Elementary Number Theory (MATH 324) – v 1.02

Paul Buckingham

These notes provide the core material for a course on elementary number theory taught at the University of Alberta. The reader is assumed to have had a first course in basic ring theory, covering (i) basic modular arithmetic, (ii) basic properties of rings, (iii) the notions of integral domain and field, and (iv) unique factorization domains.

All proofs are given, although many are in the Appendix instead of the main text.

Notation

- The floor of a real number x, i.e., the greatest integer less than or equal to x, will be denoted $\lfloor x \rfloor$.
- The symbol $\approx$ will mean *is approximately equal to*.
- Occasionally, when required to because of limited space, we will abbreviate a congruence $x \equiv y \bmod m$ to $x \equiv y \, (m)$.
- The set of units in a unital ring R will be denoted $R^\times$.

Contents

(I) Preliminaries

I – 1 Basic properties of integers

We recall some basic concepts and properties concerning the integers. If $a, b \in \mathbb{Z}$, then b is said to *divide* a if there is $c \in \mathbb{Z}$ such that $a = bc$. The following fact is fundamental to the integers:

> Let $a, b \in \mathbb{Z}$, and assume that $b \neq 0$. Then there are unique integers q and r with $0 \leq r < |b|$ such that $a = qb + r$.

The process of finding q and r as above is called *division with remainder*.

If a, b are integers, not both zero, then a *greatest common divisor* of a and b is a positive common divisor of a and b that is divisible by all common divisors. The following is proven in MATH 228.

Theorem 1.1 (G.C.D. Theorem)**.** *Let $a, b \in \mathbb{Z}$, not both zero.*

(i) A greatest common divisor of a and b exists and is unique. We denote it $\gcd(a, b)$.

(ii) There exist integers m and n such that $\gcd(a, b) = ma + nb$.

Example.

a	b	$\gcd(a, b)$
4	6	2
-4	6	2
0	11	11
5	10	5
15	35	5
126	147	21

The Euclidean algorithm

We briefly recall the Euclidean algorithm from MATH 228. Suppose that $a, b \in \mathbb{Z}$ where $b \neq 0$, and write $a = qb + r$ with $0 \leq r < |b|$. Then any common divisor of b and r divides a as well and so is a common divisor of a and b. Also, because $r = a - qb$, any common divisor of a and b is a common divisor of b and r. Thus, $\gcd(a, b) = \gcd(b, r)$. The Euclidean algorithm takes advantage of this fact to compute greatest common divisors.

Example. Find $\gcd(14\,161, 11\,011)$, and find integers m and n such that $\gcd(14\,161, 11\,011) = m \cdot 14\,161 + n \cdot 11\,011$.

Solution: We repeatedly apply division with remainder:

$$\begin{aligned} 14\,161 &= 11\,011 + 3150 \\ 11\,011 &= 3 \cdot 3150 + 1561 \\ 3150 &= 2 \cdot 1561 + 28 \\ 1561 &= 55 \cdot 28 + 21 \end{aligned}$$

$$28 = 21 + 7$$

Because $7 \mid 21$, we see that $\gcd(14\,161, 11\,011) = 7$. Further,

$$\begin{aligned}
7 &= 28 - 21 \\
&= 28 - (1561 - 55 \cdot 28) \\
&= 56 \cdot 28 - 1561 \\
&= 56(3150 - 2 \cdot 1561) - 1561 \\
&= 56 \cdot 3150 - 113 \cdot 1561 \\
&= 56 \cdot 3150 - 113(11\,011 - 3 \cdot 3150) \\
&= 395 \cdot 3150 - 113 \cdot 11\,011 \\
&= 395(14\,161 - 11\,011) - 113 \cdot 11\,011 \\
&= 395 \cdot 14\,161 - 508 \cdot 11\,011
\end{aligned}$$

The two main steps above—finding the greatest common divisor, and then expressing it in terms of the two original integers—together form the Euclidean algorithm.

The Fundamental Theorem of Arithmetic

Theorem 1.2 (Fundamental Theorem of Arithmetic)**.** *Every positive integer can be factorized into a product of primes, and the factorization is unique up to the order of the prime factors. (We allow* 1 *to be considered the* empty product *of primes, i.e., the product of no primes.)*

The theorem is proven in MATH 228. Here are the prime factorizations of the first 20 positive integers:

n	Prime factorization of n	n	Prime factorization of n
1	1	11	11
2	2	12	$2^2 \cdot 3$
3	3	13	13
4	2^2	14	$2 \cdot 7$
5	5	15	$3 \cdot 5$
6	$2 \cdot 3$	16	2^4
7	7	17	17
8	2^3	18	$2 \cdot 3^2$
9	3^2	19	19
10	$2 \cdot 5$	20	$2^2 \cdot 5$

Valuations

Because of the Fundamental Theorem of Arithmetic, every non-zero rational number can be expressed uniquely as a product

$$a = \epsilon(a) \prod_p p^{v_p(a)},$$

where $\epsilon(a) \in \{1, -1\}$, p runs through the primes, $v_p(a) \in \mathbb{Z}$ for each prime p, and $v_p(a) = 0$ for all but finitely many p. The integer $v_p(a)$ is called the *p-adic valuation* of a.

Example. If $a = 35/169 = 5^1 \cdot 7^1 \cdot 13^{-2}$, then

$$\epsilon(a) = 1, \quad v_5(a) = 1, \quad v_7(a) = 1, \quad v_{13}(a) = -2,$$

and $v_p(a) = 0$ for all other primes.

Example. If $a = -100/21 = 2^2 \cdot 3^{-1} \cdot 5^2 \cdot 7^{-1}$, then

$$\epsilon(a) = -1, \quad v_2(a) = 2, \quad v_3(a) = -1, \quad v_5(a) = 2, \quad v_7(a) = -1,$$

and $v_p(a) = 0$ for all other primes.

If we define $v_p(0) = \infty$, then the function $v_p : \mathbb{Q} \to \mathbb{Z} \cup \{\infty\}$ satisfies the following properties, where $a, b \in \mathbb{Q}$:

$$v_p(a + b) \geq \min(v_p(a), v_p(b))$$
$$v_p(ab) = v_p(a) + v_p(b)$$

From these properties, a third follows, namely, that

$$v_p(a + b) = \min(v_p(a), v_p(b)) \quad \text{if } v_p(a) \neq v_p(b).$$

We leave these facts regarding v_p as exercises.

Remark. We make a cautionary remark regarding the last fact, which applies only when $v_p(a) \neq v_p(b)$. If, instead, $v_p(a) = v_p(b)$, then both the following can occur: $v_p(a + b) = \min(v_p(a), v_p(b))$, and $v_p(a + b) > \min(v_p(a), v_p(b))$. For example, if $p = 3$, $a = 3$, and $b = 12$, then $v_3(a) = v_3(b) = 1$, and $v_3(a + b) = v_3(15) = 1 = \min(v_3(a), v_3(b))$. By contrast, if $p = 3$, $a = 3$, and $b = 6$, then $v_3(a) = v_3(b) = 1$ again, but this time, $v_3(a + b) = v_3(9) = 2 > \min(v_3(a), v_3(b))$.

Exercise. Let X be a non-empty set and $f : \mathbb{Z}_{\geq 0} \to X$ a periodic function. Let m be the minimum period of f, and let n be any period of f. Show that the remainder on dividing m into n is zero, so that m in fact exactly divides n. (If $n = qm + r$, consider $f(a + r) = f(a + n - qm)$ for integers $a \geq 0$.)

I – 2 Induction

We briefly recall the two forms of induction.

First form

Let $n_0 \in \mathbb{Z}$, and for each $n \geq n_0$, let $P(n)$ be a statement depending on n. Assume that

(i) $P(n_0)$ is true, and

(ii) for all $n \geq n_0$, if $P(n)$ is true, then $P(n+1)$ is true.

Then $P(n)$ is true for all $n \geq n_0$.

Example. Show by induction that $\sum_{k=0}^{n} \frac{k}{(k+1)!} = 1 - \frac{1}{(n+1)!}$ for all $n \geq 0$.

Solution: The statement holds when $n = 0$, because both sides are zero in that case. Now let $n \geq 0$, and assume that $\sum_{k=0}^{n} \frac{k}{(k+1)!} = 1 - \frac{1}{(n+1)!}$. Then

$$\begin{aligned}\sum_{k=0}^{n+1} \frac{k}{(k+1)!} &= \frac{n+1}{(n+2)!} + \sum_{k=0}^{n} \frac{k}{(k+1)!} \\ &= \frac{n+1}{(n+2)!} + 1 - \frac{1}{(n+1)!} \quad \text{by the inductive hypothesis} \\ &= 1 + \frac{(n+1)-(n+2)}{(n+2)!} = 1 - \frac{1}{(n+2)!},\end{aligned}$$

and the induction is complete.

Second form

Let $n_0 \in \mathbb{Z}$, and for each $n \geq n_0$, let $P(n)$ be a statement depending on n. Assume that

(i) $P(n_0)$ is true, and

(ii) for all $n \geq n_0$, if $P(k)$ is true for all $k \in \{n_0, \ldots, n\}$, then $P(n+1)$ is true.

Then $P(n)$ is true for all $n \geq n_0$.

Example. Show by induction that every positive integer is a product of primes. (This is one part of the statement of the Fundamental Theorem of Arithmetic.)

Solution: The case $n = 1$ holds, because 1 is the empty product of primes. Now let $n \geq 1$, and assume that every $k \in \{1, \ldots, n\}$ is a product of primes. There are two cases: (i) $n+1$ is prime, in which case we are done immediately. (ii) $n+1$ is not prime, meaning that $n+1 = ab$ where $a, b \in \{1, \ldots, n\}$. In this case, by the inductive hypothesis, each of a and b is a product of primes, so the same is true of $ab = n+1$. This completes the induction.

(II) Congruences

II – 1 The integers mod m

Recall from MATH 228 that if m is a positive integer, and if $a, b \in \mathbb{Z}$, we write $a \equiv b \bmod m$ if $m \mid a-b$. The relationship $a \equiv b \bmod m$ is read "a is congruent to b mod m".

It is often convenient to reformulate a congruence mod m as an equality in the ring $\mathbb{Z}/m\mathbb{Z}$ of integers mod m. Let us recall $\mathbb{Z}/m\mathbb{Z}$. For each $a \in \mathbb{Z}$, let $[a]_m$ be its residue class mod m, i.e.,

$$[a]_m = \{b \in \mathbb{Z} \mid b \equiv a \bmod m\} = \{a + km \mid k \in \mathbb{Z}\}.$$

Example. In the case $m = 3$,

$$\begin{aligned}
[0]_3 &= \{\ldots, -3, 0, 3, 6, 9, \ldots\}, \\
[1]_3 &= \{\ldots, -2, 1, 4, 7, 10, \ldots\}, \\
[2]_3 &= \{\ldots, -1, 2, 5, 8, 11, \ldots\}.
\end{aligned}$$

The set of residue classes mod m is denoted $\mathbb{Z}/m\mathbb{Z}$, i.e.,

$$\mathbb{Z}/m\mathbb{Z} = \{[a]_m \mid a \in \mathbb{Z}\} = \{[0]_m, [1]_m, \ldots, [m-1]_m\}.$$

Observe, now, the relationship between congruence of integers and equality in $\mathbb{Z}/m\mathbb{Z}$, namely, $a \equiv b \bmod m$ if and only if $[a]_m = [b]_m$.

The set $\mathbb{Z}/m\mathbb{Z}$ is in fact a ring with respect to the operations

$$\begin{aligned}
[a]_m + [b]_m &= [a+b]_m, \\
[a]_m \cdot [b]_m &= [ab]_m,
\end{aligned}$$

both well-defined. The ring $\mathbb{Z}/m\mathbb{Z}$ is commutative and unital, the identity being $[1]_m$.

Remark. If the modulus m is understood, we will usually omit the subscript on $[a]_m$ and write simply $[a]$.

The congruence $ax \equiv b \bmod m$

Proposition 1.1. *Let m be a positive integer, let $a, b \in \mathbb{Z}$, and let $d = \gcd(a, m)$. Then the congruence $ax \equiv b \bmod m$ has a solution if and only if $d \mid b$. In this case, there are exactly d solutions mod m.*

Proof. Suppose that x is a solution, i.e., $ax \equiv b \bmod m$. Then $m \mid ax-b$, so also $d \mid ax-b$ because $d \mid m$. But then $d \mid b$ because $d \mid a$.

Conversely, suppose that $d \mid b$. Write $a = da'$, $b = db'$, and $m = dm'$. Then $ax \equiv b \bmod m$ if and only if $a'x \equiv b' \bmod m'$. But $\gcd(a', m') = 1$, so there is $c \in \mathbb{Z}$ such that $ca' \equiv 1 \bmod m'$, and hence

$$a'x \equiv b' \bmod m' \iff ca'x \equiv cb' \bmod m' \iff x \equiv cb' \bmod m'.$$

Thus, the set of solutions is $\{cb' + km' \mid k \in \mathbb{Z}\}$. Now, two solutions $cb' + km'$ and $cb' + lm'$ are congruent mod m if and only if

$$\begin{aligned} & m \,|\, (cb' + km') - (cb' + lm') = (k-l)m' \\ \iff & dm' \,|\, (k-l)m' \\ \iff & d \,|\, k - l. \end{aligned}$$

Therefore, the solutions $cb' + km'$ with $0 \leq k \leq d-1$ represent all the solutions mod m. □

Example. Find all solutions to $55x \equiv 10 \bmod 105$, and give your answer first as a single congruence $x \equiv a \bmod n$ for appropriate a and n, and second as a set of congruences $x \equiv a_1, \ldots, a_{k-1}$, or $a_k \bmod 105$.

Solution:

$$\begin{aligned} 55x \equiv 10 \bmod 105 & \iff 11x \equiv 2 \bmod 21 \\ & \iff 2 \cdot 11x \equiv 2 \cdot 2 \bmod 21 \\ & \iff x \equiv 4 \bmod 21 \quad \text{because } 2 \cdot 11 \equiv 1 \bmod 21 \\ & \iff x \equiv 4,\ 25,\ 46,\ 67, \text{ or } 88 \bmod 105. \end{aligned}$$

Prime residue classes

A residue class $[a] \in \mathbb{Z}/m\mathbb{Z}$ is called a *prime residue class* if it is a unit (i.e., is invertible) in the ring, i.e., if there is $[b] \in \mathbb{Z}/m\mathbb{Z}$ such that $[a][b] = [1]$. The set of prime residue classes is denoted $(\mathbb{Z}/m\mathbb{Z})^\times$. It follows from the G.C.D. Theorem that

$$(\mathbb{Z}/m\mathbb{Z})^\times = \{[a] \in \mathbb{Z}/m\mathbb{Z} \mid \gcd(a, m) = 1\}.$$

The number of prime residue classes mod m is denoted $\phi(m)$. The function $\phi : \mathbb{Z}_{\geq 1} \to \mathbb{Z}_{\geq 1}$ is called Euler's *totient function.*

Example.

$$(\mathbb{Z}/21\mathbb{Z})^\times = \{[1], [2], [4], [5], [8], [10], [11], [13], [16], [17], [19], [20]\},$$

and $\phi(21) = 12$.

Remark. One always has $[1] \in (\mathbb{Z}/m\mathbb{Z})^\times$. If $\alpha, \beta \in (\mathbb{Z}/m\mathbb{Z})^\times$, then $\alpha\beta$ and α^{-1} are in $(\mathbb{Z}/m\mathbb{Z})^\times$ as well.

If $\alpha \in \mathbb{Z}/m\mathbb{Z}$ and $n \in \mathbb{Z}_{\geq 0}$, then $\alpha^n = \underbrace{\alpha \cdots \alpha}_{n}$, where $\alpha^0 = [1]$ by definition. In the special case where $\alpha \in (\mathbb{Z}/m\mathbb{Z})^\times$, we may even define $\alpha^n = (\alpha^{-n})^{-1}$ when $n < 0$.

We have the following rules of exponentiation:

$$\alpha^{n_1+n_2} = \alpha^{n_1}\alpha^{n_2}, \quad (\alpha^{n_1})^{n_2} = \alpha^{n_1 n_2}.$$

For an arbitrary $\alpha \in \mathbb{Z}/m\mathbb{Z}$, these rules hold for $n_1, n_2 \geq 0$, and they hold for all $n_1, n_2 \in \mathbb{Z}$ when $\alpha \in (\mathbb{Z}/m\mathbb{Z})^\times$.

II – 2 The Chinese Remainder Theorem

We recall the Chinese Remainder Theorem from MATH 228. Let $m_1, \ldots, m_r$ be pairwise-coprime positive integers, and let $m = m_1 \cdots m_r$. Then the map

$$\begin{aligned} \mathbb{Z}/m\mathbb{Z} &\to \mathbb{Z}/m_1\mathbb{Z} \times \cdots \times \mathbb{Z}/m_r\mathbb{Z} \\ [a]_m &\mapsto ([a]_{m_1}, \ldots, [a]_{m_r}) \end{aligned}$$

is a ring isomorphism. In particular, for any $a_1, \ldots, a_r \in \mathbb{Z}$, there is $a \in \mathbb{Z}$ such that $a \equiv a_i \bmod m_i$ for all i, and a is determined uniquely mod m.

In the case where $r = 2$, the simultaneous congruences

$$\left.\begin{aligned} x &\equiv a_1 \quad \bmod m_1 \\ x &\equiv a_2 \quad \bmod m_2 \end{aligned}\right\} \tag{2.1}$$

are solved as follows. First, write $1 = s_1 m_1 + s_2 m_2$ with $s_1, s_2 \in \mathbb{Z}$—via the Euclidean algorithm, for example. Then the general solution to (2.1) is

$$x \equiv a_1 s_2 m_2 + a_2 s_1 m_1 \bmod m_1 m_2.$$

Example. Solve the simultaneous congruences

$$\begin{aligned} x &\equiv 14 \quad \bmod 19 \\ x &\equiv 4 \quad \bmod 8 \end{aligned}$$

Solution: From the Euclidean algorithm, one finds that $1 = 3 \cdot 19 - 7 \cdot 8$, so the solution is

$$\begin{aligned} x &\equiv 14(-7) \cdot 8 + 4 \cdot 3 \cdot 19 \quad && \bmod 19 \cdot 8 \\ &= -556, \\ \text{i.e.,} \quad x &\equiv 52 && \bmod 152. \end{aligned}$$

The Chinese Remainder Theorem for prime residue classes

The map $\mathbb{Z}/m\mathbb{Z} \to \mathbb{Z}/m_1\mathbb{Z} \times \cdots \times \mathbb{Z}/m_r\mathbb{Z}$ in the Chinese Remainder Theorem gives a map

$$\begin{aligned} (\mathbb{Z}/m\mathbb{Z})^\times &\to (\mathbb{Z}/m_1\mathbb{Z})^\times \times \cdots \times (\mathbb{Z}/m_r\mathbb{Z})^\times \\ [a]_m &\mapsto ([a]_{m_1}, \ldots, [a]_{m_r}) \end{aligned}$$

that is again a bijection. In other words, if a_i is coprime to m_i for all i, then the solutions $a \in \mathbb{Z}$ to the system $x \equiv a_i \bmod m_i$, $i = 1, \ldots, r$, are coprime to m.

Example. Consider the coprime moduli 11 and 14. Because $\gcd(3, 11) = \gcd(5, 14) = 1$, the solutions to the system

$$\begin{aligned} x &\equiv 3 \quad \bmod 11 \\ x &\equiv 5 \quad \bmod 14 \end{aligned}$$

are coprime to $11 \cdot 14 = 154$. We leave it as an exercise to find the solutions and verify that they are indeed coprime to 154.

$\phi(m)$ via prime factorization

Proposition 2.1. *Let m be a positive integer, and let $m = p_1^{a_1} \cdots p_r^{a_r}$ be its prime factorization, where $a_i > 0$ for all i. Then*

$$\phi(m) = \phi(p_1^{a_1}) \cdots \phi(p_r^{a_r}) = (p_1 - 1)p_1^{a_1 - 1} \cdots (p_r - 1)p_r^{a_r - 1}.$$

Proof. By the Chinese Remainder Theorem for prime residue classes,

$$\begin{aligned} \phi(m) = |(\mathbb{Z}/m\mathbb{Z})^\times| &= |(\mathbb{Z}/p_1^{a_1}\mathbb{Z})^\times \times \cdots \times (\mathbb{Z}/p_r^{a_r}\mathbb{Z})^\times| \\ &= |(\mathbb{Z}/p_1^{a_1}\mathbb{Z})^\times| \cdots |(\mathbb{Z}/p_r^{a_r}\mathbb{Z})^\times| = \phi(p_1^{a_1}) \cdots \phi(p_r^{a_r}). \end{aligned}$$

It remains to show that $\phi(p^a) = (p-1)p^{a-1}$ when p is prime and $a > 0$. The residue classes that are *not* coprime to p^a are represented by the multiples kp of p satisfying $0 \leq kp < p^a$. There are p^{a-1} of these, corresponding to $k \in \{0, \ldots, p^{a-1} - 1\}$, so there are $p^a - p^{a-1} = (p-1)p^{a-1}$ prime residue classes. □

Example.

$$\begin{aligned} \phi(36) &= \phi(4 \cdot 9) = \phi(4)\phi(9) = 2 \cdot (2 \cdot 3) = 12 \\ \phi(525) &= \phi(3 \cdot 5^2 \cdot 7) = \phi(3)\phi(5^2)\phi(7) = 2 \cdot (4 \cdot 5) \cdot 6 = 240 \end{aligned}$$

Notice how the formula $\phi(m) = \phi(p_1^{a_1}) \cdots \phi(p_r^{a_r})$ in Proposition 2.1 shows that $\phi(mn) = \phi(m)\phi(n)$ when m and n are coprime positive integers. This fact will be significant later when we come to study arithmetic functions.

II – 3 Definition of order and first properties

If $\alpha \in (\mathbb{Z}/m\mathbb{Z})^\times$, then there is a positive integer n such that $\alpha^n = [1]$. Indeed, because there are only finitely many prime residue classes, there are integers l_1, l_2 with $l_1 < l_2$ such that $\alpha^{l_1} = \alpha^{l_2}$. Hence,

$$[1] = \alpha^{-l_1}\alpha^{l_2} = \alpha^{l_2 - l_1} = \alpha^n \quad \text{where } n = l_2 - l_1 > 0.$$

We define the *order* of α to be the least positive integer n such that $\alpha^n = [1]$. Also, if $a \in \mathbb{Z}$ and $\gcd(a, m) = 1$, then we write

$$\mathrm{ord}_m(a) = \mathrm{ord}([a]_m).$$

Example. Let us find $\mathrm{ord}_9(2)$ by brute force, simply multiplying successively by 2 until we obtain 1 mod 9: $2^1 = 2$, $2^2 = 4$, $2^3 = 8$, $2^4 \equiv 7 \bmod 9$, $2^5 \equiv 5 \bmod 9$, $2^6 \equiv 1 \bmod 9$. The least positive integer n such that $2^n \equiv 1 \bmod 9$ is 6, so $\mathrm{ord}_9(2) = 6$.

Example. Similarly, multiplying successively by 7, we find that $\mathrm{ord}_{11}(7) = 10$: $7^1 = 7$, $7^2 = 49 \equiv 5 \bmod 11$, $7^3 \equiv 7 \cdot 5 \bmod 11 \equiv 2 \bmod 11$, and so on, until we find that $7^{10} \equiv 1 \bmod 11$, with no smaller positive n satisfying $7^n \equiv 1 \bmod 11$. Thus, $\mathrm{ord}_{11}(7) = 10$.

We will develop tools to make the calculation of orders quicker.

Proposition 3.1 (Ord-1)**.** *If $\alpha \in (\mathbb{Z}/m\mathbb{Z})^\times$ and $l \in \mathbb{Z}$, then $\alpha^l = [1]$ if and only if* $\mathrm{ord}(\alpha) \mid l$.

Proof. Let $n = \mathrm{ord}(\alpha)$, and write $l = qn + r$ where $0 \le r < n$. Then

$$\begin{aligned}
\alpha^l = [1] &\iff \alpha^{qn+r} = [1] \\
&\iff (\alpha^n)^q \alpha^r = [1] \\
&\iff [1]\alpha^r = [1] \quad \text{because } \alpha^n = [1] \\
&\iff \alpha^r = [1] \\
&\iff r = 0 \quad \text{by the minimality of } n \\
&\iff n \mid l.
\end{aligned}$$

□

Proposition 3.2 (Ord-2)**.** *Let $\alpha \in (\mathbb{Z}/m\mathbb{Z})^\times$. Then*

(i) $\alpha^{\phi(m)} = [1]$, *and*

(ii) $\mathrm{ord}(\alpha) \mid \phi(m)$.

Proof. By Proposition Ord-1, (i) and (ii) are equivalent, so it is enough to prove (i). We leave it as an exercise to show that the map

$$\begin{aligned}
(\mathbb{Z}/m\mathbb{Z})^\times &\to (\mathbb{Z}/m\mathbb{Z})^\times \\
\beta &\mapsto \alpha\beta
\end{aligned}$$

is bijective. Hence,

$$\prod_{\beta\in(\mathbb{Z}/m\mathbb{Z})^\times} \beta = \prod_{\beta\in(\mathbb{Z}/m\mathbb{Z})^\times} (\alpha\beta) = \alpha^{\phi(m)} \prod_{\beta\in(\mathbb{Z}/m\mathbb{Z})^\times} \beta,$$

the last equality because $|(\mathbb{Z}/m\mathbb{Z})^\times| = \phi(m)$. Multiplying both sides by the inverse of $\prod_{\beta\in(\mathbb{Z}/m\mathbb{Z})^\times} \beta$ leaves $[1] = \alpha^{\phi(m)}$. □

Proposition Ord-2 is often referred to by the name *Euler's Theorem.*

Example. A widely used special case of Proposition Ord-2 is Fermat's Little Theorem:

$$a^{p-1} \equiv 1 \bmod p \quad \text{when } p \text{ is prime and } p \nmid a.$$

Let us prove this. If $p \nmid a$, then $\gcd(a, p) = 1$ because p is prime. Therefore, applying Proposition Ord-2 with $m = p$ gives $[a]^{\phi(p)} = [1]$, i.e., $a^{\phi(p)} \equiv 1 \bmod p$. But $\phi(p) = |(\mathbb{Z}/p\mathbb{Z})^\times| = p - 1$, again because p is prime.

Example. Find the remainder on dividing 3^{2023} by 17.

Solution: Note that $2023 = 126 \cdot 16 + 7$, so

$$3^{2023} = 3^{126\cdot 16+7} = (3^{126})^{16} \cdot 3^7 \equiv 1 \cdot 3^7 \bmod 17 \quad \text{by Fermat's Little Theorem.}$$

To calculate $3^7 \bmod 17$, let $\equiv$ denote congruence mod 17, and observe that

$$3^7 = 3^3 \cdot 3^4 = 27 \cdot 81 \equiv (-7)(-4) = 28 \equiv 11.$$

Thus, the remainder is 11.

Example. Find the remainder on dividing $13^{1\,234\,567}$ by 36.

Solution 1: One finds that $\phi(36) = 12$, so $a^{12} \equiv 1 \bmod 36$ for all a coprime to 36. Now, division with remainder shows that $1\,234\,567 = 12q + 7$ for some $q \in \mathbb{Z}$, so

$$\begin{aligned} 13^{1\,234\,567} = 13^{12q+7} = (13^q)^{12} \cdot 13^7 &\equiv 13^7 \bmod 36 \quad \text{by Proposition Ord-2} \\ &= 62\,748\,517 \equiv 13 \bmod 36, \end{aligned}$$

so the remainder is 13.

Solution 2: Computing the first few powers of 13, we find that $\operatorname{ord}_{36}(13) = 3$. Therefore, we need only consider the remainder on dividing $1\,234\,567$ by 3, not 12. It is easy to see that the remainder is 1, i.e., $1\,234\,567 = 3q + 1$ for some $q \in \mathbb{Z}$, so

$$\begin{aligned} 13^{1\,234\,567} = 13^{3q+1} = (13^3)^q \cdot 13 &\equiv 1^q \cdot 13 \bmod 36 \\ &= 13. \end{aligned}$$

II – 4 Further results concerning order

Proposition 4.1 (Ord-3)**.** *Let $\alpha \in (\mathbb{Z}/m\mathbb{Z})^\times$, and let $n = \operatorname{ord}(\alpha)$. Then for any $k \in \mathbb{Z}$,*

$$\operatorname{ord}(\alpha^k) = \frac{n}{\gcd(n,k)}.$$

Proof. Let $d = \gcd(n,k)$, and let $n' = n/d$ and $k' = k/d$. Then for any $l \in \mathbb{Z}$,

$$\begin{aligned}(\alpha^k)^l = [1] &\iff \alpha^{kl} = [1] \iff n \mid kl \quad \text{by Proposition Ord-1} \\ &\iff n'd \mid k'dl \iff n' \mid k'l \iff n' \mid l\end{aligned}$$

because $\gcd(n',k') = 1$. Thus, $\operatorname{ord}(\alpha^k) = n' = n/d$. □

Example. In $(\mathbb{Z}/27\mathbb{Z})^\times$, $\operatorname{ord}([2]) = 18$ (exercise). Use this fact to do the following:

(i) Find $\operatorname{ord}([8])$.

(ii) Find an element $\beta \in (\mathbb{Z}/27\mathbb{Z})^\times$ of order 18 that is not equal to $[2]$.

Solution: To find $\operatorname{ord}([8])$, observe that $[8] = [2]^3$. Therefore, remembering that $\operatorname{ord}([2]) = 18$, we may apply Proposition Ord-3 as follows:

$$\operatorname{ord}([8]) = \operatorname{ord}([2]^3) = \frac{18}{\gcd(18,3)} = \frac{18}{3} = 6.$$

For the second part of the problem, let k be any integer coprime to 18, and let $\beta = [2]^k = [2^k]$. Then by Proposition Ord-3,

$$\operatorname{ord}(\beta) = \frac{18}{\gcd(18,k)} = \frac{18}{1} = 18.$$

For example, we could take $k = 5$, and then $\beta = [2^5] = [32] = [5]$.

