p^{v_p(m)}$ divides both m and n, so also $m + n$.

(2e) If $v_p(m) \neq v_p(n)$, then $v_p(m + n) = \min(v_p(m), v_p(n))$. Indeed, assume $\alpha = v_p(m) < v_p(n)$. Then $p^\alpha | m+n$, by (2d). Also, by assumption, $p^{\alpha+1} | n$. If $p^{\alpha+1} | (m+n)$, then $p^{\alpha+1} | m$, which contradicts the definition of $v_p(m)$. Hence, $v_p(m + n) = \alpha = v_p(m)$, as claimed.

(2f) It follows by induction from (2e) that if $v_p(m_1) < v_p(m_i)$ for $i = 2, \ldots, r$, then $v_p(\sum_{i=1}^{r} m_i) = v_0(m_1)$.

(2g) $m | n$ if and only if $v_p(m) \leq v_p(n)$ for every prime number p.

(2h) If $v_p(m) = v_p(m)$ for all prime numbers p, then $m = n$.

Next we extend v_p to a function from the set of all positive rational numbers to $\mathbb{Z}$ by defining $v_p\left(\frac{m}{n}\right) = v_p(m) - v_p(n)$. If $\frac{m'}{n'} = \frac{m}{n}$, then $m'n = n'm$. By (2c), $v_p(m') + v_p(n) = v_p(n') + v_p(m)$, hence $v_p(n) - v_p(m) = v_p(n') - v_p(m')$, so $v_p\left(\frac{m}{n}\right)$ is well defined.

Next we define $v_p(u)$ for each negative rational number as $v_p(-u)$. Finally, we add the symbol ∞ to $\mathbb{Z}$ with the following rules:

(3a) $\infty > z$ for each $z \in \mathbb{Z}$.

(3b) $z + \infty = \infty + \infty = \infty$ for each $z \in \mathbb{Z}$

and define $v_p(0) = \infty$. Then it is not difficult to prove that the function $v_p\colon \mathbb{Q} \to \mathbb{Z} \cup \{\infty\}$ satisfies (2b)–(2e) (but not (2a)) for all $m, n \in \mathbb{Q}$. It is called the ***p*-adic valuation** of $\mathbb{Q}$ and satisfies $u = \prod_p p^{v_p(u)}$ for all nonzero $u \in \mathbb{Q}$, where now $v_p(u) \in \mathbb{Z}$ for each p and $v_p(u) = 0$ for almost all n.

The original definition of $\gcd(m, n)$ for two natural numbers m and n gives

$$\gcd(m, n) = \prod_p p^{\min(v_p(m), v_p(n))}. \tag{4}$$

Indeed, let $d = \gcd(m, n)$ and $d' = \prod_p p^{\min(v_p(m), v_p(n))}$. Then, $d|m$ and $d|n$, so for every prime number p we have $v_p(d) \le v_p(n)$ and $v_p(d) \le v_p(n)$, hence $v_p(d) \le \min(v_p(m), v_p(n)) = v_p(d')$. Therefore $d|d'$.

Conversely, for each p we have $v_p(d') = \min(v_p(m), v_p(n))$. Hence, $v_p(d') \le v_p(m)$ and $v_p(d') \le v_p(n)$. Therefore, $d'|m$ and $d'|n$, so $d'|\gcd(m, n) = d$. Combining the latter relation with the consequence of the proceeding paragraph, we conclude that $d = d'$, as asserted.

The **least common multiple** $\operatorname{lcm}(m, n)$ of natural numbers m and n is defined to be the unique natural number l such that

(5a) $m, n|l$, and

(5b) if $m, n|l'$, then $l|l'$.

Similarly to (4), we have

$$\operatorname{lcm}(m, n) = \prod_p p^{\max(v_p(m), v_p(n))}. \tag{6}$$

For example, $\operatorname{lcm}(12, 18) = \operatorname{lcm}(2^2 \cdot 3, 2 \cdot 3^2) = 2^2 \cdot 3^2 = 36$. It follows from (4) and (6) that

$$\gcd(m, n)\operatorname{lcm}(m, n) = mn. \tag{7}$$

This follows from the identity $\min(v_p(m), v_p(n)) + \max(v_p(m), v_p(n)) = v_p(m) + v_p(n)$ that holds for each prime number p. If m and n are relatively prime, then $\gcd(m, n) = 1$, so $\operatorname{lcm}(m, n) = mn$.

PADa input, 144

EXERCISE 5.1: *Prove that if a* **reduced quotient** $x = \frac{a}{b}$ *(thus,* $\gcd(a, b) = 1$*) of integers is a root of an equation*

$$c_n X^n + c_{n-1} X^{n-1} \cdots + c_0 = 0$$

with $c_0, \ldots, c_n \in \mathbb{Z}$ *and* $c_0, c_n \neq 0$*, then* $a|c_0$ *and* $b|c_n$*. In particular, if* $c_n = 1$*, then* $x \in \mathbb{Z}$*. Conclude that if* k *is an integer and* $\sqrt[n]{k} \in \mathbb{Q}$ *then* $\sqrt[n]{k} \in \mathbb{Z}$*. In particular, conclude that* $\sqrt{2}$ *is not a rational number.*

6. Infinitude of Prime Numbers

INFI input, 9

Observation shows that the prime numbers become rare as one goes up in the sequence of natural numbers. For example, 40% of the first ten numbers are primes, 30% of the first fifty numbers are primes, but only 25% of the first hundred numbers are primes. One may ask whether from some point on, the prime numbers stop to exist. The followig result of Euclid from two thousand years ago says this is not the case.

THEOREM 6.1: *There are infinitely many prime numbers.* INFa input, 20

Proof by contradiction: Assume there are only finitely many prime numbers $p_1, p_2, \dots, p_n$, with $p_1 = 2$, $p_2 = 3$, and so on. Consider the natural number $m = 1 + p_1 p_2 \cdots p_n$ and observe that $m \neq 1$. Hence, by Lemma 4.6, m has a prime factor p. Then $p = p_i$ for some $1 \leq i \leq n$. It follows that $p|1$, which is a contradiction. Consequently, there are infinitely many prime numbers. ∎

We may partition $\mathbb{N} \smallsetminus 4\mathbb{N}$ into three disjoint sets: $\{2(2k-1) \mid k = 1, 2, 3, \dots\}$, $\{4k+1 \mid k = 1, 2, 3, \dots\}$, $\{4k-1 \mid k = 1, 2, 3, \dots\}$. It is true that each of the two latter sets contains infinitely many primes. The proof that there are infinitely many primes of the form $4k+1$ applies more tools and will be given later. The proof that the third set contains infinitely many primes is easier.

THEOREM 6.2: *There are infinitely many prime numbers of the form* $4k-1$. INFb input, 47

Proof: Assume there are only finitely many prime numbers $p_1, \dots, p_r$ of the form $4k-1$. Let $m = 4p_1 p_2 \cdots p_r - 1$. Write m as a product of prime numbers $m = q_1 \cdots q_s$. Now note that $(4a+1)(4b+1) = 4(4ab+a+b)+1$ for all $a, b \in \mathbb{Z}$, so there exists $1 \leq j \leq s$ such that q_j has the form $4k-1$. Our assumption gives $1 \leq i \leq r$ such that $q_j = p_i$, so $p_i|m$ and $p_i|4p_1 p_2 \cdots p_r$. Hence, $p_i|1$. This contradiction to our initial assumption prove that it is false. ∎

It turns out that much more is true:

THEOREM 6.3 (Dirichlet): *If $a, b \in \mathbb{Z}$ are relatively prime, then there are infinitely many prime numbers of the form $ak+b$.* INFc input, 69

Dirichlet's proof of Theorem 6.3 applies tools from the theory of complex numbers and is beyond the framework of our course.

Exercise 6.4: Find for each $n \in \mathbb{N}$ a sequence of n consecutive composite natural numbers. Hint: Use the **faculty operation** $n! = 1 \cdot 2 \cdots n$. ▮ INFd input, 78

7. Congruences

CONGR input, 8

Our next goal is to prove Fermat's little theorem saying that $p|(a^{p-1} - 1)$ for each prime number p and every $a \in \mathbb{Z}$ not divisible by p. For example, $2^{3-1} - 1 = 3$, $2^{5-1} - 1 = 15 = 3 \cdot 5$, $2^{7-1} - 1 = 63 = 3^2 \cdot 7$ and $2^{11-1} = 1023 = 11 \cdot 93$. The most convenient way to prove the theorem is to use the notion of congruences that we develope in this section.

Definition 7.1: Let n be a natural number and $a, b \in \mathbb{Z}$. We say that a **and** b **are congruent modulo** n and write $a \equiv b \mod n$ if $n|(a-b)$. ▮

CONa input, 19

The congruence relation satisfies the following rules:

(1a) $a \equiv a \mod n$.

(1b) If $a \equiv b \mod n$, then $b \equiv a \mod n$.

(1c) If $a \equiv b \mod n$ and $b \equiv c \mod n$, then $a \equiv c \mod n$.

(1d) If $a \equiv a' \mod n$ and $b \equiv b' \mod n$, then $a+b \equiv a'+b' \mod n$ and $ab \equiv a'b' \mod n$.

(1e) If $a \equiv b \mod n$ and $f \in \mathbb{Z}[X]$ (that is, $f(X) = a_r X^r + a_{r-1}X^{r-1} + \cdots + a_0$ with $a_0, \ldots, a_{r-1}, a_r \in \mathbb{Z}$) is a **polynomial with integral coefficients**), then $f(a) \equiv f(b) \mod n$.

Conditions (1a), (1b), and (1c) mean that the congruence relation modulo n is an **equivalent relation**.

A special case of (1d) is that if $a \equiv b \mod n$, then $ka \equiv kb \mod n$ for all a, b, k, n. However, $ka \equiv kb \mod n$ does not imply $a \equiv b \mod n$. For example, $4 \cdot 3 \equiv 4 \cdot 6 \mod 12$. However $3 \not\equiv 6 \mod 12$. Nevertheless, the following complementary rule is valid:

(2a) If $ka \equiv kb \mod kn$, then $a \equiv b \mod n$ for each $k \in \mathbb{N}$.

(2b) If $a \equiv b \mod n$, then $n|a$ if and only if $n|b$. In particular, $a \equiv 0 \mod n$ means that $n|a$.

(2c) If $a \equiv b \mod n$, $k|b$, and $k|n$, then $k|a$.

Here are applications of congruences to detect divisibility by 3, 9, and 11:

Examples:

(a) A natural number a is divisible by 3 if and only if the sum of its digits in decimal representation is divisible by 3.

Indeed, there exists a polynomial $f(X) = \sum_{i=0}^{r} c_i X^i$ such that

$$f(10) = \sum_{i=0}^{r} c_i (10)^i = a.$$

Since $10 \equiv 1 \mod 3$, we have $a \equiv f(1) \mod 3$ (by (1e)). Thus, $a \equiv \sum_{i=0}^{r} c_i \mod 3$, which implies our statement.

For example $101 = 1 \cdot 10^2 + 0 \cdot 10 + 1$ is not divisible by 3 because $2 = 1 + 0 + 1$ is not.

(b) Since $10 \equiv 1 \mod 9$, divisiblity by 9 satisfies the same divisibility rule as divisibility by 3.

(c) A natural number a is divisible by 11 if and only if the sum of its even digits in its decimal representation minus the sum of its odd digits is divisible by 11. To this end note that the unit digit stands in the zeroth place.

Indeed, $10 \equiv -1 \mod 11$. Consider the number 862983 and observe that

$$(3 + 9 + 6) - (8 + 2 + 8) = 0,$$

so $11|862983$. ∎

LEMMA 7.2: *If* $ab \equiv ab' \mod n$ *and* $\gcd(a, n) = 1$, *then* $b \equiv b' \mod n$. CONb input, 109

Proof: By Proposition 3.3, there exist $x, y \in \mathbb{Z}$ such that $ax + yn = 1$, so $ax \equiv 1 \mod n$. Thus, $axb \equiv axb' \mod n$ implies that $b \equiv b' \mod n$. ∎

Here is another application of congruences that avoids tedious computation.

LEMMA 7.3: *If a is relatively prime to b_i for $i = 1, \ldots, r$, then a is relatively prime to $b = b_1 \cdots b_r$.* CONc input, 123

Proof: Proposition 3.3 gives for each i integers x_i, y_i such that $ax_i + b_i y_i = 1$. Thus, $b_i y_i \equiv 1 \mod a$ for $i = 1, \ldots, r$, so

$$by_1 \cdots y_r = (b_1 y_1) \cdots (b_r y_r) \equiv 1 \mod a.$$

Consequently, $\gcd(b, a) = 1$. ∎

Exercise 7.4: Prove that the sum of two squeres of odd integers is never a square. Hint: consider the integers involved modulo 4. ▮ CONd input, 141

Exercise 7.5: Prove that the unit digit of each square of an integer in its octal representation (that is, basis 8) is 0, 1, or 4. ▮ CONe input, 147

Exercise 7.6: Prove that the unit digit of each fourth power of an integer in its decimal representation is 0, 1, 5, or 6. ▮ CONf input, 153

Exercise 7.7: Prove that 19 divides $4n^2 + 4$ for no $n \in \mathbb{N}$. ▮ CONg input, 158

Exercise 7.8: What is the unit digit in the decimal representation of 3^{400}? Hint: Present 400 as a sum of powers of 2 with coefficients in $\{0, 1\}$. Then successively compute the value of 3^{2^k} modulo 10 for $k = 0, 1, 2, \ldots$. ▮ CONh input, 162

Exercise 7.9: Prove that $\prod_{i=1}^{100}(x + i) \equiv 0 \mod 100$ for each $x \in \mathbb{Z}$. ▮ CONj input, 172

8. The Ring $\mathbb{Z}/n\mathbb{Z}$

RINGS input, 8

Properties (1a), (1b), and (1c) of Section 7 say that the congruence relation modulo n is an **equivalent relation** on $\mathbb{Z}$. Thus, $\mathbb{Z}$ is the disjoint union of the (finitely many) congruence classes. We define addition and multiplication on these classes, making them the elements of a "ring" that we denote by $\mathbb{Z}/n\mathbb{Z}$.

We start with the notation

$$a + n\mathbb{Z} = \{a + nz \mid z \in \mathbb{Z}\}$$

and note that $a + n\mathbb{Z} = \{b \in \mathbb{Z} \mid a \equiv b \mod n\}$. We call $a + n\mathbb{Z}$ a **congruence class modulo** n. By definition, $a + n\mathbb{Z} = a' + n\mathbb{Z}$ if and only if $a \equiv a' \mod n$. Each a' satisfying the latter condition is said to be a **representative** of $a + n\mathbb{Z}$.

LEMMA 8.1: $\mathbb{Z} = \dot{\bigcup}_{a=0}^{n-1}(a + n\mathbb{Z})$. RINa input, 29

Proof: Indeed, if $b \in \mathbb{Z}$, then $b = a + nq$ with $0 \le a \le n-1$ and $q \in \mathbb{Z}$, so $b \in a + n\mathbb{Z}$. If $0 \le a < a' \le n-1$, then $0 < a' - a \le n-1$, so $n \nmid a - a'$, hence $a \not\equiv a' \mod n$, hence $a + n\mathbb{Z}$ and $a' + n\mathbb{Z}$ are disjoint. ∎

Thus, the set $\mathbb{Z}/n\mathbb{Z} = \{a + n\mathbb{Z} \mid a = 0, \ldots, n-1\}$ of all congruence classes modulo n consists of n elements. We define addition and multiplication on $\mathbb{Z}/n\mathbb{Z}$ by the rules:

$$(a + n\mathbb{Z}) + (a' + n\mathbb{Z}) = (a + a') + n\mathbb{Z}, \qquad (a + n\mathbb{Z})(a' + n\mathbb{Z}) = aa' + n\mathbb{Z},$$

If $a \equiv b \mod n$ and $a' \equiv b' \mod n$, then $a + a' \equiv b + b' \mod n$ and $aa' \equiv bb' \mod n$, so both addition and multiplication of congruence classes modulo n is well defined. One checks that Conditions (4), (5), and (6) of Section 1 hold for $\mathbb{Z}/n\mathbb{Z}$. Thus, $\mathbb{Z}/n\mathbb{Z}$ is a commutative ring with $0 + n\mathbb{Z}$ (that can also be written as $n\mathbb{Z}$) as the zero element and $1 + n\mathbb{Z}$ as the one element.

Example 8.2: In contrast to the situation in $\mathbb{Z}$, it may happen that the product of two nonzero elements of $\mathbb{Z}/n\mathbb{Z}$ is zero. For example, $(2 + 6\mathbb{Z})(3 + 6\mathbb{Z}) = 6\mathbb{Z}$. A nonzero element a of a commutative ring R is a **zero divisor** if there exists a nonzero $b \in R$ such that $ab = 0$. Thus, both $2 + 6\mathbb{Z}$ and $3 + 6\mathbb{Z}$ are zero divisors of $\mathbb{Z}/6\mathbb{Z}$. ∎ RINb input, 64

LEMMA 8.3: *Let R be a finite commutative ring with 1 and let a be a nonzero divisor of R. Then the map α: $R \to R$ defined by $\alpha(x) = ax$ is bijective. In particular, a is* **invertible** *in R, i.e. there exists $a' \in R$ such that $aa' = 1$.* RINc input, 76

Proof: If $ax = ay$, then $a(x - y) = 0$, so $x - y = 0$, hence $x = y$. It follows that α is injective. Since R has been assumed to be finite, α is also surjective, hence bijective.

In particular, there exists $a' \in R$ such that $\alpha(a') = 1$, so $aa' = 1$. ■

Definition 8.4: A **field** is a commutative ring F with 1 in which $1 \neq 0$ and every nonzero element is invertible. In more details, a set F with distinguished elements 0 and 1 and two operations $+$ and $\cdot$ is a field if all x, y, z satisfy the following conditions. RINd input, 96

(1a) $x + y = y + x$.

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

(1c) $0 + x = x$.

(1d) There exists an element $-x \in F$ such that $x + (-x) = 0$.

(2a) $xy = yx$.

(2b) $(xy)z = x(yz)$.

(2c) $1 \neq 0$ and $1 \cdot x = x$.

(2d) If $x \neq 0$, then there exists an element $x^{-1} \in F$ such that $xx^{-1} = 1$.

(3) $x(y + z) = xy + xz$. ■

Examples 8.5: (a) The set $\mathbb{Q}$ of all quotients of integers with nonzero denominators is a field called the **field of rational numbers**. RINe input, 127

(b) If p is a prime number, then $\mathbb{Z}/p\mathbb{Z}$ has no zero divisors. Indeed, let $a, a' \in \mathbb{Z}$ such that $a + p\mathbb{Z} \neq p\mathbb{Z}$ and $a' + p\mathbb{Z} \neq \mathbb{Z}$. Then $p \nmid a$ and $p \nmid a'$, so p is relatively prime to both a and a'. It follows from Lemma 4.7(a) that $p \nmid aa'$, so $(a + p\mathbb{Z})(a' + p\mathbb{Z}) = aa' + p\mathbb{Z} \neq p\mathbb{Z}$. Thus, neither $a + p\mathbb{Z}$ nor $a' + p\mathbb{Z}$ are zero divisors of $\mathbb{Z}/n\mathbb{Z}$.

It follows from Lemma 8.3, that every nonzero element of $\mathbb{Z}/p\mathbb{Z}$ is invertible. Therefore, $\mathbb{Z}/p\mathbb{Z}$ is a field. Whenever we want to emphasize that $\mathbb{Z}/p\mathbb{Z}$ is a field, we denote it by $\mathbb{F}_p$.

(c) If n is composite, then $n = ab$ with $1 < a, b < n$. Thus $(a + n\mathbb{Z})(b + n\mathbb{Z}) = n\mathbb{Z}$, so both $a + n\mathbb{Z}$ and $b + n\mathbb{Z}$ are zero divisors of $\mathbb{Z}/n\mathbb{Z}$.

In general, an element $a + n\mathbb{Z}$ of $\mathbb{Z}/n\mathbb{Z}$ is invertible, if and only if $\gcd(a, n) = 1$. Indeed, in this case there exist $x, y \in \mathbb{Z}$ such that $ax+ny = 1$, so $ax \equiv 1 \mod n$. Hence, $(a + n\mathbb{Z})(x + n\mathbb{Z}) = 1 + n\mathbb{Z}$. Conversely, if $\gcd(a, n) = d$ and $d \neq 1$, then there exist $a', n' \in \mathbb{Z}$ such that $da' = a$ and $dn' = n$. Then, $1 < n' < n$ and $an' = da'n' = na' \equiv 0 \mod n$, so $a + n\mathbb{Z}$ is a zero divisor in $\mathbb{Z}/n\mathbb{Z}$.

(d) We may take $0, 1, 2, \ldots, 11$ as representatives to the elements of $\mathbb{Z}/12\mathbb{Z}$. The nonzero elements among them that are not relatively prime to 12 are $2, 3, 4, 6, 8, 9, 10$. Each of them represents a zero divisor of $\mathbb{Z}/12\mathbb{Z}$.

On the other hand, $1, 5, 7, 11$ are relatively prime to 12. They satisfy $1 \cdot 1 \equiv 1 \mod 12$, $5\cdot 5 = 25 \equiv 1 \mod 12$, $7\cdot 7 = 49 \equiv 1 \mod 12$, and $11\cdot 11 = 121 \equiv 1 \mod 12$, so each of those numbers represents an invertible element of $\mathbb{Z}/12\mathbb{Z}$. ▮

Definition 8.6: A set A of integers is said to be a **system of representatives** modulo RINf input, 185 n if the map $f\colon A \to \mathbb{Z}/n\mathbb{Z}$ defined by $f(a) = a + n\mathbb{Z}$ for $a \in A$ is bijective. With other words, for each integer $0 \le b \le n-1$ there exists a unique $a \in A$ such that $a \equiv b \mod n$. For example $0, 1, \ldots, n-1$ and $1, 2, \ldots, n$ are systems of representatives modulo n. Note that $n \equiv 0 \mod n$. Also, $\{-1, 0, 1, 2\}$ is a system of represenatives modulo 4. ▮

Exercise 8.7: Let $f(X)$ be a polynomial with integral coefficients. Denote the num- CONi input, 199 ber of solutions modulo m of the equation $f(X) \equiv k \mod m$ by $N(k)$. Prove that $\sum_{k=1}^{m} N(k) = m$. ▮

9. Fermat's Little Theorem

FERMAT
input, 9

In addition to rings and fields, it is convenient to introduce "groups".