Proposition 4.2 (Ord-4)**.** *If m and n are positive integers with $m \mid n$, then for any $a \in \mathbb{Z}$ coprime to n, $\operatorname{ord}_m(a) \mid \operatorname{ord}_n(a)$.*

Proof. Let $k = \operatorname{ord}_n(a)$. By definition, $a^k \equiv 1 \bmod n$, so $a^k \equiv 1 \bmod m$ because $m \mid n$, and so $\operatorname{ord}_m(a) \mid k$ by Proposition Ord-1. □

Example. Consider $m = 7$ and $n = 35$, and note that $m \mid n$. We tabulate $\operatorname{ord}_{35}(a)$ and $\operatorname{ord}_7(a)$ for the first few positive integers a coprime to 35:

a coprime to 35	$\operatorname{ord}_{35}(a)$	$\operatorname{ord}_7(a)$
1	1	1
2	12	3
3	12	6
4	6	3
6	2	2
8	4	1
⋮	⋮	⋮

Proposition 4.3 (Ord-5)**.** *Suppose that $m_1, \ldots, m_r$ are pairwise-coprime positive integers, and let $a \in \mathbb{Z}$ be coprime to $m = m_1 \cdots m_r$. Then*

$$\mathrm{ord}_m(a) = \mathrm{lcm}(\mathrm{ord}_{m_1}(a), \ldots, \mathrm{ord}_{m_r}(a)).$$

Proof. We use the fact that, because the m_i are pairwise coprime, a given integer is divisible by all of $m_1, \ldots, m_r$ if and only if it is divisible by their product. This is a consequence of the G.C.D. Theorem.

Now, let $l \in \mathbb{Z}$. Then by the fact just mentioned,

$$\begin{aligned} a^l \equiv 1 \bmod m &\iff a^l \equiv 1 \bmod m_i \quad \text{for all } i \\ &\iff \mathrm{ord}_{m_i}(a) \mid l \quad \text{for all } i \text{ by Proposition Ord-1} \\ &\iff \mathrm{lcm}(\mathrm{ord}_{m_1}(a), \ldots, \mathrm{ord}_{m_r}(a)) \text{ divides } l. \end{aligned}$$

□

Example. Find $\mathrm{ord}_{105}(17)$ by considering 17 mod $3, 5, 7$.

Solution: Note that $105 = 3 \cdot 5 \cdot 7$. The following orders are easily computed:

$$\begin{aligned} \mathrm{ord}_3(17) &= \mathrm{ord}_3(2) = 2 \\ \mathrm{ord}_5(17) &= \mathrm{ord}_5(2) = 4 \\ \mathrm{ord}_7(17) &= \mathrm{ord}_7(3) = 6 \end{aligned}$$

Hence, because $3, 5, 7$ are pairwise coprime, $\mathrm{ord}_{105}(17) = \mathrm{lcm}(2, 4, 6) = 12$.

Example. Given that $\mathrm{ord}_9(2) = 6$ and $\mathrm{ord}_{22}(7) = 10$, find an integer a coprime to $198 = 9 \cdot 22$ such that $\mathrm{ord}_{198}(a) = 30$.

Solution: Suppose we can find $a \in \mathbb{Z}$ satisfying

$$\left.\begin{aligned} a &\equiv 2 \bmod 9 \\ a &\equiv 7 \bmod 22 \end{aligned}\right\} \tag{4.1}$$

Then because $\gcd(9, 22) = 1$,

$$\mathrm{ord}_{198}(a) = \mathrm{lcm}(\mathrm{ord}_9(a), \mathrm{ord}_{22}(a)) = \mathrm{lcm}(\mathrm{ord}_9(2), \mathrm{ord}_{22}(7)) = \mathrm{lcm}(6, 10) = 30,$$

and we will be done. We use the Chinese Remainder Theorem to solve (4.1). The Euclidean algorithm proceeds as follows: $22 = 2 \cdot 9 + 4$, $9 = 2 \cdot 4 + 1$, so

$$1 = 9 - 2 \cdot 4 = 9 - 2(22 - 2 \cdot 9) = 5 \cdot 9 - 2 \cdot 22.$$

Hence, the solution to the system of congruences is

$$\begin{aligned} a &\equiv 2(-2) \cdot 22 + 7 \cdot 5 \cdot 9 \bmod 198 \\ &= 227 \\ &\equiv 29 \bmod 198. \end{aligned}$$

Thus, $\mathrm{ord}_{198}(29) = 30$.

II – 5 Wilson's Theorem

Theorem 5.1 (Wilson's Theorem)**.** *If p is prime, then $(p-1)! \equiv -1 \bmod p$.*

Proof. Several proofs exist. We give one that uses Fermat's Little Theorem and some basic theory of polynomial rings. Assume that p is odd, the case $p = 2$ being immediate. Consider the polynomial $f(x) = x^{p-1} - [1] \in (\mathbb{Z}/p\mathbb{Z})[x]$, viewed as a polynomial with coefficients in the field $\mathbb{F}_p = \mathbb{Z}/p\mathbb{Z}$. By Fermat's Little Theorem,

$$f(\alpha) = \alpha^{p-1} - [1] = [0] \quad \text{for all } \alpha \in (\mathbb{Z}/p\mathbb{Z})^\times,$$

so because $f(x)$ has degree $p-1$ and there are exactly $p-1$ prime residue classes α, it follows that

$$f(x) = \prod_{\alpha \in (\mathbb{Z}/p\mathbb{Z})^\times} (x - \alpha) = \prod_{a=1}^{p-1} (x - [a]),$$

$$\text{i.e.,} \quad x^{p-1} - [1] = \prod_{a=1}^{p-1} (x - [a]).$$

Taking $x = [0]$ gives

$$\begin{aligned} [-1] = \prod_{a=1}^{p-1} (-[a]) &= \prod_{a=1}^{p-1} [a] \quad \text{because } p-1 \text{ is even} \\ &= [(p-1)!]. \end{aligned}$$

□

Example. Find the remainder on dividing 40! by $1763 = 41 \cdot 43$.

Solution: Observe that 41 and 43 are prime. Therefore, by Wilson's Theorem, $40! \equiv -1 \bmod 41$, and

$$\begin{aligned} -1 &\equiv 42! && \bmod 43 \\ &= 42 \cdot 41 \cdot 40! \\ &\equiv (-1)(-2) \cdot 40! && \bmod 43 \\ &= 2 \cdot 40!, \end{aligned}$$

so inverting 2 mod 43 we obtain $-22 \equiv 40! \bmod 43$. Now we use the Chinese Remainder Theorem to solve the system

$$\begin{aligned} x &\equiv -1 && \bmod 41 \\ x &\equiv -22 && \bmod 43 \end{aligned}$$

We have $1 = 21 \cdot 41 - 20 \cdot 43$ (by the Euclidean algorithm, for example), so

$$\begin{aligned} 40! &\equiv (-1)(-2) \cdot 43 - 22 \cdot 21 \cdot 41 && \bmod 41 \cdot 43 \\ &= -18\,082, \\ \text{i.e.,} \quad 40! &\equiv 1311 && \bmod 1763. \end{aligned}$$

II – 6 Quadratic residues

Let p be a prime. An integer a is called a *quadratic residue* mod p if there is $b \in \mathbb{Z}$ such that $a \equiv b^2 \bmod p$, i.e., if $[a]$ is the square of some element of $\mathbb{Z}/p\mathbb{Z}$.

Example. In the case $p = 5$,

$$\begin{aligned} 0^2 &= 0 \\ 1^2 &= 1 \\ 2^2 &= 4 \\ 3^2 &\equiv 4 \bmod 5 \\ 4^2 &\equiv 1 \bmod 5, \end{aligned}$$

so the squares mod 5 are $0, 1, 4$.

Especially, we will be interested in the *prime* quadratic residues, i.e., the quadratic residues that represent prime residue classes, or, put even more simply, the quadratic residues not divisible by p.

Proposition 6.1. *Let p be an odd prime. Then there are exactly $(p-1)/2$ prime quadratic residues mod p. Thus, half of the prime residue classes are quadratic residues.*

Proof. Since p is odd, $\alpha \neq -\alpha$ when α is a prime residue class, so $(\mathbb{Z}/p\mathbb{Z})^\times$ can be partitioned into $(p-1)/2$ pairs $\{\alpha, -\alpha\}$. Let S be the set of such pairs. Then the map

$$\begin{aligned} S &\to (\mathbb{Z}/p\mathbb{Z})^\times \\ \{\alpha, -\alpha\} &\mapsto \alpha^2 \end{aligned}$$

is injective (exercise), and its image is the set of quadratic prime residue classes, by definition. Therefore, the number of prime quadratic residues mod p is the cardinality of S, which is $(p-1)/2$. □

The Legendre Symbol

Let p be an odd prime, and let $a \in \mathbb{Z}$. Then we define the *Legendre symbol* $\left(\frac{a}{p}\right)$ by

$$\left(\frac{a}{p}\right) = \begin{cases} 1 & \text{if } a \text{ is coprime to } p \text{ and is a quadratic residue,} \\ -1 & \text{if } a \text{ is coprime to } p \text{ but is not a quadratic residue,} \\ 0 & \text{otherwise, i.e., if } p \mid a. \end{cases}$$

It satisfies the following:

(i) $\left(\frac{a}{p}\right)$ depends only on the class of a mod p, i.e., $\left(\frac{a+kp}{p}\right) = \left(\frac{a}{p}\right)$ for all $k \in \mathbb{Z}$.

(ii) $\left(\frac{ab}{p}\right) = \left(\frac{a}{p}\right)\left(\frac{b}{p}\right)$ for all $a, b \in \mathbb{Z}$ (exercise).

Proposition 6.2. *If p is an odd prime and a an integer, then*

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

Proof. If $p \mid a$, then the statement is obvious, so assume that $p \nmid a$.

Case (i): $\left(\frac{a}{p}\right) = 1$. In this case, $a \equiv b^2 \bmod p$ for some $b \in \mathbb{Z}$, so mod p we have

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

Case (ii): $\left(\frac{a}{p}\right) = -1$. In this case, if $\beta \in (\mathbb{Z}/p\mathbb{Z})^\times$, then the unique $\beta' \in (\mathbb{Z}/p\mathbb{Z})^\times$ such that $\beta\beta' = [a]$ cannot be equal to β, because $[a]$ is not square in $(\mathbb{Z}/p\mathbb{Z})^\times$, so $(\mathbb{Z}/p\mathbb{Z})^\times$ can be partitioned into $(p-1)/2$ pairs $\{\beta, \beta'\}$ satisfying $\beta\beta' = [a]$ with $\beta \neq \beta'$. Hence,

$$[a]^{\frac{p-1}{2}} = \prod_{\{\beta,\beta'\}} (\beta\beta') = \prod_{b=1}^{p-1} [b] = [(p-1)!] = [-1]$$

by Wilson's Theorem. Thus, mod p we have

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

□

Example. Consider $p = 29$, $a = 3$. Here,

$$a^{\frac{p-1}{2}} = 3^{14} = 4\,782\,969 \equiv -1 \bmod 29 \quad \text{(direct calculation)},$$

so $\left(\frac{3}{29}\right) = -1$, i.e., 3 is not square mod 29.

Example. Consider now $p = 29$, $a = 5$. This time,

$$a^{\frac{p-1}{2}} = 5^{14} = 6\,103\,515\,625 \equiv 1 \bmod 29 \quad \text{(direct calculation)},$$

so $\left(\frac{5}{29}\right) = 1$, i.e., 5 is square mod 29.

II – 7 Relationship to quadratic congruences

Proposition 7.1. *Let p be an odd prime, and let $a, b, c \in \mathbb{Z}$ with $p \nmid a$. Then the congruence $ax^2 + bx + c \equiv 0 \bmod p$ has*

$$\begin{cases} 2 \text{ solutions mod } p & \text{if } \left(\frac{b^2-4ac}{p}\right) = 1, \\ \text{no solutions} & \text{if } \left(\frac{b^2-4ac}{p}\right) = -1, \\ 1 \text{ solution mod } p & \text{if } \left(\frac{b^2-4ac}{p}\right) = 0. \end{cases}$$

The solutions, if any exist, are given by

$$[x]_p = [2a]_p^{-1}(-[b]_p \pm [\sqrt{b^2 - 4ac}]_p),$$

where $\sqrt{b^2 - 4ac}$ denotes any integer whose square is congruent to $b^2 - 4ac \bmod p$.

Proof. Let $\alpha = [a], \beta = [b], \gamma = [c]$ in $\mathbb{Z}/p\mathbb{Z}$. By assumption, $\alpha \neq [0]$. Then for $X \in \mathbb{Z}/p\mathbb{Z}$,

$$\begin{aligned} \alpha X^2 + \beta X + \gamma = [0] &\iff X^2 + \frac{\beta}{\alpha} X + \frac{\gamma}{\alpha} = [0] \\ &\iff \left(X + \frac{\beta}{2\alpha}\right)^2 - \frac{\beta^2}{4\alpha^2} + \frac{\gamma}{\alpha} = [0] \\ &\iff \left(X + \frac{\beta}{2\alpha}\right)^2 = \frac{\beta^2}{4\alpha^2} - \frac{\gamma}{\alpha} = \frac{1}{4\alpha^2}(\beta^2 - 4\alpha\gamma). \end{aligned} \tag{7.1}$$

This has a solution X if and only if $\beta^2 - 4\alpha\gamma$ is square in $\mathbb{Z}/p\mathbb{Z}$, specifically, two solutions mod p when $\beta^2 - 4\alpha\gamma$ is the square of a *non-zero* residue class, and one solution when it is zero. Solving for X in (7.1) shows that the solutions are as claimed. □

Example. The discriminant of the polynomial $x^2 + x + 1$ is $1^2 - 4 \cdot 1 \cdot 1 = -3$, so if p is odd, the congruence $x^2 + x + 1 \equiv 0 \bmod p$ has a solution if and only if $\left(\frac{-3}{p}\right) \in \{0, 1\}$. Obviously, $\left(\frac{-3}{p}\right) = 0$ if and only if $p = 3$, so there is only one solution mod p in that case. We will see shortly how to determine for which primes p the Legendre symbol $\left(\frac{-3}{p}\right)$ takes the value 1. For now, simply verify directly that $\left(\frac{-3}{p}\right) = 1$ if $p \in \{7, 13, 19\}$ (the congruence then having two solutions mod p), and that $\left(\frac{-3}{p}\right) = -1$ if $p \in \{5, 11, 17\}$ (no solutions to the congruence).

Remark. The case $p = 2$ is excluded from Proposition 7.1, but this case is easily worked out. If $x \in \mathbb{Z}$, then $x^2 \equiv x \bmod 2$, so

$$\begin{aligned} x^2 + bx + c &\equiv x + bx + c \bmod 2 \\ &= (b + 1)x + c. \end{aligned}$$

Therefore, the solutions to the congruence $x^2 + bx + c \equiv 0 \bmod 2$ are as follows:

- If b is odd and c is even, every integer x is a solution.
- If b and c are both odd, there are no solutions.
- If b is even, there is only the solution $x \equiv c \bmod 2$.

II – 8 Quadratic reciprocity

Theorem 8.1 (Quadratic reciprocity)**.** *If p and q are distinct odd primes, then*

$$\left(\frac{p}{q}\right)\left(\frac{q}{p}\right) = (-1)^{\frac{p-1}{2}\frac{q-1}{2}}, \quad \textit{i.e.,} \quad \left(\frac{p}{q}\right) = (-1)^{\frac{p-1}{2}\frac{q-1}{2}}\left(\frac{q}{p}\right).$$

Additionally,

$$\left(\frac{2}{p}\right) = (-1)^{\frac{p^2-1}{8}} \quad \textit{and} \quad \left(\frac{-1}{p}\right) = (-1)^{\frac{p-1}{2}}.$$

For a proof, see Section 2 of the Appendix.

Remark. The determination of the signs, i.e., $(-1)^{?}$, in the theorem can be sped up via the following observations:

- $\frac{p-1}{2}\frac{q-1}{2}$ is even if and only if at least one of p or q is congruent to 1 mod 4.
- $\frac{p^2-1}{8}$ is even if and only if $p \equiv \pm 1 \bmod 8$.
- $\frac{p-1}{2}$ is even if and only if $p \equiv 1 \bmod 4$.

Example. Find $\left(\frac{3}{29}\right)$ and $\left(\frac{5}{29}\right)$ using quadratic reciprocity.

Solution:

$$\begin{aligned}\left(\frac{3}{29}\right) &= \left(\frac{29}{3}\right) \quad (29 \equiv 1 \bmod 4)\\ &= \left(\frac{2}{3}\right) \quad (29 \equiv 2 \bmod 3)\\ &= -1 \quad (3 \not\equiv \pm 1 \bmod 8),\end{aligned}$$

and

$$\begin{aligned}\left(\frac{5}{29}\right) &= \left(\frac{29}{5}\right) \quad (29 \equiv 1 \bmod 4)\\ &= \left(\frac{4}{5}\right) = \left(\frac{2^2}{5}\right) = \left(\frac{2}{5}\right)^2 = 1.\end{aligned}$$

Example. Find $\left(\frac{7}{23}\right)$ using quadratic reciprocity.

Solution:

$$\begin{aligned}\left(\frac{7}{23}\right) &= -\left(\frac{23}{7}\right) \quad (7 \equiv 23 \equiv -1 \bmod 4)\\ &= -\left(\frac{2}{7}\right)\\ &= -(1) \quad (7 \equiv -1 \bmod 8)\\ &= -1.\end{aligned}$$

Example. Find $\left(\frac{302}{541}\right)$ using quadratic reciprocity.

Solution:

$$\begin{aligned}
\left(\frac{302}{541}\right) &= \left(\frac{2}{541}\right)\left(\frac{151}{541}\right) \quad \text{(factorize 302)} \\
&= (-1)\left(\frac{541}{151}\right) \quad (541 \equiv 5 \bmod 8,\ 541 \equiv 1 \bmod 4) \\
&= -\left(\frac{88}{541}\right) = -\left(\frac{8}{151}\right)\left(\frac{11}{151}\right) = -\left(\frac{2}{151}\right)^3\left(\frac{11}{151}\right) = -\left(\frac{2}{151}\right)\left(\frac{11}{151}\right) \\
&= -(1)\cdot(-1)\left(\frac{151}{11}\right) \quad (151 \equiv 7 \bmod 8,\ 11 \equiv 151 \equiv -1 \bmod 4) \\
&= \left(\frac{8}{11}\right) = \left(\frac{2}{11}\right)^3 = \left(\frac{2}{11}\right) = -1 \quad (11 \equiv 3 \bmod 8).
\end{aligned}$$

Example. For which primes p does the congruence $x^2+6x+2 \equiv 0 \bmod p$ have (a) two distinct solutions mod p, (b) a unique solution mod p, and (c) no solutions?

Solution: For $p = 2$, the congruence becomes $x^2 \equiv 0 \bmod 2$, which has the unique solution $x \equiv 0 \bmod 2$. Assume, henceforth, that $p > 2$. The discriminant of the given quadratic polynomial is $\Delta = 36 - 4\cdot 1\cdot 2 = 28 = 4\cdot 7$. If $p = 7$, then $p \mid \Delta$, so there is a unique solution mod 7.

Now assume that $p \notin \{2, 7\}$. Then

$$\left(\frac{\Delta}{p}\right) = \left(\frac{4\cdot 7}{p}\right) = \left(\frac{2}{p}\right)^2\left(\frac{7}{p}\right) = \left(\frac{7}{p}\right) = (-1)^{\frac{7-1}{2}\frac{p-1}{2}}\left(\frac{p}{7}\right) = (-1)^{\frac{p-1}{2}}\left(\frac{p}{7}\right).$$

Let $A = (-1)^{\frac{p-1}{2}}$ and $B = \left(\frac{p}{7}\right)$, and note that $\left(\frac{\Delta}{p}\right) = 1$ if and only if $A = B = 1$ or $A = B = -1$. We consider these two cases separately.

(i) $A = B = 1$ if and only if $p \equiv 1 \bmod 4$ and $p \equiv 1$, 2, or 4 mod 7. Now use the Chinese Remainder Theorem:

$$\begin{aligned}
a \equiv 1 \bmod 4,\ a \equiv 1 \bmod 7 &\iff a \equiv 1 \bmod 28 \\
a \equiv 1 \bmod 4,\ a \equiv 2 \bmod 7 &\iff a \equiv 9 \bmod 28 \\
a \equiv 1 \bmod 4,\ a \equiv 4 \bmod 7 &\iff a \equiv 25 \bmod 28
\end{aligned}$$

(ii) $A = B = -1$ if and only if $p \equiv 3 \bmod 4$ and $p \equiv 3$, 5, or 6 mod 7. Use the Chinese Remainder Theorem again:

$$\begin{aligned}
a \equiv 3 \bmod 4,\ a \equiv 3 \bmod 7 &\iff a \equiv 3 \bmod 28 \\
a \equiv 3 \bmod 4,\ a \equiv 5 \bmod 7 &\iff a \equiv 19 \bmod 28 \\
a \equiv 3 \bmod 4,\ a \equiv 6 \bmod 7 &\iff a \equiv 27 \bmod 28
\end{aligned}$$

In summary, the congruence $x^2+6x+2 \equiv 0 \bmod p$ has (a) two distinct solutions mod p if $p \equiv 1$, 3, 9, 19, 25, or 27 mod 28, has (b) a unique solution mod p if $p \in \{2, 7\}$, and has (c) no solutions otherwise, i.e., if $p \equiv 5$, 11, 13, 15, 17, or 23 mod 28.

II – 9 Primitive roots

Let m be a positive integer. A *primitive root* mod m is an integer a coprime to m such that every prime residue class mod m is a power of $[a]$.

Example. The integer 2 is a primitive root mod 9 because each of the six prime residue classes mod 9, namely, $[1], [2], [4], [5], [7], [8]$, is a power of $[2]$:

$$[2]^0 = [1], \quad [2]^1 = [2], \quad [2]^2 = [4], \quad [2]^3 = [8], \quad [2]^4 = [7], \quad [2]^5 = [5].$$

Proposition 9.1. *If $[a] \in (\mathbb{Z}/m\mathbb{Z})^\times$, then a is primitive mod m if and only if $\operatorname{ord}_m(a) = \phi(m)$.*

Proof. If $\operatorname{ord}_m(a) = \phi(m)$, then all $\phi(m)$ powers $[a]^0, [a]^1, [a]^2, \ldots, [a]^{\phi(m)-1}$ are distinct, so since there are $\phi(m)$ of them, they must consitute all the elements of $(\mathbb{Z}/m\mathbb{Z})^\times$.

Conversely, if every prime residue class is a power of $[a]$, then there can be no repetitions among $[a]^0, [a]^1, [a]^2, \ldots, [a]^{\phi(m)-1}$, so none of these is equal to $[1]$ except the zeroth power, and so the first positive k with $[a]^k = [1]$ is $k = \phi(m)$. □

Example. In $(\mathbb{Z}/12\mathbb{Z})^\times$, each element has order dividing 2 (check for yourself), so there is no prime residue class of order $\phi(12) = 4$, and so there is no primitive root mod 12.

Proposition 9.2. *Assume that the modulus m has a primitive root, suppose that $b \in \mathbb{Z}$ is coprime to m, and let n be a positive integer. Then the congruence*

$$x^n \equiv b \bmod m$$

has a solution if and only if $b^{\phi(m)/d} \equiv 1 \bmod m$, where $d = \gcd(\phi(m), n)$. In this case, there are d solutions mod m.

Proof. Let $\alpha = [a]$ be a primitive root mod m, let $\beta = [b]$, and write $\beta = \alpha^l$ with $l \in \mathbb{Z}$. Observe that the congruence is equivalent to the equation

$$[x]^n = \beta \tag{9.1}$$

in $\mathbb{Z}/m\mathbb{Z}$. Further, if (9.1) has a solution x, then x must be coprime to m, because $[x]^{n-1}\beta^{-1}$ is a multiplicative inverse to $[x]$ in $\mathbb{Z}/m\mathbb{Z}$. Therefore, we can replace $[x]$ in (9.1) with α^k for an integer $k \in \{0, \ldots, \phi(m)-1\}$. But $(\alpha^k)^n = \beta$ if and only if $\alpha^{nk} = \alpha^l$, if and only if

$$nk \equiv l \bmod \phi(m), \tag{9.2}$$

and we know from Proposition 1.1 that the congruence (9.2) has a solution if and only if

$$\begin{aligned} & \gcd(\phi(m), n) \mid l \\ \iff & \gcd(\phi(m), n) \mid \gcd(\phi(m), l) \\ \iff & \frac{\phi(m)}{\gcd(\phi(m), l)} \,\Big|\, \frac{\phi(m)}{\gcd(\phi(m), n)} \end{aligned}$$

$$\iff \operatorname{ord}_m(b) \,\Big|\, \frac{\phi(m)}{\gcd(\phi(m), n)} \quad \text{by Proposition Ord-3}$$
$$\iff b^{\phi(m)/d} \equiv 1 \bmod m \quad \text{by Proposition Ord-1,}$$

where $d = \gcd(\phi(m), n)$. Further, if (9.2) has a solution, the number of solutions $k \in \{0, \ldots, \phi(m) - 1\}$ is d by Proposition 1.1. □

Example. Solve $x^{21} \equiv 8 \bmod 27$ using the fact that 2 is a primitive root mod 27.

Solution: Note that $\gcd(\phi(27), 21) = \gcd(18, 21) = 3$, and

$$8^{18/3} = (2^3)^{18/3} = 2^{18} \equiv 1 \bmod 27 \quad \text{by Proposition Ord-2,}$$

so there are solutions. In fact, there are 3 solutions mod 27. Let us find them.

The congruence is equivalent to $[x]^{21} = [8]$ in $\mathbb{Z}/27\mathbb{Z}$. Let $[x] = [2]^k$ with $k \in \{0, \ldots, 17\}$. (The congruence implies that $[x]$ is a prime residue class, as explained in the proof of Proposition 9.2, and therefore is a power of [2].) Then

$$\begin{aligned}
[x]^{21} = [8] &\iff [2]^{21k} = [2]^3 \\
&\iff [2]^{21k-3} = [1] \\
&\iff \operatorname{ord}_{27}(2) \mid 21k - 3 \quad \text{by Proposition Ord-1} \\
&\iff 18 \mid 21k - 3 \\
&\iff 21k \equiv 3 \bmod 18 \\
&\iff 7k \equiv 1 \bmod 6 \\
&\iff k \equiv 1 \bmod 6 \\
&\iff k \equiv 1,\ 7, \text{ or } 13 \bmod 18 \\
&\iff [2]^k = [2]^1,\ [2]^7, \text{ or } [2]^{13}.
\end{aligned}$$

Remembering, then, that $[2]^k = [x]$, we see that x is a solution to the original congruence if and only if

$$\begin{aligned}
[x] &= [2]^1,\ [2]^7, \text{ or } [2]^{13} \\
&= [2],\ [2]^6 \cdot [2], \text{ or } [2]^{12} \cdot [2] \\
&= [2],\ [10] \cdot [2], \text{ or } [10]^2 \cdot [2] \quad \text{because } [2]^6 = [64] = [10] \\
&= [2],\ [20], \text{ or } [11] \quad \text{(multiply by [10] each time).}
\end{aligned}$$

Thus, the solutions to $x^{21} \equiv 8 \bmod 27$ are $x \equiv 2,\ 11$, or $20 \bmod 27$.

II – 10 Primitive roots continued

We study another congruence that can be handled via primitive roots, if one exists.

Proposition 10.1. *Assume that m has a primitive root, and let $b, c \in \mathbb{Z}$ be coprime to m. Then the congruence*

$$b^x \equiv c \bmod m$$

has a solution $x \in \mathbb{Z}_{\geq 0}$ if and only if $\operatorname{ord}_m(c) \mid \operatorname{ord}_m(b)$. In this case, there is a unique solution $x \in \{0, \dots, \operatorname{ord}_m(b) - 1\}$.

Proof. Let $\alpha \in (\mathbb{Z}/m\mathbb{Z})^\times$ be primitive, let $\beta = [b]$ and $\gamma = [c]$, and write

$$\beta = \alpha^k, \quad \gamma = \alpha^l \quad \text{with } k, l \in \mathbb{Z}.$$

Then

$$\begin{aligned}
\beta^x = \gamma \text{ has a solution} &\iff \alpha^{kx} = \alpha^l \text{ has a solution} \\
&\iff \alpha^{kx-l} = [1] \text{ has a solution} \\
&\iff kx \equiv l \bmod \operatorname{ord}(\alpha) \text{ has a solution} \\
&\iff \gcd(\phi(m), k) \mid l \quad \text{by Proposition 1.1} \\
&\iff \gcd(\phi(m), k) \mid \gcd(\phi(m), l) \quad \text{(exercise)} \\
&\iff \frac{\phi(m)}{\gcd(\phi(m), l)} \;\Big|\; \frac{\phi(m)}{\gcd(\phi(m), l)} \\
&\iff \operatorname{ord}_m(c) \mid \operatorname{ord}_m(b) \quad \text{by Proposition Ord-3.}
\end{aligned}$$

The uniqueness of a solution $x \in \{0, \dots, \operatorname{ord}(b) - 1\}$ follows from the observation that $\beta^x = \beta^y$ if and only if $x \equiv y \bmod \operatorname{ord}(\beta)$. □

Example. Using the fact that 17 has a primitive root, decide whether the congruence $9^x \equiv 4 \bmod 17$ has a solution. If so, find the least non-negative solution.

Solution: Observe that $4^2 \equiv -1 \bmod 17$, so $4^4 \equiv 1 \bmod 17$ and $\operatorname{ord}_{17}(4) = 4$. Now,

$$\begin{aligned}
&9^2 = 81 \equiv -4 \bmod 17, \\
\text{so} \quad &9^4 \equiv (-4)^2 \equiv -1 \bmod 17, \\
\text{and} \quad &9^8 \equiv (-1)^2 \equiv 1 \bmod 17.
\end{aligned}$$

Thus, $\operatorname{ord}_{17}(9) \mid 8$, but the foregoing calculations rule out $1, 2, 4$ as possibilities for the order, so in fact $\operatorname{ord}_{17}(9) = 8$. Therefore, $\operatorname{ord}_{17}(4) \mid \operatorname{ord}_{17}(9)$, so by the proposition, the congruence $9^x \equiv 4 \bmod 17$ has a solution.

To find the least non-negative solution, observe from the congruences $9^2 \equiv -4 \bmod 17$ and $9^4 \equiv -1 \bmod 17$ found above that $9^6 \equiv 4 \bmod 17$, so $x = 6$ is a solution. Since $6 < 8 = \operatorname{ord}_{17}(9)$, there can be no smaller solution than this.