Definition 9.1: A **group** is a set G equipped with a distinguished element 1 and a binary operation $\cdot$ (called **multiplication**) such that the following rules hold for all $x, y, z \in G$:

FERa
input, 13

(1a) $(xy)z = x(yz)$.

(1b) $1 \cdot x = x \cdot 1 = x$.

(1c) There exists $x' \in G$ such that $x'x = xx' = 1$.

We say that G is **commutative** (or **abelian**) if in addition

(1d) $xy = yx$.

In this case one sometimes use additive notation with addition replacing multiplication and 0 replacing 1. We repeat the definition in this case.

An (additive) abelian group is a set A equipped with a distinguished element 0 and a binary operation + (called **addition**) such that the following rules hold for all $x, y, z \in A$.

(2a) $(x + y) + z = x + (y + z)$.

(2b) $0 + x = x$.

(2c) There exists $x' \in A$ such that $x' + x = 0$.

(2d) $x + y = y + x$.

For example, $\mathbb{Z}$ is an abelian group with respect to addition (but not with respect to multiplication).

If a group G is finite, we call its cardinality the **order** of G. ▮

Remark 9.2: Every group G satisfies the following cancellation rules: $ax = bx$ implies $a = b$ and $xa = xb$ implies $a = b$. ▮

CAN
input, 54

Remark 9.3: The element x' satisfying (1c) is unique. Indeed, if x'' satisfies $x''x = xx'' = 1$, then $x' = x'(xx'') = (x'x)x'' = x''$. Therefore, we write x^{-1} instead of x'. In the additive case we write $-x$ for the element x' that satisfies (2c). ▮

INV
input, 60

Examples 9.4:

FERb
input, 70

(a) The collection of all invertible matrices of order $n \times n$ over a field F is a group denoted by $\mathrm{GL}_n(F)$. If $n \geq 2$, then $\mathrm{GL}_n(F)$ is not commutative. For example,

$$\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} = \begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix} \text{ but } \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ 1 & 1 \end{pmatrix}.$$

(b) The set $R^\times$ of all invertible elements of a commutative ring R is a commutative group with respect to multiplication.

(c) Given a (commutative) ring R, we may forget the multiplication of R and remain only with its addition. Then we are left with an additive abelian group. ∎

LEMMA 9.5: *Let G be a group and $a \in G$.* FERc input, 95

(a) *The map $\alpha\colon G \to G$ defined by $\alpha(x) = ax$ is bijective.*

(b) *If G is a finite commutative group of order n, then $a^n = 1$*[*].

(c) *If R is a finite commutative ring with 1 and $|R^\times| = m$, then then $a^m = 1$ for all $a \in R^\times$.*

Proof of (a): The map α has an inverse map $\alpha'\colon G \to G$ defined by $\alpha'(x) = a^{-1}x$, so α is bijective.

Proof of (b): Let $G = \{x_1, x_2, \ldots, x_n\}$. By (a), $(ax_1, ax_2, \ldots, ax_n)$ is a permutation of $(x_1, x_2, \ldots, x_n)$. Hence, $ax_1 \cdot ax_2 \cdots ax_n = x_1 \cdot x_2 \cdots x_n$. Since G is commutative, we therefore have $a^n(x_1x_2\cdots x_n) = (ax_1)(ax_2)\cdots(ax_n) = (x_1x_2\cdots x_n)$. Hence, $a^n = 1$.

Proof of (c): Apply (b) to $R^\times$ rather than to G. ∎

Definition 9.6: Given $n \in \mathbb{N}$, we denote the number of natural numbers between 1 and n that are relatively prime to n by $\varphi(n)$. By Example 8.5(c), this is also the order of the group $(\mathbb{Z}/n\mathbb{Z})^\times$. We call φ **the Euler φ-function**. By Example 8.5(b), $\varphi(p) = p - 1$ if p is a prime number. ∎ FERd input, 128

We apply Lemma 9.5(c) to the ring $\mathbb{Z}/n\mathbb{Z}$, where $n \in \mathbb{N}$.

THEOREM 9.7 (Euler): *If a is relatively prime to a natural number n, then* FERe input, 143

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

In particular, if n is a prime number p we get:

[*] This result is also true for arbitrary finite groups G.

THEOREM 9.8 (Fermat's little theorem): *If p is a prime number and $p \nmid a$, then* FERf input, 153

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

Example 9.9: FERg input, 160

(a) By Fermat's little theorem $2^{100} \equiv 1 \mod 101$. We verify this congruence by computation. First we write $100 = 64+32+4$ as a sum of powers of 2. Then we square those powers modulo 101 consecutively:

$$\begin{aligned}
2^1 &= 2\\
2^2 &= 4\\
2^4 &= 16\\
2^8 &= 256 \equiv 54 \mod 101\\
2^{16} &\equiv 54^2 \equiv 2916 \equiv 88 \equiv -13 \mod 101\\
2^{32} &\equiv (-13)^2 \equiv 169 \equiv 68 \equiv -33 \mod 101\\
2^{64} &\equiv (-33)^2 \equiv 1089 \equiv 79 \equiv -22 \mod 101\\
2^{100} &\equiv 2^{64} \cdot 2^{32} \cdot 2^4 \equiv (-22)(-33)(16) = 11616 = 45 \cdot 101 + 1 \equiv 1 \mod 101
\end{aligned}$$

Note that the number of steps in this procedure is of order of magnitude of $\log_2 101$.

(b) Again, by Fermat's little theorem $a \cdot a^{p-2} \equiv 1 \mod p$, if $p \nmid a$. Hence, a^{p-2} is the inverse of a modulo p. We may compute a representative between 0 and $p-1$ modulo p of a^{p-1} by applying the procedure used in (a). Alternatively, we may use the Euclid algorithm to find $x, y \in \mathbb{Z}$ such that $ax + py = 1$ and then to divide x by p with a remainder between 0 and $p-1$. Both procedures are effective. ∎

Another form of Fermat's little theorem is:

THEOREM 9.10: *If p is a prime number and $a \in \mathbb{Z}$, then $a^p \equiv a \mod p$.* FERh input, 199

Exercise 9.11: Prove that if $\gcd(k, \varphi(m)) = 1$, then for each a relatively prime to m there exists $x \in \mathbb{N}$ such that $x^k \equiv a \mod m$. Hint: Prove that the map $x \mapsto x^k$ maps $(\mathbb{Z}/m\mathbb{Z})^\times$ bijectively onto itself. ∎ FERi input, 204

Exercise 9.12: Prove that $42 | n^7 - n$ for each $n \in \mathbb{Z}$. ∎ FERj input, 212

Exercise 9.13: Prove that if $7 \nmid n$, then $n^{12} \equiv 1 \mod 7$. ▮ FERk input, 216

Exercise 9.14: Prove that if a and b are relatively prime to 91, then $a^{12} \equiv b^{12} \mod 91$. FERl input, 220

Hint: $91 = 7 \cdot 13$. ▮

10. Polynomials

POLYNOM
input, 9

A commutative ring R with 1 without zero divisors is called **integral domain**. For example, $\mathbb{Z}$ is an integral domain, every field is an integral domain, but $\mathbb{Z}/n\mathbb{Z}$ is not an integral domain if n is composite.

We consider an integral domain R and denote the set of all polynomials in X with coefficients in R by $R[X]$. Each $f \in R[X]$ has the form

$$f(X) = a_n X^n + a_{n-1} X^{n-1} + \cdots + a_0$$

with **coefficients** $a_0, \dots, a_{n-1}, a_n \in R$, alternatively $f(X) = \sum_{i=0}^{n} a_i X^i$. If $a_n \neq 0$, we say that the **degree** of f is n, a_n is the **leading coefficient** of f, and write $\deg(f) = n$. We say that f is **monic** if its leading coefficient is 1. An element $a \in R$ is a **zero** of f if $f(a) = 0$.

If f above has degree n and $f'(X) = \sum_{j=0}^{n'} a'_j X^j$ has degree n', then $f(X) + f'(X) = \sum_{j=0}^{\max(n,n')} (a_j + a'_j) X^j$, so

$$\text{(1)} \qquad \deg(f + f') \le \max(\deg(f), \deg(f')).$$

Also, $f(X) f'(X) = \sum_{k=0}^{n+n'} c_k X^k$, where $c_k = \sum_{i+j=k} a_i a'_j$. In particular, $c_{n+n'} = a_n a'_n \neq 0$, because R has no zero divisors. It follows that $ff' \neq 0$, so $R[X]$ is again an integral domain. Moreover,

$$\text{(2)} \qquad \deg(ff') = \deg(f) + \deg(f').$$

LEMMA 10.1: *Let R be an integral domain and $f \in R[X]$ a polynomial of degree n. Suppose $a \in R$ is a zero of f. Then there exists $g \in R[X]$ of degree $n - 1$ such that $f(X) = (X - a) g(X)$.*

POLa
input, 54

Proof: We observe that

$$\begin{aligned} f(X) = f(X) - f(a) = \sum_{i=0}^{n} a_i (X^i - a^i) &= \sum_{i=0}^{n} a_i (X - a)(X^{i-1} + X^{i-2} a + \cdots + a^{i-1}) \\ &= (X - a) \sum_{i=0}^{n} a_i (X^{i-1} + X^{i-2} a + \cdots + a^{i-1}) \\ &= (X - a) g(X) \end{aligned}$$

and note that the highest power of x in g is X^{n-1} and that its coefficient is a_n. Thus, $\deg(g) = n - 1$. ▮

LEMMA 10.2: *A nonzero polynomial f of degree n with coefficients in an integral domain R has at most n zeros in R.* POLb input, 81

Proof: If $n = 0$, then f is a nonzero constant, so it has no zeros. Suppose $n \geq 1$ and a is a zero of f and let g be as in Lemma 10.1. An induction assumption on n implies that g has at most $n - 1$ zeros in R. If $b \in R$ is a zero of f, then $(b - a)g(b) = f(b) = 0$, so either $b = a$ or b is a zero of g. Therefore, f has at most n zeros in R. ▮

Note that $(X - 1)^n$ is a polynomial of degree n in $\mathbb{Z}[X]$ but has only one zero in $\mathbb{Z}$. Similarly, $X^2 - 2$ is a polynomial of degree 2 in $\mathbb{Z}[X]$ but has no zero in $\mathbb{Z}$ neither in $\mathbb{Q}$.

Like natural numbers, polynomials have division with a remainder.

LEMMA 10.3: *Let K be a field and $g \in K[X]$ a nonzero polynomial. Then, for each $f \in K[X]$ there exist unique $q, r \in K[X]$ such that $f = qg + r$ and either $r = 0$ or $\deg(r) < \deg(g)$.* POLc input, 108

Proof: Dividing g by its leading coefficient, we may assume that g is monic. If $\deg(g) = 0$, then $g = 1$, so $f = f \cdot 1 + 0$. Otherwise $d = \deg(g) \geq 1$ and $g(X) = X^d + g'(X)$, where $\deg(g') \leq d - 1$. Let $f(X) = a_n X^n + a_{n-1} X^{n-1} + \cdots + a_0$ be a polynomial of degree n with $a_0, \ldots, a_{n-1}, a_n \in K$. If $n < d$, then $f(X) = 0 \cdot g(X) + f(X)$ is the desired representation. Otherwise, the degree $f_1(X) = f(X) - a_n X^{n-d} g(X)$ is at most $n - 1$. Applying an induction hypothesis on n, we find $q_1, r \in K[X]$ such that $f_1(X) = q_1(X)g(X) + r(X)$ and either $r = 0$ or $\deg(r) < \deg(g)$. It follows that $f(X) = a_n X^{n-d} g(X) + q_1(X)g(X) + r(X) = (a_n X^{n-d} + q_1(X))g(X) + r(X)$ is the desired representation.

In order to prove the uniqueness of the representation suppose that $q', r' \in K[X]$ are polynomials such that $f = q'g + r'$ and either $r' = 0$ or $\deg(r') < \deg(g)$. Substracting the latter representation from the former one, we get $0 = (q - q')g + (r - r')$, so

$r' - r = (q - q')g$. If $r' - r \neq 0$, then $q - q' \neq 0$, so by (1) and (2),

$$\deg(g) > \deg(r' - r) = \deg(q - q') + \deg(g) \geq \deg(g) > \deg(r).$$

We conclude from this contradiction that $r = r'$ and $q = q'$, as claimed. ∎

Exercise 10.4: Let R be a commutative ring with 1 and g a monic polynomial. Prove that for every $f \in R[X]$ there exist $q, r \in R[X]$ such that $f = qg + r$ and either $r = 0$ or $\deg(r) < \deg(g)$. Hint: Make the necessary changes in the proof of Lemma 10.3. ∎ POLd input, 152

11. Wilson's Theorem

WILSON input, 7

We prove in this section that $(p-1)! \equiv -1 \mod p$ if p is an odd prime number and conclude that there are infinitely many prime numbers $p \equiv 1 \mod 4$.

LEMMA 11.1: *Let F be a finite field. Then the product of all elements of $F^\times$ is -1.* WILa input, 12

Proof: To each $a \in F^\times$ we associate its inverse a^{-1}. If $a = a^{-1}$, then $a^2 = 1$, so $(a-1)(a+1) = 0$, hence $a = 1$ or $a = -1$. We choose for each $a \in F^\times \smallsetminus \{1, -1\}$ an element $b \in \{a, a^{-1}\}$ and denote the set of all those b's by B. Then,

$$\prod_{a \in F^\times} a = 1 \cdot (-1) \cdot \prod_{\substack{a \in F^\times \\ a \neq \pm 1}} a = -\prod_{b \in B} bb^{-1} = -1,$$

as claimed. ∎

For every prime number p, the numbers $1, 2, \ldots, p-1$ represent the elements of $\mathbb{F}_p^\times$. Thus, the following result follows from Lemma 11.1 for $F = \mathbb{F}_p$.

THEOREM 11.2 (Wilson): *Every prime number p satisfies $(p-1)! \equiv -1 \mod p$.* WILb input, 40

As an addendum to Wilson's theorem we have:

PROPOSITION 11.3: *If $n \geq 6$ is a composite number, then $(n-1)! \equiv 0 \mod n$.* WILc input, 47

Proof: Suppose $n = kl$ with $1 < k \leq l \leq n$. If $k < l$, then both k and l appear among the factors of $(n-1)!$, so $n|(n-1)!$. If $k = l$, then $k^2 = n \geq 6$. Hence, $k \geq \sqrt{6} > \sqrt{4} = 2$, so $1 < k < 2k < k^2$. It follows that both k and $2k$ appear as factors of $(n-1)!$. Therefore, $n|(n-1)!$. ∎

For example

$$(6-1)! = 5! = 1 \cdot 2 \cdot 3 \cdot 4 \cdot 5 = 6 \cdot 4 \cdot 5 \equiv 0 \mod 6$$

$$(8-1)! = 7! = 1 \cdot 2 \cdot 3 \cdot 4 \cdot 5 \cdot 6 \cdot 7 = 8 \cdot 3 \cdot 5 \cdot 6 \cdot 7 \equiv 0 \mod 8$$

$$(9-1)! = 8! = 1 \cdot 2 \cdot 3 \cdot 4 \cdot 5 \cdot 6 \cdot 7 \cdot 8 = 18 \cdot 2 \cdot 4 \cdot 5 \cdot 7 \cdot 8 \equiv 0 \mod 9$$

The only exceptional case is $n = 4$. In this case $(4-1)! = 2 \cdot 3 = 6 \equiv 2 \not\equiv -1 \mod 4$. Combining the latter observation to Theorem 11.1 and Proposition 11.2, we get a primality criterion for a natural number, which is however not a quick one.

PROPOSITION 11.4: *An integer $p \geq 2$ is prime if and only if $(p-1)! \equiv -1 \mod p$.* WILd input, 85

More important is the following consequence of Wilson's theorem:

LEMMA 11.5: *Let p be an odd prime. Then $p \equiv 1 \mod 4$ if and only if there exists* WILe input, 93 *$x \in \mathbb{Z}$ such that $x^2 \equiv -1 \mod p$.*

Proof: First suppose that $p \equiv 1 \mod 4$. Then $2|\frac{p-1}{2}$. Hence, by Wilson's theorem,

$$-1 \equiv (p-1)! = \prod_{a=1}^{p-1} a = \prod_{a=1}^{\frac{p-1}{2}} a(p-a) = (-1)^{\frac{p-1}{2}} \Big(\prod_{a=1}^{\frac{p-1}{2}} a^2\Big) \equiv \Big(\prod_{a=1}^{\frac{p-1}{2}} a\Big)^2 \mod p.$$

Conversely, suppose there exists $x \in \mathbb{Z}$ such that $x^2 \equiv -1 \mod p$. Then, by Fermat's little theorem,

$$(1) \qquad (-1)^{\frac{p-1}{2}} \equiv (x^2)^{\frac{p-1}{2}} \equiv x^{p-1} \equiv 1 \mod p.$$

It follows that $p \equiv 1 \mod 4$. Otherwise, $\frac{p-1}{2}$ is odd, so (1) implies that $-1 \equiv (-1)^{\frac{p-1}{2}} \equiv 1 \mod p$, hence $p|2$, which contradicts our assumption. ∎

We are now in a position to supplement Theorem 6.2.

THEOREM 11.6: *There are infinitely many prime numbers $p \equiv 1 \mod 4$.* WILf input, 130

Proof: As in Euclid's proof, we assume that there are only finitely many prime numbers that are congruent 1 modulo 4 and list them as $p_1, p_2, \ldots, p_n$. Let $x = p_1 p_2 \ldots p_n$ and let $m = 1 + 4x^2$. Then $m \geq 3$ (because $5 \equiv 1 \mod 4$), so m has a prime divisor p. It satisfies $p|(1+4x^2)$, so $(2x)^2 \equiv -1 \mod p$. By Lemma 11.5, $p \equiv 1 \mod 4$. Hence, $p = p_i$ for some $1 \leq i \leq n$. However, this implies the contradiction $p|1$. It follows that there infinitely many prime numbers $p \equiv 1 \mod 4$. ∎

12. Density of Prime Numbers

DENSITY
input, 7

No rule is known for the distribution of the prime numbers among the natural numbers. One can only give approximation formulas for the number of prime numbers less than a given real number x as x tends to infinity. In this short section we survey the main results in the area. The proofs apply complex analytic mehtods and are beyond the scope of this course.

We write $f(x) = g(x) + o(h(x))$ for real functions f, g, h if

$$\lim_{x\to\infty} \frac{f(x) - g(x)}{h(x)} = 0.$$

For example, $\log x = o(x)$, because by L'hopital's rule,

$$\lim_{x\to\infty} \frac{\log x}{x} = 0. \tag{1}$$

We write $f(x) \sim g(x)$ if

$$\lim_{x\to\infty} \frac{f(x)}{g(x)} = 1.$$

We also denote the number of prime numbers $p \le x$ by $\pi(x)$. The prime number theorem says that

$$\pi(x) = \frac{x}{\log x} + o\left(\frac{x}{\log x}\right). \tag{2}$$

This implies the weaker form of the theorem

$$\pi(x) \sim \frac{x}{\log x}. \tag{3}$$

Next we consider relatively prime natural numbers a, n and write $\pi(x, a, n)$ for the number of prime numbers $p \le x$ that satisfy $p \equiv a \mod n$. Then Dirichlet's theorem says that

$$\pi(x, a, n) = \frac{1}{\varphi(n)} \frac{x}{\log x} + o\left(\frac{x}{\log x}\right). \tag{4}$$

It follows from (2) and (3) that

$$\pi(x, a, n) = \frac{1}{\varphi(n)}\pi(x) + o(\pi(x)).$$

Thus, $\lim_{x\to\infty}\frac{\pi(x,a,n)}{\pi(x)} = \frac{1}{\varphi(n)}$. This may be interpreted as "the natural density of the prime numbers $p \equiv a \mod n$ is $\frac{1}{\varphi(n)}$". Note that the natural density $\frac{1}{\varphi(n)}$ is independent of a. It is equal to the discrete probability of each element of $(\mathbb{Z}/n\mathbb{Z})^\times$ in the whole set.

By (1), $\lim_{x\to\infty}\frac{x}{\log x} = \infty$. Hence, by (4), $\lim_{x\to\infty}\pi(x,a,n) = \infty$. This implies that there are infinitely many prime numbers $p \equiv a \mod n$ if $\gcd(a,n) = 1$. The latter result is referred to as the "qualitative Dirichlet's theorem". So far we have proved it for $n = 4$ and $a = \pm 1$. It is possible to prove it by elementary means for $a = 1$ and arbitrary n. But the proof of the general case goes through the proof of the approximation formula (4).

13. Chinese Remainder Theorem

CHIN
input, 8

The chinese remainder theorem allows us to solve systems of congruencial equations with relatively prime in pairs moduli. Although it is possible to prove the theorem directly, we prefer to use this opportunity in order to introduce 'direct product of rings' and to apply this concept to an easy proof of the theorem.

A map φ: $R \to S$ of two commutative rings with 1 is an **isomorphism** if φ is bijective and φ preserves addition and multiplication. That is, $\varphi(x+y) = \varphi(x) + \varphi(y)$ and $\varphi(xy) = \varphi(x)\varphi(y)$. It follows that $\varphi(0) = 0$ and $\varphi(1) = 1$. In order to prove the latter rule we choose $e \in R$ such that $\varphi(e) = 1$. Then, $1 = \varphi(e) = \varphi(1 \cdot e) = \varphi(1)\varphi(e) = \varphi(1) \cdot 1 = \varphi(1)$. It follows that the restriction of φ to $R^\times$ is an isomorphism of groups onto $S^\times$. We write $R \cong S$ and $R^\times \cong S^\times$ to denote the respective rings and groups isomorphisms.

The **direct product** of R and S is the set $R \times S$ of all pairs (x, y) with $x \in R$ and $y \in S$ in which addition and multiplication is defined componentwise. Thus

$$(x, y) + (x', y') = (x + x', y + y') \text{ and } (x, y)(x', y') = (xx', yy').$$

The zero element of $R \times S$ is $(0, 0)$ and the one element is $(1, 1)$. Note that $(1, 0)(0, 1) = (0, 0)$, so if $1 \neq 0$, $R \times S$ is not an integral domain.

Note that if both R and S are finite, then $|R \times S| = |R| \cdot |S|$.

Lemma 13.1: *Suppose that m and n are relatively prime natural numbers. Then the map φ: $\mathbb{Z}/mn\mathbb{Z} \to \mathbb{Z}/m\mathbb{Z} \times \mathbb{Z}/n\mathbb{Z}$ defined by $\varphi(a + mn\mathbb{Z}) = (a + m\mathbb{Z}, a + n\mathbb{Z})$ is an isomorphism of rings.*

CHIa
input, 51

Proof: If $a \equiv a' \mod mn$, then $a \equiv a' \mod m$ and $a \equiv a' \mod n$, so φ is well defined. It follows that φ preserves addition and multiplication. Moreover, φ is injective. Indeed, if $a + m\mathbb{Z} = 0$ and $a + n\mathbb{Z} = 0$, then $m|a$ and $n|a$. Since $\gcd(m, n) = 1$, Lemma 4.7(c) implies that $mn|a$, so $a + mn\mathbb{Z} = 0$. Next observe that $|\mathbb{Z}/mn\mathbb{Z}| = mn = |\mathbb{Z}/m\mathbb{Z}| \cdot |\mathbb{Z}/n\mathbb{Z}| = |\mathbb{Z}/m\mathbb{Z} \times \mathbb{Z}/n\mathbb{Z}|$. Since both sets are finite and φ is injective, it follows that φ is surjective. Consequently, φ is an isomorphism. ∎

THEOREM 13.2: *Let $m_1, m_2, \ldots, m_r$ be natural numbers satisfying* $\gcd(m_i, m_j) = 1$ *for all $i \neq j$. Then, for each r-tuple $(a_1, a_2, \ldots, a_r)$ of integers there exists an $a \in \mathbb{Z}$, unique modulo $m_1 m_2 \cdots m_r$, such that $a \equiv a_i \mod m_i$, $i = 1, \ldots, r$.* CHIc input, 81

Proof: We proceed by induction on r. The case $r = 1$ is trivial and the case $r = 2$ is the surjectivity of φ of Lemma 13.1 (for $m = m_1$ and $n = m_2$). So assume that $r \geq 3$ and that the Theorem holds for $r - 1$. Let $m' = m_1 \cdots m_{r-1}$. The induction hypothesis gives $a' \in \mathbb{Z}$ such that $a' \equiv a_i \mod m_i$ for $i = 1, \ldots, r - 1$. By Lemma 7.3, $\gcd(m', m_r) = 1$. Hence, by the case $r = 2$, there exists $a \in \mathbb{Z}$ satisfying $a \equiv a' \mod m'$ and $a \equiv a_r \mod m_r$. It follows that $a \equiv a_i \mod m_i$ for $i = 1, \ldots, r$.

Finally, if b is an additional integer with $b \equiv a_i \mod m_i$, then $b \equiv a \mod m_i$, $i = 1, \ldots, r$. Since the m_i's are relatively prime in pairs, $b \equiv a \mod m_1 m_2 \cdots m_r$. ∎

Exercise 13.3: Solve the following system of equations: CHId input, 110

$$x \equiv 1 \mod 2$$

$$x \equiv 1 \mod 3$$

$$x \equiv 3 \mod 4$$

$$x \equiv 4 \mod 5$$

∎

14. Euler's Phi Function and Cryptography

CRIPT input, 7

As an application of Lemma 13.1, we may compute the Euler phi function at a natural number n provided we know how to decompose n into a product of prime numbers.

LEMMA 14.1: *If* $\gcd(m,n)=1$, *then* $\varphi(mn)=\varphi(m)\varphi(n)$. CHIb input, 12

Proof: By Lemma 13.1, $\mathbb{Z}/mn\mathbb{Z} \cong \mathbb{Z}/m\mathbb{Z} \times \mathbb{Z}/n\mathbb{Z}$. Hence,

$$(\mathbb{Z}/mn\mathbb{Z})^\times \cong (\mathbb{Z}/m\mathbb{Z})^\times \times (\mathbb{Z}/n\mathbb{Z})^\times.$$

It follows that $\varphi(mn) = |(\mathbb{Z}/mn\mathbb{Z})^\times| = |(\mathbb{Z}/m\mathbb{Z})^\times| \cdot |(\mathbb{Z}/n\mathbb{Z})^\times| = \varphi(m)\varphi(n)$, as claimed ■

PROPOSITION 14.2: *For each positive integer n we have $\varphi(n) = \prod_{p|n} p^{v_p(n)-1}(p-1)$. In particular, $\varphi(p^k) = p^{k-1}(p-1)$ for each prime p and every natural number k.* CRIa input, 34

Proof: First observe that an integer a is not relatively prime to p^k if and only if $p|a$. Hence, the integers between 1 and p^k that are not relatively prime to p^k are $1{\cdot}p, 2{\cdot}p, \ldots, p^{k-1}p$. Their number is p^{k-1}. Since $\varphi(p^k)$ is the number of integers between 1 and p^k that are relatively prime to p^k, we have $\varphi(p^k) = p^k - p^{k-1} = p^{k-1}(p-1)$.

Next conclude by induction on r from Lemma 14.1 that if $m_1, m_2, \ldots, m_r$ are relatively prime in pairs, than $\varphi(m_1 m_2 \cdots m_r) = \varphi(m_1)\varphi(m_2)\cdots\varphi(m_r)$.

It follows from the first two paragraphs of the proof that

$$\varphi(n) = \varphi(\prod_{p|n} p^{v_p(n)}) = \prod_{p|n} \varphi(p^{v_p(n)}) = \prod_{p|n} p^{v_p(n)-1}(p-1),$$

as claimed. ■

Examples 14.3: $\varphi(1)=1$, $\varphi(2)=1$, $\varphi(3)=2$, $\varphi(4)=2$, $\varphi(5)=4$, $\varphi(6)=2$, $\varphi(7)=6$, $\varphi(8)=4$, $\varphi(9)=4$, $\varphi(10)=4$, $\varphi(11)=10$, $\varphi(12)=4$. As representatives modulo 12 for $(\mathbb{Z}/12\mathbb{Z})^\times$, one may take $1,5,7,11$. Next, $\varphi(60) = \varphi(2^2)\varphi(3)\varphi(5) = 2\cdot 2\cdot 4 = 16$. Representatives modulo 60 for $(\mathbb{Z}/60\mathbb{Z})^\times$ are $1,7,11,13,17,19,23,29,31,37,41,43,47,49,53,59$.■ CRIb input, 66

■

Euler's theorem and lemma 14.1 have been applied to public key cryptography. One uses this tool in order to exchange encrypted information through public channels such that although all of the data is public, only the two parties involved can decode the information.

Procedure 14.4: Public key cryptography. Reuven and Shimeon want to exchange encrypted information. CIRc input, 85

Step A: *Preperations.* Reuven chooses large distinct prime numbers p and q and a large natural mumbers l which is relatively prime to both $p-1$ and $q-1$ (e.g. another prime number greater than p and q). He computes the product $n = pq$ and sends both n and l to Shimeon.

Step B: *Transmission of data.* Having n and l Shimeon wants to send Reuven data encrypted in a large natural number $a < n$ which is relatively prime to n. Instead of sending a to Reuven, Shimeon computes a number $b < n$ satisfying

$$b \equiv a^l \mod n \tag{1}$$

(e.g. using the Euclid's algorithm) and sends b to Reuven.

Step C: *Decoding the data.* Using the decomposition $n = pq$ and applying Lemma 14.1, Reuven computes

$$\varphi(n) = \varphi(pq) = (p-1)(q-1) \tag{2}$$

Then he uses Euclid's algorithm's again to solve the equation $ll' \equiv 1 \mod \varphi(n)$. Thus, he computes k such that $ll' = 1 + k\varphi(n)$. Then Reuven use Procedrue 3.8 to 'quickly' compute $b^{l'}$ modulo n.

Claim: $b^{l'} \equiv a \mod n$. Indeed, by Euler's theorem,

$$b^{l'} \equiv a^{ll'} \equiv a^{1+k\varphi(n)} \equiv a(a^{\varphi(n)})^k \equiv a \mod n.$$

This gives Reuven the desired information sent by Shimeon.

Note that nobody else can compute $\varphi(n)$ 'quickly', because nobody else knows the factorization $n = pq$ and factorization of natural numbers into product of prime numbers takes a 'long' time (relativ to the computations that Reuven and Simeon do). Thus, nobody can 'quickly' compute a. ∎

Exercise 14.5: Find the number of natural numbers ≤ 7200 that are relatively prime to 3600. ■

FERm input, 144

15. Primitive Roots

PRIMITIVE
input, 9

We prove that there exist $\varphi(p-1)$ 'primitive roots' of p for each prime number p.

By Lemma 9.5(b), each element a of a commutative group of order n satisfies $a^n = 1$. The smallest natural number d for which $a^d = 1$ is called the **order** of a and is denoted by $\mathrm{ord}(a)$. Note that if $d = \mathrm{ord}(a)$, then $1, a, \ldots, a^{d-1}$ are distinct. Thus, if $\mathrm{ord}(a) = |G|$, then $G = \{1, a, \ldots, a^{n-1}\}$. In this case we say that a **generates** G and that G is a **cyclic group**.

LEMMA 15.1: *Let a be an element of a finite commutative group G of order n and let $d = \mathrm{ord}(a)$. Then*

PRMa
input, 23

(a) *If $a^l = 1$, then $d|l$. In particular, $d|n$.*

(b) *$a^l = a^{l'}$ if and only if $l \equiv l' \mod d$.*

(c) *For each natural number k we have* $\mathrm{ord}(a^k) = \frac{d}{\gcd(d,k)}$.

(d) $\mathrm{ord}(a^k) = d$ *if and only if* $\gcd(d,k) = 1$. *In particular, the group* $\{1, a, \ldots, a^{d-1}\}$ *has exactly $\varphi(d)$ generators.*

Proof of (a): Write $l = qd + r$ with $0 \le r \le d-1$. Then $1 = a^l = (a^d)^q a^r = a^r$. It follows from the minimality of d that $r = 0$. Hence, $d|l$.

Since, as mentioned above, $a^n = 1$, we have that $d|n$.

Proof of (b): Multiply the equality $a^l = a^{l'}$ by $a^{-l'}$ to get $a^{l-l'} = a^{l'-l'} = a^0 = 1$. Now apply (a).

Proof of (c): Let $m = \mathrm{ord}(a^k)$ and $c = \gcd(d,k)$. Then $(a^k)^{\frac{d}{c}} = (a^d)^{\frac{k}{c}} = 1$, hence $m \le \frac{d}{c}$.

Next observe that $a^{km} = (a^k)^m = 1$. Hence, by (a), $d|km$, so $\frac{d}{c}|\frac{k}{c}m$. By Proposition 3.3, there exist $x, y \in \mathbb{Z}$ with $c = dx + ky$. Hence, $1 = \frac{d}{c}x + \frac{k}{c}y$, so $\gcd\left(\frac{d}{c}, \frac{k}{c}\right) = 1$. It follows from Lemma 4.7(a) that $\frac{d}{c}|m$. Combining this relation with the one obtained in the preceding paragraph, we obtain $\mathrm{ord}(a^k) = \frac{d}{c}$, as claimed.

Proof of (d): Immediate consequence of (c). ∎

Let p be a prime number and a an integer not divisible by p. Thus, $a + p\mathbb{Z}$ is an element of the finite commutative group $(\mathbb{Z}/p\mathbb{Z})^\times$ of order $p-1$. We define $\mathrm{ord}_p a$ to be

$\mathrm{ord}(a + p\mathbb{Z})$. Thus, $\mathrm{ord}_p a$ is the smallest natural number that satisfies $a^d \equiv 1 \mod p$. By definition $\mathrm{ord}_p a = \mathrm{ord}_p a'$ if $a \equiv a' \mod p$.

Applying Lemma 15.1 to the group $(\mathbb{Z}/p\mathbb{Z})^\times$, we get:

LEMMA 15.2: *Let p be a prime number and a an integer not divisible by p. Set $d = \mathrm{ord}_p a$. Then:* PRMb input, 89

(a) $d|(p-1)$.

(b) $a^l \equiv a^{l'} \mod p$ *if and only if* $l \equiv l' \mod d$.

(c) *For each natural number k we have* $\mathrm{ord}_p(a^k) = \frac{d}{\gcd(d,k)}$.

(d) $\mathrm{ord}_p(a^k) = d$ *if and only if* $\gcd(d,k) = 1$. *In particular, the group $\{1, a, \dots, a^{d-1}\}$ has exactly $\varphi(d)$ generators.*

LEMMA 15.3: *Let p be a prime number and d a natural number. If there exists $a \in \mathbb{Z}$ such that $p \nmid a$ and $\mathrm{ord}_p(a) = d$, then there exists exactly $\varphi(d)$ elements x modulo p such that $\mathrm{ord}_p(x) = d$.* PRMc input, 107

Proof: Recall that $\mathbb{F}_p = \mathbb{Z}/p\mathbb{Z}$ is not only a ring but even a field (Example 8.5(b)). For each $x \in \mathbb{Z}$ consider the element $\bar{x} = x + p\mathbb{Z}$ of $\mathbb{F}_p$ and note that the map $x \mapsto \bar{x}$ preserves addtion and multiplication and $\bar{1}$ is the one element of $\mathbb{F}_p$, that we also denote by 1. We have $x \not\equiv y \mod p$ if and only if $\bar{x} \neq \bar{y}$. In particular, $1, \bar{a}, \bar{a}^2, \dots, \bar{a}^{d-1}$ are distinct roots in $\mathbb{F}_p$ of the equation $X^d - 1 = 0$. By Lemma 10.2, they are all of the roots of that equation in $\mathbb{F}_p$.

If a natural number x satisfies $\mathrm{ord}_p(x) = d$, then $\mathrm{ord}(\bar{x}) = d$, so $\bar{x}^d = 1$. It follows from the preceding paragraph that $\bar{x} = \bar{a}^k$ for some k between 0 and $d-1$. By Lemma 15.2(c), $\gcd(k,d) = 1$. Conversely, each of the elements $\bar{a}^k$ with $\gcd(d,k) = 1$ satisfies $\mathrm{ord}(\bar{a}^k) = 1$. Therefore, there are $\varphi(d)$ elements in $\mathbb{F}_p$ whose order is d. Consequently, there are $\varphi(d)$ integers modulo p whose order modulo p is d. ∎

LEMMA 15.4: *For every natural number n we have $\sum_{d|n} \varphi(d) = n$.* PRMd input, 142

Proof: We observe that

$$\begin{aligned}\{1,2,\dots,n\} &= \dot{\bigcup_{d|n}}\{1 \le a \le n \mid \gcd(a,n)=d\}\\ &= \dot{\bigcup_{d|n}}\{dk \mid 1 \le k \le \frac{n}{d} \text{ and } \gcd\bigl(k,\frac{n}{d}\bigr)=1\},\end{aligned}$$

where the dot in the union symbol means disjoint union. Hence, $n = \sum_{d|n}\varphi\bigl(\frac{n}{d}\bigr) = \sum_{d|n}\varphi(d)$. ∎

THEOREM 15.5: *For every prime number p and every divisor d of $p-1$ there exist exactly $\varphi(d)$ integers modulo p whose order is d.* PRMe input, 163

Proof: Given a divisor d of $p-1$, we denote the set of all integers a between 1 and $p-1$ with $\mathrm{ord}_p(a) = d$ by A_d. By Lemma 15.3, either A_d is empty or A_d has exactly $\varphi(d)$ elements. In each case $|A(d)| \le \varphi(d)$. By Lemma 15.2(a), $\{1,\dots,p-1\} = \dot{\bigcup}_{d|p-1} A_d$. Hence, by Lemma 15.4,

$$p-1 = \sum_{d|n}|A_d| \le \sum_{d|p-1}\varphi(d) = p-1.$$

It follows that $|A_d| = \varphi(d)$ for each divisor d of $p-1$. ∎

A natural number g whose order modulo p is $p-1$ is a **primitive root** of p. In this case, for every $1 \le a \le p-1$ there exists a unique k modulo $p-1$ such that $g^k \equiv a \mod p$.

Applying Theorem 15.5 to the divisor $p-1$ of $p-1$ we get:

THEOREM 15.6: *Let p be a prime number.* PRMf input, 194

(a) *There are exactly $\varphi(p-1)$ primitive roots of p.*

(b) *The group $\mathbb{F}_p^\times$ is cyclic of order $p-1$. It has $\varphi(p-1)$ generators.*

Example 15.7: The number 2 is a primitive root of $3, 5, 11, 13$ but not of $7, 17$. PRMg input, 204

Open Problem 15.8: It is unknown whether 2 is a primitive root of infinitely many prime numbers. Emil Artin conjectured it is. Moreover, he conjectured that any non-square natural number is a primitive root of infinitely many prime numbers. Roger PRMh input, 208

Heath-Brown proved that one of the numbers $2, 3, 5$ is a primitive root of infinitely many prime numbers. Indeed, he proved the same statement for every triple p, q, r of distinct prime numbers. However, the original conjecture of Artin is still open. ∎

Exercise 15.9: Let g, h be elements of a finite commutative group such that $m = \operatorname{ord}(g)$ PRMi input, 221
and $n = \operatorname{ord}(h)$ are relatively prime. Prove that $\operatorname{ord}(gh) = mn$. ∎

Exercise 15.10: Let a be an element of a finite commutative group. Prove that PRMj input, 227
$\operatorname{ord}(a^{-1}) = \operatorname{ord}(a)$. ∎

Exercise 15.11: Let $a \in \mathbb{Z}$ with $23 \nmid a$. Prove that a is a primitive root modulo 23 if PRMk input, 232
and only if $a^{11} \equiv -1 \mod 23$. ∎

Exercise 15.12: Prove that if p is a prime number, $p \equiv 1 \mod 4$, and g is a primitive PRMl input, 238
root of p, then so is $-g$. ∎

Exercise 15.13: Prove that if p is an odd prime number, $a \in \mathbb{Z}$ is not divisible by p, PRMm input, 243
and $n = \operatorname{ord}_p(a) > 1$, then $\sum_{k=0}^{n-1} a^k \equiv 0 \mod p$. Hint: multiply the left hand side by a. ∎

16. Quadratic Equations Modulo p

EQUATIONS
input, 6

We analyze in this section what does it take to solve linear and quadratic equations modulo a prime number p.

We start from a prime number p and an equation

$$ax + b \equiv 0 \mod p \tag{1}$$

with integral coefficients a, b and a variable x. We ask about the conditions for the existence of a solution $x \in \mathbb{Z}$ to (1). In case these conditions are satisfied, we ask further for the set of solutions. Finally we seek an algorithm to find x.

There are two cases.

Case A1: $p|a$. Then (1) becomes $b \equiv 0 \mod p$. If $p \nmid b$, then (1) has no solution. If $p|b$, then every $x \in \mathbb{Z}$ is a solution of (1).

Case A2: $p \nmid a$. Then there exists $a' \in \mathbb{Z}$ such that $a'a \equiv 1 \mod p$ (Example 8.5(b)). Moreover, by Example 8.5(c), we may use Euclid algorithm to effectively compute a'. It follows that the solutions of (1) are $x \equiv -a'b \mod p$. In particular, (1) has a unique solution modulo p.

Next we consider a quadratic equation

$$ax^2 + bx + c \equiv 0 \mod p. \tag{2}$$

We distinguish between three cases:

Case B1: $p|a$. Then (2) reduces to (1) and we may use the analysis above.

Case B2: $p = 2$ *and* $2 \nmid a$. Then we may multiply (2) by a' (of Case A2) to bring (2) to the form $x^2 + bx + c \equiv 0 \mod 2$. Recall that $\mathbb{F}_2 = \{0, 1\}$. Hence, there are four subcases for this equation: $x^2+x+1 \equiv 0 \mod 2$, which is unsolvable; $x^2+x \equiv 0 \mod 2$, with $0, 1$ as solutions; $x^2 \equiv 0 \mod 2$ with 0 as its unique solution; and $x^2+1 \equiv 0 \mod 2$ with 1 as its unique solution modulo 2.

Case B3: *$p \neq 2$ and $p \nmid a$.* In this case we solve (2) by "completing the square". We multiply (2) by $4a$ and add $b^2 - b^2$ to the left hand side to get

$$(2ax)^2 + 2 \cdot abx + b^2 + 4ac - b^2 \equiv 0 \mod p.$$

Thus, (2) is equivalent to

$$(2ax + b)^2 \equiv b^2 - 4ac \mod p. \tag{3}$$

We introduce a new variable $y = 2ax + b$ and set $d = b^2 - 4ac$. Then (3) takes the form

$$y^2 \equiv d \mod p. \tag{4}$$

Given a solution y of (4), we may then solve the linear equation $2ax + b - y \equiv 0$ as above. To solve (4), we may consequtively substitute $y = 0, 1, 2, \dots, \frac{p-1}{2}$ and check whether (4) holds. This is however a non-effective procedure, because it takes an order of magnitute of p steps. In the next sections we introduce the "Legendre symbol", prove the "quadratic reciprocity law" and use it to establish an effective procedure to solve (4).