Criterion for the existence of a primitive root

Theorem 10.2. *Let $m \geq 2$ be an integer.*

(i) There is a primitive root mod m if and only if either of the following holds:

(a) m is equal to 2 or 4.

(b) m is a power of an odd prime or twice such a power.

(ii) If a is a primitive root mod p, where p is an odd prime, then either a or $a + p$ is a primitive root mod p^2.

(iii) If b is primitive mod p^2, where p is an odd prime, then b is primitive mod p^j for all $j \geq 1$.

A proof is given in Section 3 of the Appendix.

Example. Show that 2 is primitive mod 5^k for all positive integers k.

Solution: By Theorem 10.2, it is enough to show that 5 is primitive mod $5^2 = 25$. We know by Proposition Ord-2 that $\mathrm{ord}_{25}(2) \,|\, \phi(25) = 20$. Now, because $20 = 2^2 \cdot 5$, every positive divisor of 20 less than 20 divides either $2 \cdot 5 = 10$ or $2^2 \cdot 5^0 = 4$. But

$$2^{10} = 1024 \not\equiv 1 \bmod 25,$$
$$\text{and} \quad 2^4 = 16 \not\equiv 1 \bmod 25,$$

so if $d \,|\, 20$ and $1 \leq d < 20$, then $2^d \not\equiv 1 \bmod 25$. Thus, $\mathrm{ord}_{25}(2) = 20$.

Example. It is a fact that 18 is primitive mod 37 but not mod 37^2. (Verify these assertions for yourself.) Therefore, by Theorem 10.2, a primitive root mod 37^2 is $18 + 37 = 55$, and then this is a primitive root mod 37^k for all $k \geq 1$.

II – 11 Polynomial congruences and Hensel's Lemma

We now make a systematic study of congruences of the form $f(x) \equiv 0 \bmod m$, where $f(x) \in \mathbb{Z}[x]$, i.e., $f(x)$ is a polynomial with integer coefficients, beginning with the case where m is a power of a prime p. The key idea is to use already-known solutions to the congruence $f(x) \equiv 0 \bmod p^k$ to find solutions to $f(x) \equiv 0 \bmod p^{k+1}$. This process is known as *lifting*.

Example. Consider the congruence $x^2 + 1 \equiv 0 \bmod 5$, which has the solutions $x \equiv 2$ or $3 \bmod 5$. Let us focus on the solution $x \equiv 2 \bmod 5$, for example. Note that 2 is not a solution mod 25. However, we may hope that $2 + 5t$ is a solution mod 25 for some $t \in \mathbb{Z}$. In fact, $t = 1$ works in this case:

$$(2 + 5)^2 + 1 = 7^2 + 1 = 50 \equiv 0 \bmod 25.$$

Thus, $2 + 5 = 7$ is a *lift* of the mod 5 solution 2 to a mod 5^2 solution.

Similarly, $3 + 3 \cdot 5 = 18$ is a lift of the mod 5 solution 3 to a mod 5^2 solution: $18^2 + 1 = 325 \equiv 0 \bmod 25$.

Lemma 11.1. *If $f(x) \in \mathbb{Z}[x]$ and $a \in \mathbb{Z}$, then there is $g(x) \in \mathbb{Z}[x]$ such that*

$$f(x) = f(a) + f'(a)(x - a) + (x - a)^2 g(x).$$

Proof. Let $F(x) = f(x) - f(a) - f'(a)(x - a)$, and note that $F(a) = 0$ and $F'(a) = 0$. Because $F(a) = 0$ and $x - a$ is monic, polynomial division shows that $F(x) = (x - a)h(x)$ where $h(x) \in \mathbb{Z}[x]$. Then

$$F'(x) = h(x) + (x - a)h'(x),$$

so $h(a) = F'(0) = 0$, and so polynomial division used again, this time on $h(x)$, gives $h(x) = (x - a)g(x)$ for some $g(x) \in \mathbb{Z}[x]$. Thus, $F(x) = (x - a)^2 g(x)$. □

Theorem 11.2 (Hensel's Lemma)**.** *Let $f(x) \in \mathbb{Z}[x]$, let p be a prime, let $k \geq 1$, and suppose that $a \in \mathbb{Z}$ is a solution to $f(x) \equiv 0 \bmod p^k$. Consider the congruence*

$$f(x) \equiv 0 \bmod p^{k+1}. \tag{11.1}$$

(i) If $p \nmid f'(a)$, then there is a unique solution to (11.1) of the form $a + tp^k$ with $0 \leq t \leq p - 1$. It is found by solving

$$\frac{f(a)}{p^k} + f'(a)t \equiv 0 \bmod p \quad \text{for } t \in \{0, \ldots, p - 1\}.$$

(ii) If $p \mid f'(a)$ and a is a solution to (11.1), then $a + tp^k$ is a solution for all t.

(iii) If $p \mid f'(a)$ and a is not a solution to (11.1), then (11.1) has no solution of the form $a + tp^k$.

Proof. Let $t \in \mathbb{Z}$. By Lemma 11.1,

$$\begin{aligned} f(a+tp^k) &= f(a) + f'(a)tp^k + t^2p^{2k}c \quad \text{where } c = g(a+tp^k) \in \mathbb{Z} \\ &\equiv f(a) + f'(a)tp^k \bmod p^{k+1}, \end{aligned}$$

so

$$\begin{aligned} f(a+tp^k) \equiv 0 \bmod p^{k+1} &\iff f(a) + f'(a)tp^k \equiv 0 \bmod p^{k+1} \\ &\iff \frac{f(a)}{p^k} + f'(a)t \equiv 0 \bmod p. \end{aligned}$$

If $f'(a) \not\equiv 0 \bmod p$, there is a unique solution $t \in \{0, \ldots, p-1\}$, found by inverting $f'(a) \bmod p$. If $f'(a) \equiv 0 \bmod p$, then any t works as long as $f(a)/p^k \equiv 0 \bmod p$, i.e., $f(a) \equiv 0 \bmod p^{k+1}$. Otherwise, no t works. □

Example. Let $f(x) = x^3 + 2x + 3$ and $p = 5$. Observe that the congruence $f(x) \equiv 0 \bmod 5$ has the solution $x \equiv 2 \bmod 5$: $f(2) = 15 \equiv 0 \bmod 5$. Let us apply Hensel's Lemma in the case $p = 5$ and $k = 1$ to see whether this solution can be lifted to a solution mod 5^2:

$$f'(x) = 3x^2 + 2, \quad \text{so} \quad f'(2) = 14 \not\equiv 0 \bmod 5.$$

This is case (i) of the theorem, so there is a unique lift $2 + 5t$ with $0 \le t \le 4$, found by solving

$$\begin{aligned} \frac{f(2)}{5} + f'(2)t &\equiv 0 \bmod 5, \\ \text{i.e.,} \quad 3 + 14t &\equiv 0 \bmod 5, \\ \text{i.e.,} \quad t &= 3. \end{aligned}$$

Thus, $2 + 3 \cdot 5 = 17$ is a solution mod 5^2, and is in fact the unique one lifting the mod 5 solution 2.

Example. Let $f(x) = x^4 + 5x^2 + 337$ and $p = 7$, and note that $x \equiv 1 \bmod 7$ is a solution to the congruence $f(x) \equiv 0 \bmod 7$, since $f(1) = 343 = 7^3$. Now, $f'(x) = 4x^3 + 10x$, so $f'(1) = 14 \equiv 0 \bmod 7$. The additional fact that $f(1) = 7^3 \equiv 0 \bmod 7^2$ puts us in case (ii) of Hensel's Lemma rather than case (iii), so every $1 + 7t$ with $t \in \mathbb{Z}$ is a solution to the congruence $f(x) \equiv 0 \bmod 7^2$.

Example. Let $f(x) = x^3 - x^2 + 4x + 2$ and $p = 11$. The congruence $f(x) \equiv 0 \bmod 11$ has the solution $x \equiv 4 \bmod 11$. Now, $f'(x) = 3x^2 - 2x + 4$, so $f'(4) = 44 \equiv 0 \bmod 11$. However, this time $f(4) = 66 \not\equiv 0 \bmod 11^2$, so we are in case (iii) of Hensel's Lemma, and there are no solutions to the mod 11^2 congruence of the form $4 + 11t$. We further observe that $x \equiv 4 \bmod 11$ is the only solution to the mod 11 congruence (exercise), so the failure of this solution to be lifted means that there are no solutions whatsoever to the congruence $f(x) \equiv 0 \bmod 11^2$.

II – 12 Hensel's Lemma continued

Proposition 12.1. *Assume that we are in case (i) or (ii) of Hensel's Lemma. If $b = a + tp^k$ is a solution mod p^{k+1} lifting the mod p^k solution a, then $f'(b) \equiv f'(a) \bmod p$. In particular, $f'(b) \equiv 0 \bmod p$ if and only if $f'(a) \equiv 0 \bmod p$.*

Proof. Applying Lemma 11.1 to $f'(x)$ instead of $f(x)$, we have

$$f'(b) = f'(a) + f''(a)tp^k + t^2p^{2k}c' \quad \text{for some } c' \in \mathbb{Z},$$

so $f'(b) \equiv f'(a) \bmod p$. □

A consequence of this proposition is that, if we are lifting repeatedly, we need calculate the derivative mod p only once. Case (i) is always followed by case (i), and case (ii) can be followed only by case (ii) or case (iii).

Example. Here is an example of repeated lifting in case (i). Let $f(x) = x^3 - 10x + 5$ and $p = 29$, and note that $f(4) = 29 \equiv 0 \bmod 29$. Now, $f'(x) = 3x^2 - 10$, so $f'(4) = 38 \equiv 9 \bmod 29$, so we are indeed in case (i), and the unique lift of the form $4 + 29t$ with $t \in \{0, \ldots, 28\}$ is found by solving

$$\frac{f(4)}{29} + f'(4)t \equiv 0 \bmod 29,$$

i.e., $1 + 9t \equiv 0 \bmod 29$. Inverting 9 mod 29 (do this yourself for practice), we find that $t = 16$, so $4 + 16 \cdot 29 = 468$ is the unique lift of the mod 29 solution 4 to 29^2.

By Proposition 12.1, if we wish to lift further to a solution mod 29^3, we know already that we are in case (i) again, so we proceed straight to finding the unique $t \in \{0, \ldots, 28\}$ such that

$$\frac{f(468)}{29^2} + f'(468)t \equiv 0 \bmod 29.$$

Further, by the same proposition, $f'(468) \equiv f'(4) \equiv 9 \bmod 29$, i.e., we do not need to recalculate $f'(468)$, so we have only to solve

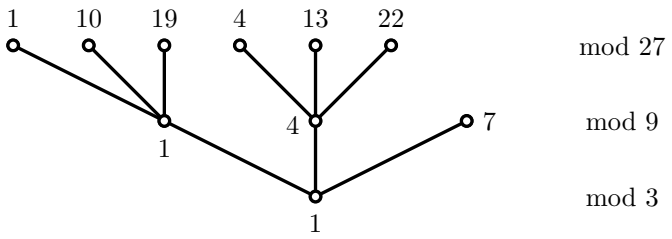
$$\frac{f(468)}{29^2} + 9t \equiv 0 \bmod 29, \quad \text{i.e.,} \quad 121\,877 + 9t \equiv 0 \bmod 29,$$

for $t \in \{0, \ldots, 28\}$, and we find easily that $t = 14$. Thus, $468 + 14 \cdot 29^2 = 12\,242$ is the unique lift to 29^3.

Example. Here is an example of repeated lifting in case (ii). Let $f(x) = x^2 + 4x + 22$ and $p = 3$. Observe that 1 is a solution mod 3. Find the mod 9 and mod 27 solutions lifting 1.

Solution: We find easily that $f'(1) = 6 \equiv 0 \bmod 3$ and $f(1) = 27 \equiv 0 \bmod 9$, so in attempting to lift from 3 to 9, we are in case (ii), and $1, 4, 7$ are all solutions mod 9.

Now, $f(1) \equiv 0 \bmod 27$, so case (ii) applies again to show that $1, 10, 19$ are all solutions mod 27. Also, $f(4) = 54 \equiv 0 \bmod 27$, so here as well we remain in case (ii). Thus, $4, 13, 22$ are all solutions mod 27. However, $f(7) = 99 \not\equiv 0 \bmod 27$, so case (iii) applies, and there are no solutions mod 27 of the form $7 + 9t$. A summary of the solutions mod 9 and 27 that lift 1 is therefore as follows:



A weakened hypothesis for indefinite lifting

We know from Proposition 12.1 that solutions may be lifted indefinitely when $p \nmid f'(a)$. In fact, we can weaken this assumption and still be guaranteed to be able to lift indefinitely, as in the following result.

Proposition 12.2. *If $f(x) \in \mathbb{Z}[x]$, p is prime, and $a \in \mathbb{Z}$ satisfies $v_p(f(a)) > 2v_p(f'(a))$, then there are integers $a_0, a_1, a_2, \ldots$, with $a_0 = a$, such that*

$$f(a_n) \equiv 0 \bmod p^{n+1} \quad \textit{and} \quad a_{n+1} \equiv a_n \bmod p^{n+1} \quad \textit{for all } n \geq 0.$$

For a proof, see Section 4 of the Appendix, where it is also explained how to construct $a_0, a_1, a_2, \ldots$ in this situation.

Example. Consider $f(x) = x^2 - x + 25$ and $p = 3$. Observe that $f(-1) = 27$ and $f'(-1) = -3$, so $v_3(f(-1)) = 3 > 2 = 2v_3(f'(-1))$. Therefore, the hypotheses of Proposition 12.2 are met, and one can lift indefinitely to obtain solutions to $f(x) \equiv 0 \bmod 3^k$ for all $k \geq 1$. For example, $x = 2+2\cdot 3^2+3^3+2\cdot 3^6+2\cdot 3^8+3^9+2\cdot 3^{11}+2\cdot 3^{13}$ is a solution mod 3^{14}.

Arbitrary moduli

Hensel's Lemma applies to moduli that are prime powers. For a general modulus, we can combine Hensel's Lemma with the Chinese Remainder Theorem.

For example, consider the polynomial $f(x) = x^2 - x + 25$ again, and suppose we wish to find solutions to $f(x) \equiv 0 \bmod 10\,125$. For this, we note that $10\,125 = 81 \cdot 125 = 3^4 \cdot 5^3$. We know from the previous example that solutions mod 3^4 can be found via Hensel's Lemma, and we leave it as an exercise to show, via Hensel's Lemma, that solutions exist mod 5^3 as well (in fact, modulo any given power of 5). Hence, if a is a solution mod 81 and b is a solution mod 125, then we have only to solve the system

$$\begin{aligned} x &\equiv a \quad \bmod 81 \\ x &\equiv b \quad \bmod 125 \end{aligned}$$

so obtain a solution mod $10\,125$.

(III) Gaussian Methods

III – 1 Sums of two squares: introduction

In number theory, a *square* is an integer that is the square of some integer. The squares are thus $0, 1, 4, 9, 16, \ldots$. It is natural to ask which positive integers n are sums of two squares, and for each n that is such a sum, in how many ways n can be so expressed. We represent this problem by the equation

$$x^2 + y^2 = n, \tag{1.1}$$

where we are trying to solve for $x, y \in \mathbb{Z}$.

A fruitful line of attack is via the observation that $x^2 + y^2 = (x + yi)(x - yi)$, where i is a fixed square root of -1 in $\mathbb{C}$, i.e., $i^2 = -1$. Thus, if $\alpha = x + yi$, then (1.1) is equivalent to

$$\alpha\,\overline{\alpha} = n,$$

where the bar denotes complex conjugation.

The Gaussian integers and uniqueness of factorization

Recall from MATH 228 the ring $\mathbb{Z}[i] = \{a + bi \in \mathbb{C} \mid a, b \in \mathbb{Z}\}$, the ring of *Gaussian integers*. It is a *unique factorization domain*. Let us review this notion from MATH 228.

In an integral domain R, a *prime element* is an element $a \in R$ satisfying all of the following: (i) $a \neq 0$, (ii) $a \notin R^\times$, and (iii) for all $b, c \in R$ such that a divides bc, either a divides b or a divides c.

In an integral domain R, an *irreducible element* is an element $a \in R$ satisfying all of the following: (i) $a \neq 0$, (ii) $a \notin R^\times$, and (iii) for all $b, c \in R$ such that $a = bc$, either $b \in R^\times$ or $c \in R^\times$.

If $a, b \in R$, we say that a is *associate* to b, written $a \sim b$, if there is $u \in R^\times$ such that $a = ub$. In this case, of course, $b = u^{-1}a$, so $b \sim a$ as well.

A *unique factorization domain (UFD)* is an integral domain R such that, for every non-zero $a \in R$ that is not a unit, the following both hold:

(i) a is a product of irreducible elements.

(ii) The factorization of a into irreducibles is essentially unique, in the sense that if $\pi_1 \cdots \pi_m$ and $\pi'_1 \cdots \pi'_n$ are two such factorizations, then $m = n$ and, after a reordering of the factors if necessary, $\pi_i \sim \pi'_i$ for all i.

Remark.

- In any integral domain, every prime element is irreducible (short exercise).
- In a unique factorization domain, it is conversely true that every irreducible element is a prime element (see MATH 228). This fact will be crucial later on.

The norm map

A useful tool in the ring $\mathbb{Z}[i]$ of Gaussian integers is the *norm map*,

$$\begin{aligned} N : \mathbb{Z}[i] &\to \mathbb{Z}_{\geq 0} \\ a + bi &\mapsto (a+bi)(a-bi) = a^2 + b^2. \end{aligned}$$

The following facts, both left as exercises, are crucial:

(i) $N(\alpha\beta) = N(\alpha)N(\beta)$ for all $\alpha, \beta \in \mathbb{Z}[i]$, and

(ii) $\alpha \in \mathbb{Z}[i]^\times$ if and only if $N(\alpha) = 1$.

Example. Observe that $5 = (1+2i)(1-2i)$. We may use the norm map to show that $1+2i$ and $1-2i$ are irreducible in $\mathbb{Z}[i]$, so 5 cannot be factorized in $\mathbb{Z}[i]$ any further than this factorization. Indeed, if $1+2i = \alpha\beta$ for some $\alpha, \beta \in \mathbb{Z}[i]$, then applying N to both sides of this equation gives $N(1+2i) = N(\alpha\beta)$, i.e., $5 = N(\alpha)N(\beta)$, so because 5 is a prime number and $N(\alpha), N(\beta) \in \mathbb{Z}_{\geq 1}$, either $N(\alpha) = 1$, in which case α is a unit, or $N(\beta) = 1$, in which case β is. The same argument shows that $1-2i$ is irreducible as well.

We leave it as a short exercise to show that $1+2i$ and $1-2i$ are not associate in $\mathbb{Z}[i]$, i.e., there is no unit $u \in \mathbb{Z}[i]^\times = \{1, i, -1, -i\}$ such that $1+2i = u(1-2i)$. Therefore, the factorization $5 = (1+2i)(1-2i)$ is a factorization of 5 into non-associate irreducibles.

Example. We have just seen that 5 is not irreducible in $\mathbb{Z}[i]$. Let us show that, by contrast, 3 is irreducible. Suppose that $3 = \alpha\beta$ where $\alpha, \beta \in \mathbb{Z}[i]$. Then

$$9 = N(3) = N(\alpha\beta) = N(\alpha)N(\beta),$$

so by uniqueness of factorization in $\mathbb{Z}$, the only three possibilities are

- $N(\alpha) = 1$, $N(\beta) = 9$,
- $N(\alpha) = N(\beta) = 3$,
- $N(\alpha) = 9$, $N(\beta) = 1$.

In fact, the middle option cannot occur, for if $\alpha = x + yi$, where $x, y \in \mathbb{Z}$, then $N(\alpha) = x^2 + y^2$, and there are no integers x, y such that $x^2 + y^2 = 3$. Therefore, either $N(\alpha) = 1$, in which case α is a unit, or $N(\beta) = 1$, in which case β is.

Some prime numbers, then, are irreducible in $\mathbb{Z}[i]$, such as 3, and some are not, such as 5. In fact, there is a simple rule to determine which prime numbers are irreducible in $\mathbb{Z}[i]$, which we will see shortly. The question is intimately linked to the equation $x^2 + y^2 = n$.

Exercise. The prime number 2 exhibits a special property with regard to factorization in $\mathbb{Z}[i]$. Show that $2 = \pi\pi'$ where π and π' are irreducible Gaussian integers that are associate to each other, i.e., $\pi = u\pi'$ for some $u \in \mathbb{Z}[i]^\times = \{1, i, -1, -i\}$.

III – 2 Gaussian splitting

Lemma 2.1 (Gaussian splitting lemma)**.** *Let p be a prime. Then p is a sum of two squares if and only if $p = 2$ or $p \equiv 1 \bmod 4$.*

Proof. Assume first that $p = x^2 + y^2$ where $x, y \in \mathbb{Z}$. Because the squares mod 4 are 0 and 1, the only possibilities for p mod 4 are (i) $0 + 0 = 0$, (ii) $0 + 1 = 1$, (iii) $1 + 0 = 1$, and (iv) $1 + 1 = 2$. The first is impossible because $4 \nmid p$, and the last implies that $p = 2$.

Conversely, assume that $p = 2$ or $p \equiv 1 \bmod 4$. Since $2 = 1^2 + 1^2$, we may assume immediately that $p \equiv 1 \bmod 4$. Then $\left(\frac{-1}{p}\right) = 1$ by Theorem 8.1 in Section II, so -1 is square mod p, i.e., there is $c \in \mathbb{Z}$ such that $-1 \equiv c^2 \bmod p$, and so there is $k \in \mathbb{Z}$ such that $c^2 + 1 = kp$. Hence,

$$(c + i)(c - i) = kp,$$

so p divides the product $(c+i)(c-i)$ in $\mathbb{Z}[i]$. But one verifies easily that p divides neither $c + i$ nor $c - i$ in $\mathbb{Z}[i]$, so p is not a prime element of the Gaussian integers. Therefore, because $\mathbb{Z}[i]$ is a unique factorization domain, p is not an irreducible element either, so

$$p = \alpha\beta$$

for some $\alpha, \beta \in \mathbb{Z}[i] \smallsetminus \mathbb{Z}[i]^\times$. Hence,

$$p^2 = N(p) = N(\alpha\beta) = N(\alpha)N(\beta).$$

But $N(\alpha), N(\beta) \neq 1$, so $N(\alpha) = N(\beta) = p$. Thus, if $\alpha = x + yi$, then

$$p = N(\alpha) = x^2 + y^2.$$

□

Gaussian irreducibles

If p is a prime number, define

$$\pi_p = \begin{cases} 1 + i & \text{if } p = 2, \\ x + yi & \text{if } p \equiv 1 \bmod 4, \text{ where } x^2 + y^2 = p \text{ and } 0 < x < y, \\ p & \text{if } p \equiv 3 \bmod 4. \end{cases}$$

By Lemma 2.1, x and y as above exist if $p \equiv 1 \bmod 4$. Further, one verifies that, with the extra constraint $0 < x < y$, the integers x and y are uniquely determined by p (exercise).

For example,

$$\pi_3 = 3, \quad \pi_5 = 1 + 2i, \quad \pi_7 = 7, \quad \pi_{11} = 11, \quad \pi_{13} = 2 + 3i.$$

Let

$$\Pi = \{\pi_2\} \cup \{\pi_p \mid p \equiv 1 \bmod 4\} \cup \{\overline{\pi}_p \mid p \equiv 1 \bmod 4\} \cup \{\pi_q \mid q \equiv 3 \bmod 4\}$$

$$= \{1+i, 3, 1+2i, 1-2i, 7, 11, 2+3i, 2-3i, 1+4i, 1-4i, \ldots\}.$$

Recall that if $\alpha, \beta \in \mathbb{Z}[i]$, then we say that α is associate to β (and write $\alpha \sim \beta$) if there is $u \in \mathbb{Z}[i]^\times = \{1, i, -1, -i\}$ such that $\alpha = u\beta$.

Proposition 2.2 (Gaussian irreducibles)**.**

(i) Every element of Π *is irreducible in* $\mathbb{Z}[i]$*.*

(ii) Every irreducible element of $\mathbb{Z}[i]$ *is associate to exactly one element of* Π*.*

We refer the reader to Section 5 of the Appendix for a proof. Lemma 2.1 plays a crucial role.

As a consequence of Proposition 2.2 and uniqueness of factorization in $\mathbb{Z}[i]$, we see that for every non-zero $\alpha \in \mathbb{Z}[i]$, there are unique non-negative integers r, s_p, s'_p ($p \equiv 1 \bmod 4$), and t_q ($q \equiv 3 \bmod 4$), and a unique unit u, satisfying

$$\alpha = u\pi_2^r \left(\prod_{p \equiv 1\,(4)} \left(\pi_p^{s_p} \overline{\pi}_p^{s'_p} \right) \right) \left(\prod_{q \equiv 3\,(4)} \pi_q^{t_q} \right).$$

Example. Here are some factorizations in $\mathbb{Z}[i]$:

$$\begin{aligned} -13 + 4i &= (-1)\pi_5 \overline{\pi}_{37} \\ 123 &= \pi_3 \pi_{41} \overline{\pi}_{41} \\ 190 + 1729i &= \pi_{17}^2 \pi_{19} \overline{\pi}_{29} \\ 54\,390 - 84\,770i &= i\pi_2^3 \pi_5^4 \overline{\pi}_5 \pi_7^2 \overline{\pi}_{13}^2 \end{aligned}$$

In each case, one could verify the factorization by expanding out the right-hand side and ensuring that we obtain the left-hand side, but in fact there is a method for taking any non-zero Gaussian integer and producing its factorization into a unit times a product of elements of Π.

To learn about splitting in a more general context, see Neukirch's *Algebraic Number Theory* [5], for example.

III – 3 Counting solutions to the equation $x^2 + y^2 = n$

Recall the p-adic valuation v_p defined in Section I – 1.

Theorem 3.1. *Let n be a positive integer.*

(i) The equation $x^2 + y^2 = n$ has a solution $(x, y) \in \mathbb{Z} \times \mathbb{Z}$ if and only if $v_q(n)$ is even for all primes $q \equiv 3 \bmod 4$.

(ii) If solutions exist, the number of solutions is

$$4 \prod_{p \equiv 1\,(4)} (v_p(n) + 1).$$

Proof. (i) We use the observation that an integer is a sum of two squares if and only if it is equal to $N(\alpha)$ for some $\alpha \in \mathbb{Z}[i]$, because $x^2 + y^2 = N(\alpha)$ where $\alpha = x + yi$.

Suppose, then, that $n = x^2 + y^2 = N(\alpha)$, where $\alpha = x + yi$. Write

$$\alpha = u\pi_2^r \left(\prod_{p \equiv 1\,(4)} \left(\pi_p^{s_p} \overline{\pi}_p^{s'_p} \right) \right) \left(\prod_{q \equiv 3\,(4)} \pi_q^{t_q} \right),$$

where $u \in \mathbb{Z}[i]^\times$ and $r, s_p, s'_p, t_q \in \mathbb{Z}_{\geq 0}$. Then

$$n = N(\alpha) = 2^r \left(\prod_{p \equiv 1\,(4)} p^{s_p + s'_p} \right) \left(\prod_{q \equiv 3\,(4)} q^{2t_q} \right),$$

so $v_q(n) = 2t_q$ for all primes $q \equiv 3 \bmod 4$.

Conversely, if $v_q(n)$ is even for all such q, then

$$\begin{aligned} n &= 2^{v_2(n)} \left(\prod_{p \equiv 1\,(4)} p^{v_p(n)} \right) \left(\prod_{q \equiv 3\,(4)} q^{v_q(n)} \right) \quad \text{by definition} \\ &= N(\alpha) \end{aligned}$$

where

$$\alpha = \pi_2^{v_2(n)} \left(\prod_{p \equiv 1\,(4)} \pi_p^{v_p(n)} \right) \left(\prod_{q \equiv 3\,(4)} \pi_q^{v_q(n)/2} \right),$$

so n is a sum of two squares.

(ii) Assume that solutions exist. By part (i), $v_q(n)$ is even for all primes $q \equiv 3 \bmod 4$. Define

$$\begin{aligned} X_0 &= \{0, 1, 2, 3\} \\ X_p &= \{0, \ldots, v_p(n)\} \quad \text{for } p \equiv 1 \bmod 4 \\ X &= X_0 \times \prod_{\substack{p \equiv 1\,(4) \\ p \,|\, n}} X_p \end{aligned}$$

If $\xi = (k, (s_p)_p) \in X$, let

$$\alpha_\xi = i^k \pi_2^{v_2(n)} \left(\prod_{\substack{p \equiv 1\,(4) \\ p \mid n}} \left(\pi_p^{s_p} \overline{\pi}_p^{v_p(n) - s_p} \right) \right) \left(\prod_{\substack{q \equiv 3\,(4) \\ q \mid n}} \pi_q^{v_q(n)/2} \right).$$

Then

$$N(\alpha_\xi) = 2^{v_2(n)} \left(\prod_{\substack{p \equiv 1\,(4) \\ p \mid n}} \left(p^{s_p} p^{v_p(n) - s_p} \right) \right) \left(\prod_{\substack{q \equiv 3\,(4) \\ q \mid n}} q^{v_q(n)} \right) = n,$$

so $(x, y) = (\mathrm{Re}(\alpha_\xi), \mathrm{Im}(\alpha_\xi))$ is a solution to $x^2 + y^2 = n$. Therefore, if

$$Y = \{(x, y) \in \mathbb{Z} \times \mathbb{Z} \mid x^2 + y^2 = n\},$$

we have a map

$$\begin{aligned} f : X &\to Y \\ \xi &\mapsto (\mathrm{Re}(\alpha_\xi), \mathrm{Im}(\alpha_\xi)). \end{aligned}$$

To finish our proof, it is enough to show that f is bijective, because the cardinality of Y is the number of solutions to $x^2 + y^2 = n$, and the cardinality of X is $4 \prod_{p \equiv 1\,(4)} (v_p(n) + 1)$. For injectivity, note that if $\alpha_{\xi_1} = \alpha_{\xi_2}$, then $\xi_1 = \xi_2$ by Proposition 2.2. As for surjectivity, suppose that $x^2 + y^2 = n$, i.e., $N(\alpha) = \alpha\,\overline{\alpha} = n$ where $\alpha = x + yi$. Then writing

$$\alpha = u \pi_2^r \left(\prod_{p \equiv 1\,(4)} \left(\pi_p^{s_p} \overline{\pi}_p^{s'_p} \right) \right) \left(\prod_{q \equiv 3\,(4)} \pi_q^{t_q} \right) \quad (u \in \mathbb{Z}[i]^\times),$$

we see from the equality $N(\alpha) = n$ that $r = v_2(n)$, $s_p + s'_p = v_p(n)$ when $p \equiv 1 \bmod 4$, and $2t_q = v_q(n)$ when $q \equiv 3 \bmod 4$, so $\alpha = \alpha_\xi$ where $\xi = (k, (s_p)_p) \in X$, k here being the unique integer in $\{0, 1, 2, 3\}$ such that $u = i^k$. Thus,

$$(x, y) = (\mathrm{Re}(\alpha), \mathrm{Im}(\alpha)) = (\mathrm{Re}(\alpha_\xi), \mathrm{Im}(\alpha_\xi)) = f(\xi).$$

□

Example. Let $n = 2^9 \cdot 17^5 \cdot 29^2 \cdot 23^4 \cdot 31^6$. Decide whether the equation $x^2 + y^2 = n$ has any integral solutions, and determine how many if so.