17. Legendre Symbol

LEGENDRE
input, 9

Given an odd prime number p and a relatively prime integer a, we say that a is a **quadratic residue** modulo p if there exists $x \in \mathbb{Z}$ such that $x^2 \equiv a \mod p$, otherwise we say that a is a **quadratic non-residue** modulo p. Using this terminology we define the **Legendre symbol** $\left(\frac{a}{p}\right)$ by the following rule:

$$\left(\frac{a}{p}\right) = \begin{cases} 1 & \text{if } a \text{ is a quadratic residue modulo } p \\ -1 & \text{if } a \text{ is a quadratic non-residue modulo } p. \end{cases}$$

For example $1, 2, 4$ are quadratic residues modulo 7 while $3, 5, 6$ are quadratic non-residues modulo 7. By definition,

(1a) $\left(\frac{a^2}{p}\right) = 1$ and

(1b) $a \equiv b \mod p$ implies $\left(\frac{a}{p}\right) = \left(\frac{b}{p}\right)$ for all a, b not divisible by p.

LEMMA 17.1 (Euler's criterion): *If* $p \nmid a$, *then* $\left(\frac{a}{p}\right) \equiv a^{\frac{p-1}{2}} \mod p$.

LEGa
input, 35

Proof: First suppose that $\left(\frac{a}{p}\right) = 1$. Then there exists $x \in \mathbb{N}$ such that $x^2 \equiv a \mod p$. Hence, by Fermat's little theorem, $a^{\frac{p-1}{2}} \equiv x^{p-1} \equiv 1 \equiv \left(\frac{a}{p}\right) \mod p$.

Conversely suppose $\left(\frac{a}{p}\right) = -1$. By Fermat's little theorem $(a^{\frac{p-1}{2}})^2 \equiv a^{p-1} \equiv 1 \mod p$. Hence, $a^{\frac{p-1}{2}} \equiv \pm 1 \mod p$. Assume that $a^{\frac{p-1}{2}} \equiv 1 \mod p$. By Theorem 15.6, there exists a primitive root g modulo p. Hence, there exists $k \in \mathbb{N}$ such that $g^k \equiv a \mod p$. By (1a), $2 \nmid k$. On the other hand, $1 \equiv a^{\frac{p-1}{2}} \equiv g^{k\frac{p-1}{2}} \mod p$. Hence, by Lemma 15.2(b), $p-1 | k\frac{p-1}{2}$, so $2|k$. We conclude from this contradiction that $a^{\frac{p-1}{2}} \equiv -1 \equiv \left(\frac{a}{p}\right)$. ∎

Procedure 17.2: Euler's criterion, allows us to effectively compute the Legendre symbol. To this end we write

LEGb
input, 64

$$\frac{p-1}{2} = \sum_{i=0}^{m} a_i 2^i, \tag{3}$$

with $a_i \in \{0, 1\}$ for $i = 0, \dots, m$. Then we successively compute the powers 2^i modulo p and apply (2) to compute $a^{\frac{p-1}{2}}$ modulo p. Finally we apply Lemma 17.1. ∎

Lemma 17.3: *If $p \nmid a, b$, then $\left(\frac{ab}{p}\right) = \left(\frac{a}{p}\right)\left(\frac{b}{p}\right)$.* LEGc input, 78

Proof: By Euler's criterion,

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

Since both sides of the congruence are ±1 and $p > 2$, we have $\left(\frac{ab}{p}\right) = \left(\frac{a}{p}\right)\left(\frac{b}{p}\right)$, as claimed. ∎

The latter argument combined again with Euler's criterion also yields statement (a) of the following lemma: Statement (b) of that lemma follows then from statement (a).

Lemma 17.4: LEGd input, 101
(a) $\left(\frac{-1}{p}\right) = (-1)^{\frac{p-1}{2}}$.
(b) -1 *is a quadratic residue modulo p if and only if $p \equiv 1 \mod 4$.*

Note that Statement (b) is a repetition of Proposition 11.4.

We say that a quadratic residue a modulo p is **reduced** if $1 \le a \le p-1$. We say that a quadratic nonresidue a modulo p is **reduced** if $1 \le a \le p-1$.

Lemma 17.5: *There are exactly $\frac{p-1}{2}$ reduced quadratic residues and $\frac{p-1}{2}$ reduced quadratic non-residues modulo p.* LEGe input, 118

Proof: If a is a reduced quadratic residue, then there exists $x \in \mathbb{Z}$ not divisible by p such that $a \equiv x^2 \mod p$. Replacing x by its remainder modulo p, we may assume that $1 \le x \le p-1$. If $x > \frac{p-1}{2}$, we replace x by $p - x$. Thus, all of the reduced quadratic residues modulo p are $1^2, 2^2, \dots, \left(\frac{p-1}{2}\right)^2$. If $1 \le x < y \le \frac{p-1}{2}$ and $x^2 \equiv y^2 \mod p$, then $(x-y)(x+y) \equiv 0 \mod p$. Hence either $x \equiv y \mod p$ or $x + y \equiv 0 \mod p$. The latter possibility does not occur, because $1 \le x + y \le 2\frac{p-1}{2} = p - 1$. Thus, there are exactly $\frac{p-1}{2}$ reduced quadratic residues modulo p. All the other reduced residues modulo p are quadratic non-residues modulo p. Their number is therefore also $\frac{p-1}{2}$. ∎

A set B of integers is a **reduced system of representatives** modulo n if there exists a bijective map g: $B \to (\mathbb{Z}/n\mathbb{Z})^\times$ such that $g(b) = b + n\mathbb{Z}$ for every $b \in B$. The cardinality of B is therefore $\varphi(n)$. For example, $\{1 \le a \le n \mid \gcd(a,n) = 1\}$ is a reduced system of representatives modulo n.

LEMMA 17.6: *Let A be a reduced set of representatives modulo p. Then $\sum_{a\in A}\left(\frac{a}{p}\right) = 0$.* LEGf input, 152

Proof: By Lemma 17.5, there exists a quadratic non-residue x modulo p. The map $a \mapsto ax$ maps A bijectively onto another reduced system of residues modulo p. Hence, by (1b) and Lemma 17.3,

$$\sum_{a\in A}\left(\frac{a}{p}\right) = \sum_{a\in A}\left(\frac{ax}{p}\right) = \sum_{a\in A}\left(\frac{a}{p}\right)\left(\frac{x}{p}\right) = -\sum_{a\in A}\left(\frac{a}{p}\right).$$

Therefore, $\sum_{a\in A}\left(\frac{a}{p}\right) = 0$. ∎

Exercise 17.7: Prove that the product of the quadratic residues modulo an odd prime number p is congruent to 1 or to -1 modulo p depending on whether $p \equiv -1 \mod 4$ or $p \equiv 1 \mod 4$. Hint: Use Wilson's theorem. ∎ LEGg input, 172

18. More on Polynomials

GAUSS input, 8

The **quadratic reciprocity law** due to Gauss says that $\left(\frac{p}{q}\right) = (-1)^{\frac{p-1}{2}\frac{q-1}{2}}\left(\frac{q}{p}\right)$ if p and q are distinct odd prime numbers. Gauss gave about 50 different proofs of this fundamental theorem. We chose the proof that uses ‘cyclotomic fields’, because it gives a glimse into the arithmetic of algebraic number fields. For that we need to know more about polynomials.

We recall that the valuation v_p associated with a prime number p is defined on $\mathbb{Q}$ by the rule $v_p(\frac{a}{b}p^k) = k$, where a, b are nonzero integers, not divisible by p, and $k \in \mathbb{Z}$. It satisfies the rule (2) of Section 5.

We extend v_p to a valuation of $\mathbb{Q}[X]$ having the same name v_p by the following rule:

$$v_p(\sum_{i=0}^{n} a_i X^i) = \min(v_p(a_0), \ldots, v_p(a_n)), \tag{1}$$

where $a_0, \ldots, a_n \in \mathbb{Q}$. The rule (2c) of Section 5 saying $v_p(a + b) \geq \min(v_p(a), v_p(b))$ for $a, b \in \mathbb{Q}$ extends by definition to polynomials:

(2) $v_p(f + g) \geq \min(v_p(f), v_p(g))$ for all $f, g \in \mathbb{Q}[X]$.

The rule (2c) of Section 5 for multiplication is trickier.

Lemma 18.1: *Let i, j, r, s be real numbers such that $i + j = r + s$ and $(i, j) \neq (r, s)$. Then either $i > r$ or $j > s$.* TRC input, 42

Proof: Otherwise, $i \leq r$ or $j \leq s$, so $i + j \leq r + s = i + j$, Hence, $i = r$ and $j = s$, which is a contradiction. ∎

Lemma 18.2: *For each prime number p and for all $f, g \in \mathbb{Q}[X]$ we have $v_p(fg) = v_p(f) + v_p(g)$.* GAUa input, 55

Proof: Let $f(X) = \sum_{i=0}^{m} a_i X^i$ and $g(X) = \sum_{j=0}^{n} b_j X^j$ where $a_i, b_j \in \mathbb{Q}$ and $a_m, b_n \neq 0$. Then $h(X) = f(X)g(X) = \sum_{k=0}^{n} c_k X^k$ with $c_k = \sum_{i+j=k} a_i b_j$ for $k = 0, \ldots, m + n$. Then, $v_p(c_k) \geq \min_{i+j=k}(v_p(a_i) + v_p(b_j)) \geq v_p(f) + v_p(g)$ for each $0 \leq k \leq m + n$. Hence,

$$v_p(h) = \min(v_p(c_0), \ldots, v_p(c_{m+n})) \geq v_p(f) + v_p(g). \tag{3}$$

Now let r be the greatest integer between 1 and m with $v_p(a_r) = v_p(f)$. Then, $v_p(a_i) \geq v_p(a_r)$ for each $0 \leq i \leq m$, and $v_p(a_i) \geq v_p(a_r)$ for each $r < i \leq m$. Similarly, let s be the greatest integer between 0 and n satisfying $v_p(b_s) = v_p(g)$. Then, $v_p(b_j) \geq v_p(b_s)$ for each $0 \leq j \leq n$, and $v_p(b_j) > v_p(b_s)$ for each $j \leq s \leq n$.

If $i + j = r + s$ and $(i, j) \neq (r, s)$, then either $i > r$ or $j > s$ (Lemma 18.1). In both cases $v_p(a_i b_j) = v_p(a_i) + v_p(b_j) > v_p(f) + v_p(g)$. Since

$$c_{r+s} = a_r b_s + \sum_{\substack{i+j=r+s \\ (i,j)\neq(r,s)}} a_i b_j,$$

it follows from (2e) of Section 5 that

$$v_p(c_{r+s}) = v_p(a_r b_s) = v_p(a_r) + b_p(b_s) = v_p(f) + v_p(g).$$

Hence, $v_p(h) \leq v_p(f) + v_p(g)$. Adding this conclusion to (3), we have $v_p(fg) = \min_{0\leq k\leq m+n} v_k(c_k) = v_p(f) + v_p(g)$, as claimed. ∎

Rule (1) and Lemma 18.2 imply that the v_p is a **discrete valuation** of $\mathbb{Q}[X]$ and we can further extend it to a valuation of the field $\mathbb{Q}(X)$ of rational functions over $\mathbb{Q}$ by the rule $v_p\left(\frac{f}{g}\right) = v_p(f) - v_p(g)$.

Our next concept depends on the set of all valuations v_p. We define the **content** of a nonzero polynomial $f(X) = \sum_{i=0}^n a_i X^i$ with coefficients in $\mathbb{Z}$ by

$$\text{(3)} \qquad \operatorname{cont}(f) = \gcd(a_0, a_1, \ldots, a_n).$$

Note that $\operatorname{cont}(bf) = b \cdot \operatorname{cont}(f)$ for each nonzero $b \in \mathbb{Z}$. If $\operatorname{cont}(f) = 1$, we say that f is **primitive**. This is the case if and only if $v_p(f) = \min(v_p(a_0), \ldots, v_p(a_n)) = 0$ for all prime numbers p.

LEMMA 18.3 (Gauss's lemma): *All nonzero polynomials $f, g \in \mathbb{Z}[X]$ satisfy* $\operatorname{cont}(fg) = \operatorname{cont}(f)\operatorname{cont}(g)$. *In particular, if f and g are primitive, then so is fg.* GAUb input, 128

Proof: We first suppose that both f and g are primitive. Then, for each prime number p we have $v_p(f) = 0$ and $v_p(g) = 0$. By 18.2, $v_p(fg) = v_p(f) + v_p(g) = 0$. Hence, fg is also primitive.

In the general case let b (resp. c) be the greatest common divisor of the coefficients of f (resp. g). Then there exist primitive nonzero $f', g' \in \mathbb{Z}[X]$ such that $f(X) = bf'(X)$ and $g(X) = cg'(X)$. It follows from the first paragraph that $\mathrm{cont}(fg) = \mathrm{cont}(bcf'g') = bc \cdot \mathrm{cont}(f'g') = bc = \mathrm{cont}(f)\mathrm{cont}(g)$, as claimed. ∎

Let R be an integral domain. Recall that element $u \in R$ is **invertible** if there exists $u' \in R$ such that $uu' = 1$ (see Lemma 8.3). For example, the only invertible elements in $\mathbb{Z}$ are ± 1. We say that an element $a \in R$ is **reducible** if there exist non-invertible elements $b, c \in R$ such that $a = bc$. Otherwise, a is said to be **irreducible**.

For a field K, every nonzero element of K is invertible. If $u \in K[X]$ is invertible, then there exists $u' \in K[X]$ such that $uu' = 1$. Hence, by rule (2) of Section 10, $\deg(u) + \deg(u') = 0$. Therefore, $\deg(u) = \deg(u') = 0$, so $u, u' \in K^\times$.

It follows that a polynomial $f \in K[X]$ is reducible if and only if $f = gh$ with $g, h \in K[X]$ satisfying $0 < \deg(g), \deg(h) < \deg(f)$.

COROLLARY 18.4: *If a primitive polynomial $f \in \mathbb{Z}[X]$ is irreducible in $\mathbb{Z}[X]$, then f is irreducible in $\mathbb{Q}[X]$.* GAUc input, 173

Proof: It suffices to prove that if f is reducible in $\mathbb{Q}[X]$, then f is reducible in $\mathbb{Z}[X]$.

Indeed, suppose $f = gh$ with $g, h \in \mathbb{Q}[X]$ satisfying $0 < \deg(g), \deg(h) < \deg(f)$. Then there exist nonzero $b, b', c, c' \in \mathbb{Z}$ and primitive polynomials $g', h' \in \mathbb{Z}[X]$ such that $\gcd(b, b') = 1$, $\gcd(c, c') = 1$, $g = \frac{b}{b'}g'$, and $h = \frac{c}{c'}h'$. Thus, $b'c'f = bcg'h'$. Taking into account that f is primitive, we conclude from Gauss' lemma that $b'c' = \mathrm{cont}(b'c'f) = bc \cdot \mathrm{cont}(g'h') = bc$. By Lemma 4.7(a), $b'|c$ and $c'|b$. Hence, $g'' = \frac{b}{c'}g' \in \mathbb{Z}[X]$, $h'' = \frac{c}{b'}h' \in \mathbb{Z}[X]$, $f = g''h''$, $\deg(g'') = \deg(g)$, and $\deg(h'') = \deg(h)$, so $0 < \deg(g''), \deg(h'') < \deg(f)$. We conclude that $f = \frac{b}{c'}g \cdot \frac{c}{b'}h'$ is reducible. ∎

LEMMA 18.5 (Eisenstein's criterion): *Let $f(X) = a_nX^n + a_{n-1}X^{n-1} + \cdots + a_0 \in \mathbb{Z}[X]$.* GAUd input, 206 *Suppose there exists a prime number p such that $p \nmid a_n$, $p|a_i$ for $i = 0, \ldots, n-1$, and $p^2 \nmid a_0$. Then f is irreducible in $\mathbb{Q}[X]$.*

Proof: Dividing f by its content, we may assume that f is primitive. By Corollary 18.4, it suffices to prove that f is irreducible in $\mathbb{Z}[X]$.

If f is reducible, then $f = gh$, where

$$g(X) = b_k X^k + b_{k-1} X^{k-1} + \cdots + b_0 \text{ and } h(X) = c_l X^l + c_{l-1} X^{l-1} + \cdots + c_0,$$

are polynomials in $\mathbb{Z}[X]$ with $0 < \deg(g), \deg(h) < n$. Then, $a_0 = b_0 c_0$. Hence, for example, $p \nmid b_0$ and $p | c_0$. Also, $a_n = c_k b_l$ and $p \nmid a_n$, so $p \nmid c_l$. Let r be the smallest integer for which $p \nmid c_r$. Then, $0 < r \leq l < n$. Now observe that

$$a_r = b_0 c_r + \sum_{\substack{i+j=r \\ i \neq 0}} b_i c_j.$$

If $i > 0$ and $i + j = r$, then $j < r$, so $p | c_j$. However, $p \nmid b_0 c_r$, hence $p \nmid a_r$. It follows from this contradiction that f is irreducible in $\mathbb{Z}[X]$. ∎

19. The Binomial Formula

NEWTON
input, 7

For integers $0 \le k \le n$ one defines the **binomial coefficient** by the following formula

$$\binom{n}{k} = \frac{n!}{k!(n-k)!}.$$

In particular $\binom{n}{0} = \binom{n}{n} = 1$ for all $n \ge 0$. Direct computation shows that

$$\binom{n}{k} + \binom{n}{k+1} = \binom{n+1}{k+1}. \tag{1}$$

It follows by induction on n, that $\binom{n}{k}$ is a natural number. Also, we may compute the first binomial coefficients by the **Pascal triangle**:

$$\begin{array}{ccccccccccccc}
 & & & & & & 1 & & & & & & \\
 & & & & & 1 & & 1 & & & & & \\
 & & & & 1 & & 2 & & 1 & & & & \\
 & & & 1 & & 3 & & 3 & & 1 & & & \\
 & & 1 & & 4 & & 6 & & 4 & & 1 & & \\
 & 1 & & 5 & & 10 & & 10 & & 5 & & 1 & \\
1 & & 6 & & 15 & & 20 & & 15 & & 6 & & 1
\end{array}$$

In this triangle $\binom{n}{k}$ stands in the kth place of the nth row (where the triangle starts from the 0's row and each row starts at the 0th place).

If $1 < k < p$ and p is a prime number, then $p|p!$, while $p \nmid k!$ and $p \nmid (p-k)!$. Hence, p divides

$$\binom{p}{k} = \frac{p!}{k!(p-k)!}\ . \tag{2}$$

LEMMA 19.1 (The binomial theorem): *In a commutative ring R with 1 every $x, y \in R$ satisfy $(x+y)^n = \sum_{k=0}^{n} \binom{n}{k} x^k y^{n-k}$.*

NEWa
input, 55

Proof: We apply induction on n and first observe that the lemma trivially holds for

$n = 0$. Now assume that it holds for n. Then, by (1),

$$\begin{aligned}
(x+y)^{n+1} &= (x+y)(x+y)^n \\
&= (x+y)\sum_{k=0}^{n}\binom{n}{k}x^k y^{n-k} \\
&= \sum_{k=0}^{n}\binom{n}{k}x^{k+1}y^{n-k} + \sum_{k=0}^{n}\binom{n}{k}x^k y^{n+1-k} \\
&= \sum_{l=1}^{n+1}\binom{n}{l-1}x^l y^{n+1-l} + \sum_{l=0}^{n}\binom{n}{l}x^l y^{n+1-l} \\
&= x^{n+1} + \sum_{l=1}^{n}\left[\binom{n}{l-1}+\binom{n}{l}\right]x^l y^{n+1-l} + y^{n+1} \\
&= \sum_{l=0}^{n+1}\binom{n+1}{l}x^l y^{n+1-l}
\end{aligned}$$

and the induction is complete. ∎

LEMMA 19.2: *For each prime number p the polynomial* NEWb input, 89

$$\Phi_p(X) = X^{p-1} + X^{p-2} + \cdots + 1$$

is irreducible in $\mathbb{Q}[X]$.

Proof: It suffices to prove that $g(Y) = \Phi_p(Y+1)$ is irreducible in $\mathbb{Q}[Y]$. Indeed, $\Phi_p(X) = \frac{X^p-1}{X-1}$, so

$$\begin{aligned}
\Phi_p(Y+1) = \frac{(Y+1)^p-1}{Y} &= \frac{1}{Y}\Big[\sum_{k=0}^{p}\binom{p}{k}Y^k - 1\Big] \\
&= \frac{1}{Y}\sum_{k=1}^{p}\binom{p}{k}Y^k = \sum_{k=1}^{p}\binom{p}{k}Y^{k-1} = \sum_{l=0}^{p-1}\binom{p}{l+1}Y^l.
\end{aligned}$$

The right hand side is a monic polynomial of degree $p-1$, the coefficient of Y^l is divisible by p for each $0 \le l \le p-2$ (by (2)) and its free coefficient is p, so it is not divisible by p^2. By Eisentstein's criterion (Lemma 18.5) $g(Y)$ is irreducible in $\mathbb{Q}[Y]$. ∎

20. The Quadratic Reciprocity Law

CYCLOTOMIC
input, 8

We use "Gauss's sum" to prove the quadratic reciprocity law.