Solution: The primes congruent to 3 mod 4 that divide n are 23 and 31, which both occur to even powers, so integral solutions exist. The number of solutions is

$$4(v_{17}(n) + 1)(v_{29}(n) + 1) = 4 \cdot 6 \cdot 3 = 72.$$

Example. Repeat the preceding problem with $n = 37^2 \cdot 7^4 \cdot 43^7$ instead.

Solution: The prime divisor 43 is congruent to 3 mod 4 and occurs to an odd power, so the equation $x^2 + y^2 = n$ has no integral solutions.

III−4 The equation $y^2 = x^n - 1$

Let n be a positive integer, and consider the equation $y^2 = x^n - 1$, where x and y are integers to be solved for. The case $n = 1$ is trivial: there are infinitely many solutions in this case, because for each $y \in \mathbb{Z}$, we may let $x = y^2 + 1$. The case $n = 2$ can be solved by observing that $y^2 = x^2 - 1$ if and only if $(x + y)(x - y) = 1$, if and only if $x + y = x - y = 1$ or $x + y = x - y = -1$. The solutions in the case $n = 2$ are therefore $x = \pm 1$, $y = 0$.

For higher values of n, one may approach the equation by considering factorization in the Gaussian integers. A key ingredient is the following, which is proven in Section 6 of the Appendix.

Lemma 4.1. *Suppose that*

- *R is a unique factorization domain,*
- *$a, b \in R \smallsetminus \{0\}$,*
- *n is a positive integer.*

If a, b are coprime and $ab = c^n$ for some $c \in R$, then there are units u, v in R and elements $a', b' \in R$ such that $a = u(a')^n$ and $b = v(b')^n$.

Let us illustrate the relevance of this lemma for solving the equation $y^2 = x^n - 1$ by considering the case $n = 3$ as an example; see also [2, Sect. 1.5] for this case. We show that the only solution in integers to the equation $y^2 = x^3 - 1$ is $(x, y) = (1, 0)$. To see this, assume that $x, y \in \mathbb{Z}$ satisfy the equation, and rearrange it to read

$$\begin{aligned} (y + i)(y - i) &= x^3, \\ \text{i.e.,} \quad \alpha\,\overline{\alpha} &= x^3 \end{aligned}$$

where $\alpha = y + i \in \mathbb{Z}[i]$ and the bar denotes complex conjugation, as usual. We claim that the Gaussian integers α and $\overline{\alpha}$ are coprime. Suppose, for a contradiction, that π is a Gaussian irreducible that divides both α and $\overline{\alpha}$. Then π divides also $\alpha - \overline{\alpha} = 2i = (1+i)^2$, so π is associate to $1 + i$ by uniqueness of factorization in $\mathbb{Z}[i]$. Therefore, $1 + i$ divides α, so $(1 + i)(1 - i)$ divides $\alpha\,\overline{\alpha}$, i.e., 2 divides x^3—in $\mathbb{Z}[i]$, but also then in $\mathbb{Z}$. Hence, x is even, so x^3 is divisible by 8, and we arrive at the congruence $y^2 \equiv -1 \bmod 8$. But this is impossible, because -1 is not square mod 8. Thus, α and $\overline{\alpha}$ have no common Gaussian irreducible divisor, so they are coprime as claimed.

We may now apply Lemma 4.1, remembering that $\mathbb{Z}[i]$ is a unique factorization domain. By the lemma, the fact that $\alpha\,\overline{\alpha} = x^3$, a cube, implies that α is associate to a cube (and $\overline{\alpha}$ is also). That is, $\alpha = \eta\beta^3$ where $\beta \in \mathbb{Z}[i]$ and $\eta \in \mathbb{Z}[i]^\times$. In fact, since every unit in $\mathbb{Z}[i]$ is a cube, we have $\eta = (\eta')^3$ for some unit η', and then $\alpha = (\eta'\beta)^3$, so we may in fact assume that $\eta = 1$. Thus, $\alpha = \beta^3$ for some $\beta \in \mathbb{Z}[i]$.

Write $\beta = a + bi$ with $a, b \in \mathbb{Z}$. Then

$$y + i = (a + bi)^3 = a^3 - 3ab^2 + (3a^2b - b^3)i,$$

so

$$\begin{aligned} y &= a^3 - 3ab^2 = a(a^2 - 3b^2), \\ \text{and} \quad 1 &= 3a^2b - b^3 = b(3a^2 - b^2). \end{aligned}$$

This last line shows that $b = \pm 1$ and, subsequently, that $3a^2 - 1 = \pm 1$. We cannot have $3a^2 - 1 = 1$, because $3 \nmid 2$, so $3a^2 - 1 = -1$, i.e., $a = 0$. But then $y = 0$, so $x = 1$.

Exercise. Solve the equation $y^2 = x^4 - 1$ for $x, y \in \mathbb{Z}$ by considering the solutions to the equation $y^2 = x^2 - 1$ discussed above.

Exercise. Solve the equation $y^2 = x^5 - 1$ for $x, y \in \mathbb{Z}$ by following the method illustrated above in the case $n = 3$.

Exercise. Solve the equation $y^2 = x^6 - 1$ for $x, y \in \mathbb{Z}$ by following the method alluded to in the exercise above concerning the case $n = 4$. How do the cases $n = 4$ and $n = 6$ generalize?

Exercise. How might one solve $y^2 = x^{21} - 1$ by using an example we have already considered, but without any reasoning any further in terms of Gaussian integers?

Remark. An exercise in Washington's book [11, Chap. 1] considers the related equation $y^2 = x^3 - 5$. The method used above for the equation $y^2 = x^3 - 1$ may be extended to handle this related equation, although further considerations come into play that are beyond the scope of this course. The difficulty, which can be overcome, lies in the fact that the ring $\mathbb{Z}[\sqrt{-5}]$ does not have uniqueness of factorization. The subject of *algebraic number theory* provides tools for overcoming this difficulty.

Remark. A more recent approach to the study of equations such as $y^2 = x^3 + k$, where k is an integer, is via the theory of *elliptic curves*, which in fact concerns even more general equations than this. See Silverman's book [9] for an introduction.

III – 5 Pythagorean triples

A *Pythagorean triple* is a triple (x, y, z) of positive integers satisfying $x^2 + y^2 = z^2$.

Observe that if $x, y, z, d \in \mathbb{Z}_{\geq 1}$, then

$$x^2 + y^2 = z^2 \iff d^2(x^2 + y^2) = d^2 z^2 \iff (dx)^2 + (dy)^2 = (dz)^2,$$

so (x, y, z) is a Pythagorean triple if and only if (dx, dy, dz) is. Therefore, to find all Pythagorean triples, it is enough to find all those for which the only positive common divisor of the three numbers is 1, and then scale them.

Lemma 5.1. *Let (x, y, z) be a Pythagorean triple such that the only positive common divisor of x, y, z is 1. Then*

(i) x, y, z are pairwise coprime,

(ii) exactly one of x and y is even, and

(iii) z is odd.

Proof. (i) We show that x and y are coprime, the proof for the other two pairs being similar. Let d be a positive divisor of x and y. Then d^2 divides x^2 and y^2, so d^2 divides $x^2 + y^2 = z^2$ as well, and so d divides z. (Exercise: If $a, b \in \mathbb{Z}$, then $a \mid b$ if and only if $a^2 \mid b^2$.) Therefore, $d \mid x, y, z$, so $d = 1$.

(ii) Because x and y are coprime, they cannot both be even. If they were both odd, then $z^2 = x^2 + y^2$ would be congruent to 2 mod 4, which is impossible because 2 is not square mod 4.

(iii) This follows immediately from (ii) and the equality $z^2 = x^2 + y^2$. □

In light of Lemma 5.1, if (x, y, z) is a Pythagorean triple such that the only positive common divisor of x, y, z is 1, then either x is even or y is even, but not both, and there is nothing lost in assuming that y is the even one. We will therefore define a *primitive* Pythagorean triple to be one such that the numbers are coprime and y is even.

Theorem 5.2. *Let $x, y, z \in \mathbb{Z}_{\geq 1}$. Then (x, y, z) is a primitive Pythagorean triple if and only if there are coprime positive integers u and v such that*

$$u > v, \quad u \not\equiv v \bmod 2, \quad \textit{and} \quad (x, y, z) = (u^2 - v^2, 2uv, u^2 + v^2).$$

Proof. We will employ a proof that uses uniqueness of factorization in $\mathbb{Z}[i]$, although an elementary proof may be given as well; see Section 7 of the Appendix.

Suppose first that (x, y, z) is a primitive Pythagorean triple. Then $\alpha\,\overline{\alpha} = z^2$ where $\alpha = x + yi$. We claim that α and $\overline{\alpha}$ are coprime in $\mathbb{Z}[i]$. Indeed, if not, then both would be divisible by some Gaussian irreducible π, and then π would divide both $\alpha + \overline{\alpha} = 2x$ and $\alpha - \overline{\alpha} = 2yi$. If $\pi \not\sim \pi_2$, then we would have $\pi \mid x$ and $\pi \mid y$, and hence $N(\pi) \mid x^2$ and $N(\pi) \mid y^2$, which is impossible because x and y are coprime. Therefore, $\pi \sim \pi_2$, but from this we deduce that 2 divides $N(\alpha) = z^2$ in $\mathbb{Z}$, contradicting the fact that z is odd.

Knowing now that α and $\overline{\alpha}$ are coprime in $\mathbb{Z}[i]$, we return to the equation $\alpha\,\overline{\alpha} = z^2$ and realize, then, that it implies via Lemma 4.1 that $\alpha = \varepsilon\beta^2$ for some $\beta \in \mathbb{Z}[i]$ and some $\varepsilon \in \mathbb{Z}[i]$. The fact that x is odd rules out $\varepsilon \in \{i, -i\}$, and if $\varepsilon = -1$, then $\alpha = (i\beta)^2$, so we may in fact assume that $\varepsilon = 1$, that is, $\alpha = \beta^2$ for some $\beta = u + vi \in \mathbb{Z}[i]$.

Hence, $x + yi = (u + vi)^2 = u^2 - v^2 + 2uvi$, so

$$x = u^2 - v^2, \quad y = 2uv.$$

Because $y > 0$, we may assume, replacing both u and v by their negatives as necessary, that $u, v > 0$. Next, because $x > 0$, we deduce that $u > v$. Also, the fact that x is odd implies that $u \not\equiv v \bmod 2$. Finally, if d is some positive common divisor of u and v, then d divides $u^2 - v^2 = x$ and $2uv = y$, so $d = 1$. Thus, u and v are coprime.

Conversely, suppose that u and v are coprime positive integers satisfying the properties in the theorem, and let $x = u^2 - v^2$, $y = 2uv$, and $z = u^2 + v^2$, all positive. Then

$$x^2 + y^2 = (u^2 - v^2)^2 + (2uv)^2 = (u^2 + v^2)^2 = z^2.$$

Note that x is odd. To complete the proof, it is sufficient to show that x and z are coprime. To that end, let d be a positive common divisor of x and z. Then d divides $z + x = 2u^2$ and also divides $z - x = 2v^2$, so because d is odd, it divides u^2 and v^2. But u and v are coprime, so u^2 and v^2 are coprime, and so $d = 1$. □

We use Theorem 5.2 to tabulate the first few primitive Pythagorean triples. It suffices to run through pairs (u, v) as in the theorem.

u	v	$x = u^2 - v^2$	$y = 2uv$	$z = u^2 + v^2$
2	1	3	4	5
3	2	5	12	13
4	1	15	8	17
4	3	7	24	25
5	2	21	20	29
5	4	9	40	41
6	1	35	12	37
6	5	11	60	61

Exercise. Let (x, y, z) be a primitive Pythagorean triple, and let (u, v) be the pair corresponding to it via Theorem 5.2. Show that u^2 and v^2 are positive coprime integers such that $u^2 > v^2$ and $u^2 \not\equiv v^2 \bmod 2$, and find the primitive Pythagorean triple (x', y', z') corresponding to (u^2, v^2) via the theorem, expressing each of x', y', z' in terms of x, y, z.

(IV) Arithmetic Functions

IV – 1 Definitions and first examples

An *arithmetic function* is any function $f : \mathbb{Z}_{\geq 1} \to \mathbb{C}$. The following are all arithmetic functions:

(i) $f : m \mapsto 1$

(ii) $f : m \mapsto m$

(iii) $f : m \mapsto m^2$

(iv) $f : m \mapsto \left(\frac{m}{p}\right)$, where p is a fixed odd prime

(v)
$$f : m \mapsto \begin{cases} 0 & \text{if } m \equiv 0 \bmod 5 \\ 1 & \text{if } m \equiv 1 \bmod 5 \\ i & \text{if } m \equiv 2 \bmod 5 \\ -i & \text{if } m \equiv 3 \bmod 5 \\ -1 & \text{if } m \equiv 4 \bmod 5 \end{cases}$$

(vi) $f : m \mapsto \phi(m)$ (Euler's ϕ-function)

(vii) $f : m \mapsto m + 1$

There are two important types of arithmetic function, one a special case of the other:

- A *completely multiplicative* arithmetic function is an arithmetic function f that is not identically zero and that satisfies $f(mn) = f(m)f(n)$ for all positive integers m and n. Examples (i)–(v) above are all completely multiplicative.
- A *multiplicative* arithmetic function is an arithmetic function f that is not identically zero and that satisfies $f(mn) = f(m)f(n)$ for all *coprime* positive integers m and n. Example (vi) above, Euler's ϕ-function, is multiplicative, as we saw in Proposition 2.1 in Section II.

Every completely multiplicative arithmetic function is multiplicative, but not conversely, ϕ being a multiplicative arithmetic function that is not completely multiplicative.

The Möbius function

Define

$$\begin{array}{rcl} \mu : \mathbb{Z}_{\geq 1} & \to & \mathbb{C} \\ m & \mapsto & \begin{cases} (-1)^r & \text{if } m = p_1 \cdots p_r \text{ where } p_1, \ldots, p_r \text{ are distinct primes,} \\ 0 & \text{otherwise.} \end{cases} \end{array}$$

Note that the positive integer 1 is viewed as being a product $p_1 \cdots p_r$ where $r = 0$, so $\mu(1) = (-1)^0 = 1$. It is a short exercise—one well worth doing—to show that μ is multiplicative. However, it is not completely multiplicative, because $\mu(4) = 0$ while $\mu(2)\mu(2) = (-1)^2 = 1$.

Example. $\mu(21) = (-1)^2 = 1$, $\mu(105) = (-1)^3 = -1$, $\mu(99) = 0$ (because $3^2 \mid 99$).

Example. If p is a prime, then $\mu(p^k)$ is 1 if $k = 0$, is -1 if $k = 1$, and is 0 otherwise.

Some useful properties

(i) If f is multiplicative and $m \in \mathbb{Z}_{\geq 1}$ has prime factorization $m = p_1^{a_1} \cdots p_r^{a_r}$, then $f(m) = f(p_1^{a_1}) \cdots f(p_r^{a_r})$ because the factors $p_j^{a_j}$ are pairwise coprime.

(ii) If f is multiplicative, then $f(1) = 1$. Let us prove this. Observe that $f(1) = f(1 \cdot 1) = f(1)f(1)$ because 1 is coprime to itself, so $f(1)(f(1) - 1) = 0$, and so either $f(1) = 0$ or $f(1) = 1$. If the former, we would have, for all $m \in \mathbb{Z}_{\geq 1}$,

$$f(m) = f(1 \cdot m) = f(1)f(m) = 0f(m) = 0,$$

contradicting the assumption that f is not the zero function. Thus, $f(1) = 1$.

Of course, an arithmetic function f for which $f(1) = 1$ need not be multiplicative. An easy example is the function f given by $f(1) = 1$, $f(2) = f(3) = f(6) = 2$, and $f(m) = 0$ for all other $m \in \mathbb{Z}_{\geq 1}$.

(iii) If f and g are multiplicative arithmetic functions, then the pointwise product fg of f and g, defined by $(fg)(m) = f(m)g(m)$, is multiplicative. The same fact holds if one replaces *multiplicative* by *completely multiplicative* in both places in the assertion.

Dirichlet convolution

Let f and g be arithmetic functions, not necessarily multiplicative. The *Dirichlet convolution* $f * g$ is the arithmetic function defined by $(f * g)(m) = \sum_{d \mid m} f(d)g(m/d)$, the sum running over all positive divisors of m.

Example. If $f(m) = m^2$ and $g(m) = m$, then $(f * g)(m) = \sum_{d \mid m} d^2(m/d) = m \sum_{d \mid m} d$. Here is a table of the first few values of $f * g$:

m	1	2	3	4	5	6
$(f * g)(m)$	1	6	12	28	30	72

The set of arithmetic functions is a commutative unital ring where addition is the pointwise addition of functions and the product is Dirichlet convolution. We highlight some of the key properties, left as exercises along with the other ring properties:

- $(f * g) * h = f * (g * h)$ (associativity of $*$)
- $f * g = g * f$ (commutativity of $*$)
- $f * (g + h) = f * g + f * h$ (distributivity)
- $\iota * f = f$ where ι is the arithmetic function defined by $\iota(1) = 1$ and $\iota(m) = 0$ for $m > 1$

IV – 2 Inverses and sums

Inverses

An arithmetic function f is called *invertible* if there exists an arithmetic function g such that $f * g = \iota$.

Proposition 2.1. *An arithmetic function f is invertible if and only if $f(1) \neq 0$.*

Proof. If $f * g = \iota$, then in particular, $(f * g)(1) = \iota(1) = 1$, i.e., $f(1)g(1/1) = 1$, so $f(1) \neq 0$.

Conversely, assume that $f(1) \neq 0$, and define $g(m)$ recursively for $m \geq 1$ by

$$\begin{aligned} g(1) &= 1/f(1) \\ g(m) &= -\frac{1}{f(1)} \sum_{\substack{d \mid m, \\ d \neq 1}} f(d)g(m/d) \quad \text{if } m > 1. \end{aligned}$$

By construction, $(f * g)(1) = 1$ and $(f * g)(m) = 0$ if $m > 1$, so $f * g = \iota$. □

If f is invertible, then the arithmetic function g such that $f * g = \iota$ is unique (exercise). It is called the *Dirichlet inverse* of f, and is denoted f^{-1}.

Example. Let $f : m \mapsto m^2$, and let $g = f^{-1}$. Find $g(m)$ for $m \in \{1, \ldots, 6\}$.

Solution:

$$\begin{aligned} g(1) &= 1/f(1) = 1 \\ g(2) &= -\frac{1}{1} f(2)g(1) = -4 \\ g(3) &= -\frac{1}{1} f(3)g(1) = -9 \\ g(4) &= -\frac{1}{1}(f(4)g(1) + f(2)g(2)) = -(16 + (-16)) = 0 \\ g(5) &= -\frac{1}{1} f(5)g(1) = -25 \\ g(6) &= -\frac{1}{1}(f(6)g(1) + f(3)g(2) + f(2)g(3)) = -(36 - 36 - 36) = 36 \end{aligned}$$

Theorem 2.2.

(i) If f and g are multiplicative arithmetic functions, then so is $f * g$.

(ii) If f is a multiplicative arithmetic function, then so is f^{-1}, its Dirichlet inverse.

Proof. Let us prove (i) here and leave the proof of (ii) to the Appendix (Section 8).

Let $m, n \in \mathbb{Z}_{\geq 1}$ be coprime. Then the positive divisors of mn correspond bijectively to the pairs $(d, e) \in \mathbb{Z}_{\geq 1} \times \mathbb{Z}_{\geq 1}$ such that $d \mid m$ and $e \mid n$, with (d, e) corresponding to de. Then because f and g are multiplicative,

$$(f * g)(mn) = \sum_{d \mid m} \sum_{e \mid n} f(de)g\left(\tfrac{mn}{de}\right) = \sum_{d \mid m} \sum_{e \mid n} f(d)f(e)g\left(\tfrac{m}{d}\right)g\left(\tfrac{n}{e}\right)$$

$$= \left(\sum_{d \,|\, m} f(d) g\left(\tfrac{m}{d}\right) \right) \left(\sum_{e \,|\, n} f(e) g\left(\tfrac{n}{e}\right) \right) = (f * g)(m)\,(f * g)(n).$$

□

Sums

If f is an arithmetic function, let $\widehat{f}$ be the arithmetic function defined by

$$\widehat{f}(m) = \sum_{d \,|\, m} f(d).$$

Note that $\widehat{f} = \mathbf{1} * f$ where $\mathbf{1}$ is the constant arithmetic function, i.e., $\mathbf{1} : m \mapsto 1$. Indeed

$$\widehat{f}(m) = \sum_{d \,|\, m} f(d) = \sum_{d \,|\, m} f(d)\mathbf{1}(m/d) = (f * \mathbf{1})(m) = (\mathbf{1} * f)(m).$$

Observe the following:

> If f is a multiplicative arithmetic function, then so is $\widehat{f}$.

Indeed, $\mathbf{1}$ is multiplicative, so if f is multiplicative, then the Dirichlet convolution $\mathbf{1} * f = \widehat{f}$ is multiplicative.

Example. Let τ be the arithmetic function that counts the number of positive divisors of a given positive integer, that is,

$$\tau(m) = \#\{d \mid d \geq 1 \text{ and } d \,|\, m\} = \sum_{d \,|\, m} 1 = \widehat{\mathbf{1}}(m).$$

Then being equal to $\widehat{\mathbf{1}}$, the arithmetic function τ is multiplicative. Now, if p is prime and $a \geq 0$, then the positive divisors of p^a are $1, p, \ldots, p^a$, so $\tau(p^a) = a + 1$. Therefore, because τ is multiplicative, if a positive integer m has prime factorization $p_1^{a_1} \cdots p_r^{a_r}$, then

$$\tau(m) = \prod_{j=1} (a_j + 1).$$

Example. Let σ be the arithmetic function that adds the positive divisors of a given positive integer. Thus, if $f : m \mapsto m$, then

$$\sigma(m) = \sum_{d \,|\, m} d = \widehat{f}(m).$$

Hence, because f is multiplicative, the same is true of σ. We may therefore compute σ via prime powers, as we did for τ. Specifically, if p is prime and $a \geq 0$, then

$$\sigma(p^a) = 1 + p + \cdots + p^a = \frac{p^{a+1} - 1}{p - 1},$$

so if $m \in \mathbb{Z}_{\geq 1}$ has prime factorization $p_1^{a_1} \cdots p_r^{a_r}$,

$$\sigma(m) = \prod_{j=1}^{r} \frac{p_j^{a_j+1} - 1}{p_j - 1}.$$

Exercise. Use the above to show that $\sigma(14\,175) = 30\,008$.

IV – 3 Möbius inversion

Proposition 3.1 (Möbius inversion)**.**

(i) $\mu * \mathbf{1} = \iota$*, i.e.,* $\mathbf{1}^{-1} = \mu$*. Equivalently,* $\widehat{\mu} = \iota$*.*

(ii) If f *is an arithmetic function, then* $\mu * \widehat{f} = f$*.*

Proof. (i) Because μ and $\mathbf{1}$ are multiplicative, so is $\mu * \mathbf{1}$, so it is enough to show that $\mu * \mathbf{1}$ and ι agree on prime powers, ι being multiplicative as well, of course. Therefore, it is enough to show that $(\mu * \mathbf{1})(p^a) = \iota(p^a)$ for all positive integers a. (Why is the case $a = 0$ automatically true?) Now, $\mu(p^k)$ is equal to 1 if $k = 0$, to -1 if $k = 1$, and to 0 otherwise, so for $a \geq 1$,

$$(\mu * \mathbf{1})(p^a) = 1 - \mathbf{1}(p) = 1 - 1 = 0 = \iota(p^a),$$

as claimed.

(ii) This follows from (i):

$$\mu * \widehat{f} = \mu * (\mathbf{1} * f) = (\mu * \mathbf{1}) * f = \iota * f = f.$$

□

Example. Consider the von Mangoldt function, the arithmetic function defined by

$$\Lambda : m \mapsto \begin{cases} \log(p) & \text{if } m = p^k, \text{ where } p \text{ is prime and } k \geq 1, \\ 0 & \text{otherwise.} \end{cases}$$

(Note that Λ is not multiplicative, because $\Lambda(1) = 0$.) We will use Möbius inversion to give an alternative description of Λ. Let $m \in \mathbb{Z}_{\geq 1}$ have prime factorization $p_1^{a_1} \cdots p_r^{a_r}$. Then because $\Lambda(n) = 0$ when n is not a prime power,

$$\widehat{\Lambda}(m) = \sum_{d \,|\, m} \Lambda(d) = \sum_{j=1}^{r} \sum_{k=1}^{a_j} \Lambda(p_j^k) = \sum_{j=1}^{r} \sum_{k=1}^{a_j} \log(p_j) = \sum_{j=1}^{r} a_j \log(p_j) = \log(m).$$

Hence, by Möbius inversion,

$$\begin{aligned} \Lambda(m) = (\mu * \widehat{\Lambda})(m) = \sum_{d \,|\, m} \mu(d) \widehat{\Lambda}\left(\tfrac{m}{d}\right) &= \sum_{d \,|\, m} \mu(d) \log\left(\tfrac{m}{d}\right) \\ &= \sum_{d \,|\, m} \mu(d) \log(m) - \sum_{d \,|\, m} \mu(d) \log(d) \\ &= \log(m) \widehat{\mu}(m) - \sum_{d \,|\, m} \mu(d) \log(d) \\ &= - \sum_{d \,|\, m} \mu(d) \log(d), \end{aligned}$$

because $\log(m) = 0$ if $m = 1$, and $\widehat{\mu}(m) = \iota(m) = 0$ if $m > 0$.

IV – 4 A strategy for computing a multiplicative function

Suppose that f is a multiplicative arithmetic function. A common strategy to find an explicit formula for $f(m)$ is as follows:

(i) Justify first that f really is multiplicative. Refer to any relevant facts: If f and g are multiplicative arithmetic functions, then $f * g$, f^{-1}, $\widehat{f}$, and fg, for example, are all multiplicative.

(ii) Find $f(p^a)$ where p is prime and $a \in \mathbb{Z}_{\geq 1}$.

(iii) Put the above together to obtain an expression for $f(m)$ via the fact that $m = \prod_{p \,|\, m} p^{v_p(m)}$.

Example. Show that $\widehat{\phi}(m) = m$ for all positive integers m, where ϕ is Euler's ϕ-function.

Solution: We know already that ϕ is multiplicative, so $\widehat{\phi}$ is also multiplicative. Now, if p is prime and $a \in \mathbb{Z}_{\geq 1}$, then

$$\widehat{\phi}(p^a) = \sum_{k=0}^{a} \phi(p^k) = 1 + \sum_{k=1}^{a} \phi(p^k) = 1 + \sum_{k=1}^{a} (p^k - p^{k-1}) = p^a, \tag{4.1}$$

the last sum being a telescoping one. Therefore,

$$\begin{aligned} \widehat{\phi}(m) = \widehat{\phi}\left(\prod_{p \,|\, m} p^{v_p(m)} \right) &= \prod_{p \,|\, m} \widehat{\phi}(p^{v_p(m)}) \quad \text{because } \widehat{\phi} \text{ is multiplicative} \\ &= \prod_{p \,|\, m} p^{v_p(m)} \quad \text{by (4.1)} \\ &= m. \end{aligned}$$

Example. Let $f = \mu\phi$, the pointwise product of μ and ϕ, that is, $f(m) = \mu(m)\phi(m)$, and consider $\widehat{f} = \mathbf{1} * f$. Show that for all $m \in \mathbb{Z}_{\geq 1}$,

$$\widehat{f}(m) = \prod_{p \,|\, m} (2 - p),$$

the product running over all primes p that divide m.

Solution: The arithmetic functions μ and ϕ are multiplicative, so their pointwise product $f = \mu\phi$ is as well, and so $\widehat{f}$ is multiplicative. Now, if p is prime and $a \in \mathbb{Z}_{\geq 1}$,

$$\widehat{f}(p^a) = \sum_{k=0}^{a} f(p^k) = \sum_{k=0}^{a} \mu(p^k)\phi(p^k) = 1 - (p-1) = 2 - p.$$

Hence, because $\widehat{f}$ is multiplicative,

$$\widehat{f}(m) = \prod_{p \,|\, m} \widehat{f}(p^{v_p(m)}) = \prod_{p \,|\, m} (2 - p).$$

IV – 5 Bell series

A *formal power series* (over $\mathbb{C}$) is an expression of the form $F(x) = \sum_{n=0}^{\infty} a_n x^n$ where the coefficients a_n are in $\mathbb{C}$ and x is an indeterminate. We add and multiply formal power series as follows:

$$\sum_{n=0}^{\infty} a_n x^n + \sum_{n=0}^{\infty} b_n x^n = \sum_{n=0}^{\infty} (a_n + b_n) x^n,$$
$$\left(\sum_{n=0}^{\infty} a_n x^n\right)\left(\sum_{n=0}^{\infty} b_n x^n\right) = \sum_{n=0}^{\infty}\left(\sum_{k=0}^{n} a_k b_{n-k}\right) x^n.$$

These operations make the set of formal power series a commutative unital ring.

It is useful to be able to *reindex*: If $k + l \geq 0$, then

$$\sum_{n=k}^{\infty} a_n x^{n+l} = \sum_{n=k+l}^{\infty} a_{n-l} x^n.$$

For example, $\sum_{n=3}^{\infty} n^2 \sin(n) x^{n+2} = \sum_{n=5}^{\infty} (n-2)^2 \sin(n-2) x^n$.

Formal differentiation of power series is defined term by term: If $F(x) = \sum_{n=0}^{\infty} a_n x^n$, then

$$F'(x) = \sum_{n=1}^{\infty} n a_n x^{n-1} = \sum_{n=0}^{\infty} (n+1) a_{n+1} x^n.$$

The product rule holds for formal differentiation: $(FG)' = F'G + FG'$. If $G(x)$ is a power series with zero constant term (i.e., the coefficient of x^0 is zero), then for any power series $F(x)$, the substitution $F(G(x))$ is possible. The power series $F(G(x))$ is denoted $(F \circ G)(x)$. In this situation, the chain rule holds for formal differentiation: $(F \circ G)'(x) = G'(x) F'(G(x))$.

One often uses the power series $(1 - tx)^{-1}$, where t is some fixed complex number. As an explicit power series, it is given by

$$\frac{1}{1 - tx} = \sum_{n=0}^{\infty} t^n x^n.$$

To see this, we multiply $\sum_{n=0}^{\infty} t^n x^n$ by $1 - tx$:

$$\begin{aligned}(1 - tx) \sum_{n=0}^{\infty} t^n x^n &= \sum_{n=0}^{\infty} t^n x^n - \sum_{n=0}^{\infty} t^{n+1} x^{n+1} \\ &= \sum_{n=0}^{\infty} t^n x^n - \sum_{n=1}^{\infty} t^n x^n \quad \text{(reindexing)} \\ &= 1.\end{aligned}$$

Example. Let us find $\frac{1}{1-2x} \frac{1}{1-\frac{1}{2}x}$ as an explicit power series:

$$\frac{1}{1-2x} \frac{1}{1-\frac{1}{2}x} = \left(\sum_{n=0}^{\infty} 2^n x^n\right)\left(\sum_{n=0}^{\infty} \left(\frac{1}{2}\right)^n x^n\right) = \sum_{n=0}^{\infty}\left(\sum_{k=0}^{n} 2^k \left(\frac{1}{2}\right)^{n-k}\right) x^n$$

$$= \sum_{n=0}^{\infty} \left(\sum_{k=0}^{n} 2^{2k-n} \right) x^n$$
$$= \sum_{n=0}^{\infty} \frac{4^{n+1}-1}{3 \cdot 2^n} \, x^n,$$

the last step using the observation that

$$\sum_{k=0}^{n} 2^{2k} = \sum_{k=0}^{n} 4^k = \frac{4^{n+1}-1}{4-1} = \frac{4^{n+1}-1}{3}.$$

Now let f be an arithmetic function and p a prime number. The *Bell series* of f at p is the formal power series

$$B_{f,p}(x) = \sum_{n=0}^{\infty} f(p^n) x^n.$$

Example.

$$B_{\iota,p}(x) = \sum_{n=0}^{\infty} \iota(p^n) x^n = 1$$
$$B_{\mathbf{1},p}(x) = \sum_{n=0}^{\infty} \mathbf{1}(p^n) x^n = \sum_{n=0}^{\infty} x^n = \frac{1}{1-x}$$
$$B_{\mu,p}(x) = \sum_{n=0}^{\infty} \mu(p^n) x^n = 1 - x$$

Example. The computation of $B_{\phi,p}(x)$ requires a little more manipulation:

$$\begin{aligned} B_{\phi,p}(x) = \sum_{n=0}^{\infty} \phi(p^n) x^n &= 1 + \sum_{n=1}^{\infty} \phi(p^n) x^n \\ &= 1 + \sum_{n=1}^{\infty} (p^n - p^{n-1}) x^n \\ &= 1 + \sum_{n=1}^{\infty} p^n x^n - \sum_{n=1}^{\infty} p^{n-1} x^n \\ &= \sum_{n=0}^{\infty} p^n x^n - \sum_{n=0}^{\infty} p^n x^{n+1} \\ &= (1-x) \sum_{n=0}^{\infty} p^n x^n = \frac{1-x}{1-px}. \end{aligned}$$

Example. Let N be the arithmetic function $m \mapsto m$, i.e., $N(m) = m$. Then

$$B_{N,p}(x) = \sum_{n=0}^{\infty} N(p^n) x^n = \sum_{n=0}^{\infty} p^n x^n = \frac{1}{1-px}.$$

IV – 6 Bell series continued

Proposition 6.1. *If f and g are arithmetic functions, and if p is prime, then*

$$B_{f*g,p}(x) = B_{f,p}(x)B_{g,p}(x).$$

Proof.

$$\begin{aligned} B_{f,p}(x)B_{g,p}(x) = \left(\sum_{n=0}^{\infty} f(p^n)x^n\right)\left(\sum_{n=0}^{\infty} g(p^n)x^n\right) &= \sum_{n=0}^{\infty}\left(\sum_{k=0}^{n} f(p^k)g(p^{n-k})\right)x^n \\ &= \sum_{n=0}^{\infty}\left(\sum_{d\,|\,p^n} f(d)g\Big(\frac{p^n}{d}\Big)\right)x^n \\ &= \sum_{n=0}^{\infty}(f*g)(p^n)x^n = B_{f*g,p}(x). \end{aligned}$$

□

Example. Show that $\tau = \mathbf{1} * \mathbf{1}$, and use this fact to show that

$$B_{\tau,p}(x) = \frac{1}{1-2x+x^2}.$$

Solution: By definition,

$$\tau(m) = \sum_{d\,|\,m} 1 = \sum_{d\,|\,m} \mathbf{1}(d)\mathbf{1}\Big(\frac{m}{d}\Big) = (\mathbf{1}*\mathbf{1})(m).$$

Hence, by Proposition 6.1,

$$B_{\tau,p}(x) = B_{\mathbf{1}*\mathbf{1},p}(x) = B_{\mathbf{1},p}(x)^2 = \frac{1}{(1-x)^2} = \frac{1}{1-2x+x^2}.$$

Example. Show that $\sigma = N * \mathbf{1}$, where $N : m \mapsto m$, and use this fact to show that

$$B_{\sigma,p}(x) = \frac{1}{1-(p+1)x+px^2}.$$

Solution: For all $m \geq 1$,

$$\sigma(m) = \sum_{d\,|\,m} d = \sum_{d\,|\,m} N(d)\mathbf{1}\Big(\frac{m}{d}\Big) = (N*\mathbf{1})(m),$$

so $\sigma = N * \mathbf{1}$ and

$$B_{\sigma,p}(x) = B_{N*\mathbf{1},p}(x) = B_{N,p}(x)B_{\mathbf{1},p}(x) = \frac{1}{1-px}\,\frac{1}{1-x} = \frac{1}{1-(p+1)x+px^2}.$$

Proposition 6.2. *If f and g are multiplicative arithmetic functions (note especially the word* multiplicative *here), then $f = g$ if and only if $B_{f,p}(x) = B_{g,p}(x)$ for all primes p.*

Proof. If $B_{f,p}(x) = B_{g,p}(x)$, then $\sum_{n=0}^{\infty} f(p^n)x^n = \sum_{n=0}^{\infty} g(p^n)x^n$, so equating coefficients, we obtain $f(p^n) = g(p^n)$ for all $n \geq 0$. If this is true for all primes p, it follows from the assumption that f and g are multiplicative that $f = g$. □

Example. If $m \in \mathbb{Z}_{\geq 1}$ has prime factorization $p_1^{a_1} \cdots p_r^{a_r}$, where $p_1, \ldots, p_r$ are distinct primes and $a_i > 0$ for all i, let $\omega(m) = r$ (and define $\omega(1) = 0$). Then the arithmetic function $f : m \mapsto 2^{\omega(m)}$ is multiplicative (exercise). Use Bell series to express f as a convolution of well-known functions.

Solution: We begin with the definition of $B_{f,p}(x)$ and then manipulate the series:

$$B_{f,p}(x) = \sum_{n=0}^{\infty} f(p^n)x^n = 1 + \sum_{n=1}^{\infty} 2x^n = -1 + 2\sum_{n=0}^{\infty} x^n = -1 + \frac{2}{1-x}$$
$$= \frac{1+x}{1-x} = \frac{1}{1-x}(1+x).$$

Now, we know that $\frac{1}{1-x} = B_{\mathbf{1},p}(x)$, and it is easy to verify that $1 + x = B_{\mu^2,p}(x)$, where μ^2 is the pointwise product of μ with itself, i.e., $(\mu^2)(m) = \mu(m)^2$, so

$$B_{f,p}(x) = B_{\mathbf{1},p}(x)B_{\mu^2,p}(x) = B_{\mathbf{1}*(\mu^2),p}(x).$$

Hence, because f and $1 * (\mu^2)$ are multiplicative, it follows from Proposition 6.2 that $f = \mathbf{1} * (\mu^2)$.

IV – 7 The Möbius function and roots of unity

The Möbius function μ, which we introduced in Section 1 and subsequently saw is the Dirichlet inverse of the constant function $\mathbf{1}$, has an elegant description in terms of roots of unity.

A complex number ζ is called a *root of unity* if $\zeta^n = 1$ for some positive integer n. In this case, ζ is called an nth root of unity.

If ζ is a root of unity, then the least positive integer n such that $\zeta^n = 1$ is called its *order*. A root of unity of order n is called a *primitive* nth root of unity. For example, -1 is a root of unity of order 2, so it is a primitive 2nd root of unity. It is also a 4th root of unity, but not a primitive one. The primitive 4th roots of unity are i and $-i$.

Proposition 7.1. *For each positive integer n, there are n distinct nth roots of unity in $\mathbb{C}$.*

The proof of this proposition is straightforward, resting on only some basic trigonometry, and is given in Section 1 of the Appendix.

Lemma 7.2. *If ζ is an nth root of unity, then its order divides n.*

Proof. Let m be the order of ζ, and write $n = qm + r$ where $q, r \in \mathbb{Z}$ and $0 \leq r < m$. Then

$$\begin{aligned} \zeta^r &= \zeta^{n-qm} \\ &= \zeta^n \quad \text{because } \zeta^m = 1 \\ &= 1 \quad \text{because } \zeta^n = 1. \end{aligned}$$

Hence, by the minimality of m, r must be zero. □

Proposition 7.3. *For each positive integer n, $\mu(n)$ is the sum of all primitive nth roots of unity in $\mathbb{C}$.*

Proof. Define arithmetic functions F and f by letting $F(n)$ be the sum of all nth roots of unity in $\mathbb{C}$ and letting $f(n)$ be the sum of all primitive nth roots of unity, i.e., roots of unity of order n. Because the order of an nth root of unity divides n by Lemma 7.2, we see that

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

i.e., $F = \mathbf{1} * f$.

We show that $F = \iota$. It is clear that $F(1) = 1$, because the only 1st root of unity is 1. Now we let $n > 1$ and show that $F(n) = 0$. By Proposition 7.1, there are n distinct nth roots of unity in $\mathbb{C}$, say $\zeta_1, \ldots, \zeta_n$, so the polynomial $x^n - 1$ factorizes as

$$x^n - 1 = (x - \zeta_1) \cdots (x - \zeta_n).$$

Expanding the right-hand side out, we see that the coefficient of x^{n-1} is $-(\zeta_1 + \cdots + \zeta_n)$, while on the left-hand side the corresponding coefficient is 0 because $n > 1$. Thus, $F(n) = 0$, as desired.

In summary, we have $\mathbf{1} * f = F = \iota$, so Möbius inversion (Proposition 3.1) gives $f = \mu$. □

Example. The sum of the primitive cubic (3rd) roots of unity in $\mathbb{C}$ is $\mu(3)$, which is -1 according to the definition of μ in Section 1. More generally, if p is a prime, then the sum of the primitive pth roots of unity in $\mathbb{C}$ is $\mu(p) = -1$.

Example. The sum of the primitive 15th roots of unity in $\mathbb{C}$ is $\mu(15) = 1$. More generally, if p and q are distinct primes, then the sum of the primitive (pq)th roots of unity in $\mathbb{C}$ is $\mu(pq) = 1$.

Example. If n is divisible by the square of some prime, then the sum of the primitive nth roots of unity in $\mathbb{C}$ is $\mu(n) = 0$. For example, the sum of the primitive 45th roots of unity in $\mathbb{C}$ is 0.

An application to sums of cosines

As discussed in Section 1 of the Appendix, the nth roots of unity in $\mathbb{C}$ are

$$\cos(\tfrac{2\pi k}{n}) + i\sin(\tfrac{2\pi k}{n})$$

with $k \in \{0, \ldots, n-1\}$. By the same argument as in the proof of Proposition Ord-3, the primitive nth roots of unity are obtained by restricting k to be coprime to n. For example, the primitive 15th roots of unity are

$$\cos(\tfrac{2\pi k}{15}) + i\sin(\tfrac{2\pi k}{15}) \quad \text{with } k \in \{1, 2, 4, 7, 8, 11, 13, 14\}.$$

But by Proposition 7.3, the sum of these is $\mu(15) = 1$, so

$$\sum_{\substack{k=0,\\ \gcd(k,15)=1}}^{14} \cos(\tfrac{2\pi k}{15}) + i \sum_{\substack{k=0,\\ \gcd(k,15)=1}}^{14} \sin(\tfrac{2\pi k}{15}) = 1.$$

In particular, equating real parts we obtain

$$\sum_{\substack{k=0,\\ \gcd(k,15)=1}}^{14} \cos(\tfrac{2\pi k}{15}) = 1. \tag{7.1}$$

Using the identity $\cos(2\pi - x) = \cos(x)$, we rewrite four of the eight terms in the sum as follows:

$$\cos(\tfrac{14\pi}{15}) = \cos(\tfrac{16\pi}{15}), \quad \cos(\tfrac{22\pi}{15}) = \cos(\tfrac{8\pi}{15}), \quad \cos(\tfrac{26\pi}{15}) = \cos(\tfrac{4\pi}{15}), \quad \cos(\tfrac{28\pi}{15}) = \cos(\tfrac{2\pi}{15}).$$

Hence, making these substitutions in (7.1), and then dividing by 2, we arrive at

$$\cos(\tfrac{2\pi}{15}) + \cos(\tfrac{4\pi}{15}) + \cos(\tfrac{8\pi}{15}) + \cos(\tfrac{16\pi}{15}) = \frac{1}{2}.$$

(V) Pell's Equation and Continued Fractions

V – 1 Pell's equation: introduction

Let d be a positive integer that is not a square. Pell's equation for the integer d can refer to either of the following equations:

$$x^2 - dy^2 = 1,$$
$$x^2 - dy^2 = -1.$$

The second is often called the *negative* Pell equation. In either case, we seek positive integral solutions, i.e., solutions in which both x and y are positive integers.

It is a fact (see Section 14 of the Appendix for a proof) that the equation $x^2 - dy^2 = 1$, where d is a positive integer that is not a square, always has positive integral solutions. We will see a method to find the solutions, involving the theory of *continued fractions*. Before we start on that theory, we will make some preliminary observations.

The norm map

We have already seen, in Section III, the norm map on the Gaussian integers, i.e., the map $N : \mathbb{Z}[i] \to \mathbb{Z}_{\geq 0}$ sending $x + yi$ to $x^2 + y^2$. More generally, if d is an integer that is not a square, and if $\sqrt{d}$ is a fixed square root of d in $\mathbb{C}$, then we have the map

$$\begin{aligned} N_d : \mathbb{Z}[\sqrt{d}] &\to \mathbb{Z} \\ x + y\sqrt{d} &\mapsto x^2 - dy^2, \end{aligned}$$

where $\mathbb{Z}[\sqrt{d}] = \{x + y\sqrt{d} \mid x, y \in \mathbb{Z}\}$. The relevance of N_d to Pell's equation lies in the fact that $x^2 - dy^2 = 1$ (or -1) if and only if $N_d(x + y\sqrt{d}) = 1$ (or -1). If there is no confusion, we will omit the subscript d and write just N for the norm map on $\mathbb{Z}[\sqrt{d}]$.

Example. In the case $d = 5$, $N(7 + 3\sqrt{5}) = 7^2 - 5 \cdot 3^2 = 4$.

As with the map N on the Gaussian integers, we have $N_d(\alpha\beta) = N_d(\alpha)N_d(\beta)$ for all $\alpha, \beta \in \mathbb{Z}[\sqrt{d}]$. We leave this observation as a short exercise and use it to prove the following.

Proposition 1.1. *Let d be a positive integer that is not a square, and suppose that $x, y \in \mathbb{Z}_{\geq 1}$ satisfy $x^2 - dy^2 = 1$. For each $n \geq 1$, write $(x + y\sqrt{d})^n = x_n + y_n\sqrt{d}$ with $x_n, y_n \in \mathbb{Z}_{\geq 1}$. Then the pairs (x_n, y_n) constitute infinitely many integral solutions to $x^2 - dy^2 = 1$.*

Proof. Suppose that $\alpha = a + b\sqrt{d}$ and $\alpha' = a' + b'\sqrt{d}$ correspond to solutions to Pell's equation for d, i.e., $N(\alpha) = N(\alpha') = 1$ where $N = N_d$. Then because N respects multiplication, $N(\alpha\alpha') = N(\alpha)N(\alpha') = 1 \cdot 1 = 1$, so $\alpha\alpha'$ also corresponds to a solution to the equation. Therefore, if $x^2 - dy^2 = 1$, i.e., $N(x + y\sqrt{d}) = 1$, then induction on n shows that for all $n \geq 1$,

$$N((x + y\sqrt{d})^n) = 1,$$

$$\text{i.e.,} \quad N(x_n + y_n\sqrt{d}) = 1 \quad \text{in the notation of the proposition,}$$
$$\text{i.e.,} \quad x_n^2 - dy_n^2 = 1.$$

It remains only to show that the (x_n, y_n) yield infinitely many pairs. But

$$x_{n+1} + y_{n+1}\sqrt{d} = (x + y\sqrt{d})(x_n + y_n\sqrt{d}) = xx_n + dyy_n + (xy_n + yx_n)\sqrt{d},$$

so $x_{n+1} = xx_n + dyy_n > x_n$. □

Example. One solution to the equation $x^2 - 3y^2 = 1$ is $(x, y) = (2, 1)$, which corresponds to $2 + \sqrt{3} \in \mathbb{Z}[\sqrt{3}]$. Therefore, another solution is obtained from

$$(2 + \sqrt{3})^2 = 7 + 4\sqrt{3},$$

i.e., $(x, y) = (7, 4)$ is another solution. (Check: $7^2 - 3 \cdot 4^2 = 49 - 48 = 1$.) Yet another solution can be found by multiplying by $2 + \sqrt{3}$ again:

$$(2 + \sqrt{3})(7 + 4\sqrt{3}) = 26 + 15\sqrt{3},$$

so $(x, y) = (26, 15)$ is a solution. (Check: $26^2 - 3 \cdot 15^2 = 676 - 675 = 1$.) Of course, this process may be iterated as many times as one wishes.

Exercise. If (x, y) is a solution to the negative Pell equation for d, i.e., $x^2 - dy^2 = -1$, and if $(x + y\sqrt{d})^n = x_n + y_n\sqrt{d}$, show that $x_n^2 - dy_n^2 = (-1)^n$. One may use the same argument as in the proof of Proposition 1.1.

Note that Proposition 1.1 does not guarantee the existence of solutions $(x, y) \in \mathbb{Z}_{\geq 1} \times \mathbb{Z}_{\geq 1}$ to Pell's equation, only that if at least one such solution exists, then there are infinitely many. To tackle the existence of solutions, we turn to the theory of continued fractions.

Historical note

The problem of solving $x^2 - dy^2 = 1$ goes back well over a thousand years, with the mathematician Brahmagupta making important progress on it in the 7th century [10, Sect. 5.4]. Brahmagupta's contributions include the discovery of the equality $N_d(\alpha\beta) = N_d(\alpha)N_d(\beta)$, in different notation. While many mathematicians subsequently found and refined methods for obtaining solutions to the equation, it was Lagrange, around 1768, who provided the first rigorous proof of the validity of a method, employing continued fractions to do so [10, Sect. 3.4]. Lagrange's proof can be found in his collected works [3].

V – 2 Definition of continued fractions and first examples

Let $(a_0; a_1, a_2, \ldots, a_n)$ be an $(n+1)$-tuple of real numbers $a_0, a_1, \ldots, a_n$ with $a_i > 0$ for all $i \geq 1$. (There is no restriction on a_0.) The *continued fraction* associated to $(a_0; a_1, \ldots, a_n)$ is the real number

$$[a_0; a_1, a_2, \ldots, a_n] = a_0 + \cfrac{1}{a_1 + \cfrac{1}{a_2 + \cfrac{1}{a_3 + \cfrac{1}{\ddots + \cfrac{1}{a_n}}}}}$$

Example.

$$[3; 1, 4, 7] = 3 + \cfrac{1}{1 + \cfrac{1}{4 + \cfrac{1}{7}}}$$

Example.

$$[1; 2, 3, 4, 5] = 1 + \cfrac{1}{2 + \cfrac{1}{3 + \cfrac{1}{4 + \cfrac{1}{5}}}}$$

Observe that if $a_0, a_1, \ldots, a_n \in \mathbb{Z}$ and $a_i > 0$ for $i \geq 1$, then $[a_0; a_1, \ldots, a_n] \in \mathbb{Q}$. Conversely, we have:

Proposition 2.1. *Every rational number can be expressed uniquely in the form* $[a_0; a_1, \ldots, a_n]$ *where* $a_i \in \mathbb{Z}$ *for all* i, $a_i \in \mathbb{Z}_{\geq 1}$ *for* $i \geq 1$, *and* $a_n \geq 2$ *if* $n \geq 1$.

Proof. Let $c, d \in \mathbb{Z}$ with $d > 0$, and consider the rational number c/d. Let $c_0 = c$ and $c_1 = d$, and perform the Euclidean algorithm on c_0 and c_1:

$$c_0 = a_0 c_1 + c_2 \quad (0 \leq c_2 < c_1) \tag{2.1}$$
$$c_1 = a_1 c_2 + c_3 \quad (0 \leq c_3 < c_2) \tag{2.2}$$
$$c_2 = a_2 c_3 + c_4 \quad (0 \leq c_4 < c_3) \tag{2.3}$$
$$\vdots$$
$$c_{n-1} = a_{n-1} c_n + c_{n+1} \quad (0 \leq c_{n+1} < c_n)$$
$$c_n = a_n c_{n+1}$$

Hence,

$$\frac{c}{d} = \frac{c_0}{c_1} \overset{(2.1)}{=} a_0 + \frac{c_2}{c_1} = a_0 + \frac{1}{c_1/c_2} \overset{(2.2)}{=} a_0 + \frac{1}{a_1 + \frac{c_3}{c_2}}$$

$$= a_0 + \cfrac{1}{a_1 + \cfrac{1}{c_2/c_3}} \overset{(2.3)}{=} a_0 + \cfrac{1}{a_1 + \cfrac{1}{a_2 + \frac{c_4}{c_3}}}$$

$$= \cdots = a_0 + \cfrac{1}{a_1 + \cfrac{1}{a_2 + \cfrac{1}{a_3 + \cfrac{1}{\ddots + \cfrac{1}{a_n}}}}} = [a_0; a_1, \ldots, a_n].$$

If $n \geq 1$, then because $c_{n+1} < c_n$ and $c_n = a_n c_{n+1}$, we have $a_n \geq 2$.

The uniqueness is proven in Section 9 of the Appendix. □

If q is a rational number, the representation $q = [a_0; a_1, \ldots, a_n]$ provided by Proposition 2.1 will be called the *canonical continued-fraction representation* of q.

Example. Find the canonical continued-fraction representation of 121/84.

Solution:

$$\begin{aligned} 121 &= 1 \cdot 84 + 37 \\ 84 &= 2 \cdot 37 + 10 \\ 37 &= 3 \cdot 10 + 7 \\ 10 &= 1 \cdot 7 + 3 \\ 7 &= 2 \cdot 3 + 1 \\ 3 &= 3 \cdot 1, \end{aligned}$$

so $121/84 = [1; 2, 3, 1, 2, 3]$.

V – 3 Explicit computation of $[a_0; a_1, \ldots, a_n]$

If $\alpha = (a_0; a_1, \ldots, a_n)$, where $a_k \in \mathbb{R}$ for all k and $a_k > 0$ for $k \geq 1$, then for each $k \in \{0, \ldots, n\}$, let

$$C_k(\alpha) = [a_0; a_1, \ldots, a_k] \in \mathbb{R}.$$

Define numbers $p_k(\alpha)$ and $q_k(\alpha)$ recursively by

$$\begin{aligned} p_0(\alpha) &= a_0 & q_0(\alpha) &= 1 \\ p_1(\alpha) &= a_1 a_0 + 1 & q_1(\alpha) &= a_1 \\ p_k(\alpha) &= a_k p_{k-1}(\alpha) + p_{k-2}(\alpha) & q_k(\alpha) &= a_k q_{k-1}(\alpha) + q_{k-2}(\alpha) \quad \text{for } k \geq 2 \end{aligned}$$

Theorem 3.1. *If $\alpha = (a_0; a_1, \ldots, a_n)$, where $a_k \in \mathbb{R}$ for all k and $a_k > 0$ for $k \geq 1$, then*

$$C_k(\alpha) = \frac{p_k(\alpha)}{q_k(\alpha)} \quad \textit{for all } k \in \{0, \ldots, n\}.$$

Before proving the theorem, let us make some observations, which are left as short exercises.

(i) If $\alpha = (a_0; a_1, a_2, \ldots, a_{k+1})$ and $\beta = (a_0; a_1, a_2, \ldots, a_{k-1}, a_k + \frac{1}{a_{k+1}})$, then $C_{k+1}(\alpha) = C_k(\beta)$.

(ii) If $\alpha = (a_0; a_1, \ldots, a_m)$ and $\beta = (b_0; b_1, \ldots, b_n)$, and if $k \leq \min(m, n)$ is such that $a_i = b_i$ for all $i \leq k$, then

  (a) $C_i(\alpha) = C_i(\beta)$ for all $i \leq k$, and

  (b) $p_i(\alpha) = p_i(\beta)$ and $q_i(\alpha) = q_i(\beta)$ for all $i \leq k$.

Proof. (Theorem 3.1) We prove by induction on $k \geq 0$ the statement that if $n \geq k$ and $\alpha = (a_0; a_1, \ldots, a_n)$, then $C_k(\alpha) = p_k(\alpha)/q_k(\alpha)$. The cases $k = 0, 1$ are obvious. Let us treat the case $k = 2$ separately. Let $\beta = (a_0; a_1 + \frac{1}{a_2})$. Then

$$\begin{aligned} C_2(\alpha) &= C_1(\beta) \quad \text{by (ii)(a) and (i)} \\ &= \frac{p_1(\beta)}{q_1(\beta)} \quad \text{by the } k = 1 \text{ case applied to } \beta \\ &= \frac{(a_1 + \frac{1}{a_2})a_0 + 1}{a_1 + \frac{1}{a_2}} = \frac{a_2(a_1 a_0 + 1) + a_0}{a_2 a_1 + 1} = \frac{a_2 p_1(\alpha) + p_0(\alpha)}{a_2 q_1(\alpha) + q_0(\alpha)} = \frac{p_2(\alpha)}{q_2(\alpha)}. \end{aligned}$$

Now let $k \geq 2$, and assume that the statement is true for this k. Let $\alpha = (a_0; a_1, \ldots, a_n)$ where $n \geq k + 1$, and let

$$\begin{aligned} \alpha' &= (a_0; a_1, a_2, \ldots, a_{k+1}) \\ \beta &= (a_0; a_1, a_2, \ldots, a_{k-1}, a_k + \tfrac{1}{a_{k+1}}) \end{aligned}$$

Then

$$C_{k+1}(\alpha) = C_{k+1}(\alpha') \quad \text{by (ii)(a)}$$

$$
\begin{aligned}
&= C_k(\beta) \quad \text{by (i)} \\
&= \frac{p_k(\beta)}{q_k(\beta)} \quad \text{by the inductive hypothesis applied to } \beta \\
&= \frac{(a_k + \frac{1}{a_{k+1}})p_{k-1}(\beta) + p_{k-2}(\beta)}{(a_k + \frac{1}{a_{k+1}})q_{k-1}(\beta) + q_{k-2}(\beta)} \\
&= \frac{(a_k + \frac{1}{a_{k+1}})p_{k-1}(\alpha) + p_{k-2}(\alpha)}{(a_k + \frac{1}{a_{k+1}})q_{k-1}(\alpha) + q_{k-2}(\alpha)} \quad \text{by (ii)(b)} \\
&= \frac{a_{k+1}(a_k p_{k-1}(\alpha) + p_{k-2}(\alpha)) + p_{k-1}(\alpha)}{a_{k+1}(a_k q_{k-1}(\alpha) + q_{k-2}(\alpha)) + q_{k-1}(\alpha)} \\
&= \frac{a_{k+1}p_k(\alpha) + p_{k-1}(\alpha)}{a_{k+1}q_k(\alpha) + q_{k-1}(\alpha)} \\
&= \frac{p_{k+1}(\alpha)}{q_{k+1}(\alpha)}.
\end{aligned}
$$

This completes the induction. □

The algorithm for finding the numbers C_k via the p_k and q_k will be called the (p, q)-*algorithm*.

Example. Use the (p, q)-algorithm to find $[4; 2, 5, 3]$ as an explicit rational number.

Solution: Let $\alpha = (4; 2, 5, 3)$. Then

$$
\begin{aligned}
p_0(\alpha) &= 4 & q_0(\alpha) &= 1 \\
p_1(\alpha) &= 2 \cdot 4 + 1 = 9 & q_1(\alpha) &= 2 \\
p_2(\alpha) &= 5 \cdot 9 + 4 = 49 & q_2(\alpha) &= 5 \cdot 2 + 1 = 11 \\
p_3(\alpha) &= 3 \cdot 49 + 9 = 156 & q_3(\alpha) &= 3 \cdot 11 + 2 = 35
\end{aligned}
$$

Thus, $[4; 2, 5, 3] = 156/35$. Note that this fraction is in lowest terms. We will soon see why.

Example. Find $[2; 3, 5, 7, 11]$ as an explicit rational number.