Consider a prime number p and the complex number

$$\zeta = e^{\frac{2\pi i}{p}} = \cos\left(\frac{2\pi}{p}\right) + i\sin\left(\frac{2\pi}{p}\right).$$

Then the order of ζ in the multiplicative group $\mathbb{C}^\times$ of the field of complex number $\mathbb{C}$ is p. Thus, $\zeta^p = 1$ and $\zeta^k \neq 0$ for each $1 \le k \le p-1$, so ζ generates a subgroup of order p of $\mathbb{C}^\times$. In particular, ζ is a root of the irreducible polynomial $\Phi_p(X) = \frac{X^p-1}{X-1}$ over $\mathbb{Q}$ (Lemma 19.2). We say that ζ is a **primitive root of unity of order** p.

Lemma 20.1: *The element ζ has the following properties:*

CYCa
input, 29

(a) $i \equiv j \mod p$ *if and only if* $\zeta^i = \zeta^j$.

(b) ζ^k *is a primitive root of unity of order p if and only if $p \nmid k$.*

(c) *Let A be a reduced set of representatives modulo p. Then* $\sum_{a\in A} \zeta^a = -1$.

(d) *The elements $1, \zeta, \ldots, \zeta^{p-2}$ are linearly independent over $\mathbb{Q}$.*

Proof of (a): Statement (a) is a special case of Lemma 15.1(b).

Proof of (b): This is a special case of Lemma 15.1(d).

Proof of (c): By definition, there is a bijective map φ: $A \to \{1, \ldots, p-1\}$ such that $\varphi(a) \equiv a \mod p$ for each $a \in A$. By what we wrote above, $\zeta^{p-1} + \cdots + \zeta + 1 = \frac{\zeta^p-1}{\zeta-1} = 0$. Hence, by (a), $\sum_{a\in A} \zeta^a = \sum_{i=1}^{p-1} \zeta^i = -1$.

Proof of (d): Assume there exist $a_0, a_1, \ldots, a_{p-2} \in \mathbb{Q}$, not all of them equal to 0, such that $a_0 + a_1\zeta + \cdots + a_{p-2}\zeta^{p-2} = 0$. Then the polynomial $h \in \mathbb{Q}[X]$ of smallest degree that satisies $h(\zeta) = 0$ has degree $\le p-2$. By Lemma 10.3 there exist $q, r \in \mathbb{Q}[X]$ such that $\Phi_p(X) = q(X)h(X) + r(X)$ and either $r = 0$ or $\deg(r) < \deg(h)$. In the latter case, $0 = \Phi_p(\zeta) = q(\zeta)h(\zeta) + r(\zeta) = r(\zeta)$. This contradiction to the minimality of $\deg(h)$ proves that $r = 0$, so $h|\Phi_p$. But this contradicts the irreducibility of Φ_p (Lemma 19.2). ▮

Definition 20.2: Congruence relation modulo q on R. Next we consider the subset CYCb input, 78

$$R = \{\sum_{i=0}^{p-2} a_i\zeta^i \mid a_0, \ldots, a_{p-2} \in \mathbb{Z}\}$$

of $\mathbb{C}$. It contains $\mathbb{Z}$ and is closed under addition. Moreover, R is closed under multiplication, because $\zeta^{p-1} = -1 - \zeta - \cdots - \zeta^{p-2}$ (by Lemma 20.1(c)), hence by induction, $\zeta^k \in R$ for all $k \geq 0$. Thus, R is a **subring** of $\mathbb{C}$.

Given a prime number q and $\alpha, \beta \in R$, we say that α is **congruent to β modulo q** and write $\alpha \equiv \beta \mod q$ if there exists $\gamma \in R$ such that $\alpha = \beta + \gamma q$. This is an equivalence relation on R compatible with addition and multiplication. Thus, it is a **congruence relation**. It follows that if $\alpha \equiv \beta \mod q$ and $f \in R[X]$, then $f(\alpha) \equiv f(\beta) \mod q$.

Next consider $a, b \in \mathbb{Z}$ that are congruent modulo q in R. Then, there exists $\gamma \in R$ with $a = b + \gamma q$. Write $\gamma = c_0 + c_1\zeta + \cdots + c_{p-2}\zeta^{p-2}$ with $c_0, c_1, \ldots, c_{p-2} \in \mathbb{Z}$. Then, by Lemma 20.1(d), $a = b + c_0 q$, so a is congruent to b modulo q in $\mathbb{Z}$. Thus, the congruence modulo q relation on R is an extension of the congruence modulo q relation on $\mathbb{Z}$.

The same argument implies that $R \cap \mathbb{Q} = \mathbb{Z}$.

By the binomial theorem, $(\alpha+\beta)^q = \sum_{k=0}^{q} \binom{q}{k}\alpha^k\beta^{q-k}$. By (2) of Section 19, $q \big| \binom{q}{k}$ for each $1 \leq k \leq q-1$. Hence,

$$(\alpha + \beta)^q \equiv \alpha^q + \beta^q \mod q. \qquad \blacksquare \tag{2}$$

Let p be an odd prime and let A be a reduced system of representatives modulo p. We introduce the following **Gauss sum**:

$$\tau = \sum_{a \in A} \left(\frac{a}{p}\right)\zeta^a.$$

By Lemma 20.1(a) and Condition (1b) of Section 17, τ does not depend on A.

LEMMA 20.3: $\tau^2 = \left(\frac{-1}{p}\right)p$. CYCc input, 134

Proof: Using Lemma 17.3, we have:

$$\tau^2 = \sum_{a \in A} \left(\frac{a}{p}\right)\zeta^a \sum_{b \in A} \left(\frac{b}{p}\right)\zeta^b = \sum_{a \in A}\sum_{b \in A} \left(\frac{ab}{p}\right)\zeta^{a+b}. \tag{3}$$

For each $c \in \mathbb{Z}$ satisfying $1+c \not\equiv 0 \mod p$, ζ^{1+c} is a primitive root of unity of order p (Lemma 20.1(b). Hence by Lemma 20.1(c), $\sum_{a\in A}(\zeta^{1+c})^a = -1$. For each $a \in A$ we choose $a' \in \mathbb{Z}$ such that $aa' \equiv 1 \mod p$. Then, as b ranges over A, $c = a'b$ ranges over a reduced system of representatives A_a modulo p and $b \equiv ac \mod p$, so $\left(\frac{ab}{p}\right) = \left(\frac{a^2c}{p}\right)$ (Statement (1a) of Section 17) and $\zeta^{a+b} = \zeta^{a+ac}$ (Lemma 20.1(a)). It follows from (3) and from Lemma 17.6 that

$$\begin{aligned}\tau^2 = \sum_{a\in A}\sum_{c\in A_a}\left(\frac{a^2c}{p}\right)\zeta^{a+ac} &= \sum_{a\in A}\sum_{c\in A}\left(\frac{a^2c}{p}\right)\zeta^{a+ac}\\ &= \sum_{c\in A}\left(\frac{c}{p}\right)\sum_{a\in A}(\zeta^{1+c})^a\\ &= \left(\frac{-1}{p}\right)(p-1) - \sum_{c\not\equiv -1 \bmod p}\left(\frac{c}{p}\right)\\ &= \left(\frac{-1}{p}\right)p - \sum_{c\in A}\left(\frac{c}{p}\right) = \left(\frac{-1}{p}\right)p.\end{aligned}$$

as claimed. ▮

THEOREM 20.4 (Quadratic reciprocity law): *If p and q are distinct odd prime numbers, then* CYCd input, 178

$$\left(\frac{q}{p}\right) = (-1)^{\frac{p-1}{2}\frac{q-1}{2}}\left(\frac{p}{q}\right). \tag{4}$$

Proof: We compute τ^q modulo p in two ways. In the first one we apply Lemma 20.3, Euler's criterion (Lemma 17.1), and Lemma 17.4(a).

$$\tau^q = \tau(\tau^2)^{\frac{q-1}{2}} = \tau\left(\frac{-1}{p}\right)^{\frac{q-1}{2}} p^{\frac{q-1}{2}} \equiv \tau(-1)^{\frac{p-1}{2}\frac{q-1}{2}}\left(\frac{p}{q}\right) \mod q \tag{5}$$

In the second computation we use Statement (2) in Definition 20.2, and Lemma 17.3,

$$\tau^q \equiv \sum_{\nu=1}^{p-1}\left(\frac{\nu}{p}\right)\zeta^{\nu q} \equiv \sum_{\nu=1}^{p-1}\left(\frac{\nu}{p}\right)\left(\frac{q^2}{p}\right)\zeta^{\nu q} \equiv \left(\frac{q}{p}\right)\sum_{\nu=1}^{p-1}\left(\frac{\nu q}{p}\right)\zeta^{\nu q} \equiv \left(\frac{q}{p}\right)\tau \mod q \tag{6}$$

It follows from (5) and (6) that

$$\left(\frac{q}{p}\right)\tau \equiv (-1)^{\frac{p-1}{2}\frac{q-1}{2}}\left(\frac{p}{q}\right)\tau \mod q \tag{7}$$

Multiplying both sides of (7) by τ, we get by Lemma 20.3 that

$$(8)\qquad \left(\frac{q}{p}\right)\left(\frac{-1}{p}\right)p \equiv -1^{\frac{p-1}{2}\frac{q-1}{2}}\left(\frac{p}{q}\right)\left(\frac{-1}{p}\right)p \mod q$$

Now we use that $p \neq q$ and divide both sides of (8) by $(\frac{-1}{p})p$ to get

$$(9)\qquad \left(\frac{q}{p}\right) \equiv -1^{\frac{p-1}{2}\frac{q-1}{2}}\left(\frac{p}{p}\right) \mod q$$

Since both sides of (9) are ± 1 and q is odd, we get (4). ∎

Theorem 20.4 leaves out the case where $q = 2$. The next result handles this case.

THEOREM 20.5: *Each odd prime p satisfies the following formula:* CYCe input, 238

$$(10)\qquad \left(\frac{2}{p}\right) = (-1)^{\frac{p^2-1}{8}}. = \begin{cases} 1 & \textit{if } p \equiv 1,7 \mod 8 \\ -1 & \textit{if } p \equiv 3,5 \mod 8 \end{cases}$$

Proof: We set $\sigma = (1+i)^p$, where $i = \sqrt{-1}$. Then $\sigma^2 = (2i)^p$. By (2), $\sigma \equiv 1+i^p \mod p$. Next note that $p^2 \equiv 1 \mod 4$, hence

$$(11)\qquad \sigma^p \equiv 1 + i^{p^2} \equiv 1 + i \mod p.$$

On the other hand, by Euler's criterion

$$(12)\quad \sigma^p = \sigma(\sigma^2)^{\frac{p-1}{2}} \equiv \sigma(2i)^{p\cdot\frac{p-1}{2}} \equiv \sigma\left(\frac{2}{p}\right)^p i^{p\cdot\frac{p-1}{2}} \equiv \left(\frac{2}{p}\right)i^{\frac{p(p-1)}{2}} + \left(\frac{2}{p}\right)i^{\frac{p(p+1)}{2}} \mod p.$$

It follows from (11) and (12) that

$$(13)\qquad 1 + i \equiv \left(\frac{2}{p}\right)\left(i^{\frac{p(p-1)}{2}} + i^{\frac{p(p+1)}{2}}\right) \mod p.$$

Finally we let p run over the reduced residues $1, 3, 5, 7$ modulo 8 and conclude (10) from (13). ∎

Exercise 20.6: Let p be a prime number and let R be a commutative ring with 1 of **characteristic** p, that is $p \cdot 1 = 0$ in R. Then $(a+b)^{p^k} = a^{p^k} + b^{p^k}$ for all $a, b \in R$ and each $k \in \mathbb{N}$. ∎ CYCf input, 279

21. The Jacobi Symbol

JACOBI
input, 9

We establish an effective procedure to compute the Legendre symbol $\left(\frac{a}{p}\right)$ for every odd prime number p and every integer a not divisible by p. The basic tool in this procedure is an extension of the Legendre symbol called the **Jacobi symbol**.

Given relatively prime integers a, b such that b is odd, we decompose b into a product of prime numbers, $b = p_1 p_2 \cdots p_r$ and define

$$(1) \qquad \left(\frac{a}{b}\right) = \left(\frac{a}{p_1}\right)\left(\frac{a}{p_2}\right)\cdots\left(\frac{a}{p_r}\right),$$

where the factors on the right hand side of (1) are the respective Legendre symbols. In particular, if b is an odd prime number, then the Jacobi symbol $\left(\frac{a}{b}\right)$, coincides with the corresponding Legendre symbol.

Proposition 21.1: *The Jacobi symbol has the following properties for all $a, a_1, a_2 \in \mathbb{Z}$ and all odd $b, b_1, b_2 \in \mathbb{Z}$.*

JACa
input, 30

(a) $\left(\frac{a_1 a_2}{b}\right) = \left(\frac{a_1}{b}\right)\left(\frac{a_2}{b}\right)$.

(b) $\left(\frac{a}{b_1 b_2}\right) = \left(\frac{a}{b_1}\right)\left(\frac{a}{b_2}\right)$.

(c) $a_1 \equiv a_2 \mod b$ *implies* $\left(\frac{a_1}{b}\right) = \left(\frac{a_2}{b}\right)$.

(d) $\left(\frac{-1}{b}\right) = (-1)^{\frac{b-1}{2}} = \begin{cases} 1 & \textit{if } b \equiv 1 \mod 4 \\ -1 & \textit{if } b \equiv -1 \mod 4 \end{cases}$.

(e) $\left(\frac{2}{b}\right) = (-1)^{\frac{b^2-1}{8}} = \begin{cases} 1 & \textit{if } b \equiv 1, 7 \mod 8 \\ -1 & \textit{if } b \equiv 3, 5 \mod 8 \end{cases}$.

(f) $\left(\frac{a}{b}\right)\left(\frac{b}{a}\right) = (-1)^{\frac{a-1}{2}\cdot\frac{b-1}{2}} = \begin{cases} 1 & \textit{if } a \textit{ or } b \textit{ are congruent } 1 \textit{ modulo } 4 \\ -1 & \textit{if both } a \textit{ and } b \textit{ are congruent } -1 \textit{ modulo } 4 \end{cases}$ *if both a and b are odd.*

Proof of (d): Since (d) holds for b prime (Lemma 17.4) and its left hand side is multiplicative, it suffices to prove that its right hand side is multiplicative. This follows from the following congruence:

$$(1) \qquad \frac{rs-1}{2} \equiv \frac{r-1}{2} + \frac{s-1}{2} \mod 2$$

for $r = 2r' + 1$ and $s = 2s' + 1$. To prove (1) note that $rs - 1 \cong 2r' + 2s' \mod 4$.

In order to verify the second equality, note that $\frac{b-1}{2}$ is even if $b \equiv 1 \mod 4$ and odd if $b \equiv -1 \mod 4$.

Proof of (e): Again, by Theorem 20.5, it suffices to prove that the right hand side of (e) is multiplicative. Indeed, if r and s are odd, then $r^2 \equiv 1 \mod 8$ and $s^2 \equiv 1 \mod 8$. Hence, $(r^2-1)(s^2-1) \equiv 0 \mod 64$, so $(rs)^2-1 \equiv (r^2-1)+(s^2-1) \mod 64$. Therefore, $\frac{(rs)^2-1}{8} \equiv \frac{r^2-1}{8} + \frac{s^2-1}{8} \mod 8$, which implies the claimed multiplicity.

To prove the second equality of (e) observe that $b^2 \equiv 1 \mod 16$ if $b \equiv \pm 1 \mod 8$ and $b^2 \equiv 9 \mod 16$ if $b \equiv \pm 3 \mod 8$. Hence, $\frac{b^2-1}{8}$ is even if $b \equiv 1, 7 \mod 8$ and $\frac{b^2-1}{8}$ is odd if $b \equiv 3, 5 \mod 8$.

Proof of (f): Here the left hand side is multiplicative in both a and b (by Lemma and by definition). The right hand side of (f) is multiplicative in both variables by the proof of (d). Hence, (f) is a consequence of the quadratic reciprocity law.

Finally, oveserve that $\frac{a-1}{2} \cdot \frac{b-1}{2}$ is even if a or b are congruent 1 modulo 4 and odd if both of them are congruent -1 modulo 4. ∎

Procedure 21.2: Effective procedure to compute Jacobi symbols. Let $a, b \in \mathbb{Z}$ with b JACb input, 109 odd. We show how to compute $\left(\frac{a}{b}\right)$ effectively. If a is even, we divide a by 2 several times till we find a representation $a = 2^k a'$ with $k \in \mathbb{N}$ and a' odd. Now we use Proposition 21.1(a) to write $\left(\frac{a}{b}\right) = \left(\frac{2}{b}\right)^k \left(\frac{a'}{b}\right)$ and compute $\left(\frac{2}{b}\right)$ by Proposition 21.1(d), taking into account that $(-1)^{\frac{b-1}{2}}$ depends only on the residue of b modulo 4.

This allows us to assume that a is odd. If $a > b$ we write $a = qb + r$ with $0 \le r < b$ and use Proposition 21.1(c) to replace a by r.

If $a < b$, we use Proposition 21.1 to replace $\left(\frac{a}{b}\right)$ by $(-1)^{\frac{a^2-1}{8} \cdot \frac{b^2-1}{8}} \left(\frac{b}{a}\right)$. Note that $(-1)^{\frac{a^2-1}{8} \cdot \frac{b^2-1}{8}}$ depends only on a and b modulo 8, so it can be effectively computed.

We repeat the former steps to decrease a and b to half of their size or below. The whole procedure will be carried out in $c \cdot \log(\max(|a|, |b|))$ steps, for some computable constant c.

We may expedite the procedure by using the following tables:

$$(-1)^{\frac{b-1}{2}} = \begin{cases} 1 & \text{if } b \equiv 1 \mod 4 \\ -1 & \text{if } b \equiv 3 \mod 4 \end{cases} \qquad (-1)^{\frac{b^2-1}{8}} = \begin{cases} 1 & \text{if } b \equiv 1 \text{ or } b \equiv 7 \mod 8 \\ -1 & \text{if } b \equiv 3 \text{ or } b \equiv 5 \mod 8 \end{cases}$$

∎

Example 21.3: We demonstrate Procedure 21.2 by the following example. JACc input, 159

$$\begin{aligned}\left(\frac{2819}{4177}\right) &= \left(\frac{4177}{2819}\right) = \left(\frac{1358}{2819}\right) = \left(\frac{2}{2819}\right)\left(\frac{679}{2819}\right) = -\left(\frac{679}{2819}\right)\\ &= \left(\frac{2819}{679}\right) = \left(\frac{103}{679}\right) = -\left(\frac{679}{103}\right) = -\left(\frac{61}{103}\right) = -\left(\frac{103}{61}\right) = -\left(\frac{42}{61}\right)\\ &= -\left(\frac{2}{61}\right)\left(\frac{21}{61}\right) = \left(\frac{21}{61}\right) = \left(\frac{61}{21}\right) = \left(\frac{19}{21}\right) = \left(\frac{21}{19}\right) = \left(\frac{2}{19}\right) = -1.\end{aligned}$$

Since 4177 is a prime number, we conclude that 2819 is a quadratic non-residue modulo 2819. ∎

We conclude this section with a theoretical application of the Jacobi symbol.

THEOREM 21.4: *If an integer a is a quadratic residue modulo every prime number, then a is a square in $\mathbb{Z}$.* JACe input, 193

Proof: If $a = a'm^2$ for some $a', m \in \mathbb{Z}$, then a' is also a quadratic residue modulo every prime number. If we prove that a' is a square, then a is also a square. Thus, we may assume that $a = (-1)^e 2^f p_1 \cdots p_r$, where $e, f \in \{0,1\}$, $r \geq 0$, and $p_1, \ldots, p_r$ are distinct odd prime numbers. We also assume that $a \neq 1$ and draw a contradiction by finding in each possible case an integer b relatively prime to a such that $\left(\frac{a}{b}\right) = -1$.

CASE A: $r \geq 1$. By Lemma 17.5 there exists a quadratic non-residue c modulo p_1. By the Chinese remainder theorem, there exists $b \in \mathbb{N}$ such that $b \equiv 1 \mod 8$, $b \equiv c \mod p_1$, and $b \equiv 1 \mod p_i$ for $i = 2, \ldots, r$. In particular $p_i \nmid b$ for $i = 1, \ldots, r$, so $\gcd(a, b) = 1$. By Proposition 21.1,

$$\left(\frac{a}{b}\right) = (-1)^{\frac{b-1}{2}e}(-1)^{\frac{b^2-1}{8}f}\left(\frac{p_1}{b}\right)\left(\frac{p_2}{b}\right)\cdots\left(\frac{p_r}{b}\right) = \left(\frac{b}{p_1}\right)\left(\frac{b}{p_2}\right)\cdots\left(\frac{b}{p_r}\right) = -1.$$

CASE B: $r = 0$ *and* $f = 1$. Then $a = \pm 2$. In this case $\left(\frac{\pm 2}{5}\right) = -1$ gives the desired contradiction.

CASE C: $r = 0$ *and* $f = 0$. Then $a = -1$ and $\left(\frac{-1}{3}\right) = -1$ gives the desired contradiction. ∎

Exercise 21.5: Compute $\left(\frac{751}{919}\right)$. ∎ JACf input, 242

Exercise 21.6: Compute the **prime divisors** of the polynomial $9X^2 + 6X + 16$, that is the prime numbers p for which there exists $x \in \mathbb{Z}$ such that $9x^2 + 6x + 16 \equiv 0 \mod p$. JACg input, 246

∎

Exercise 21.7: Prove that there exist infinitely many prime numbers of the form $12n-1$. JACh input, 252
Hint: Find the prime divisors of the polynomial $12X^2 - 1$. ∎

Exercise 21.8: Prove that there exist infinitely many prime numbers of the form $12n+5$. JACi input, 258
Hint: Find the prime divisors of the polynomial $(6X + 1)^2 + 4$. ∎

Exercise 21.9: Prove that there exist infinitely many prime numbers of the form $12n+7$. JACj input, 265
Hint: Find the prime divisors of the polynomial $3(2X + 1)^2 + 4$. ∎

22. Characters of Finite Abelian Groups

CHAR
input, 7

The proof of Dirichlet's theorem about the existence of infinitely many prime numbers in every reduced arithmetic progression uses the theory of charecters of finite cyclic groups. Nevertheless, we apply the structure theorem of finite abelian groups to describe the characters of all finite abelian groups.