Solution: Let $\alpha = (2; 3, 5, 7, 11)$. Then

$$
\begin{aligned}
p_0(\alpha) &= 2 & q_0(\alpha) &= 1 \\
p_1(\alpha) &= 3 \cdot 2 + 1 = 7 & q_1(\alpha) &= 3 \\
p_2(\alpha) &= 5 \cdot 7 + 2 = 37 & q_2(\alpha) &= 5 \cdot 3 + 1 = 16 \\
p_3(\alpha) &= 7 \cdot 37 + 7 = 266 & q_3(\alpha) &= 7 \cdot 16 + 3 = 115 \\
p_4(\alpha) &= 11 \cdot 266 + 37 = 2963 & q_4(\alpha) &= 11 \cdot 115 + 16 = 1281
\end{aligned}
$$

Thus,

$$[2; 3, 5, 7, 11] = \frac{2963}{1281}.$$

V − 4 Towards infinite continued fractions

Proposition 4.1. *Let $\alpha = (a_0; a_1, \ldots, a_n)$ where $a_k \in \mathbb{Z}$ for all k and $a_k > 0$ for $k \geq 1$. Then*

$$p_k(\alpha)q_{k-1}(\alpha) - p_{k-1}(\alpha)q_k(\alpha) = (-1)^{k-1} \quad \textit{for all } k \geq 1.$$

Proof. This is done by induction on k. We abbreviate $p_k(\alpha), q_k(\alpha)$ to p_k, q_k. The case $k = 1$ is immediate: $p_1q_0 - p_0q_1 = a_1a_0 + 1 - a_0a_1 = 1$. Now let $k \geq 1$ and assume the equality for this k. Then

$$\begin{aligned} p_{k+1}q_k - p_kq_{k+1} = (a_{k+1}p_k + p_{k-1})q_k - p_k(a_{k+1}q_k + q_{k-1}) &= p_{k-1}q_k - p_kq_{k-1} \\ &= (-1)^k, \end{aligned}$$

the last equality by the inductive hypothesis. □

Corollary 4.2 (C-cor 1)**.** *In the notation of the proposition, $p_k(\alpha)$ and $q_k(\alpha)$ are coprime for all $k \geq 0$.*

Proof. The case $k = 0$ is obvious because $q_0 = 1$. If $k \geq 1$, then any positive common divisor of p_k and q_k divides $p_kq_{k-1} - p_{k-1}q_k = (-1)^{k-1}$ and so must be 1. □

Corollary 4.3 (C-cor 2)**.** *In the notation of the proposition,*

$$C_k(\alpha) - C_{k-1}(\alpha) = \frac{(-1)^{k-1}}{q_k(\alpha)q_{k-1}(\alpha)} \quad \textit{for all } k \geq 1,$$
$$C_k(\alpha) - C_{k-2}(\alpha) = \frac{(-1)^k a_k}{q_k(\alpha)q_{k-2}(\alpha)} \quad \textit{for all } k \geq 2.$$

Proof. Again omitting the α's, we have, for $k \geq 1$,

$$C_k - C_{k-1} = \frac{p_k}{q_k} - \frac{p_{k-1}}{q_{k-1}} = \frac{p_kq_{k-1} - p_{k-1}q_k}{q_kq_{k-1}} = \frac{(-1)^{k-1}}{q_kq_{k-1}}$$

by the proposition. If $k \geq 2$, then again forming a common denominator, we have

$$\begin{aligned} C_k - C_{k-2} = \frac{p_kq_{k-2} - p_{k-2}q_k}{q_kq_{k-2}} &= \frac{(a_kp_{k-1} + p_{k-2})q_{k-2} - p_{k-2}(a_kq_{k-1} + q_{k-2})}{q_kq_{k-2}} \\ &= \frac{a_k(p_{k-1}q_{k-2} - p_{k-2}q_{k-1})}{q_kq_{k-2}} = \frac{(-1)^k a_k}{q_kq_{k-2}} \end{aligned}$$

by the proposition again. □

Corollary 4.4 (C-cor 3)**.** *In the notation of the proposition,*

(i) $C_0(\alpha) < C_2(\alpha) < C_4(\alpha) < \cdots$,

(ii) $C_1(\alpha) > C_3(\alpha) > C_5(\alpha) > \cdots$,

(iii) $C_{2k}(\alpha) < C_{2j+1}(\alpha)$ *for all* $k, j \geq 0$ *(every odd is greater than every even).*

Proof. For the first two assertions, observe that, by C-cor 2,

$$C_k - C_{k-2} = \frac{(-1)^k a_k}{q_k q_{k-2}},$$

which is positive when k is even and negative when k is odd. For the last assertion, note that

$$C_{2m+1} - C_{2m} = \frac{(-1)^{2m}}{q_{2m+1}q_{2m}} = \frac{1}{q_{2m+1}q_{2m}} > 0,$$

i.e.,

$$C_{2m+1} > C_{2m}. \tag{4.1}$$

Hence, if $k, j \geq 0$,

$$C_{2j+1} \overset{\text{(ii)}}{\geq} C_{2j+2k+1} \overset{(4.1)}{>} C_{2j+2k} \overset{\text{(i)}}{\geq} C_{2k}.$$

□

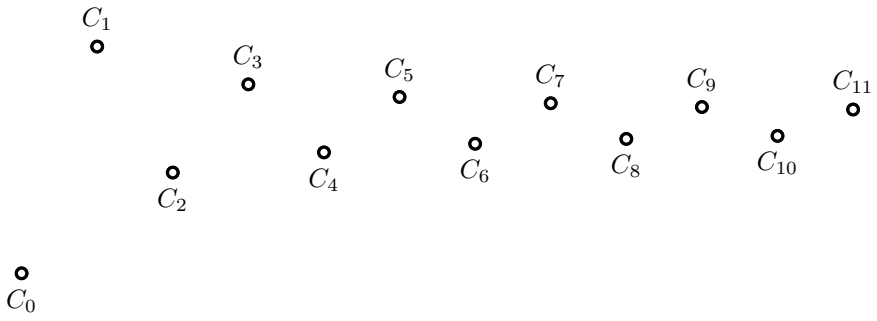
Theorem 4.5. *Let $(a_n)_{n\geq 0}$ be a sequence of integers with $a_n > 0$ for all $n \geq 1$. For each $k \geq 0$, let $C_k = [a_0; a_1, \ldots, a_k] \in \mathbb{Q}$. Then the sequence $(C_k)_{k\geq 0}$ converges in $\mathbb{R}$. We denote its limit by $[a_0; a_1, a_2, \ldots]$.*

Proof. By C-cor 3, the sequence $C_0, C_2, C_4, \ldots$ is monotone increasing and has an upper bound (e.g., C_1), so it has a limit x_- in $\mathbb{R}$. Similarly, the sequence $C_1, C_3, C_5, \ldots$, monotone decreasing and bounded below, has a limit x_+. Further, the proof of C-cor 3 shows that $C_{2m+1} - C_{2m} = 1/(q_{2m+1}q_{2m}) \to 0$ as $m \to \infty$, so

$$0 = \lim_{m\to\infty}(C_{2m+1} - C_{2m}) = \lim_{m\to\infty} C_{2m+1} - \lim_{m\to\infty} C_{2m} = x_+ - x_-.$$

Therefore, $(C_k)_{k\geq 0}$ converges to $x_+ = x_-$. □

Here is a visual illustration of the convergence of the $C_k(\alpha)$:



V – 5 Infinite continued fractions

Recall that if $(a_0; a_1, a_2, \ldots)$ is a sequence of integers with $a_k > 0$ for all $k \geq 1$, then $\lim_{n\to\infty}[a_0; a_1, a_2, \ldots, a_n]$ exists and is denoted $[a_0; a_1, a_2, \ldots]$. The following facts are proven in Section 10 of the Appendix:

(i) $[a_0; a_1, a_2, \ldots]$ is irrational.

(ii) Every irrational real number can be expressed as $x = [a_0; a_1, a_2, \ldots]$ for some sequence as above.

(iii) The integers $a_0, a_1, a_2, \ldots$ are uniquely determined by the irrational number x. That is, if $[a_0; a_1, a_2, \ldots] = [b_0; b_1, b_2, \ldots]$, then $a_n = b_n$ for all $n \geq 0$.

There is therefore a bijection between $\mathbb{R} \smallsetminus \mathbb{Q}$ and the set of integer sequences of the above kind. The rational numbers $[a_0; a_1, a_2, \ldots, a_k]$ are called the *convergents* to the irrational number $[a_0; a_1, a_2, \ldots]$.

The existence part of Section 10 of the Appendix shows that the continued-fraction representation of an irrational number x is $[a_0; a_1, a_2, \ldots]$ where

$$x_0 = x, \quad a_k = \lfloor x_k \rfloor, \quad x_{k+1} = \frac{1}{x_k - a_k} \quad (k \geq 0).$$

Example. The continued-fraction representation $[a_0; a_1, a_2, \ldots]$ of π^2 begins $\pi^2 = [9; 1, 6, 1, \ldots]$. Indeed, the first few iterations of the above algorithm are

$$\begin{aligned} x_0 &= \pi^2 \approx 9.870, & a_0 &= 9 \\ x_1 &= \frac{1}{x_0 - 9} \approx 1.150, & a_1 &= 1 \\ x_2 &= \frac{1}{x_1 - 1} \approx 6.669, & a_2 &= 6 \\ x_3 &= \frac{1}{x_2 - 6} \approx 1.495, & a_3 &= 1 \end{aligned}$$

Quadratic irrationals and periodic continued fractions

Let $x \in \mathbb{R} \smallsetminus \mathbb{Q}$. We will say that x has a *periodic* continued-fraction representation $[a_0; a_1, a_2, \ldots]$ if it takes the form $[a_0; a_1, \ldots, a_{k-1}, b_0, \ldots, b_{l-1}, b_0, \ldots, b_{l-1}, \ldots]$ for some repeating sequence $b_0, \ldots, b_{l-1}$, where $k \geq 0$ and $l \geq 1$. A common notation in this situation is to indicate the repeating sequence by a bar:

$$[a_0; a_1, \ldots, a_{k-1}, b_0, \ldots, b_{l-1}, b_0, \ldots, b_{l-1}, \ldots] = [a_0; a_1, \ldots, a_{k-1}, \overline{b_0, \ldots, b_{l-1}}].$$

Example.

$$\begin{aligned} [\overline{1}] &= [1; 1, 1, 1, \ldots] \\ [\overline{1; 2}] &= [1; 2, 1, 2, \ldots] \\ [1; \overline{2, 3, 4}] &= [1; 2, 3, 4, 2, 3, 4, \ldots] \\ [1; 2, \overline{3, 4, 5}] &= [1; 2, 3, 4, 5, 3, 4, 5, \ldots] \end{aligned}$$

The periodic continued fractions have a number-theoretic description. A real number is called a *quadratic irrational* if it is a root of a polynomial $x^2 + bx + c$ where $b, c \in \mathbb{Q}$ and $b^2 - 4c$ is not the square of a rational number. The quadratic irrationals are the real numbers of the form $(u \pm \sqrt{d})/v$ where $u, v, d \in \mathbb{Z}$, $v, d > 0$, and d is not square.

Theorem 5.1. *A real number is a quadratic irrational if and only if it has a periodic continued-fraction representation.*

The proof that a quadratic irrational has a periodic continued-fraction representation is given in Section 12 of the Appendix. The other direction, that $[a_0; a_1, a_2, \ldots]$ is a quadratic irrational if the sequence of integers is periodic, is easier. For this, suppose that $x = [a_0; a_1, \ldots, a_{k-1}, \overline{b_0, \ldots, b_{l-1}}]$, and observe that

$$x = [a_0; a_1, \ldots, a_{k-1}, y] \tag{5.1}$$

where

$$y = [\overline{b_0; b_1, \ldots, b_{l-1}}] = [b_0; b_1, \ldots, b_{l-1}, \overline{b_0, \ldots, b_{l-1}}] = [b_0; b_1, \ldots, b_{l-1}, y]. \tag{5.2}$$

We can now use the (p, q)-algorithm (Section 3) to express y in terms of itself. Specifically, if $\beta = (b_0; b_1, \ldots, b_{l-1}, y)$, then

$$y = [b_0; b_1, \ldots, b_{l-1}, y] = \frac{p_l(\beta)}{q_l(\beta)} = \frac{y p_{l-1}(\beta) + p_{l-2}(\beta)}{y q_{l-1}(\beta) + q_{l-2}(\beta)},$$

and rearranging gives

$$q_{l-1}(\beta) y^2 + (q_{l-2}(\beta) - p_{l-1}(\beta)) y - p_{l-2}(\beta) = 0,$$

showing that y is a rational or a quadratic irrational. We now use the (p, q)-algorithm again, but this time on $\alpha = (a_0; a_1, \ldots, a_{k-1}, y)$:

$$x = [a_0; a_1, \ldots, a_{k-1}, y] = \frac{p_k(\alpha)}{q_k(\alpha)} = \frac{y p_{k-1}(\alpha) + p_{k-2}(\alpha)}{y q_{k-1}(\alpha) + q_{k-2}(\alpha)}.$$

This quotient is either a rational or a quadratic irrational because the same is true of the numerator and denominator. However, x cannot be rational, because it has an infinite continued-fraction representation. Thus, x is a quadratic irrational.

Remark. Our proof tacitly assumed that $k, l \geq 2$. We leave it as an exercise to make the necessary modifications in the remaining cases.

Remark. The justification for the equality at (5.1) and the final equality in (5.2) is given by Proposition 11.2 in the Appendix.

The *period length* of the continued-fraction representation of a quadratic irrational is the minimum period of the repeating part of the sequence. For example, the period length of the quadratic irrational $[5; 2, 6, \overline{10, 9, 8, 7}]$ is 4. The period length of $[5; 2, 6, \overline{3, 7, 3, 7}]$ may also appear to be 4 at first sight, but it is in fact 2, as the periodic sequence (3,7,3,7,...) has minimum period 2. In fact, we would usually prefer to write this second quadratic irrational as $[5; 2, 6, \overline{3, 7}]$.

V – 6 Examples of quadratic irrationals as continued fractions

Let us give an example of the fact that a quadratic irrational has a periodic continued-fraction representation.

Example. Find the continued-fraction representation of $(23 - \sqrt{37})/12$.

Solution: We construct the x_k and a_k as in the usual algorithm for finding the continued-fraction representation of an irrational number. If $x = (23 - \sqrt{37})/12$, then

$$\begin{aligned}
x_0 &= x = \frac{23 - \sqrt{37}}{12} \approx 1.410, && a_0 = 1 \\
x_1 &= \frac{1}{\frac{23-\sqrt{37}}{12} - 1} = \frac{12}{11 - \sqrt{37}} \\
&= \frac{12(11 + \sqrt{37})}{84} = \frac{11 + \sqrt{37}}{7} \approx 2.440, && a_1 = 2 \\
x_2 &= \frac{1}{\frac{11+\sqrt{37}}{7} - 2} = \cdots = \frac{3 + \sqrt{37}}{4} \approx 2.271, && a_2 = 2 \\
x_3 &= \frac{1}{\frac{3+\sqrt{37}}{4} - 2} = \cdots = \frac{5 + \sqrt{37}}{3} \approx 3.694, && a_3 = 3 \\
x_4 &= \frac{1}{\frac{5+\sqrt{37}}{3} - 3} = \cdots = \frac{4 + \sqrt{37}}{7} \approx 1.440, && a_4 = 1 \\
x_5 &= \frac{1}{\frac{4+\sqrt{37}}{7} - 1} = \cdots = \frac{3 + \sqrt{37}}{4} = x_2.
\end{aligned}$$

Since $x_5 = x_2$, it follows that we have the repeating pattern

$$(a_2, a_3, a_4, a_5, a_6, \ldots) = (2, 3, 1, 2, 3, 1, 2, 3, 1, \ldots),$$

so $(23 - \sqrt{37})/12 = [1; 2, \overline{2, 3, 1}]$. The period length, incidentally, is 3.

Remark. Note that we do not carry forward the approximations from each step to the next, instead using each approximation only to obtain the current integer a_k. This approach eliminates the possibility of rounding errors accumulating. Once we have found a_k, we compute x_{k+1} using a_k and the exact value of x_k, not any approximation of x_k.

Next, we illustrate how to obtain a quadratic irrational explicitly from a periodic continued fraction.

Example. Find $[\overline{2; 1}]$ as an explicit quadratic irrational.

Solution: Let $x = [\overline{2; 1}]$, and note that $x = [2; 1, x]$. The (p, q)-algorithm gives

$$\begin{aligned}
p_0 &= 2 & q_0 &= 1 \\
p_1 &= 3 & q_1 &= 1 \\
p_2 &= 3x + 2 & q_2 &= x + 1
\end{aligned}$$

Hence,

$$\begin{aligned} x &= \frac{3x+2}{x+1}, \\ \text{i.e.,} \quad x^2 + x &= 3x + 2, \\ \text{i.e.,} \quad x^2 - 2x - 2 &= 0, \\ \text{i.e.,} \quad x &= \frac{1}{2}(2 \pm \sqrt{12}) \\ &= 1 \pm \sqrt{3}. \end{aligned}$$

But $x = [\overline{2;1}] > 2$, so we must have $x = 1 + \sqrt{3}$.

Example. Find $[-2; 3, \overline{5, 2, 2}]$ as an explicit quadratic irrational.

Solution: First, let $y = [\overline{5; 2, 2}]$. Then $y = [5; 2, 2, y]$, so the (p, q)-algorithm gives

$$\begin{aligned} p_0 &= 5 & q_0 &= 1 \\ p_1 &= 11 & q_1 &= 2 \\ p_2 &= 27 & q_2 &= 5 \\ p_3 &= 27y + 11 & q_3 &= 5y + 2 \end{aligned}$$

Therefore,

$$y = [5; 2, 2, y] = \frac{27y + 11}{5y + 2},$$

i.e., $5y^2 - 25y - 11 = 0$, i.e., $y = \frac{1}{10}(25 \pm \sqrt{845})$. But $y = [\overline{5; 2, 2}] > 5$, so we must have $y = \frac{1}{10}(25 + \sqrt{845})$. For $x = [-2; 3, y]$, we use the (p, q)-algorithm again:

$$\begin{aligned} p_0 &= -2 & q_0 &= 1 \\ p_1 &= -5 & q_1 &= 3 \\ p_2 &= -5y - 2 & q_2 &= 3y + 1 \end{aligned}$$

Then

$$\begin{aligned} x = [-2; 3, y] = -\frac{5y+2}{3y+1} &= -\frac{10(5y+2)}{10(3y+1)} \\ &= -\frac{125 + 5\sqrt{845} + 20}{75 + 3\sqrt{845} + 10} \\ &= -\frac{145 + 5\sqrt{845}}{85 + 3\sqrt{845}} \\ &= -\frac{(145 + 5\sqrt{845})(85 - 3\sqrt{845})}{85^2 - 845 \cdot 3^2} \\ &= -\frac{35 + \sqrt{845}}{38}. \end{aligned}$$

V – 7 Quadratic irrationals defined by regular expressions

Consider the following problem: For an integer $d \geq 2$, show that

$$\sqrt{d^2-1} = [d-1; \overline{1, 2d-2}].$$

We may prove this by beginning with $[d-1; \overline{1, 2d-2}]$ and using the (p, q)-algorithm to obtain the desired equality. To that end, let $y = [\overline{1; 2d-2}]$, and note that $y = [1; 2d-2, y]$. The (p, q)-algorithm applied to the sequence $(1; 2d-2, y)$ gives

$$\begin{aligned} p_0 &= 1 & q_0 &= 1 \\ p_1 &= 2d-1 & q_1 &= 2d-2 \\ p_2 &= (2d-1)y+1 & q_2 &= (2d-2)y+1 \end{aligned}$$

Therefore,

$$y = \frac{(2d-1)y+1}{(2d-2)y+1},$$

and rearranging this equation yields $y^2 - y - \frac{1}{2d-2} = 0$, so that

$$y = \frac{1}{2}\left(1 \pm \sqrt{1 + \frac{2}{d-1}}\right). \tag{7.1}$$

But $y = [\overline{1; 2d-2}] > 1$, so the sign in (7.1) is a plus. Hence,

$$\begin{aligned} [d-1; \overline{1, 2d-2}] &= [d-1; y] \\ &= d-1+\frac{1}{y} \\ &= d-1+\frac{2}{1+\sqrt{1+\frac{2}{d-1}}} \\ &= d-1+\frac{2\left(\sqrt{1+\frac{2}{d-1}}-1\right)}{2/(d-1)} \quad \text{(rationalizing the denominator)} \\ &= d-1+(d-1)\left(\sqrt{1+\tfrac{2}{d-1}}-1\right) \\ &= \sqrt{(d-1)^2+2(d-1)} \\ &= \sqrt{d^2-1}. \end{aligned}$$

An alternative method

Suppose that, in the above problem, we had not been given an equality to prove, but instead had been given only the expression $\sqrt{d^2-1}$ and been asked to find its continued-fraction representation in terms of d. Let us illustrate how one might tackle such a problem. We will use a different expression for our example.

Example. Let d be a positive integer. Find the continued-fraction representation of $\sqrt{d^2+2}$ in terms of d.

Solution: We use the usual algorithm for finding the continued-fraction representation of a real number, beginning with $x_0 = x = \sqrt{d^2+2}$. It is clear that $x_0 > d$, but also $x_0 < d+1$, as we may see as follows:

$$\begin{aligned}\sqrt{d^2+2} < d+1 &\iff d^2+2 < (d+1)^2 = d^2+2d+1,\\ &\iff 1 < 2d.\end{aligned}$$

Because the last assertion is true and we have $\iff$ the whole way, the original claim is true as well. Therefore, $a_0 = \lfloor x_0 \rfloor = d$.

Next, let

$$x_1 = \frac{1}{x_0 - a_0} = \frac{1}{\sqrt{d^2+2} - d} = \frac{\sqrt{d^2+2}+d}{2} > d.$$

We claim that $x_1 < d+1$:

$$\begin{aligned}\frac{\sqrt{d^2+2}+d}{2} < d+1 &\iff \sqrt{d^2+2} < d+2,\\ &\iff d^2+2 < d^2+4d+4,\\ &\iff 0 < 4d+2.\end{aligned}$$

Hence, because $d < x_1 < d+1$, we have $a_1 = \lfloor x_1 \rfloor = d$.

Continuing, we let

$$\begin{aligned}x_2 = \frac{1}{x_1 - a_1} = \frac{1}{\frac{\sqrt{d^2+2}+d}{2} - d} &= \frac{2}{\sqrt{d^2+2} - d}\\ &= \frac{2(\sqrt{d^2+2}+d)}{2}\\ &= \sqrt{d^2+2}+d,\end{aligned}$$

whose floor is $d+d = 2d$, so $a_2 = 2d$.

Finally, we let

$$x_3 = \frac{1}{x_2 - a_2} = \frac{1}{\sqrt{d^2+2} - d} = x_1.$$

Thus, $\sqrt{d^2+2} = [d; \overline{d, 2d}]$. Observe that the period length is 2.

V – 8 The solutions to Pell's equation

We may now solve Pell's equation. Let d be a positive integer that is not a square.

Theorem 8.1. *Let $\epsilon \in \{1, -1\}$, let n be the period length of the continued-fraction representation of $\sqrt{d}$, and let p_k, q_k be the numbers appearing in the (p, q)-algorithm for the continued-fraction representation of $\sqrt{d}$. Then the solutions $(x, y) \in \mathbb{Z}_{\geq 1} \times \mathbb{Z}_{\geq 1}$ to the equation $x^2 - dy^2 = \epsilon$ are the pairs (p_{rn-1}, q_{rn-1}) for which the positive integer r satisfies $(-1)^{rn} = \epsilon$.*

A proof is given in Section 14 of the Appendix.

Example. Find the first three positive solutions to $x^2 - 6y^2 = 1$, and decide whether the equation $x^2 - 6y^2 = -1$ has a solution.

Solution: We begin by finding the periodic continued-fraction representation of $\sqrt{6}$. In fact, since $6 = d^2 + 2$ with $d = 2$, we may use the fact that $\sqrt{d^2 + 2} = [d; \overline{d, 2d}]$, as we saw in Section 7. Thus,

$$\sqrt{6} = [2; \overline{2, 4}].$$

Observe that the period length is 2, so the solutions to $x^2 - 6y^2 = 1$ are (p_{2r-1}, q_{2r-1}) where $(-1)^{2r} = 1$, i.e., $1^r = 1$. There is consequently no restriction on r, so the solutions are simply (p_{2r-1}, q_{2r-1}) with r running through all positive integers, i.e.,

$$(p_1, q_1), (p_3, q_3), (p_5, q_5), \ldots.$$

The (p, q)-algorithm yields

$$\begin{aligned} p_0 &= 2 & q_0 &= 1 \\ p_1 &= 5 & q_1 &= 2 \\ p_2 &= 22 & q_2 &= 9 \\ p_3 &= 49 & q_3 &= 20 \\ p_4 &= 218 & q_4 &= 89 \\ p_5 &= 485 & q_5 &= 198 \end{aligned}$$

Therefore, the first three solutions are $(5, 2)$, $(49, 20)$, and $(485, 198)$.

Finally, since the period length $n = 2$ is even, there are no integers r satisfying $(-1)^{rn} = -1$, so the equation $x^2 - 6y^2 = -1$ has no integral solutions.

Example. Use the (p, q)-algorithm on the periodic continued-fraction representation $\sqrt{41} = [6; \overline{2, 2, 12}]$ to find the first positive solution to $x^2 - 41y^2 = -1$ and the first positive solution to $x^2 - 41y^2 = 1$.

Solution: We can see right away from the fact that the period length $n = 3$ is odd that the negative Pell equation does indeed have solutions, and they are (p_{3r-1}, q_{3r-1}) where $(-1)^{3r} = -1$, i.e., $(-1)^r = -1$, i.e., r is odd. The first is therefore (p_2, q_2). Further, the

solutions to $x^2 - 41y^2 = 1$ are (p_{3r-1}, q_{3r-1}) where $(-1)^{3r} = 1$, i.e., $(-1)^r = 1$, i.e., r is even. The first solution to the positive Pell equation therefore corresponds to $r = 2$, i.e., is (p_5, q_5). We now perform the (p, q)-algorithm:

$$\begin{aligned} p_0 &= 6 & q_0 &= 1 \\ p_1 &= 13 & q_1 &= 2 \\ p_2 &= 32 & q_2 &= 5 \\ p_3 &= 397 & q_3 &= 62 \\ p_4 &= 826 & q_4 &= 129 \\ p_5 &= 2049 & q_5 &= 320 \end{aligned}$$

Thus, the first solution to the negative equation is $(32, 5)$, and the first solution to the positive equation is $(2049, 320)$.

The following example illustrates a slight modification of the above strategy, one that is a little more efficient.

Example. By calculating the q's, but not the p's, in the (p, q)-algorithm for $\sqrt{31} = [5; \overline{1, 1, 3, 5, 3, 1, 1, 10}]$, find the first positive solution to the equation $x^2 - 31y^2 = 1$.

Solution: The period length is 8, so the solutions to the positive equation are (p_{8r-1}, q_{8r-1}) where $(-1)^{8r} = 1$, i.e., $1^r = 1$. There is no restriction on r, so the first solution corresponds to $r = 1$ and is therefore (p_7, q_7). Note that the equality $p_7^2 - 31q_7^2 = 1$ gives $p_7 = \sqrt{31q_7^2 + 1}$, so it suffices to find q_7 and from there obtain p_7. The (p, q)-algorithm can be applied to the q's alone, without reference to the p's, and one finds that $q_0, \ldots, q_7$ are $1, 1, 2, 7, 37, 118, 155, 273$. Hence, $p_7 = \sqrt{31 \cdot 273^2 + 1} = 1520$, so the first positive solution to $x^2 - 31y^2 = 1$ is $(1520, 273)$.

Determining whether the negative Pell equation has a solution

From Theorem 8.1, we see that the equation $x^2 - dy^2 = -1$, i.e., the negative equation, has a solution if and only if the period length of $\sqrt{d}$ is odd. However, sometimes, one can see that there is no solution to the negative Pell equation simply by reducing modulo some appropriate modulus.

Example. By choosing an appropriate modulus, show that there are no integral solutions to $x^2 - 15y^2 = -1$.

Solution: If a solution existed, then reducing mod 3, we would have $x^2 \equiv -1 \bmod 3$. But we know that -1 is not square mod 3, so no solution exists.

Exercise. One may show directly that -1 is square mod 146. Does it follow that the equation $x^2 - 146y^2 = -1$ *does* have integral solutions? Are there, in fact, any integral solutions to this equation? How can you find out one way or the other?

Appendix

Appendix: 1 Roots of unity in $\mathbb{C}$

We prove that for each positive integer n, there are n distinct nth roots of unity in $\mathbb{C}$. The well-known formulas from trigonometry for $\cos(x+y)$ and $\sin(x+y)$ show that

$$\big(\cos(x)+i\sin(x)\big)\big(\cos(y)+i\sin(y)\big)=\cos(x+y)+i\sin(x+y)$$

for all $x, y \in \mathbb{R}$, and then induction on n yields

$$\big(\cos(x)+i\sin(x)\big)^n=\cos(nx)+i\sin(nx) \tag{1.1}$$

for all $n \in \mathbb{Z}_{\geq 0}$. It then follows that (1.1) holds for all integers n, simply by the fact that $\big(\cos(x)+i\sin(x)\big)^{-1}=\cos(x)-i\sin(x)=\cos(-x)+i\sin(-x)$. The formula in (1.1) is known as *de Moivre's formula*.

Now consider the complex numbers

$$\zeta_k=\cos(\tfrac{2\pi k}{n})+i\sin(\tfrac{2\pi k}{n})$$

with $k \in \{0,\ldots,n-1\}$. They are distinct by basic properties of cos and sin, and it follows immediately from de Moivre's formula that

$$\zeta_k^n=\cos(2\pi k)+i\sin(2\pi k)=1$$

for all k. Thus, $\zeta_0,\ldots,\zeta_{n-1}$ constitute n distinct nth roots of unity in $\mathbb{C}$.

Appendix: 2 Proof of quadratic reciprocity

Many proofs of quadratic reciprocity exist. We follow the one given in Lang's *Algebraic Number Theory* [4, Chapter IV, Sect. 2]. Its value may be found not only in its demonstrating the truth of quadratic reciprocity, but also in its providing, in addition, a method for constructing square roots of integers explicitly in terms of roots of unity. We hope that this additional benefit will justify the choice of a proof of quadratic reciprocity that is longer than some others. (See Section 1 above for a construction of roots of unity in $\mathbb{C}$.)

Recall that for an odd prime p and an integer a, the Legendre symbol $\left(\frac{a}{p}\right)$ depends only on the residue class of a mod p. Therefore, we may define $\left(\frac{\alpha}{p}\right)$ for a residue class $\alpha \in \mathbb{Z}/p\mathbb{Z}$ by $\left(\frac{\alpha}{p}\right)=\left(\frac{a}{p}\right)$ for any a with $[a]=\alpha$.