As usuall, we denote the field of complex numbers by $\mathbb{C}$. A **character** of a finite abelian group A is a **homomorphism** $\chi\colon A \to \mathbb{C}^\times$, that is it satisfies $\chi(ab) = \chi(a)\chi(b)$ for all $a, b \in A$. In particular, if $n = |A|$ and $a \in A$, then $\chi(a)^n = \chi(a^n) = \chi(1) = 1$. In other words, $\chi(a)$ is a root of unity of order that devides n. Since there are only n such roots, the set of all characters of A is finite. We denote it by $\mathrm{Hom}(A, \mathbb{C}^\times)$, or more shortly, by $\hat{A}$. We make $\hat{A}$ into a commutative group by defining the product between two characters χ_1 and χ_2:

$$(\chi_1\chi_2)(a) = \chi_1(a)\chi_2(a).$$

The unit element of $\hat{A}$ is the character ε mapping each element of A onto 1. The inverse of χ is given by the formula $\chi^{-1}(a) = \chi(a)^{-1}$.

The main properties of characters that we prove in this section are the orthogonality formulas and the isomorphism $\hat{A} \cong A$.

LEMMA 22.1: *Let A be a finite abelian group and $\chi \in \hat{A}$. Then*

CHAa
input, 43

$$\sum_{a\in A} \chi(a) = \begin{cases} |A| & \text{if } \chi = \varepsilon \\ 0 & \text{if } \chi \neq \varepsilon \end{cases}$$

Proof: First note that $\sum_{a\in A} \varepsilon(a) = \sum_{a\in A} 1 = |A|$. Now suppose $\chi \neq \varepsilon$. Then there exists $b \in A$ such that $\chi(b) \neq 1$. When a ranges over all elements of A, so does ab. Hence,

$$\sum_{a\in A} \chi(a) = \sum_{a\in A} \chi(ab) = \sum_{a\in A} \chi(a) \cdot \chi(b).$$

Since $\chi(b) \neq 1$, we have $\sum_{a\in A} \chi(a) = 0$. ∎

We extend the correspondance $A \rightsquigarrow \hat{A}$ to a **functor** of the category of finite abelian groups into itself.

For each homomorphism $\alpha\colon A \to B$ of finite abelian groups we define a homomorphism $\hat\alpha\colon \hat B \to \hat A$ by $\hat\alpha(\psi) = \psi \circ \alpha$. Thus, $\hat\alpha(\psi)(a) = \psi(\alpha(a))$ for each $a \in A$. Note that the direction of the arrow is reversed. Thus, if $\beta\colon B \to C$ is an additional homomorphism between finite abelian groups, than $\widehat{\beta\circ\alpha} = \hat\alpha \circ \hat\beta$. If α is the identity map of A, then $\hat\alpha$ is the identity map of $\hat A$. It follows that the correspondance $A \rightsquigarrow \hat A$ and $\alpha \rightsquigarrow \hat\alpha$ form a **contra-variant functor** that we call the **hat functor**.

A sequence $A \xrightarrow{\alpha} B \xrightarrow{\beta} C$ of homomorphisms of abelian groups is **exact** if $\mathrm{Im}(\alpha) = \mathrm{Ker}(\beta)$. In other words, $\{\alpha(a) \mid a \in A\} = \{b \in B \mid \beta(b) = 1\}$. If $A = 1$, this means that β is injective. If $C = 1$, this means that α is surjective.

A longer sequence $\cdots \longrightarrow A \xrightarrow{\alpha} B \xrightarrow{\beta} C \longrightarrow \cdots$ is said to be **exact at** B if the sequence $A \xrightarrow{\alpha} B \xrightarrow{\beta} C$ is exact.

LEMMA 22.2: *The hat functor is* **left exact**. *In other words, if* CHAb input, 103

$$1 \longrightarrow A \xrightarrow{\alpha} B \xrightarrow{\beta} C \longrightarrow 1 \tag{1}$$

is an exact sequence of finite abelian groups, then the sequence

$$1 \longrightarrow \hat C \xrightarrow{\hat\beta} \hat B \xrightarrow{\hat\alpha} \hat A \tag{2}$$

is **exact**. *This means that $\hat\beta$ is injective and* $\mathrm{Im}(\hat\beta) = \mathrm{Ker}(\hat\alpha)$.

Proof: The proof has two parts.

PART A: *$\hat\beta$ is injective.* If $\hat\beta(\chi) = \varepsilon_B$ for some $\chi \in \hat C$, then $\chi(\beta(b)) = 1$ for all $b \in B$. Since β is surjective, this implies that $\chi(c) = 1$ for each $c \in C$. Thus, $\chi = \varepsilon_C$.

PART B: *Exactness at $\hat B$.* First notice that for all $\chi \in \hat C$ and $a \in A$ we have

$$(\hat\alpha \circ \hat\beta)(\chi)(a) = (\chi \circ \beta \circ \alpha)(a) = \chi(1) = 1.$$

Secondly, suppose $\hat\alpha(\chi) = \varepsilon_A$ for some $\chi \in \hat B$. Then, $\chi(\alpha(a)) = 1$ for all $a \in A$. Since (1) is exact in B, we have $\chi(b) = 1$ for all $b \in \mathrm{Ker}(\beta)$. Since β is surjective, the first isomorphism theorem yields a homomorphism $\psi\colon C \to \mathbb{C}^\times$ such that $\psi \circ \beta = \chi$, i.e. $\hat\beta(\psi) = \chi$. ∎

We prove in the sequel that the hat functor is also right exact. But first we has to prove that the functor preserves direct products.

LEMMA 22.3: *There is a natural isomorphism $\widehat{A\times B}\cong \hat{A}\times\hat{B}$.* CHAc input, 159

Proof: To each character $\chi\colon A\times B\to\mathbb{C}^\times$ we attach the pair (χ_A,χ_B) of characters of A and B, respectively, defined by the formulas $\chi_A(a)=\chi(a,1)$ and $\chi_B(b)=\chi(1,b)$. Thus, $\chi(a,b)=\chi_A(a)\chi_B(b)$. One checks that the correspondance $\chi\mapsto(\chi_A,\chi_B)$ is an isomorphism of $\widehat{A\times B}$ onto $\hat{A}\times\hat{B}$. ∎

LEMMA 22.4: *For each finite abelian group there exist a (non-natural) isomorphism $A\cong\hat{A}$.* CHAd input, 176

Proof: We decompose A into a direct product of cyclic groups $A=\prod_{i=1}^m A_i$ (fundamental theorem of the theory of finite abelian groups). By Lemma 22.3, $\hat{A}\cong\prod_{i=1}^m\hat{A}_i$. We may therefore assume that A is cyclic of order n and choose a generator a of A.

Let $\zeta_n=e^{2\pi i/n}$ be a primitive root of unity of order n. For each natural number k we define $\chi_k\in\hat{A}$ by $\chi_k(a)=\zeta_n^k$. Then the map $k\mapsto\chi_k$ gives an isomorphism of $\mathbb{Z}/n\mathbb{Z}$ onto $\hat{A}$. Hence, $A\cong\hat{A}$. ∎

We say that the **short exact sequence (1) is exact** if it is exact in A, in B, and in C.

COROLLARY 22.5: *The hat functor is exact. In other words, if (1) is a short exact sequence of finite abelian groups, then the following sequence is also exact:* CHAe input, 201

$$1\longrightarrow\hat{C}\xrightarrow{\hat{\beta}}\hat{B}\xrightarrow{\hat{\alpha}}\hat{A}\longrightarrow 1. \tag{3}$$

Proof: By Lemma 22.2, it suffices to prove that $\hat{\alpha}$ is surjective. Indeed, $\mathrm{Im}(\hat{\alpha})$ is a subgroup of $\hat{A}$. By Lemma 22.2, $\mathrm{Im}(\hat{\alpha})\cong\hat{B}/\hat{\beta}(\hat{C})$. In addition $\hat{\beta}(\hat{C})\cong\hat{C}$. Hence, by Lemma 22.4,

$$|\mathrm{Im}(\hat{\alpha})|=|\hat{B}|/|\hat{\beta}(\hat{C})|=|\hat{B}|/|\hat{C}|=|B|/|C|=|A|=|\hat{A}|.$$

Hence, $\mathrm{Im}(\hat{\alpha})=\hat{A}$. ∎

COROLLARY 22.6: *Let B be a finite abelian group of order n, let b be an element of B of order d, and let $\zeta\in\mathbb{C}$ satisfy $\zeta^d=1$. Then the number of $\psi\in\hat{B}$ with $\psi(b)=\zeta$ is $\frac{n}{d}$.* NUM input, 231

Proof: Let A be the subgroup of B generated by b and let $C=B/A$. Denote the inclusion of A in B by α. There exists exatly one $\chi\in\hat{A}$ such that $\chi(b)=\zeta$. Moreover, if $\psi \in \hat{B}$, then $\hat{\alpha}(\psi)$ is the restriction of ψ to A. Hence, $\psi(b) = \zeta$ if and only if $\psi = \hat{\alpha}^{-1}(\chi)$. By Corollary 22.5, the order of the latter set is $|\hat{C}|$. Hence, by Lemma 22.4,

$$|\alpha^{-1}(\chi) = |\hat{C}| = |C| = \frac{|B|}{|A|} = \frac{n}{d},$$

as claimed. ∎

LEMMA 22.7: *Let A be a finite abelian group and let $1 \neq a \in A$. Then there exists $\chi \in \hat{A}$ such that $\chi(a) \neq 1$.* CHAf input, 257

Proof: As in the proof of Lemma 22.4, we present A as a direct product $A = \prod_{j=1}^{m} A_j$ of cyclic groups. Let $n_j = |A_j|$ and choose a generator a_j of A_j. Since $a \neq 1$, we have $a = \prod_{j=1}^{m} a_j^{k_j}$, where $0 \leq k_j < n_j$ and at least one of the exponents is not 0. Assume without loss that $k_1 \neq 0$. Then define $\chi \in \hat{A}$ by $\chi(a_1) = \zeta_{n_1}$ and $\chi(a_j) = 1$ for all $j \geq 2$. Then $\chi(a) = \zeta_{n_1}^{k_1} \neq 1$. ∎

COROLLARY 22.8: *Let A be a finite abelian group and let $a \in A$. Then* CHAg input, 274

$$\sum_{\chi \in \hat{A}} \chi(a) = \begin{cases} |A| & \text{if } a = 1 \\ 0 & \text{if } a \neq 1 \end{cases}$$

Proof: If $a = 1$, then $\sum_{\chi \in \hat{A}} \chi(a) = \sum_{\chi \in \hat{A}} 1 = |\hat{A}| = |A|$, by Lemma 22.4. Otherwise $a \neq 1$ and we choose by Lemma 22.7 a character $\psi \in \hat{A}$ such that $\psi(a) \neq 1$. When χ ranges over all elements of $\hat{A}$, so does also $\chi\psi$. Hence,

$$\sum_{\chi \in A} \chi(a) = \sum_{\chi \in A} (\psi\chi)(a) = \psi(a) \sum_{\chi \in \hat{A}} \chi(a).$$

Hence, $\sum_{\chi \in \hat{A}} \chi(a) = 0$. ∎

Finally, let A be a finite commutative group. For each $a \in A$ we define $\psi_a \in \widehat{\hat{A}}$ by $\psi_a(\chi) = \chi(a)$. Then the map $a \mapsto \psi_a$ is a natural homorphism of A into $\widehat{\hat{A}}$ that we denote by Ψ.

COROLLARY 22.9: *The homomorphsim Ψ is an isomorphism.* CHAh input, 311

Proof: By Corollary 22.8, Ψ is injective. By Lemma 22.4, $|A| = |\hat{A}| = |\widehat{\hat{A}}|$. Hence, Ψ is surjective. ∎

23. Dirichlet's Series

DIRICHLET input, 8

We introduce Dirichlet series, prove that they converge in a certain half plane of the complex plane, and define there an analytic function.

We start with a discrete analog of integration by parts.

Lemma 23.1 (Abel's summation): *Let a_i, b_i, $i = 1, \dots, n$ be complex numbers. For each $1 \le i \le n$ let $A_i = a_1 + \cdots + a_i$. Then* DIRa input, 16

$$\sum_{i=1}^{n} a_i b_i = A_n b_n + \sum_{i=1}^{n-1} A_i (b_i - b_{i+1}).$$

Proof: We compute from right to left using that $A_1 = a_1$ and $A_i - A_{i-1} = a_i$ for $i = 2, \dots, n$:

$$\begin{aligned} A_n b_n + \sum_{i=1}^{n-1} A_i (b_i - b_{i+1}) &= A_n b_n + \sum_{i=1}^{n-1} A_i b_i - \sum_{i=1}^{n-1} A_i b_{i+1} \\ &= A_n b_n + \sum_{i=1}^{n-1} A_i b_i - \sum_{i=2}^{n} A_{i-1} b_i \\ &= A_n b_n + A_1 b_1 + \sum_{i=2}^{n-1} (A_i - A_{i-1}) b_i - A_{n-1} b_n \\ &= A_1 b_1 + \sum_{i=2}^{n} (A_i - A_{i-1}) b_i \\ &= \sum_{i=1}^{n} a_i b_i \qquad \blacksquare \end{aligned}$$

Instead of using $z = x + iy$ for the complex variable, $x = \mathrm{Re}(z)$ for its real part, and $y = \mathrm{Im}(z)$ for its imaginary part, it is customary in analytic number theory to use $s = \sigma + it$ for the complex variable, $\sigma = \mathrm{Re}(s)$ for its real part, and $t = \mathrm{Im}(s)$ for its imaginary part. We speak about the **half plane** $\mathrm{Re}(s) > \sigma_0$, whenever we want to refer to the set $\{s \in \mathbb{C} \mid \mathrm{Re}(s) > \sigma_0\}$.

To each sequence $a_1, a_2, a_3, \dots$ of complex numbers we associate a Dirichlet series

$$f(s) = \sum_{n=1}^{\infty} \frac{a_n}{n^s} \tag{1}$$

with the complex variable $s = \sigma + it$.

LEMMA 23.2: *If the Dirichlet series (1) converges at a point $s_0 = \sigma_0 + it_0$, then it converges at every point $s = \sigma + it$ satisfying $\sigma > \sigma_0$. Moreover, the series converges uniformly in every compact subset of the half plane* $\mathrm{Re}(s) > \sigma_0$. DIRb input, 62

Proof: Recall that a compact subset of the complex plane is a closed bounded subset of $\mathbb{C}$. Such a subset can be covered by a set of the form

$$S = \{s \in \mathbb{C} \mid \sigma_1 < \sigma - \sigma_0 \leq |s - s_0| < r\},$$

where $r, \sigma_1 > 0$. Thus, it suffices to prove that for every $\sigma_1 > 0$ and every $r > 0$ the tail of the series (1) uniformly converges to 0 on S.

To this end we recall that the formula $(x^{s_0-s})' = (s_0 - s)x^{s_0-s-1}$ for the derivative of a power of the real variable x leads to the formula

$$(s - s_0)\int_k^l \frac{dx}{x^{s-s_0+1}} = \frac{1}{k^{s-s_0}} - \frac{1}{l^{s-s_0}} \tag{2}$$

for $l > k$. Now fix a natural number m and let $n > m$ be another natuar number. We set

$$P_n(s) = \sum_{i=m+1}^{n} \frac{a_i}{i^s}.$$

We use the Abel summation and (2) to compute:

$$\begin{aligned}\sum_{k=m+1}^{n} \frac{a_k}{k^s} &= \sum_{k=m+1}^{n} \frac{a_k}{k^{s_0}} \cdot \frac{1}{k^{s-s_0}} \\ &= \frac{P_n(s_0)}{n^{s-s_0}} + \sum_{k=m+1}^{n-1} P_k(s_0)\Big[\frac{1}{k^{s-s_0}} - \frac{1}{(k+1)^{s-s_0}}\Big] \\ &= \frac{P_n(s_0)}{n^{s-s_0}} + \sum_{k=m+1}^{n-1} P_k(s_0)(s - s_0)\int_k^{k+1} \frac{dx}{x^{s-s_0+1}}\end{aligned} \tag{3}$$

Next we recall for $x > 0$ that $x^s = x^\sigma x^{it} = x^\sigma e^{it\log x}$ and $|e^{it\log x}| = 1$, so

$$|x^s| = x^\sigma. \tag{4}$$

Given an $\varepsilon > 0$, there exists an m_0 such that

$$|P_n(s_0)| < \varepsilon. \tag{5}$$

if $m > m_0$. For such an m and when $s \in S$, we deduce from (3), (4), and (5) that

$$
\begin{aligned}
(6) \qquad \Big|\sum_{k=m+1}^{n} \frac{a_k}{k^s}\Big| &\le \frac{|P_n(s_0)|}{|n^{s-s_0}|} + |s-s_0| \sum_{k=m+1}^{n-1} |P_k(s_0)| \cdot \int_k^{k+1} \frac{dx}{|x^{s-s_0+1}|} \\
&\le \frac{\varepsilon}{n^{\sigma-\sigma_0}} + |s-s_0|\varepsilon \sum_{k=m+1}^{n-1} \int_k^{k+1} \frac{dx}{x^{\sigma-\sigma_0+1}} \\
&\le \frac{\varepsilon}{n^{\sigma-\sigma_0}} + \varepsilon|s-s_0| \int_{m+1}^{n} \frac{dx}{x^{\sigma-\sigma_0+1}} \\
&= \frac{\varepsilon}{n^{\sigma-\sigma_0}} + \frac{\varepsilon|s-s_0|}{\sigma-\sigma_0}\Big[\frac{1}{(m+1)^{\sigma-\sigma_0}} - \frac{1}{n^{\sigma-\sigma_0}}\Big] \\
&\le \frac{\varepsilon}{n^{\sigma_1}} + \frac{\varepsilon r}{\sigma_1} \cdot \frac{1}{(m+1)^{\sigma_1}}
\end{aligned}
$$

Finally note that the right hand side of (6) is independent of $s \in S$ and converges to 0 as n tends to infinity. This proves our claim. ∎

A similar proof yields a condition for $f(s)$ to converge at a given point s_0.

LEMMA 23.3: *Let $a_1, a_2, a_3, \ldots$ be complex numbers and set $A_n = a_1 + a_2 + \cdots + a_n$ for each natural number n. Suppose there exist $c > 0$ and $\sigma_1 > 0$ such that $|A_n| \le cn^{\sigma_1}$ for every $n \in \mathbb{N}$. Then the Dirichlet series (1) converges in the half plane $\sigma > \sigma_1$.*

DIRc
input, 163

Proof: Consider natural numbers $m < n$ and let $B_n = A_n - A_m = a_{m+1} + \cdots + a_n$. Then

$$
(7) \qquad |B_n| \le |A_n| + |A_m| \le cn^{\sigma_1} + cm^{\sigma_1} \le 2cn^{\sigma_1}.
$$

By Abel's summation and (2),

$$
\begin{aligned}
(8) \qquad \sum_{k+1}^{n} \frac{a_k}{k^s} &= \sum_{k=m+1}^{n} a_k \cdot \frac{1}{k^s} \\
&= B_n \frac{1}{n^s} + \sum_{k=m+1}^{n-1} B_k \Big[\frac{1}{k^s} - \frac{1}{(k+1)^s}\Big] \\
&= \frac{B_n}{n^s} + \sum_{k=m+1}^{n-1} B_k s \int_k^{k+1} \frac{dx}{x^{s+1}}
\end{aligned}
$$

Given $s \in \mathbb{C}$ with $\sigma > \sigma_1$, we have:

$$
\begin{aligned}
(9) \qquad \Big|\sum_{m+1}^{n} \frac{a_k}{k^s}\Big| &\le \frac{|B_n|}{n^\sigma} + \sum_{k=m+1}^{n-1} |B_k| \cdot |s| \int_k^{k+1} \frac{dx}{x^{\sigma+1}} \\
&\le \frac{2cn^{\sigma_1}}{n^\sigma} + 2c|s| \sum_{k=m+1}^{n-1} \int_k^{k+1} \frac{k^{\sigma_1}}{x^{\sigma+1}} dx \\
&\le \frac{2cn^{\sigma_1}}{n^\sigma} + 2c|s| \sum_{k=m+1}^{n-1} \int_k^{k+1} \frac{1}{x^{\sigma-\sigma_1+1}} dx \\
&\le \frac{2cn^{\sigma_1}}{n^\sigma} + 2c|s| \int_{m+1}^{n} \frac{1}{x^{\sigma-\sigma_1+1}} dx \\
&\le \frac{2cn^{\sigma_1}}{n^\sigma} + \frac{2c|s|}{\sigma-\sigma_1}\Big[\frac{1}{(m+1)^{\sigma-\sigma_1}} - \frac{1}{n^{\sigma-\sigma_1}}\Big] \\
&\le \frac{2cn^{\sigma_1}}{n^\sigma} + \frac{2c|s|}{\sigma-\sigma_1} \frac{1}{(m+1)^{\sigma-\sigma_1}}
\end{aligned}
$$

The right hand side of (8) tends to 0 as m tends to infinity. Thus, $\sum_{n=1}^\infty \frac{a_n}{n^s}$ converges.