Lemma 2.1. *Let $n \geq 2$. If $\omega \in \mathbb{C}$ is an nth root of unity other than 1, then $\sum_{k=0}^{n-1}\omega^k=0$.*

Proof. Because $\omega^n=1$, we have

$$0=\omega^n-1=(\omega-1)\left(\sum_{k=0}^{n-1}\omega^k\right).$$

Therefore, since $\omega \neq 1$, the sum is zero. □

Fix a primitive pth root of unity $\zeta \in \mathbb{C}$, i.e., a root of unity of order p. If $\alpha \in \mathbb{Z}/p\mathbb{Z}$, define $\zeta^\alpha = \zeta^a$ where a is any integer satisfying $[a] = \alpha$. This is well defined, since $\zeta^a = \zeta^b$ whenever $a \equiv b \bmod p$. Hence, we may define

$$S_p = \sum_{\alpha \in (\mathbb{Z}/p\mathbb{Z})^\times} \left(\frac{\alpha}{p}\right) \zeta^\alpha.$$

This is called a *Gauss sum*.

Proposition 2.2. *With notation as above,*

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

Proof. For brevity, let $G = (\mathbb{Z}/p\mathbb{Z})^\times$. Then

$$\begin{aligned}
S_p^2 &= \left(\sum_{\alpha \in G} \left(\frac{\alpha}{p}\right) \zeta^\alpha\right) \left(\sum_{\beta \in G} \left(\frac{\beta}{p}\right) \zeta^\beta\right) \\
&= \sum_{\alpha \in G} \sum_{\beta \in G} \left(\frac{\alpha\beta}{p}\right) \zeta^{\alpha+\beta} \\
&= \sum_{\alpha \in G} \sum_{\beta \in G} \left(\frac{\alpha\beta^2}{p}\right) \zeta^{\alpha\beta+\beta} \quad \text{(replace } \alpha \text{ with } \alpha\beta\text{)} \\
&= \sum_{\alpha \in G} \sum_{\beta \in G} \left(\frac{\alpha}{p}\right) \zeta^{(\alpha+1)\beta} \\
&= \sum_{\beta \in G} \left(\frac{-1}{p}\right) + \sum_{\alpha \in G \smallsetminus \{[-1]\}} \left(\frac{\alpha}{p}\right) \sum_{\beta \in G} (\zeta^{\alpha+1})^\beta \\
&= \left(\frac{-1}{p}\right)(p-1) + \sum_{\alpha \in G \smallsetminus \{[-1]\}} \left(\frac{\alpha}{p}\right) \sum_{\beta \in G} (\zeta^{\alpha+1})^\beta. \qquad (2.1)
\end{aligned}$$

Now, if $\alpha \in G \smallsetminus \{[-1]\}$, then $\zeta^{\alpha+1}$ is a pth root of unity different from 1, so by Lemma 2.1, $\sum_{\beta \in G} (\zeta^{\alpha+1})^\beta = -1$. Therefore, continuing from (2.1), we have

$$\begin{aligned}
S_p^2 &= \left(\frac{-1}{p}\right)(p-1) - \sum_{\alpha \in G \smallsetminus \{[-1]\}} \left(\frac{\alpha}{p}\right) \\
&= \left(\frac{-1}{p}\right) p - \sum_{\alpha \in G} \left(\frac{\alpha}{p}\right) \\
&= \left(\frac{-1}{p}\right) p,
\end{aligned}$$

the last equality holding because $\left(\frac{\alpha}{p}\right)$ takes the value 1 as many times as it does -1 as α runs through the prime residue classes mod p; see Proposition 6.1 in Section II. □

If A is a commutative ring and q a prime number, we will write $a \equiv b \bmod qA$ for elements $a, b \in A$ if $a - b$ is in the ideal qA of A.

Lemma 2.3. *If A is a commutative ring and q a prime number, then $(a+b)^q \equiv a^q + b^q \bmod qA$ for all $a, b \in A$.*

Proof. Because q is prime, $\binom{q}{k} \equiv 0 \bmod q$ for every $k \in \{1, \ldots, q-1\}$, so expanding out $(a+b)^q$ using the binomial theorem, we obtain the result. □

Lemma 2.4. *If ω is a root of unity, then $\mathbb{Q} \cap \mathbb{Z}[\omega] = \mathbb{Z}$.*

Proof. Let $a \in \mathbb{Q} \cap \mathbb{Z}[\omega]$. Then a is an algebraic integer (because ω is) and a rational number. Being an algebraic integer, a is a root of a monic polynomial in $\mathbb{Z}[x]$. But a rational root of such a polynomial must lie in $\mathbb{Z}$. □

Theorem 2.5. *Let p and q be distinct odd primes. Then*

$$\left(\frac{-1}{p}\right) = (-1)^{\frac{p-1}{2}},$$
$$\text{and} \quad \left(\frac{p}{q}\right) = (-1)^{\frac{p-1}{2}\frac{q-1}{2}} \left(\frac{q}{p}\right).$$

Proof. The first equality follows from Proposition 6.2 in Section II: Take $a = -1$ and note that, because p divides $\left(\frac{-1}{p}\right) - (-1)^{\frac{p-1}{2}} \in \{-2, 0, 2\}$ and p is odd, we must have $\left(\frac{-1}{p}\right) = (-1)^{\frac{p-1}{2}}$.

For the second equality, we compute $S_p^q \bmod qA$ in two different ways, where $A = \mathbb{Z}[\zeta]$, with ζ being a primitive pth root of unity. On the one hand,

$$\begin{aligned}
S_p^q &= S_p (S_p^2)^{\frac{q-1}{2}} \\
&= S_p \left(\frac{-1}{p}\right)^{\frac{q-1}{2}} p^{\frac{q-1}{2}} \quad \text{by Proposition 2.2} \\
&= S_p (-1)^{\frac{p-1}{2}\frac{q-1}{2}} p^{\frac{q-1}{2}} \quad \text{by the first equality of the theorem} \\
&\equiv S_p (-1)^{\frac{p-1}{2}\frac{q-1}{2}} \left(\frac{p}{q}\right) \quad \bmod qA \quad \text{by Proposition 6.2 in Section II again.}
\end{aligned}$$

On the other hand, treating S_p^q another way, we have

$$\begin{aligned}
S_p^q &= \left(\sum_{\alpha \in G} \left(\frac{\alpha}{p}\right) \zeta^p\right)^q \\
&\equiv \sum_{\alpha \in G} \left(\frac{\alpha}{p}\right)^q \zeta^{q\alpha} \quad \bmod qA \quad \text{by Lemma 2.3} \\
&= \sum_{\alpha \in G} \left(\frac{\alpha}{p}\right) \zeta^{q\alpha} \quad \text{because } q \text{ is odd} \\
&= \left(\frac{q}{p}\right) \sum_{\alpha \in G} \left(\frac{q\alpha}{p}\right) \zeta^{q\alpha} \\
&= \left(\frac{q}{p}\right) \sum_{\alpha \in G} \left(\frac{\alpha}{p}\right) \zeta^{\alpha} \quad (\text{replace } q\alpha \text{ with } \alpha)
\end{aligned}$$

$$= \left(\frac{q}{p}\right) S_p.$$

Combining this with the above, we have

$$\begin{aligned}
S_p(-1)^{\frac{p-1}{2}\frac{q-1}{2}} \left(\frac{p}{q}\right) &\equiv S_p \left(\frac{q}{p}\right) \mod qA, \\
\text{so} \quad S_p^2(-1)^{\frac{p-1}{2}\frac{q-1}{2}} \left(\frac{p}{q}\right) &\equiv S_p^2 \left(\frac{q}{p}\right) \mod qA, \\
\text{i.e.,} \quad \left(\frac{-1}{p}\right) p(-1)^{\frac{p-1}{2}\frac{q-1}{2}} \left(\frac{p}{q}\right) &\equiv \left(\frac{-1}{p}\right) p \left(\frac{q}{p}\right) \mod qA, \\
\text{i.e.,} \quad p(-1)^{\frac{p-1}{2}\frac{q-1}{2}} \left(\frac{p}{q}\right) &\equiv p \left(\frac{q}{p}\right) \mod qA.
\end{aligned}$$

Because p is invertible mod q, we deduce that

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

i.e., $a/q \in A$ where $a = (-1)^{\frac{p-1}{2}\frac{q-1}{2}} \left(\frac{p}{q}\right) - \left(\frac{q}{p}\right)$. Hence, $a/q \in \mathbb{Q} \cap A = \mathbb{Z}$ by Lemma 2.4, so $a \in q\mathbb{Z}$. Therefore, since $a \in \{-2, 0, 2\}$, and since $q \geq 3$, $a = 0$. □

Complementing Theorem 2.5, we have the following.

Proposition 2.6. *If p is an odd prime, then*

$$\left(\frac{2}{p}\right) = (-1)^{(p^2-1)/8}.$$

Proof. We give a proof that follows similar lines to that of Theorem 2.5. Let $G = (\mathbb{Z}/8\mathbb{Z})^\times$, and define

$$\begin{aligned}
\chi : G &\to \{1, -1\} \\
\alpha &\mapsto \begin{cases} 1 & \text{if } \alpha = [1] \text{ or } [7], \\ -1 & \text{if } \alpha = [3] \text{ or } [5]. \end{cases}
\end{aligned}$$

By a slight abuse of notation, if a is an odd integer then $\chi(a)$ will mean $\chi([a])$. Note that $\chi(\alpha\beta) = \chi(\alpha)\chi(\beta)$. Now fix a primitive 8th root of unity ζ, and let

$$S_2 = \sum_{\alpha \in G} \chi(\alpha)\zeta^\alpha.$$

A short calculation shows that $S_2^2 = 8$. Therefore,

$$\begin{aligned}
S_2^p &= S_2 (S_2^2)^{\frac{p-1}{2}} \\
&= S_2 \cdot 8^{\frac{p-1}{2}} \\
&\equiv S_2 \left(\frac{8}{p}\right) \mod pA \quad \text{by Proposition 6.2 in Section II, where } A = \mathbb{Z}[\zeta]
\end{aligned}$$

$$= S_2 \left(\frac{2}{p}\right).$$

But we also have

$$\begin{aligned} S_2^p &= \left(\sum_{\alpha\in G} \chi(\alpha)\zeta^\alpha\right)^p \\ &\equiv \sum_{\alpha\in G} \chi(\alpha)^p \zeta^{p\alpha} \mod pA \quad \text{by Lemma 2.3} \\ &= \sum_{\alpha\in G} \chi(\alpha)\zeta^{p\alpha} \quad \text{because } p \text{ is odd} \\ &= \chi(p) \sum_{\alpha\in G} \chi(p\alpha)\zeta^{p\alpha} \\ &= \chi(p) \sum_{\alpha\in G} \chi(\alpha)\zeta^{\alpha} \quad (\text{replace } p\alpha \text{ with } \alpha) \\ &= \chi(p) S_2. \end{aligned}$$

Hence, $S_2\left(\frac{2}{p}\right) \equiv S_2\chi(p) \bmod pA$, so $S_2^2\left(\frac{2}{p}\right) \equiv S_2^2\chi(p) \bmod pA$, i.e., $8\left(\frac{2}{p}\right) \equiv 8\chi(p) \bmod pA$, and so $\left(\frac{2}{p}\right) \equiv \chi(p) \bmod pA$ because 8 is invertible mod p. By the same argument as appears at the end of the proof of Theorem 2.5, this congruence implies that $\left(\frac{2}{p}\right) = \chi(p)$. It is an easy matter to check, by referring to the definition of χ, that $\chi(p) = (-1)^{(p^2-1)/8}$. □

Appendix: 3 Determination of the moduli admitting primitive roots

We determine which moduli admit primitive roots, adopting the approach taken in Rosen's book [6, Chap. 9]. The proofs of Lemma 3.1 and Theorem 3.3 below, while based on proofs in Rosen's book, are also the same in essence as arguments given in Serre's book [8, Chap. I, Sect. 1.2].

Lemma 3.1. *Let p be a prime, and for each positive divisor d of $p-1$, let*

$$f(d) = \#\{\alpha \in (\mathbb{Z}/p\mathbb{Z})^\times \mid \operatorname{ord}(\alpha) = d\}.$$

Then $f(d) \le \phi(d)$.

Proof. We show more, namely, that $f(d)$ is either 0 or $\phi(d)$. Assume, then, that $f(d) > 0$, so that there is at least one $\alpha \in (\mathbb{Z}/p\mathbb{Z})^\times$ such that $\operatorname{ord}(\alpha) = d$. Then the powers $\alpha^0, \alpha^1, \ldots, \alpha^{d-1}$ are d distinct roots of the polynomial $x^d - 1 \in \mathbb{F}_p[x]$, where $\mathbb{F}_p$ is the field $\mathbb{Z}/p\mathbb{Z}$. But a polynomial of degree d over a field cannot have more than d roots in that field, so $\alpha^0, \alpha^1, \ldots, \alpha^{d-1}$ must be all of the roots of $x^d - 1$. Therefore, every $\beta \in (\mathbb{Z}/p\mathbb{Z})^\times$ of order d, being necessarily a root of $x^d - 1$, is a power of α. Hence,

$$f(d) = \#\{k \in \{0, \ldots, d-1\} \mid \operatorname{ord}(\alpha^k) = d\}$$

$$= \#\left\{ k \in \{0, \ldots, d-1\} \;\middle|\; \frac{d}{\gcd(d,k)} = d \right\} \quad \text{by Proposition Ord-3}$$
$$= \#\{k \in \{0, \ldots, d-1\} \mid \gcd(d,k) = 1\}$$
$$= \phi(d).$$

□

Lemma 3.2. *If m is a positive integer, then $\sum_{d \mid m} \phi(d) = m$, the sum running through the positive divisors of m.*

Proof. Later in the course, we will develop tools to prove this via the theory of multiplicative arithmetic functions, but for now we will adopt the approach given in [6]. A bridge between m and $\sum_{d \mid m} \phi(d)$, the two sides of the equation, is provided by the sets

$$A_d = \{a \in \{0, \ldots, m-1\} \mid \gcd(a,m) = d\},$$
$$B_d = \{b \in \{0, \ldots, \tfrac{m}{d} - 1\} \mid \gcd(b, \tfrac{m}{d}) = 1\},$$

where d is a positive divisor of m. On the one hand, the set $\{0, \ldots, m-1\}$ is partitioned by the sets A_d as d runs through the positive divisors of m, so that

$$m = \sum_{d \mid m} \#A_d.$$

On the other hand, as d runs through the positive divisors of m, so does m/d, and we consequently have

$$\sum_{d \mid m} \phi(d) = \sum_{d \mid m} \phi(m/d) = \sum_{d \mid m} \#B_d,$$

the last equality by definition of ϕ. Therefore, the proof may be completed by showing that $\#A_d = \#B_d$ for all divisors of m. We leave it as a short exercise to verify that the sets A_d and B_d are in fact in bijection via the map

$$\begin{aligned} f : A_d &\to B_d \\ a &\mapsto a/d. \end{aligned}$$

□

Theorem 3.3. *Let p be a prime, and recall from Lemma 3.1 the number $f(d)$ defined for each positive divisor d of $p-1$. Then in fact $f(d) = \phi(d)$ for every d. In particular, there are $\phi(p) = p-1$ primitive roots mod p.*

Proof. By Proposition Ord-2, $\operatorname{ord}(\alpha) \mid p-1$ for every $\alpha \in (\mathbb{Z}/p\mathbb{Z})^\times$, so the number of elements in $(\mathbb{Z}/p\mathbb{Z})^\times$ equals the sum of the numbers $f(d)$, i.e.,

$$\begin{aligned} p - 1 &= \sum_{d \mid p-1} f(d) \\ &\leq \sum_{d \mid p-1} \phi(d) \quad \text{by Lemma 3.1} \end{aligned}$$

$$= p - 1 \quad \text{by Lemma 3.2.}$$

If there were any divisor d of $p-1$ with $f(d) < \phi(d)$, then the inequality above would be a strict inequality, giving $p - 1 < p - 1$, a contradiction. □

Proposition 3.4. *Let p be an odd prime, and let a be a primitive root mod p. Then either a or $a + p$ is a primitive root mod p^2.*

Proof. By Proposition Ord-4, $\operatorname{ord}_p(a) \mid \operatorname{ord}_{p^2}(a)$, i.e., $p-1 \mid \operatorname{ord}_{p^2}(a)$. On the other hand, Proposition Ord-2 tells us that $\operatorname{ord}_{p^2}(a) \mid \phi(p^2) = p(p-1)$. Therefore,

$$\operatorname{ord}_{p^2}(a) \in \{p - 1, p(p-1)\},$$

and similarly, because the integer $b = a + p$ is also a primitive root mod p, we have $\operatorname{ord}_{p^2}(b) \in \{p - 1, p(p-1)\}$ as well.

Suppose, then, that a is not a primitive root mod p^2, so that $a^{p-1} \equiv 1 \bmod p^2$. Then

$$\begin{aligned} b^{p-1} &= (a+p)^{p-1} \\ &= a^{p-1} + p(p-1)a^{p-2} + cp^2 \quad \text{for some } c \in \mathbb{Z} \text{ by the binomial theorem} \\ &\equiv a^{p-1} - a^{p-2} \mod p^2 \\ &\equiv 1 - a^{p-2} \mod p^2 \\ &\not\equiv 1 \mod p^2 \quad \text{because } a \text{ is coprime to } p. \end{aligned}$$

Hence, $\operatorname{ord}_{p^2}(b)$ is not equal to $p - 1$ and therefore is equal to $p(p-1)$, that is, b is a primitive root mod p^2. □

Proposition 3.5. *Let p be an odd prime. If a is a primitive root mod p^2, then it is a primitive root mod p^k for all $k \geq 1$.*

Proof. Note first that if a is a primitive root mod p^2, then for any b not divisible by p, there exists by definition some integer $n \geq 0$ such that $a^n \equiv b \bmod p^2$, and then of course this congruence holds mod p as well, so a is a primitive root mod p.

To treat the case $k > 2$, we first show, by induction on k, that $a^{p^{k-2}(p-1)} \not\equiv 1 \bmod p^k$ for all $k \geq 2$. The case $k = 2$ holds because a is assumed to be a primitive root mod p, so $a^{p-1} \not\equiv 1 \bmod p^2$. Assume that the statement is true for some $k \geq 2$. Now,

$$\begin{aligned} 1 &\equiv a^{\phi(p^{k-1})} \mod p^{k-1} \quad \text{by Proposition Ord-2} \\ &= a^{p^{k-2}(p-1)}, \end{aligned}$$

so $a^{p^{k-2}(p-1)} = 1 + bp^{k-1}$ for some $b \in \mathbb{Z}$. If p divided b, then $a^{p^{k-2}(p-1)}$ would be congruent to 1 mod p^k, contradicting the inductive hypothesis, so $p \nmid b$. But then

$$\begin{aligned} a^{p^{k-1}(p-1)} &= (1 + bp^{k-1})^p \\ &\equiv 1 + bp^k \bmod p^{k+1} \quad \text{by the binomial theorem (but see the remark below)} \\ &\not\equiv 1 \bmod p^{k+1}, \end{aligned}$$

completing the induction.

Now, Proposition Ord-4 shows that $\operatorname{ord}_{p^2}(a) \mid \operatorname{ord}_{p^k}(a)$, i.e., $p(p-1) \mid \operatorname{ord}_{p^k}(a)$. On the other hand, $\operatorname{ord}_{p^k}(a) \mid p^{k-1}(p-1)$ by Proposition Ord-2. Therefore,

$$\operatorname{ord}_{p^k}(a) = p^n(p-1)$$

for some $n \in \{1, \ldots, k-1\}$. But if $n \leq k-2$, then $a^{p^{k-2}(p-1)} \equiv 1 \bmod p^k$, contradicting what we proved above, so $n = k-1$. Thus, a is a primitive root mod p^k. □

Remark. There is one step in the proof of Proposition 3.5 that would fail if p were 2, and that is the step where we used the binomial theorem. Specifically, the step is incorrect if $p = 2$ and $k = 2$. Of course, something in the proof has to go wrong in the case $p = 2$, because there is no primitive root mod 2^k when $k \geq 3$, but it is satisfying to pinpoint precisely the moment where we use the assumption that the prime p is greater than 2.

It remains to show that if an integer $m \geq 2$ admits a primitive root, and m is neither 2 nor 4, then m must be either a power of an odd prime or twice such a power.

Lemma 3.6. *If $k \in \mathbb{Z}_{\geq 3}$ and a is an odd integer, then $a^{\phi(2^k)/2} \equiv 1 \bmod 2^k$. In particular, 2^k does not admit a primitive root.*

Proof. Fix an odd integer a. We prove by induction on $k \geq 3$ the assertion that $a^{\phi(2^k)/2} \equiv 1 \bmod 2^k$. For the case $k = 3$, write $a = 1 + 2b$ with $b \in \mathbb{Z}$, and observe that

$$\begin{aligned} a^{\phi(2^3)/2} = a^2 = (1+2b)^2 = 1 + 4b + 4b^2 &= 1 + 4b(1+b) \\ &\equiv 1 \bmod 2^3, \end{aligned}$$

because $b(1+b)$ is even.

Now let $k \geq 3$, and assume that $a^{\phi(2^k)/2} \equiv 1 \bmod 2^k$, i.e., $a^{2^{k-2}} = 1 + 2^k c$ for some $c \in \mathbb{Z}$. Then

$$\begin{aligned} a^{\phi(2^{k+1})/2} = a^{2^{k-1}} = (a^{2^{k-2}})^2 = (1+2^k c)^2 &= 1 + 2^{k+1}c + 2^{2k}c^2 \\ &\equiv 1 \bmod 2^{k+1}. \end{aligned}$$

The induction is complete. □

Lemma 3.7. *Let $t_1, \ldots, t_r$ be positive integers and N their least common multiple. If the product $t_1 \cdots t_r$ divides N, then $t_1, \ldots, t_r$ are pairwise coprime.*

Proof. Let p be a prime, and observe that $v_p(N) = \max(v_p(t_1), \ldots, v_p(t_r))$. Let m_p be this maximum, and choose i such that $v_p(t_i) = m_p$. If $t_1 \cdots t_r$ divides N, then

$$\begin{gathered} v_p(t_1 \cdots t_r) \leq v_p(N), \\ \text{i.e.,} \quad v_p(t_1) + \cdots + v_p(t_r) \leq \max(v_p(t_1), \ldots, v_p(t_r)), \end{gathered}$$

$$\begin{aligned} \text{i.e.,} \quad & v_p(t_i) + \sum_{j \neq i} v_p(t_j) \leq m_p, \\ \text{i.e.,} \quad & m_p + \sum_{j \neq i} v_p(t_j) \leq m_p, \\ \text{i.e.,} \quad & \sum_{j \neq i} v_p(t_j) \leq 0, \end{aligned}$$

so each $v_p(t_j)$ with $j \neq i$ is zero. Therefore, p can divide at most one of $t_1, \ldots, t_r$. This being true for an arbitrary prime p, the claim of the lemma follows. □

Proposition 3.8. *Let $m \in \mathbb{Z}_{\geq 2}$, and assume that m has a primitive root. If m is neither 2 nor 4, then m is a power of an odd prime or twice such a power.*

Proof. Let a be a primitive root mod m, and write $m = p_1^{k_1} \cdots p_r^{k_r}$, where the p_i are pairwise distinct primes and $k_i \geq 1$ for all i. Let N be the least common multiple of $\phi(p_1^{k_1}), \ldots, \phi(p_r^{k_r})$. For each i, $a^{\phi(p_i^{k_i})} \equiv 1 \bmod p_i^{k_i}$ by Proposition Ord-2, so also $a^N \equiv 1 \bmod p_i^{k_i}$ because $\phi(p_i^{k_i}) \,|\, N$. Thus, $p_i^{k_i} \,|\, a^N - 1$ for all i. But the integers $p_i^{k_i}$ are pairwise coprime, so their product, m, divides $a^N - 1$, i.e., $a^N \equiv 1 \bmod m$. Hence, $\mathrm{ord}_m(a) \,|\, N$ by Proposition Ord-1, i.e., $\phi(m) \,|\, N$ because a is primitive mod m. Consequently, $\phi(p_1^{k_1}) \cdots \phi(p_r^{k_r}) \,|\, N$, because ϕ is multiplicative. But N is the least common multiple of the $\phi(p_i^{k_i})$, so these numbers are pairwise coprime by Lemma 3.7.

Now, if m had two distinct odd prime divisors, say p_i and p_j, then both $\phi(p_i^{k_i})$ and $\phi(p_j^{k_j})$ would be even, contradicting what we found above. Therefore, $m = 2^k p^l$, where p is an odd prime and $k, l \geq 0$. If $l = 0$, then $m = 2^k$, so Lemma 3.6 shows that $k \leq 2$. Otherwise, if $l \geq 1$, so that $\phi(p^l)$ is even, then $\phi(2^k)$ must be odd, i.e., $k \leq 1$. □

Appendix: 4 Proof of Proposition 12.2 in Section II

We recall the statement to be proven:

> If $f(x) \in \mathbb{Z}[x]$, p is prime, and $a \in \mathbb{Z}$ satisfies $v_p(f(a)) > 2v_p(f'(a))$, then there are integers $a_0, a_1, a_2, \ldots$, with $a_0 = a$, such that
>
> $$f(a_n) \equiv 0 \bmod p^{n+1} \quad \text{and} \quad a_{n+1} \equiv a_n \bmod p^{n+1} \quad \text{for all } n \geq 0.$$

We in fact prove the following, from which the above can be deduced immediately by induction.

Proposition 4.1. *Assume that $f(x) \in \mathbb{Z}[x]$, p is prime, and $a_0 \in \mathbb{Z}$ satisfies $v_p(f(a_0)) > 2v_p(f'(a_0))$. Suppose that for some $n \geq 0$, an integer a_n satisfies*

$$v_p(f(a_n)) \geq v_p(f(a_0)) + n \tag{4.1}$$

$$v_p(f'(a_n)) = v_p(f'(a_0)) \tag{4.2}$$

Then if $f(a_n) \neq 0$, there are $\lambda_n, \mu_n \in \mathbb{Z}$, not divisible by p, and $t_n \geq n+1$ such that

$$\frac{f(a_n)}{f'(a_n)} = \frac{\lambda_n}{\mu_n} p^{t_n}.$$

Further, if $b_n \in \mathbb{Z}$ satisfies $\lambda_n + \mu_n b_n \equiv 0 \bmod p$ and we let $a_{n+1} = a_n + b_n p^{t_n} \in \mathbb{Z}$, then (4.1) and (4.2) hold with n replaced by $n+1$.

Proof. Observe that

$$v_p(f(a_n)/f'(a_n)) \geq n + v_p(f(a_0)) - v_p(f'(a_0)) > n + v_p(f'(a_0)) \geq n.$$

This tells us immediately that, as long as $f(a_n) \neq 0$,

$$\frac{f(a_n)}{f'(a_n)} = \frac{\lambda_n}{\mu_n} p^{t_n}$$

for some $\lambda_n, \mu_n \in \mathbb{Z}$ coprime to p and some $t_n \geq n+1$. Let $b_n \in \mathbb{Z}$ be chosen such that $\lambda_n + \mu_n b_n \equiv 0 \bmod p$, and let $a_{n+1} = a_n + b_n p^{t_n} \in \mathbb{Z}$.

Write

$$f(x) = f(a_n) + f'(a_n)(x - a_n) + g_n(x)(x - a_n)^2$$

for some $g_n(x) \in \mathbb{Z}[x]$. Then

$$\begin{aligned} f(a_{n+1}) &= f(a_n) + f'(a_n) b_n p^{t_n} + g_n(a_{n+1}) b_n^2 p^{2t_n} \\ &= \frac{f'(a_n)}{\mu_n}(\lambda_n + \mu_n b_n) p^{t_n} + g_n(a_{n+1}) b_n^2 p^{2t_n}. \end{aligned}$$

Now, the p-adic valuation of the first main term here is at least

$$v_p(f'(a_n)) + 1 + t_n = v_p(f(a_n)) + 1$$

by definition of t_n. Further, the p-adic valuation of the second main term is at least

$$\begin{aligned} 2t_n &= 2v_p(f(a_n)) - 2v_p(f'(a_n)) \quad \text{by definition of } t_n \\ &= v_p(f(a_n)) + v_p(f(a_n)) - 2v_p(f'(a_n)) \\ &\geq v_p(f(a_n)) + v_p(f(a_0)) - 2v_p(f'(a_0)) \quad \text{because } v_p(f'(a_n)) = v_p(f'(a_0)) \\ &\geq v_p(f(a_n)) + 1. \end{aligned}$$

Hence,

$$v_p(f(a_{n+1})) \geq v_p(f(a_n)) + 1 \geq v_p(f(a_0)) + n + 1.$$

Next, we write

$$f'(x) = f'(a_n) + f''(a_n)(x - a_n) + h_n(x)(x - a_n)^2$$

for some $h_n(x) \in \mathbb{Z}[x]$. Then

$$f'(a_{n+1}) = f'(a_n) + f''(a_n) b_n p^{t_n} + h_n(a_{n+1}) b_n^2 p^{2t_n},$$

but

$$\begin{aligned} v_p(f'(a_n)) &= v_p(f(a_n)) - t_n \\ &\geq v_p(f(a_0)) - t_n \\ &> 2v_p(f'(a_0)) - t_n \\ &= 2v_p(f'(a_n)) - t_n, \end{aligned}$$

so $t_n > v_p(f'(a_n))$, and so

$$v_p(f'(a_{n+1})) = v_p(f'(a_n)) = v_p(f'(a_0)).$$

□

Appendix: 5 Proof of Proposition 2.2 in Section III

We recall the statement to be proven:

(i) Every element of Π is irreducible in $\mathbb{Z}[i]$.

(ii) Every irreducible element of $\mathbb{Z}[i]$ is associate to exactly one element of Π.

First, we prove that every element of Π is irreducible in $\mathbb{Z}[i]$. If $\pi_2 = \alpha\beta$, then

$$2 = N(\pi_2) = N(\alpha\beta) = N(\alpha)N(\beta),$$

so $N(\alpha) = 1$ or $N(\beta) = 1$.