∎

LEMMA 23.4: *If a Dirichlet series $f(s) = \sum_{n=1}^\infty \frac{a_n}{n^s}$ converges at $s_0 = \sigma_0 + it_0$, then $f(s)$ is analytic at each complex number $s = \sigma + it$ satisfying $\sigma > \sigma_0$.* DIRd input, 226

Proof: Recall that a complex function is said to be **analytic** in an open domain D of $\mathbb{C}$ if it is differentiable there. This means that the derivitive

$$f'(s) = \lim_{s \to s_0} \frac{f(s) - f(s_0)}{s - s_0}$$

exists for each $s \in D$. One proves, like in the theory of real functions, that if f_n, $n = 1, 2, 3, \ldots$, are analitic functions in D and the series $\sum_{n=1}^\infty f_n(s)$ uniformly converges in D to a function $f(s)$, then $f(s)$ is also analytic on D.

In our case, each of the functions $\frac{a_n}{n^s} = a_n e^{s \log n}$ is analytic at each point of $\mathbb{C}$. By Lemma 23.2, the series $\sum_{n=1}^\infty \frac{a_n}{n^s}$ uniformly converges to $f(s)$ over every open disc of the half plane $\mathrm{Re}(s) > \sigma_0$. Hence, $f(s)$ is analytic at each point of that half plane.

∎

LEMMA 23.5: *If a Dirichlet series $f(s) = \sum_{n=1}^\infty \frac{a_n}{n^s}$ converges at $s_0 = \sigma_0 + it_0$, then $f(s)$ is analytic at each complex number $s = \sigma + it$ satisfying $\sigma > \sigma_0$.* DIRe input, 258

Proof:

24. Riemann Zeta Function

RIEMANN input, 8

The Riemann zeta function is initially defined by the Dirichlet series

$$(1) \qquad \zeta(s) = \sum_{n=1}^{\infty} \frac{1}{n^s}.$$

It is one of the most intriguing functions of mathematics. In particular, it is an essential ingredient in the proof of Dirichlet's theorem about the primes in reduced arithmetic progressions.

It is well known that the series $\sum_{n=1}^{\infty} \frac{1}{n^\sigma}$ converges for every $\sigma > 1$. Hence, by Lemma 23.4, $\zeta(s)$ is an analytic function on the complex half plane $\mathrm{Re}(s) > 1$. The proof of Lemma 24.1 below applies the principle of "analytic continuation" to extend $\zeta(s)$ to an analytic function on the half plane $\mathrm{Re}(s) > 0$ except at $s = 1$.

Let $f(s)$ and $g(s)$ be complex functions that are analytic on domains C and D respectively. Suppose $C \subseteq D$ and $f(s) = g(s)$ for all $s \in C$. Then g is said to be an **analytic continuation** of f to D. If C and D are non-empty and open then g is unique.

Recall that a complex function is analytic at a point s_0 if and only if there exists an $r > 0$ such that f is defined on the open disc $|s - s_0| < r$ and has there a presentation as a converging power series $f(s) = \sum_{n=0}^{\infty} a_n (s - s_0)^n$, with complex coefficients a_n.

One says that f is **meromorphic** at s_0 with a **simple pole** if there exist a complex number a_{-1} and an analytic function f_0 in an open disc D around s_0 such that $f(s) = \frac{a_{-1}}{s - s_0} + f_0(s)$. The number a_{-1} is the **residue** of f at s_0 and is denoted by $\mathrm{res}_{s=s_0} f(s)$. Since $f_0(s)$ is continues at s_0, we have

$$(2) \qquad \mathrm{res}_{s=s_0} f(s) = \lim_{s \to s_0} (s - s_0) f(s).$$

LEMMA 24.1: *It is possible to continue* $\zeta(s)$ *to an analytic function on the half plane* $\mathrm{Re}(s) > 0$ *except at the point* $s = 1$, *where* $\zeta(s)$ *has a simple pole with residue* 1.

RIEa input, 57

Proof: Recall that for a real number x the symbol $[x]$ denote the greatest integer less than or equal to x. Then $\{x\} = x - [x]$ satisfies $0 \le \{x\} < 1$. Using again the Abel summation (Lemma 23.1), we rewrite the partial sums of the right hand side of (3) for an arbitrary $s = \sigma + it \in \mathbb{C}$:

$$
\begin{aligned}
(3) \qquad \sum_{k=1}^{n} \frac{1}{k^s} = \sum_{k=1}^{n} 1 \cdot \frac{1}{k^s} &= \frac{n}{n^s} + \sum_{k=1}^{n-1} k\Big(\frac{1}{k^s} - \frac{1}{(k+1)^s}\Big) \\
&= \frac{1}{n^{s-1}} + \sum_{k=1}^{n-1} ks \int_k^{k+1} \frac{dx}{x^{s+1}} \\
&= \frac{1}{n^{s-1}} + s \sum_{k=1}^{n-1} \int_k^{k+1} \frac{[x]}{x^{s+1}} dx \\
&= \frac{1}{n^{s-1}} + s \int_1^n \frac{[x]}{x^{s+1}} dx \\
&= -s \int_1^n \frac{\{x\}}{x^{s+1}} dx + \frac{1}{n^{s-1}} + s \int_1^n \frac{x}{x^{s+1}} dx \\
&= -s \int_1^n \frac{\{x\}}{x^{s+1}} dx + \frac{1}{n^{s-1}} + s \frac{1 - n^{1-s}}{s-1}
\end{aligned}
$$

If $\sigma > 1$, then $n^{1-s} \to 0$ as $n \to \infty$. It follows from (3) that

$$
(4) \qquad f(s) = \sum_{k=1}^{\infty} \frac{1}{k^s} = \frac{s}{s-1} - s \int_1^{\infty} \frac{\{x\}}{x^{s+1}} dx = \frac{1}{s-1} + 1 - s \int_1^{\infty} \frac{\{x\}}{x^{s+1}} dx.
$$

Since $0 \le \{x\} < 1$ for all x, and x^{s+1} is an analytic function on the whole complex plane, the integral uniformly converges to an analytic function on the half plane $\mathrm{Re}(s) > 0$. It follows that the right hand side of (4) is an analytic function $g(s)$ on the whole half plane $\mathrm{Re}(s) > 0$ except at $s = 1$, where it has a simple pole with residue 1. Since $f(s) = g(s)$ for all $\sigma > 1$, the function $g(s)$ is an analytic continuation of $f(s)$ as stated in the lemma. ∎

Corollary 24.2: *The function $\zeta(\sigma)$ has real values for $\sigma > 1$ and* $\lim_{\sigma \to 1^+} \zeta(\sigma) = \infty$. RIEb input, 116

Proof: By Lemma 24.1, $\zeta(\sigma) = \frac{1}{\sigma - 1} + f_0(\sigma)$ for $\sigma > 1$, where $f_0(\sigma)$ is a continuous function at 1. This implies the Corollary. ∎

25. Characters Modulo m

MODN input, 6

Along with the Riemann ζ-function, one introduces the Dirichlet's L-functions. They are defined as Dirichlet series with coefficients that are characters modulo m. Then one uses the orthogonality relation of characters to isolate the prime numbers congruent to a specific a modulo m among all prime numbers.

We fix a natural number m for the whole section. To each character χ of $(\mathbb{Z}/m\mathbb{Z})^\times$ we associate a function $\tilde{\chi}$: $\mathbb{Z} \to \mathbb{C}$:

$$\tilde{\chi}(a) = \begin{cases} \chi(a + m\mathbb{Z}) & \text{if } \gcd(a,m) = 1 \\ 0 & \text{otherwise.} \end{cases}$$

We call $\tilde{\chi}$ a **character modulo** m. The multiplicity of χ yields a **strong multiplicity** of $\tilde{\chi}$:

$$\tilde{\chi}(ab) = \tilde{\chi}(a)\tilde{\chi}(b)$$

for all $a, b \in \mathbb{Z}$. The lifting of the trivial character of $(\mathbb{Z}/n\mathbb{Z})^\times$ to $\mathbb{Z}$ is called the **principal character modulo** m and is denoted by χ_1. It is defined by $\chi_1(a) = 1$ if $\gcd(a,m) = 1$ and $\chi_1(a) = 0$ otherwise.

The orthogonality relations of characters yield orthogonality relations of the characters modulo m. To this end we denote the set of all characters modulo m by $X(m)$. The map $\chi \mapsto \tilde{\chi}$ is a bijection of $\widehat{(\mathbb{Z}/m\mathbb{Z})^\times}$ onto $X(m)$ making $X(m)$ a group.

Lemma 25.1: MODa input, 46

(a) *Let A be a set of represenatives of $\mathbb{Z}/m\mathbb{Z}$ in $\mathbb{Z}$. Then, for each $\chi \in X(m)$ we have*

$$\sum_{a\in A} \tilde{\chi}(a) = \begin{cases} \varphi(m) & \textit{if } \tilde{\chi} \textit{ is principal.} \\ 0 & \textit{if } \tilde{\chi} \textit{ is non-principal} \end{cases}$$

(b) *For each $a \in \mathbb{Z}$ we have:*

$$\sum_{\psi\in X(m)} \psi(a) = \begin{cases} \varphi(m) & \textit{if } a \equiv 1 \mod m \\ 0 & \textit{otherwise.} \end{cases}$$

Proof of (a): By Definition 9.6, $|(\mathbb{Z}/m\mathbb{Z})^\times| = \varphi(m)$. Also note that $\tilde{\chi}$ is principal if and only if χ is the trivial character of $(\mathbb{Z}/m\mathbb{Z})^\times$. Hence, by Lemma 22.1,

$$\sum_{a\in A} \tilde{\chi}(a) = \sum_{\substack{a\in A \\ \gcd(a,m)=1}} \tilde{\chi}(a) = \sum_{\bar{a}\in(\mathbb{Z}/m\mathbb{Z})^\times} \chi(\bar{a}) = \begin{cases} \varphi(m) & \text{if } \tilde{\chi} \text{ is principal} \\ 0 & \text{otherwise.} \end{cases}$$

Proof of (b): Let $\bar{a} = a + \mathbb{Z}$. If $\gcd(a, m) \neq 1$, then $\psi(a) = 0$ for all $\psi \in X(m)$, so $\sum_{\psi \in X(m)} \psi(a) = 0$. If $\gcd(a, m) = 1$, then $\bar{a} \in (\mathbb{Z}/m\mathbb{Z})^\times$, so by Corollary 22.8 we have

$$\sum_{\psi \in X(m)} \psi(a) = \sum_{\chi \in \widehat{(\mathbb{Z}/m\mathbb{Z})^\times}} \chi(\bar{a}) = \begin{cases} \varphi(m) & \text{if } a \equiv 1 \bmod m \\ 0 & \text{if } a \not\equiv 1 \bmod m. \end{cases} \qquad \blacksquare$$

26. Euler Products

EULER input, 6

We apply the fundamental theorem of arithmetic to write Dirichlet series with ‘strongly multiplicative’ coefficients as ‘Euler products’. Here we say that a function $f\colon \mathbb{Z} \to \mathbb{C}$ is **strongly multiplicative** if $f(ab) = f(a)f(b)$ for all $a, b \in \mathbb{Z}$.

LEMMA 26.1: *Let $f\colon \mathbb{Z} \to \mathbb{C}$ be a strongly multiplicative function and let s be a complex number such that the series $\sum_{n=1}^{\infty} \frac{f(n)}{n^s}$ absolutely converges. Then the infinite product $\prod_p (1 - f(p)p^{-s})^{-1}$, where p ranges over all prime number, converges and* EULa input, 14

$$\sum_{n=1}^{\infty} \frac{f(n)}{n^s} = \prod_p \frac{1}{1 - \frac{f(p)}{p^s}} \tag{1}$$

Proof: Let m be a natural number. By the fundamental theorem of number theory, every natural number n whose prime divisors are at most m has a unique representation as a product $\prod_{p \le m} p^{k(p)}$ with non-negative integral exponents $k(p)$. For such an n we have $f(n) = \prod_{p \le m} f(p)^{k(p)}$. Hence, by the absolute convergence,

$$\prod_{p \le m} \frac{1}{1 - \frac{f(p)}{p^s}} = \prod_{p \le m} (1 + \frac{f(p)}{p^s} + \frac{f(p)^2}{p^{2s}} + \frac{f(p)^3}{p^{3s}} + \cdots) = \sum_{n=1}^{m} \frac{f(n)}{n^s} + {\sum_{n>m}}' \frac{f(n)}{n^s}$$

where the prime in the latter sum indicates that n ranges over all natural number greater than m all of their prime divisors are at most m. It follows that

$$\Big| \sum_{n=1}^{\infty} \frac{f(n)}{n^s} - \prod_{p \le m} \frac{1}{1 - \frac{f(p)}{p^s}} \Big| = \Big| \sum_{n>m} \frac{f(n)}{n^s} - {\sum_{n>m}}' \frac{f(n)}{n^s} \Big| \le \sum_{n=m+1}^{\infty} \Big| \frac{f(n)}{n^s} \Big|. \tag{2}$$

By the absolut convergence, the series on the right hand side of (2) tends to 0 as m tends to infinity. This implies that the infinite product comverges and that it is equal to the infinite series as stated in the lemma. ∎

Example 26.2: EULb input, 62

(a) The Riemann zeta function $\zeta(s) = \sum_{n=1} \frac{1}{n^s}$ absolutely converges when $\mathrm{Re}(s) > 1$. Since the constant function 1 is strongly multiplicative, the Euler formula

$$\zeta(s) = \prod_p \frac{1}{1 - \frac{1}{p^s}} \tag{3}$$

holds when $\mathrm{Re}(s) > 1$.

(b) Let m be a natural number and let χ be a character modulo m. We consider the Dirichlet **L-function**

$$L(s,\chi) = \sum_{n=1}^{\infty} \frac{\chi(n)}{n^s}. \tag{4}$$

By Section 25, $\chi\colon \mathbb{Z} \to \mathbb{C}$ is strongly multiplicative. Moreover, $\chi(n)^m = 1$ if $\gcd(n,m) = 1$ and $\chi(n) = 0$ if $\gcd(n,m) \neq 1$. Thus, in each case $|\chi(n)| \le 1$, so $|\chi(1)| + \cdots + |\chi(n)| \le n$ for each n. It follows Lemma 23.3 that the Dirichlet L-function abolutely converges when $\mathrm{Re}(s) > 1$. It follows from Lemma 23.4 that $L(s,\chi)$ is an analytic function on the half plane $\mathrm{Re}(s) > 1$. By Lemma 26.1,

$$L(s,\chi) = \prod_p \frac{1}{1 - \frac{\chi(p)}{p^s}} \tag{5}$$

for each $s \in \mathbb{C}$ with $\mathrm{Re}(s) > 1$. ∎

The L-function $L(s,\chi_1)$ is closely related to the Riemann zeta function. This becomes evident in the following result.

Lemma 26.3: *The function $L(s,\chi_1)$ can be continued to an analytic function on the half plane $\mathrm{Re}(s) > 0$ exept at $s = 1$ where it has a simple pole with residue $\frac{\varphi(m)}{m}$.* EULc input, 109

Proof: We compare $L(s,\chi_1)$ to $\zeta(s)$. For a prime number p, $\chi_1(p) = 0$ if $p|m$ and $\chi_1(p) = 1$ if $p \nmid m$. Hence, by (5) and (3), we have for each $s \in \mathbb{C}$ with $\mathrm{Re}(s) > 1$ that

$$\begin{aligned} L(s,\chi_1) &= \prod_p \Big(1 - \frac{\chi_1(p)}{p^s}\Big)^{-1} = \prod_{p\nmid m} \Big(1 - \frac{1}{p^s}\Big)^{-1} \\ &= \prod_p \Big(1 - \frac{1}{p^s}\Big)^{-1} \prod_{p|m} \Big(1 - \frac{1}{p^s}\Big) = \zeta(s) \prod_{p|m} \Big(1 - \frac{1}{p^s}\Big). \end{aligned} \tag{6}$$

By Lemma 24.1, $\zeta(s)$ is an analytic function on the half plane $\mathrm{Re}(s) > 0$ except for a simple point at $s = 1$ with residue 1. Moreover, $\prod_{p|m}\big(1 - \frac{1}{p^s}\big)$ is analytic on the whole plane. Thus, the right hand side of (6) continues $L(s,\chi_1)$ to an anylitic function on the half plane $\mathrm{Re}(s) > 0$ except for a simple pole at $s = 1$. By (2) of Section 24 and by

Lemma 24.1 its residue at $s = 1$ is

$$\begin{aligned}\mathrm{res}_{s=1}L(s,\chi_1) &= \lim_{s\to 1}(s-1)\zeta(s)\prod_{p|m}\big(1-\frac{1}{p^s}\big)\\ &= \lim_{s\to 1}\big((s-1)\zeta(s)\big)\lim_{s\to 1}\prod_{p|m}(1-\frac{1}{p^s})\\ &= \big(\mathrm{res}_{s=1}\zeta(s)\big)\prod_{p|m}\big(1-\frac{1}{p}\big) = \frac{\varphi(m)}{n}\end{aligned}$$

as claimed. ∎

LEMMA 26.4: *For each charater $\chi \neq \chi_1$ modulo n the function $L(s,\chi)$ is analytic in the whole right plane* $\mathrm{Re}(s) > 0$. EULd input, 160

Proof: We write each natural number n in the form $n = qm + r$ for some $q \in \mathbb{N}$, $r \in \mathbb{Z}$, and $0 \le r \le m-1$. Then notice that

$$|\sum_{k=1}^{m}\chi(k)| = |\sum_{i=0}^{q-1}\sum_{j=0}^{m-1}\chi(im+j)+\sum_{j=0}^{r}\chi(qm+j)| \le \sum_{i=0}^{q-1}|\sum_{j=0}^{m-1}\chi(im+j)|+\sum_{j=0}^{r}|\chi(qm+j)|.$$

By Lemma 25.1(a), $\sum_{j=0}^{m-1}\chi(im+j) = 0$ for each i. Also, $|\chi(k)| \le 1$ for each k. Hence, $|\sum_{k=1}^{n}\chi(k)| \le m\cdot n^0$ for all m. It follows from Lemma 23.3 that $L(s,\chi) = \sum_{n=1}^{\infty}\frac{\chi(n)}{n^s}$ converges for each s with $\mathrm{Re}(s) > 0$. It follows from Lemma 23.4 that $L(s,\chi)$ is analytic in the whole half plane $\mathrm{Re}(s) > 0$. ∎

27. Abscissa of Convergence

ABSCISSA
input, 7

Following Lemma 23.2, we define the **abscissa of convergence** of a Dirichlet series $f(s) = \sum_{n=1}^{\infty} \frac{a_n}{n^s}$ as the infimum of all real numbers σ_0 for which there exist $t_0 \in \mathbb{R}$ such that $f(s)$ converges at $s_0 = \sigma_0 + it_0$. By that lemma, $-\infty \le \alpha \le \infty$ and $f(s)$ converges on the half plane $\mathrm{Re}(s) > \sigma_0$. By Lemma 23.4, $f(s)$ is analytic on that half plane.

ABSa
input, 19

LEMMA 27.1: *Let $f(s) = \sum_{n=1}^{\infty} \frac{a_n}{n_s}$ be a Dirichlet series with non-negative real coefficients a_n. Suppose the abscissa of convergence α of $f(s)$ is a real number. Then $s = \alpha$ is a singular point of $f(s)$.*

Proof: Let $\sigma_1 > \alpha$ and present $f(s)$ as a power series around $s = \sigma_1$:

$$f(s) = \sum_{k=0}^{\infty} \frac{f^{(k)}(\sigma_1)}{k!}(s - \sigma_1)^k.$$

By what we wrote above, the sum converges in each s of the Disc $D = \{z \in \mathbb{C} \mid |z| \le \sigma_1 - \alpha\}$ except possibly in $s = \alpha$.

Assume by contradiction that $s = \alpha$ is a regular point of $f(s)$. Then D_1 is contained in the regularity domain of $f(s)$. In particular, $f(s)$ is analytic in an open disc around α. Therefore, the radius of convergence r of $f(s)$ around σ_1 is greator than $\sigma_1 - \alpha$.

We choose $\varepsilon > 0$ such that $r > \sigma_1 - \alpha + \varepsilon$ and note that $(n^{-s})' = (-\log n)n^{-s}$, so that the k'th derivative of $f(s) = \sum_{n=1}^{\infty} \frac{a_n}{n^s}$ is

$$f^{(k)}(\sigma_1) = \sum_{n=1}^{\infty} \frac{(-\log n)^k a_n}{n^{\sigma_1}}.$$

Now we use that we may change the order of the summation in a double series with non-negative real number and observe in the next computation that the Dirichlet series

$\sum_{n=1}^{\infty}\frac{a_n}{n^{\alpha-\varepsilon}}$ converges:

$$\begin{aligned}\sum_{n=1}^{\infty}\frac{a_n}{n^{\alpha-\varepsilon}} &= \sum_{n=1}^{\infty}\frac{a_n}{n^{\sigma_1}}\cdot n^{\sigma_1-\alpha+\varepsilon}\\ &= \sum_{n=1}^{\infty}\frac{a_n}{n^{\sigma_1}}e^{(\log n)(\sigma_1-\alpha+\varepsilon)}\\ &= \sum_{n=1}^{\infty}\sum_{k=0}^{\infty}\frac{(\log n)^k(\sigma_1-\alpha+\varepsilon)^k}{k!}\\ &= \sum_{k=0}^{\infty}\sum_{n=1}^{\infty}\frac{(\log n)^k(\sigma_1-\alpha+\varepsilon)^k a_n}{k!n^{\sigma_1}}\\ &= \sum_{k=0}^{\infty}\frac{(\alpha-\varepsilon-\sigma_1)^k}{k!}\sum_{n=1}^{\infty}\frac{(-\log n)^k a_n}{n^{\sigma_1}}\\ &= \sum_{k=0}^{\infty}\frac{f^{(k)}(\sigma_1)}{k!}(\alpha-\varepsilon-\sigma_1)^k = f(\alpha-\varepsilon)\end{aligned}$$

This contradicts the minimality of α. ∎

LEMMA 27.2: *Let $\beta\in\mathbb{R}$ and let $f(s)=\sum_{n=1}^{n}\frac{a_n}{n^s}$ be a Dirichlet series with non-negative coefficients that converges on the half plane $\mathrm{Re}(s)>\beta$. Suppose $f(s)$ has an analytic continuation to the half plane $\mathrm{Re}(s)>\alpha$ for some $\alpha\le\beta$. Then $\sum_{n=1}^{\infty}\frac{a_n}{n^s}$ converges on the half plane $\mathrm{Re}(s)>\alpha$.* ABSb input, 84

Proof: Assume $\sum_{n=1}^{\infty}\frac{a_n}{n^{s'}}$ diverges for some $s'=\sigma'+it'$ with $\sigma'>\alpha$. Then the abscissa of convergence α' of $\sum_{n=1}^{\infty}\frac{a_n}{n^s}$ satisfies $\alpha'\ge\sigma'$. By Lemma 27.1, α' is a singular point of $f(s)$, in contrust to our assumption on $f(s)$. ∎

LEMMA 27.3: *Let $f(s)=\sum_{n=1}^{\infty}\frac{a_n}{n_s}$ and $g(s)=\sum_{n=1}^{\infty}\frac{b_n}{n_s}$ be Dirichlet series with non-zero non-negative real coefficients a_n. Let α and β be the abscissas of convergence of g and f respecitvely. Suppose $-\infty<\alpha\le\beta<\infty$ and let $h(s)=f(s)g(s)$ for each $s\in\mathbb{C}$ with $\mathrm{Re}(s)>\beta$. Then $h(s)=\sum_{n=1}^{\infty}\frac{c_n}{n^s}$, where $c_n=\sum_{d|n}a_db_{n/d}\ge 0$ for each n is a representation of $h(s)$ as a Dirichlet series and the abcissa of convergence of h is at most β.* ABSc input, 104

Proof: For each real $\sigma>\beta$ both series $\sum_{n=1}^{\infty}\frac{a_n}{n^\sigma}$ and $\sum_{n=1}^{\infty}\frac{b_n}{n^\sigma}$ are absolutely convergent series. By a theorem of Cauchy, their product also absolutely converges, where the nth term is $\frac{c_n}{n^\sigma}$ with c_n given as above. This conclude the proof. ∎

28. Non-Vanishing of $L(1,\chi)$

NONVANISH
input, 8

A crucial point in the proof of Dirichlet's theorem is the non-vanishing of $L(1,\chi)$ for each non-principal character $\chi \in X(m)$. Here we follow Serre's book "A course in Arithmetic". The proof circumvent delicate points of handling complex logarithm.

Again we fix a natural number m. If $p \nmid m$, we let $f(p) = \mathrm{ord}_m p$. By Lemma 15.1(a), $f(p)|\varphi(m)$. Let $g(p) = \frac{\varphi(m)}{f(p)}$.

Lemma 28.1: *For each natural number*

NONa
input, 21

$$1 - X^n = \prod_w (1 - wX), \tag{1}$$

where w ranges over all roots of $X^n - 1 = 0$ in $\mathbb{C}$.

Proof: The complex roots of $X^n - 1$ form a multiplicative group W of order n. Both sides of (1) are polynomials of order n having W as their set of roots. Indeed, when w ranges over W, so does w^{-1} and $1 - ww^{-1} = 0$.

It follows that each $w \in W$ is also the root of the polynomial

$$f(X) = (1 - X^n) - \prod_{w \in W} (1 - wX).$$

However, the free coeffient of both $1 - X^n$ and $\prod_{w\in W}(1 - mX)$ is 1. Hence, $f(X) = Xg(X)$, where $g \in \mathbb{C}[X]$ is a polynomial of degree $n-1$. Moreover, $w \neq 0$ for each $w \in W$, so $g(w) = 0$ for each $w \in S$. This contradicts Lemma 10.2 unless $g = 0$. Thus, (1) is indeed an identity. ∎

Lemma 28.2: *If $p \nmid m$, then*

NONb
input, 53

$$\prod_{\chi \in X(m)} (1 - \chi(p)X) = (1 - X^{f(p)})^{g(p)} \tag{2}$$

Proof: We apply Corollary 22.6 to the group $B = (\mathbb{Z}/m\mathbb{Z})^\times$ of order $\varphi(m)$ and to the element $\bar{p} = p + m\mathbb{Z}$ whose order is $f(p)$. By that Corollary, for each complex root w of $X^{f(p)} - 1$ the number of characters χ of $(\mathbb{Z}/m\mathbb{Z})^\times$ such that $\chi(\bar{p}) = w$ is $\frac{\varphi(m)}{f(p)}$. Hence, the number of characters $\chi \in X(m)$ such that $\chi(p) = w$ is $g(p)$. Thus, $\prod_{\chi\in X(m)}(1 - \chi(p)X) = \big(\prod_{w\in W}(1 - wX)\big)^{g(p)}$, where W now is the group of complex roots of $X^{f(p)} - 1 = 0$. By Lemma 28.1, $\prod_{w \in W}(1 - wX) = 1 - X^{f(p)}$. Combining this with the preceding equality, we get (2). ▮

LEMMA 28.3: *The function* $\zeta_m(s)$ *defined for* $\mathrm{Re}(s) > 1$ *by the formula* NONc input, 80

$$\zeta_m(s) = \prod_{\chi \in X(m)} L(s, \chi)$$

has the multiplicative representation

$$\zeta_m(s) = \prod_{p \nmid m} \frac{1}{\left(1 - \frac{1}{p^{f(p)s}}\right)^{g(p)}}.$$

Moreover, $\zeta_m(s)$ *has a representation as a Dirichlet series with non-negative integral coefficients that converges in the half plane* $\mathrm{Re}(s) > 1$.

Proof: By Examples 26.2(b) and by Lemma 28.2, we have for each s with $\mathrm{Re}(s) > 1$,

$$(3)\ \zeta_m(s) = \prod_{\chi \in X(m)} \prod_p \frac{1}{1 - \frac{\chi(p)}{p^s}} = \prod_{p \nmid m} \Big(\prod_{\chi \in X(m)} 1 - \chi(p)p^{-s} \Big)^{-1}$$
$$= \prod_{p \nmid m} \frac{1}{\left(1 - \frac{1}{p^{f(p)s}}\right)^{g(p)}} = \prod_{d \mid m} \Big(\prod_{\substack{p \nmid m \\ f(p) = d}} \frac{1}{1 - \frac{1}{p^{ds}}} \Big)^{\varphi(m)/d}$$

For each divisor d of m we define

$$h_d(n) = \begin{cases} 0 & \text{if } \gcd(m, n) > 1 \\ 1 & \text{if } \gcd(m, n) = 1. \end{cases}$$

By Lemma 26.1,

$$(4) \qquad \prod_{p \nmid m} \frac{1}{1 - \frac{1}{p^s}} = \prod_p \frac{1}{1 - \frac{h(p)}{p^s}} = \sum_{n=1}^{\infty} \frac{h(n)}{n^s} = \sum_{\gcd(m,n)=1} \frac{1}{n^s}$$

and the right hand side is a Dirichlet series with non-negative integral coefficients that converges when $\mathrm{Re}(s) > 1$. Therefore, the same statement holds for the product $\prod_{p \nmid m} \frac{1}{1 - \frac{1}{p^{ds}}}$ for each divisor d of m. It follows from Lemma 27.3 that the right hand side of (3) can be represented as a Dirichlet series with non-negative integral numbers that converges for $\mathrm{Re}(s) > 1$, as claimed. ▮

Lemma 28.4: *For each natural number m, the series $\sum_{\gcd(m,n)=1} \frac{1}{n}$ diverges.* HAR input, 144

Proof: Observe that $\gcd(km+1, m) = 1$ and $2km \geq km+1$ for each natural number k. Hence,

$$\sum_{\gcd(m,n)=1} \frac{1}{n} \geq \sum_{k=1}^{\infty} \frac{1}{km+1} \geq \frac{1}{2m} \sum_{k=1}^{\infty} \frac{1}{k}.$$

Since the the harmonic series diverges, so does the series given in the lemma. ∎

Proposition 28.5: *$L(1,\chi) \neq 0$ for every non-principal character modulo m.* NONd input, 162

Proof: By Lemma 26.3,

$$L(s,\chi_1) = \frac{\varphi(m)}{m(s-1)} + f(s) \tag{5}$$

where $f(s)$ is analytic at $s = 1$.

Assume there exists a non-principal character χ_2 modulo m such that $L(1,\chi_2) = 0$. By Lemma 26.4, $L(s,\chi_2)$ is analytic at $s = 1$. Hence

$$L(s,\chi_2) = b_1(s-1) + g(s)(s-1)^2 \tag{5}$$

where $g(s)$ is analytic at $s = 1$. Multiplying (5) and (6), we get

$$L(s,\chi_1)L(s,\chi_2) = \frac{\varphi(m)b_1}{m} + (b_1 f(s) + \frac{\varphi(m)}{m} g(s))(s-1) + f(s)g(s)(s-1)^2$$

is analytic at $s = 1$.

By Lemma 26.4, $L(s,\chi)$ is analytic at $s = 1$ for all non-principal characters modulo m. Hence,

$$\zeta_m(s) = \prod_{\chi \in X(m)} L(s,\chi) = \big(L(s,\chi_1)L(s,\chi_2)\big) \prod_{\chi \neq \chi_1,\chi_2} L(s,\chi)$$

is analytic at $s = 1$. In addition, by Lemmas 26.3 and 26.4, for each $\chi \in X(m)$, the function $L(s,\chi)$ is analytic at each $s \in \mathbb{C}$ with $\mathrm{Re}(s) > 0$ and $s \neq 1$. It follows that $\zeta_m(s)$ is analytic on the whole half plane $\mathrm{Re}(s) > 0$.

Since $\zeta_m(s)$ has a presentation as a Dirichlet series with non-negative coeffients on the half plane $\mathrm{Re}(s) > 1$, it follows from Lemma 27.2 that the Dirichlet series representing $\zeta_m(s)$ converges on the half plane $\mathrm{Re}(s) > 0$. In particular, that series converges at $\frac{1}{\varphi(m)}$.

On the other hand, for each $\sigma > 0$, the pth factor of $\zeta_m(\sigma)$ is equal to

$$\frac{1}{(1-p^{-f(p)\sigma})^{g(p)}} = (1 + p^{-f(p)\sigma} + p^{-2f(p)\sigma} + \cdots)^{g(p)} \tag{6}$$

and dominates the series

$$1 + p^{-\varphi(m)\sigma} + p^{-2\varphi(m)\sigma} + \cdots \tag{7}$$

because $f(p)g(p) = \varphi(m)$ and each of the summands of (7) is a summand of (6). It follows that the Dirichlet series representing $\zeta_m(\sigma) = \prod_{p\nmid m}(1-p^{-f(p)\sigma})^{-g(p)}$ dominates the product

$$\sum_{n=1}^{\infty} \frac{c_n}{n^\sigma} = \prod_{p\nmid m}(1 + p^{-\varphi(m)\sigma} + p^{-2\varphi(m)\sigma} + \cdots) \tag{8}$$

with non-negative integral coefficients, hence the series

$$\sum_{\gcd(m,n)=1} n^{-\varphi(m)\sigma}, \tag{9}$$

because each element of (9) appears in the series (8). However, by Lemma 28.4, the series (9) diverges at $\sigma = \frac{1}{\varphi(m)}$, so $\zeta\big(\frac{1}{\varphi(m)}\big)$ also diverges at $\sigma = \frac{1}{\varphi(m)}$. This contradiction to conclusion of the the preceding paragraph, proves that $L(1,\chi) \neq 0$ for all non-principal characters modulo m. ∎

29. Proof of Dirichlet's Theorem

LOG input, 8

We define the **Dirichlet density** of a set A of prime numbers as the limit

$$\delta(A)=\lim_{\sigma\to 1^+}\frac{\sum_{p\in A}\frac{1}{p^\sigma}}{\sum_p\frac{1}{p^\sigma}} \tag{1}$$

if it exists.

THEOREM 29.1 (Dirichlet): *Let m be a natural number and let a be a relatively prime integer to m. Then:* LLGb input, 23

(a) $\lim_{\sigma\to 1^+}\sum_p\frac{1}{p^\sigma}=\infty$.

(b) *The Dirichlet density of each finite set of prime numbers is* 0.

(c) *The Dirichlet density of the set of prime numbers $p\equiv a \mod m$ is $\frac{1}{\varphi(m)}$.*

(d) *There are infinitely many prime numbers $p\equiv a \mod m$.*

Proof: We break up the proof of the theorem into several parts.

PART A: *A branch of the logarithm function.* Recall that a **branch** of $\log z$ is a continuous function $l(z)$ defined on a connected open subset U of the complex plane such that $z=e^{l(z)}$ for each $z\in U$. If $l_1(z)$ is another branch of $\log z$ on U, then there exists $q\in\mathbb{Z}$ such that $l_1(z)=l(z)+2\pi i\cdot q$.

By Proposition 28.5 $L(1,\chi)\neq 0$ for each non-principal character χ modulo m. Hence, we may choose a ray R in the left half plane $\mathrm{Re}(z)\le 0$ starting at the origin and passing through none of the points $L(1,\chi)$, where $\chi\in X(m)$ and $\chi\neq\chi_1$. Let $D=\mathbb{C}\smallsetminus R$ and choose a branch of $\log z$ which is analytic on D and coincides with the usual real valued logarithm for each positive real number z. By the principal of analytic continuation,

$$-\log(1-z)=\sum_{n=1}^{\infty}\frac{z^n}{n} \tag{2}$$

if $|z-1|<1$ (becasue (2) holds for real numbers z satifying $|z-1|<1$). Moreover, if $z_1,z_2,z_3,\ldots$ is a sequence of complex numbers of absolute value less that 1, then

(3) $l=\sum_{n=1}^{\infty}\log(1-z_n)$ converges if and only if $\prod_{n=1}^{\infty}(1-z_n)$ converges.

In that case,

$$\prod_{n=1}^{\infty}(1-z_n)=e^{l} \tag{4}$$

(See, K. Knopp, Theory and Application of Infinite Series, Blackie & Son, London, 1928, Section 57).

Part B: *The function $l(s,\chi)$.* For each $\chi \in X(m)$, for every prime number p, and for each $\sigma > 1$ we have $\left|\frac{\chi(p)}{p^\sigma}\right| < 1$ and $L(\sigma,\chi)=\sum_{n=1}^{\infty}\frac{\chi(n)}{n^\sigma}$ converges (Example 26.2(b)). Moreover, $L(s,\chi)=\prod_p(1-\frac{\chi(p)}{p^s})^{-1}$. Since $\left|\frac{\chi(p)}{p^s}\right| = \frac{1}{p^\sigma} < 1$, we have by (3) and (4), that $l(s,\chi)=-\sum_p \log\big(1-\frac{\chi(p)}{p^s}\big)$ is well defined function on the half plane $\mathrm{Re}(s)>1$ that satisfies

$$L(s,\chi)=e^{l(s,\chi)}. \tag{5}$$

By (2),

$$\begin{aligned} l(s,\chi) &= -\sum_p \log\big(1-\frac{\chi(p)}{p^s}\big) \\ &= \sum_p \sum_{n=1}^{\infty}\frac{\chi(p^n)}{np^{ns}} = \sum_p \frac{\chi(p)}{p^s} + R_\chi(s), \end{aligned} \tag{6}$$

where

$$R_\chi(s)=\sum_p\sum_{n=2}^{\infty}\frac{\chi(p^n)}{np^{n\sigma}}.$$

We estimate $R_\chi(\sigma)$ for $\sigma>1$ close to 1:

$$\begin{aligned} |R_\chi(\sigma)| &\le \sum_p\sum_{n=2}^{\infty}\frac{1}{np^{n\sigma}} \le \frac{1}{2}\sum_p\sum_{n=2}^{\infty}\frac{1}{p^{n\sigma}} = \frac{1}{2}\sum_p\frac{1}{p^{2\sigma}}\cdot\frac{1}{1-p^{-\sigma}} \\ &\le \frac{1}{2}\cdot\frac{1}{1-2^{-\sigma}}\sum_{k=1}^{\infty}\frac{1}{k^{2\sigma}} = \frac{1}{2}\cdot\frac{1}{1-2^{-\sigma}}\zeta(2\sigma) \end{aligned} \tag{7}$$

The right hand side of (7) converges to $\zeta(2)$ as $\sigma \to 1^+$. Hence,

(8) $R_\chi(\sigma)$ is bounded in a right real neighborhood of 1.

Part C: *The limit of $l(s,\chi)$ at $s=1$.* Consider $\chi \in X(m)$. Set $a_k = \frac{k}{\log_p k}$ if $k = p^n$ with p a prime number and $n \in \mathbb{N}$. If k is not of the above form set $a_k = 0$. Then, $|a_k| \le 1$ for each k (hence $|\sum_{k=1}^{n} a_k| \le n$), and by (6), $l(s,\chi) = \sum_{k=1}^{\infty} \frac{a_k}{k^s}$. By Lemmas 23.3 and 23.4, $l(s,\chi)$ is an analytic function on the half plane $\mathrm{Re}(s) > 1$. It follows from (5) that $l(s,\chi)$ is a branch of the $\log L(s,\chi)$ on the half plane $\mathrm{Re}(s) > 1$. It follows that there exists a positive integer $q(\chi)$ such that $c(\chi) = 2pi \cdot q(\chi)$ satisfies $l(s,\chi) = \log L(s,\chi) + c(\chi)$.

If $\chi \ne \chi_1$, then, by Lemma 26.4, $L(s,\chi)$ is analytic on the half plane. Moreover, $L(1,\chi) \in D$, so $L(s,\chi) \in D$ if s is sufficiently close to 1. Hence, in this case,

(9) $\log L(\sigma,\chi)$ is analytic in an open neighborhood of 1, so $\lim_{\sigma\to 1^+} \log L(\sigma,\chi) = \log L(1,\chi)$.

In particular, if $\chi \ne \chi_1$, then by (9), $l(s,\chi)$ is analytic in an open neighborhood of 1 and

$$\lim_{\sigma\to 1^+} l(s,\chi) = l(1,\chi). \tag{10}$$

Part D: *An application of the orthogonality relation.* Next we choose an integer a' satisfying $a'a \equiv 1 \mod m$. Then $a'p \equiv 1 \mod m$ if and only if $p \equiv a \mod m$. Hence, by Lemma 25.1(b), $\sum_{\chi\in X(m)} \chi(a')\chi(p) = \varphi(m)$ if $p \equiv a \mod m$ and $\sum_{\chi\in X(m)} \chi(a')\chi(p) = 0$ if $p \not\equiv a \mod m$. Multiplying (6) by $\chi(a')$, summing up over all characters χ modulo n, we get

$$\begin{aligned} \sum_{\chi} \chi(a') l(\sigma,\chi) &= \sum_{p} \frac{1}{p^s} \sum_{\chi} \chi(a')\chi(p) + \sum_{\chi} \chi(a') R_\chi(s) \\ &= \varphi(m) \sum_{p\equiv a \mod m} \frac{1}{p^s} + R^*(s), \end{aligned} \tag{11}$$

where

(12) $R^*(s) = \sum_\chi \chi(a')R_\chi(s)$ is a bounded function in a right real neighborhood of 1.

By (6) in the proof of Lemma 26.3, $e^{l(s,\chi_1)} = L(s,\chi_1) = \zeta(\sigma)\prod_{p|m} \left(1 - \frac{1}{p^\sigma}\right)$. Since each of the numbers appearing in these equalities are real, we have, by (6), that

$$\sum_p \frac{1}{p^\sigma} + R_{\chi_1}(\sigma) = l(s,\chi_1) = \log\zeta(\sigma) + \sum_{p|m} \log\left(1 - \frac{1}{p^\sigma}\right). \tag{13}$$

By Corollary 24.2, $\zeta(\sigma)$ is a real valued function for $\sigma > 1$ that approaches ∞ as $\sigma \to 1^+$. Hence, $\lim_{\sigma\to 1^+} \log\zeta(\sigma) = \infty$. It follows from (13) and (8) that

$$\lim_{\sigma\to 1^+} \sum_p \frac{1}{p^s} = \infty. \tag{14}$$

This proves (a).

Part E: *Proofs of (b), (c), and (d).* It follows from (a) that if A is a finite set of prime numbers, than the numerator of the fraction on the right hand side of (1) tends to the finite number $\sum_{p\in A} \frac{1}{p}$, while, by (14), the denominater tends to ∞. Thus, $\delta(A) = 0$, as (b) asserts.

Now note that $\chi_1(a') = 1$ because $\gcd(a', m) = 1$. Hence, by (13) and (11),

$$\begin{aligned} \sum_p \frac{1}{p^\sigma} + R_{\chi_1}(\sigma) + \sum_{\chi\neq\chi_1} \chi(a')l(\sigma,\chi) \\ &= \sum_{\chi\in X(m)} \chi(a')l(\sigma,\chi) \\ &= \varphi(m) \sum_{p\equiv a \bmod m} \frac{1}{p^\sigma} + R^*(\sigma). \end{aligned} \tag{15}$$

We divide equality (15) by $\sum \frac{1}{p^\sigma}$ and let σ approach 1 from the right. The first summand on the left of (15) becomes 1, the second one tends to 0 by (8), the third one tends to 0, by (10). Finally the second summand on the right hand side tends to 0 in the limit because of (12). It follows that

$$\frac{\lim\limits_{\sigma\to 1^+} \sum\limits_{p\equiv a \bmod m} \frac{1}{p^\sigma}}{\lim\limits_{\sigma\to 1^+} \sum_p \frac{1}{p^\sigma}} = \frac{1}{\varphi(m)},$$

as claimed by (c). It follows from (c) that there are infinitely many prime numbers $p \equiv a \bmod m$, as (d) claims. ∎