If $p \equiv 1 \bmod 4$ and $\pi_p = \alpha\beta$, then

$$p = N(\pi_p) = N(\alpha\beta) = N(\alpha)N(\beta),$$

so $N(\alpha) = 1$ or $N(\beta) = 1$. The argument is the same for $\overline{\pi}_p$.

If $q \equiv 3 \bmod 4$ and $\pi_q = \alpha\beta$, i.e., $q = \alpha\beta$, then

$$q^2 = N(q) = N(\alpha\beta) = N(\alpha)N(\beta).$$

The Gaussian splitting lemma, i.e., Lemma 2.1 in Section III, shows that we cannot have $N(\alpha) = N(\beta) = p$, so either $N(\alpha) = 1$ or $N(\beta) = 1$.

Next, we prove that every irreducible element of $\mathbb{Z}[i]$ is associate to exactly one element of Π. Let $\pi \in \mathbb{Z}[i]$ be any irreducible element. Then $N(\pi) \in \mathbb{Z}_{\geq 2}$, so $\pi\overline{\pi} = p_1 \cdots p_r$ for some prime numbers $p_1, \ldots, p_r$. Thus, $\pi \mid p_1 \cdots p_r$, so because π is irreducible and therefore prime in the unique factorization domain $\mathbb{Z}[i]$, π divides some $p = p_j$ in $\mathbb{Z}[i]$, say $p = \pi\alpha$ with $\alpha \in \mathbb{Z}[i]$. If $\alpha \in \mathbb{Z}[i]^\times$, then p is irreducible in $\mathbb{Z}[i]$ and is therefore

congruent to 3 mod 4 by the Gaussian splitting lemma again, so $\pi_p = p \sim \pi$. Otherwise, i.e., if $\alpha \notin \mathbb{Z}[i]^\times$, then because

$$p^2 = N(p) = N(\pi\alpha) = N(\pi)N(\alpha)$$

and π is not a unit either, we must have $N(\pi) = N(\alpha) = p$, and then $p = 2$ or $p \equiv 1 \bmod 4$ by the Gaussian splitting lemma once more. Hence,

$$\pi_p\overline{\pi}_p = p = N(\pi) = \pi\overline{\pi},$$

so uniqueness of factorization in $\mathbb{Z}[i]$ implies that $\pi \sim \pi_p$ or $\pi \sim \overline{\pi}_p$.

Finally, to show that no two elements of Π are associate, suppose that π_{p_1} is associate to π_{p_2} or $\overline{\pi}_{p_2}$. Then $N(\pi_{p_1}) = N(\pi_{p_2})$, so because $N(\pi_{p_1}) \in \{p_1, p_1^2\}$ and $N(\pi_{p_2}) \in \{p_2, p_2^2\}$, it follows that $p_1 = p_2$. It therefore remains to show that $\pi_p \not\sim \overline{\pi}_p$ when $p \equiv 1 \bmod 4$. But $\pi_p = x + yi$ with $0 < x < y$, and $\overline{\pi}_p = x - yi$, so because $|x| \neq |y|$, there is no $u \in \{1, i, -1, -i\}$ such that $u\pi_p = \overline{\pi}_p$.

Appendix: 6 Products of coprime elements in a UFD

We prove a lemma that we used in Sections III–4 and III–5, regarding products of coprime elements in a unique factorization domain. We recall the statement to be proven:

Suppose that

- R is a unique factorization domain,
- $a, b \in R \smallsetminus \{0\}$,
- n is a positive integer.

If a, b are coprime and $ab = c^n$ for some $c \in R$, then there are units u, v in R and elements $a', b' \in R$ such that $a = u(a')^n$ and $b = v(b')^n$.

Proof. Write

$$a = up_1^{r_1} \cdots p_k^{r_k}$$
$$b = vq_1^{s_1} \cdots q_l^{s_l}$$

where $p_1, \ldots, p_k, q_1, \ldots, q_l$ are pairwise non-associate irreducible elements of R, and $u, v \in R^\times$. Then $uvp_1^{r_1} \cdots p_k^{r_k} q_1^{s_1} \cdots q_l^{s_l}$ is an nth power, say

$$uvp_1^{r_1} \cdots p_k^{r_k} q_1^{s_1} \cdots q_l^{s_l} = w^n \pi_1^{nt_1} \cdots \pi_m^{nt_m},$$

where $\pi_1, \ldots, \pi_m$ are pairwise coprime irreducible elements, and $w \in R^\times$. By uniqueness of factorization, p_1 is associate to one of $\pi_1, \ldots, \pi_m$, say to π_i, and then $r_1 = nt_i$. Similarly, each of the other r_j is a multiple of n, and then a is the product of u and an nth power. In the same way, we see that b is the product of v and an nth power. □

Appendix: 7 An elementary derivation of the formula for primitive Pythagorean triples

Let (x, y, z) be a primitive Pythagorean triple. We show by elementary means that

$$(x, y, z) = (u^2 - v^2, 2uv, u^2 + v^2)$$

for coprime positive integers u and v satisfying $u > v$ and $u \not\equiv v \bmod 2$. We begin by rearranging the Pythagorean equation to read

$$\begin{aligned} z^2 - x^2 &= y^2, \\ \text{i.e.,} \quad (z + x)(z - x) &= y^2, \\ \text{i.e.,} \quad \frac{z + x}{2}\,\frac{z - x}{2} &= \left(\frac{y}{2}\right)^2 \quad \text{(remember that } x \text{ and } z \text{ are odd).} \end{aligned}$$

Since $\frac{z+x}{2} + \frac{z-x}{2} = z$ and $\frac{z+x}{2} - \frac{z-x}{2} = x$, any common divisor of $\frac{z+x}{2}$ and $\frac{z-x}{2}$ divides z and x, which are coprime, so $\frac{z+x}{2}$ and $\frac{z-x}{2}$ are coprime. But if a and b are coprime positive integers whose product is square, then each of a and b is square. Therefore, there are positive integers u and v such that

$$\frac{z + x}{2} = u^2 \quad \text{and} \quad \frac{z - x}{2} = v^2.$$

Note that these equations give $x = u^2 - v^2$ and $z = u^2 + v^2$, and of course $(y/2)^2 = u^2 v^2$, so $y = 2uv$.

Now, $u > v$ because x is positive, and u^2 and v^2 are coprime, so u and v are coprime. Also, $u^2 = v^2 + x$, and x is odd, so $u^2 \not\equiv v^2 \bmod 2$, and so $u \not\equiv v \bmod 2$. Thus, u and v satisfy all the desired properties.

Appendix: 8 Proof that the Dirichlet inverse of a multiplicative arithmetic function is multiplicative

If f is a multiplicative arithmetic function, then the fact that its Dirichlet inverse, f^{-1}, is multiplicative follows immediately from the proposition below upon taking $g = f^{-1}$.

Proposition 8.1. *If f and g are arithmetic functions such that f and $f * g$ are multiplicative, then g is multiplicative.*

Proof. We prove by induction on $N \geq 1$ the assertion that, if m and n are coprime positive integers such that $mn = N$, then $g(mn) = g(m)g(n)$. For the case $N = 1$, we observe that

$$1 = (f * g)(1) = f(1)g(1) = g(1),$$

so $g(1^2) = g(1) = 1 = g(1)^2$. Now let $N > 1$, and assume the statement for all smaller values of N. Let m and n be coprime positive integers such that $mn = N$, let

$$S = \{(d, e) \in \mathbb{Z}_{\geq 1} \times \mathbb{Z}_{\geq 1} \mid d \,|\, m \text{ and } e \,|\, n\},$$

and let $S' = S \smallsetminus \{(m,n)\}$. Then

$$\begin{aligned}
(f*g)(mn) &= \sum_{(d,e)\in S} f\big(\tfrac{mn}{de}\big)g(de) \\
&= \sum_{(d,e)\in S'} f\big(\tfrac{mn}{de}\big)g(de) + g(mn) \\
&= \sum_{(d,e)\in S'} f\big(\tfrac{m}{d}\big)f\big(\tfrac{n}{e}\big)g(d)g(e) + g(mn) \quad \text{by the inductive hypothesis} \\
&= \sum_{(d,e)\in S} f\big(\tfrac{m}{d}\big)f\big(\tfrac{n}{e}\big)g(d)g(e) - g(m)g(n) + g(mn) \\
&= \left(\sum_{d\,|\,m} f\big(\tfrac{m}{d}\big)g(d)\right)\left(\sum_{e\,|\,n} f\big(\tfrac{n}{e}\big)g(e)\right) - g(m)g(n) + g(mn) \\
&= (f*g)(m)(f*g)(n) - g(m)g(n) + g(mn) \\
&= (f*g)(mn) - g(m)g(n) + g(mn),
\end{aligned}$$

so $g(mn) = g(m)g(n)$, and the induction is complete. □

Appendix: 9 The uniqueness part of Proposition 2.1 in Sec. V

We show the uniqueness of the representation of a rational number in the form $[a_0; a_1, \ldots, a_n]$, where the a_i are integers, $a_i \geq 1$ if $i \geq 1$, and $a_n \geq 2$ if $n \geq 1$.

It will help to introduce some ad hoc terminology, namely, if $n \geq 0$, a *well-formed* n-tuple will mean an n-tuple $(a_0; a_1, \ldots, a_n)$ such that $a_i \in \mathbb{Z}$ for all i, $a_i \geq 1$ for $i \geq 1$, and, if $n \geq 1$, then $a_n \geq 2$.

Given $n \geq 0$, let $P(n)$ be the following assertion: For all $n' \geq n$, if $(a_0; a_1, \ldots, a_n)$ is a well-formed n-tuple, $(b_0; b_1, \ldots, b_{n'})$ is a well-formed n'-tuple, and $[a_0; a_1, \ldots, a_n] = [b_0; b_1, \ldots, b_{n'}]$, then $n = n'$ and $a_i = b_i$ for all i. We prove $P(n)$ by induction.

Let us first prove $P(0)$. Suppose that $n' \geq 0$, $a_0 \in \mathbb{Z}$, and $(b_0; b_1, \ldots, b_{n'})$ is a well-formed n'-tuple such that $a_0 = [b_0; b_1, \ldots, b_{n'}]$. Assume, for a contradiction, that $n' \geq 1$. Then

$$a_0 = [b_0; b_1, \ldots, b_{n'}] = b_0 + \frac{1}{[b_1; b_2, \ldots, b_{n'}]},$$

so

$$|a_0 - b_0| = \frac{1}{[b_1; b_2, \ldots, b_{n'}]}.$$

If $n' = 1$, then $b_1 \geq 2$, so the left-hand side is less than 1, and if $n' \geq 2$, then $[b_1; b_2, \ldots, b_{n'}] > b_1 \geq 1$, so the left-hand side is again less than 1. Either way, the non-negative integer $|a_0 - b_0|$ is less than 1 and is therefore zero, showing that $1/[b_1; b_2, \ldots, b_{n'}] = 0$, a contradiction. Thus, $n' = 0$, and then $a_0 = b_0$.

Now let $n \geq 0$, and assume $P(n)$. Let $n' \geq n+1$, let $(a_0; a_1, \ldots, a_{n+1})$ be a well-formed $(n+1)$-tuple and $(b_0; b_1, \ldots, b_{n'})$ a well-formed n'-tuple, and assume that

$[a_0; a_1, \ldots, a_{n+1}] = [b_0; b_1, \ldots, b_{n'}]$. Then

$$|a_0 - b_0| = \left| \frac{1}{[b_1; b_2, \ldots, b_{n'}]} - \frac{1}{[a_1; a_2, \ldots, a_{n+1}]} \right|,$$

so since the right-hand side is less than 1 and the left-hand side is a non-negative integer, both sides are zero and we obtain $a_0 = b_0$ and $[a_1; a_2, \ldots, a_{n+1}] = [b_1; b_2, \ldots, b_{n'}]$. Hence, by the inductive hypothesis, $n' = n + 1$ and $a_i = b_i$ for all $i \geq 1$. The induction is complete.

Appendix: 10 On infinite continued fractions

We prove the facts (i)–(iii) concerning infinite continued fractions stated at the beginning of Section V – 5.

Irrationality

First, we show that if $a_0, a_1, a_2, \ldots$ are integers with $a_k > 0$ for all $k \geq 1$, then the real number $x = [a_0; a_1, a_2, \ldots]$ is irrational. The proof of Theorem 4.5 in Section V shows that

$$\begin{aligned} & C_{2n} < x < C_{2n+1} \quad \text{for all } n \geq 0, \\ \text{i.e.,} \quad & 0 < x - C_{2n} < C_{2n+1} - C_{2n}, \\ \text{i.e.,} \quad & 0 < x - \frac{p_{2n}}{q_{2n}} < \frac{1}{q_{2n+1} q_{2n}}, \\ \text{i.e.,} \quad & 0 < q_{2n} x - p_{2n} < \frac{1}{q_{2n+1}}. \end{aligned}$$

If x were rational, say $x = a/b$ with $a, b \in \mathbb{Z}$ and $b > 0$, then multiplying the last line by b, we would have

$$0 < q_{2n} a - p_{2n} b < \frac{b}{q_{2n+1}}.$$

But the sequence $(q_k)_{k \geq 0}$ is a monotone-increasing sequence of integers and is therefore unbounded above, so we may choose n such that $b/q_{2n+1} < 1$, contradicting the fact that $q_{2n} a - p_{2n} b \in \mathbb{Z}$.

Existence

Next, we show that if $x \in \mathbb{R}$ is irrational, then there are integers $a_0, a_1, a_2, \ldots$, with $a_k > 0$ for all $k \geq 1$, such that $x = [a_0; a_1, a_2, \ldots]$. Recall that for a real number y, $\lfloor y \rfloor$ denotes its floor, i.e., the greatest integer less than or equal to y. If y is irrational, then certainly $y \notin \mathbb{Z}$, so $y - \lfloor y \rfloor > 0$ and $1/(y - \lfloor y \rfloor)$ is still irrational. Therefore, we may define irrational numbers x_k and integers a_k recursively by

$$x_0 = x, \quad a_k = \lfloor x_k \rfloor, \quad x_{k+1} = \frac{1}{x_k - a_k} \quad (k \geq 0).$$

Note that $0 < x_k - a_k < 1$ for all $k \geq 0$, so $x_{k+1} > 1$, and so $a_{k+1} \geq 1$.

Now, since

$$\begin{aligned}[a_0; a_1, a_2, \ldots, a_k, x_{k+1}] &= [a_0; a_1, a_2, \ldots, a_k, a_{k+1} + \tfrac{1}{x_{k+2}}] \\ &= [a_0; a_1, a_2, \ldots, a_k, a_{k+1}, x_{k+2}],\end{aligned}$$

it is clear by induction on k that

$$x = [a_0; a_1, a_2, \ldots, a_k, x_{k+1}] \quad \text{for all } k \geq 0.$$

Hence, if $C_k = [a_0; a_1, a_2, \ldots]$, and if p_k and q_k are defined in terms of $a_0, a_1, a_2, \ldots$ as usual, i.e., via the (p, q)-algorithm, then for all $k \geq 1$,

$$\begin{aligned}|x - C_k| &= |[a_0; a_1, a_2, \ldots, a_k, x_{k+1}] - C_k| \\ &= \left| \frac{x_{k+1}p_k + p_{k-1}}{x_{k+1}q_k + q_{k-1}} - \frac{p_k}{q_k} \right| \quad \text{by Theorem 3.1 in Section V} \\ &= \left| \frac{x_{k+1}p_kq_k + p_{k-1}q_k - x_{k+1}p_kq_k - p_kq_{k-1}}{(x_{k+1}q_k + q_{k-1})q_k} \right| \\ &= \left| \frac{(-1)^k}{(x_{k+1}q_k + q_{k-1})q_k} \right| \quad \text{by Proposition 4.1 in Section V} \\ &= \frac{1}{(x_{k+1}q_k + q_{k-1})q_k} \quad \text{because } x_{k+1}, q_k, q_{k-1} > 0.\end{aligned}$$

But $a_{k+1} = \lfloor x_{k+1} \rfloor < x_{k+1}$, so

$$\frac{1}{(x_{k+1}q_k + q_{k-1})q_k} < \frac{1}{(a_{k+1}q_k + q_{k-1})q_k} = \frac{1}{q_{k+1}q_k}.$$

Finally, using again the fact that the sequence $(q_k)_{k \geq 0}$ is a monotone-increasing sequence of integers, we observe that $1/(q_{k+1}q_k) \to 0$ as $k \to \infty$, so $C_k \to x$.

Uniqueness

We show that if $a_0, a_1, a_2, \ldots$ and $b_0, b_1, b_2, \ldots$ are integers such that $a_k, b_k > 0$ for all $k \geq 1$ and such that $[a_0; a_1, a_2, \ldots] = [b_0; b_1, b_2, \ldots]$, then $a_k = b_k$ for all $k \geq 0$. The core idea is contained in the following lemma.

Lemma 10.1. *If $a'_0, a'_1, a'_2, \ldots$ and $b'_0, b'_1, b'_2, \ldots$ are integers such that $a'_k, b'_k > 0$ for all $k \geq 1$ and such that $[a'_0; a'_1, a'_2, \ldots] = [b'_0; b'_1, b'_2, \ldots]$, then $a'_0 = b'_0$ and $[a'_1; a'_2, \ldots] = [b'_1; b'_2, \ldots]$.*

Proof. For brevity, we omit the primes, i.e., write simply a_k and b_k instead of a'_k and b'_k. Let $x = [a_0; a_1, a_2, \ldots] = [b_0; b_1, b_2, \ldots]$, and let $\alpha = (a_0; a_1, a_2, \ldots)$ and $\beta = (b_0; b_1, b_2, \ldots)$. By the proof of Theorem 4.5,

$$\begin{gathered}C_0(\alpha) < x < C_1(\alpha), \\ \text{i.e.,} \quad a_0 < x < a_0 + \frac{1}{a_1} \leq a_0 + 1,\end{gathered}$$

so $\lfloor x \rfloor = a_0$. Similarly, $\lfloor x \rfloor = b_0$, so $a_0 = b_0$. Further,

$$\begin{aligned} x &= \lim_{n\to\infty} [a_0; a_1, a_2, \ldots, a_n] \\ &= \lim_{n\to\infty} \left(a_0 + \frac{1}{[a_1; a_2, \ldots, a_n]} \right) \\ &= a_0 + \frac{1}{\lim_{n\to\infty} [a_1; a_2, \ldots, a_n]} \\ &= a_0 + \frac{1}{[a_1; a_2, \ldots]}, \end{aligned}$$

and similarly

$$x = b_0 + \frac{1}{[b_1; b_2, \ldots]}.$$

Therefore, because $a_0 = b_0$, it follows that $[a_1; a_2, \ldots] = [b_1; b_2, \ldots]$. □

Now let $a_0, a_1, a_2, \ldots$ and $b_0, b_1, b_2, \ldots$ be as above. We show by induction on $n \geq 0$ that $a_n = b_n$ and $[a_{n+1}; a_{n+2}, a_{n+3}, \ldots] = [b_{n+1}; b_{n+2}, b_{n+3}, \ldots]$. The case $n = 0$ is simply the lemma with $a'_k = a_k$ and $b'_k = b_k$ for all $k \geq 0$. Next, let $n \geq 0$ and assume the statement for this n. Then the lemma applied with $a'_k = a_{n+1+k}$ and $b'_k = b_{n+1+k}$ for all $k \geq 0$ yields $a_{n+1} = b_{n+1}$ and $[a_{n+2}; a_{n+3}, \ldots] = [b_{n+2}; b_{n+3}, \ldots]$, completing the induction.

Appendix: 11 Substitution of infinite continued fractions

Lemma 11.1. *Let $k, l \geq 0$, let $a_0, a_1, \ldots, a_{k+l}$ be real numbers with $a_i > 0$ for $i \geq 1$, and let*

$$\begin{aligned} x &= [a_0; a_1, a_2, \ldots, a_{k+l}], \\ y &= [a_k; a_{k+1}, a_{k+2}, \ldots, a_{k+l}]. \end{aligned}$$

Then $x = [a_0; a_1, a_2, \ldots, a_{k-1}, y]$.

Proof. Fix $l \geq 0$. We proceed by induction on k. For $k \geq 0$, let $P(k)$ be the assertion that, for all real numbers $a_0, \ldots, a_{k+l}$ with $a_i > 0$ for $i \geq 1$, the claimed equality holds. The case $k = 0$ is vacuous. Now let $k \geq 0$ and assume $P(k)$. Let $a_0, \ldots, a_{k+1+l}$ be real numbers with $a_i > 0$ for $i \geq 1$. Then by the inductive hypothesis, $x' = [a_1; a_2, \ldots, a_k, y]$ where

$$\begin{aligned} x' &= [a_1; a_2, \ldots, a_{k+1+l}], \\ y &= [a_{k+1}; a_{k+2}, \ldots, a_{k+1+l}]. \end{aligned}$$

Then

$$[a_0; a_1, a_2, \ldots, a_{k+1+l}] = a_0 + \frac{1}{x'} = a_0 + \frac{1}{[a_1; a_2, \ldots, a_k, y]} = [a_0; a_1, a_2, \ldots, a_k, y],$$

and the induction is complete. □

Proposition 11.2. *Let $a_0, a_1, a_2, \ldots$ be integers with $a_i > 0$ for $i \geq 1$, and let $x = [a_0; a_1, a_2, \ldots]$. Suppose that $k \geq 0$, and let $y = [a_k; a_{k+1}, a_{k+2}, \ldots]$. Then*

$$x = [a_0; a_1, a_2, \ldots, a_{k-1}, y].$$

Proof. For each $l \geq 0$, let $y_l = [a_k; a_{k+1}, \ldots, a_{k+l}]$. Then

$$\begin{aligned} x &= \lim_{n\to\infty} [a_0; a_1, a_2, \ldots, a_n] \\ &= \lim_{l\to\infty} [a_0; a_1, a_2, \ldots, a_{k-1}, a_k, a_{k+1}, \ldots, a_{k+l}] \\ &= \lim_{l\to\infty} [a_0; a_1, a_2, \ldots, a_{k-1}, y_l] \quad \text{by Lemma 11.1} \\ &= [a_0; a_1, a_2, \ldots, a_{k-1}, \lim_{l\to\infty} y_l] \quad \text{by standard properties of limits} \\ &= [a_0; a_1, a_2, \ldots, a_{k-1}, y]. \end{aligned}$$

□

Appendix: 12 The continued-fraction representation of a quadratic irrational

We prove that the continued-fraction representation of a quadratic irrational is periodic, following the method of proof given in [6].

Lemma 12.1. *Every quadratic irrational can be expressed in the form $(S + \sqrt{d})/T$ where S, T, d are integers such that $T \neq 0$, d is positive and not a square, and $T \mid S^2 - d$.*

Proof. The formula for the roots of a quadratic polynomial shows that a quadratic irrational may be expressed in the form $(u + \sqrt{v})/w$ where u, v, w are integers such that v is positive and not a square, and $w \neq 0$. Multiplying the numerator and denominator by $|w|$ then shows that our quadratic irrational is equal to $(S + \sqrt{d})/T$, where $S = |w|u$, $T = |w|w$, and $d = w^2 v$. That d is positive and not a square is an immediate consequence of the fact that v is, and we may easily verify that $T \mid S^2 - d$:

$$S^2 - d = w^2 u - w^2 v = w^2(u - v) = \pm T(u - v).$$

□

Lemma 12.2. *Fix a positive integer d that is not a square. Suppose that, for some $k \geq 0$, S_k and T_k are integers such that $T_k \neq 0$ and $T_k \mid S_k^2 - d$. If $b_k = \lfloor (S_k + \sqrt{d})/T_k \rfloor$, $S_{k+1} = b_k T_k - S_k$, and $T_{k+1} = (d - S_{k+1}^2)/T_k$, then S_{k+1} and T_{k+1} are integers such that $T_{k+1} \neq 0$ and $T_{k+1} \mid S_{k+1}^2 - d$.*

Proof. It is obvious that $S_{k+1} \in \mathbb{Z}$. As for T_{k+1}, we have

$$T_{k+1} = \frac{d - S_{k+1}^2}{T_k} = \frac{d - (b_k T_k - S_k)^2}{T_k} = \frac{d - S_k^2}{T_k} + 2b_k S_k - b_k^2 T_k.$$

Since $T_k \mid d - S_k^2$ by assumption, it follows that $T_{k+1} \in \mathbb{Z}$. Further, the fact that d is not a square shows that T_{k+1}, by its definition, is non-zero. Finally, it is obvious from the fact that $T_k T_{k+1} = d - S_{k+1}^2$ that $T_{k+1} \mid S_{k+1}^2 - d$. □

Fix a quadratic irrational x. By Lemma 12.1, there are integers S_0, T_0, d such that d is positive and not a square, $T_0 \neq 0$, $T_0 \mid S_0^2 - d$, and $x = (S_0 + \sqrt{d})/T_0$. Hence, by Lemma 12.2, we may construct numbers y_k, b_k, S_k, T_k $(k \geq 0)$ recursively by

$$\begin{aligned} y_k &= \frac{S_k + \sqrt{d}}{T_k}, \\ b_k &= \lfloor y_k \rfloor, \\ S_{k+1} &= b_k T_k - S_k, \\ T_{k+1} &= \frac{d - S_{k+1}^2}{T_k}, \end{aligned}$$

the numbers S_k and T_k being, by the same lemma, integers such that $T_k \neq 0$ and $T_k \mid S_k^2 - d$. We will call the algorithm for constructing the integers S_k and T_k as above the (S, T)-*algorithm.*

Also, define the numbers a_k and x_k, as in Section V–5, by $x_0 = x$, $a_k = \lfloor x_k \rfloor$, and $x_{k+1} = 1/(x_k - a_k)$.

Proposition 12.3. *With notation as above, $a_k = b_k$ and $x_k = y_k$ for all $k \geq 0$. In particular, $x = [b_0; b_1, b_2, \ldots]$.*

Proof. Observe first that $y_{k+1} = 1/(y_k - b_k)$:

$$\begin{aligned} y_k - b_k &= \frac{S_k + \sqrt{d}}{T_k} - b_k \\ &= \frac{S_k + \sqrt{d} - b_k T_k}{T_k} \\ &= \frac{\sqrt{d} - S_{k+1}}{T_k} \\ &= \frac{d - S_{k+1}^2}{T_k(S_{k+1} + \sqrt{d})} = \frac{T_{k+1}}{S_{k+1} + \sqrt{d}} = \frac{1}{y_{k+1}}. \end{aligned}$$

Now, $x = [a_0; a_1, a_2, \ldots]$ by the existence part of Section 10. Since $x_0 = x = y_0$ and $a_0 = \lfloor x_0 \rfloor = \lfloor y_0 \rfloor = b_0$, and since also $x_{k+1} = 1/(x_k - a_k)$ and $y_{k+1} = 1/(y_k - b_k)$, induction shows that the pair (a_k, x_k) is equal to the pair (b_k, y_k) for all $k \geq 0$. In particular, $x = [a_0; a_1, a_2, \ldots] = [b_0; b_1, b_2, \ldots]$. □

Before completing the proof that a quadratic irrational has a periodic continued-fraction representation, we introduce the notion of *conjugation.* If b and c are rational numbers such that $b^2 - 4c$ is not the square of any rational number, then the polynomial $f(x) = x^2 + bx + c$ has two distinct roots in $\mathbb{C}$, both non-rational. These roots, say x_1, x_2, are said to be *conjugate* to each other, and we define $\overline{x}_1 = x_2$ and $\overline{x}_2 = x_1$. It is easy to see that if u, v, w are rational numbers, and if v is not the square of any rational

number, then $u + w\sqrt{v}$ and $u - w\sqrt{v}$ are the two roots of a polynomial $f(x)$ as above and are therefore conjugate to each other. We also define the conjugate of a rational number to be that rational number itself.

We will need the following fact about conjugation, which we leave as an exercise: If $u, v, w, z \in \mathbb{Q}$, y is a quadratic irrational, and w and z are not both zero, then $(uy + v)/(wy + z)$ is either a rational or a quadratic irrational, and its conjugate is $(u\overline{y} + v)/(w\overline{y} + z)$.

Now let x be a quadratic irrational, write it as $x = (S_0 + \sqrt{d})/T_0$ as in Lemma 12.1, and construct the numbers y_k, b_k, S_k, T_k $(k \geq 0)$ as above, via the (S, T)-algorithm. As we saw in Proposition 12.3, each pair (b_k, y_k) is equal to (a_k, x_k) where the latter is as in the existence part of Section 10, so then by that same section,

$$x = [b_0; b_1, b_2, \ldots, b_{k-1}, y_k]$$

for all $k \geq 1$. Hence, by the (p, q)-algorithm,

$$x = \frac{p_{k-1}y_k + p_{k-2}}{q_{k-1}y_k + q_{k-2}}$$

for all $k \geq 2$, where the p_k and q_k are the integers associated to the sequence $(b_0; b_1, b_2, \ldots)$ via the (p, q)-algorithm. Taking conjugates of both sides gives

$$\begin{aligned}
\overline{x} &= \frac{p_{k-1}\overline{y}_k + p_{k-2}}{q_{k-1}\overline{y}_k + q_{k-2}}, \\
\text{i.e.,} \quad q_{k-1}\overline{x}\,\overline{y}_k + q_{k-2}\overline{x} &= p_{k-1}\overline{y}_k + p_{k-2}, \\
\text{i.e.,} \quad (q_{k-1}\overline{x} - p_{k-1})\overline{y}_k &= p_{k-2} - q_{k-2}\overline{x}, \\
\text{i.e.,} \quad \overline{y}_k &= \frac{p_{k-2} - q_{k-2}\overline{x}}{q_{k-1}\overline{x} - p_{k-1}} = -\frac{q_{k-2}}{q_{k-1}} \frac{\overline{x} - \frac{p_{k-2}}{q_{k-2}}}{\overline{x} - \frac{p_{k-1}}{q_{k-1}}}.
\end{aligned}$$

Both $\overline{x} - \frac{p_{k-2}}{q_{k-2}}$ and $\overline{x} - \frac{p_{k-1}}{q_{k-1}}$ tend to the same non-zero real number $\overline{x} - x$ as $k \to \infty$, so since $-q_{k-2}/q_{k-1} < 0$, there is $N \geq 2$ such that $\overline{y}_k < 0$ for all $k \geq N$. Hence, because $y_k > 0$ for all $k \geq 1$, it follows that for all $k \geq N$,

$$0 < y_k - \overline{y}_k = \frac{S_k + \sqrt{d}}{T_k} - \frac{S_k - \sqrt{d}}{T_k} = \frac{2\sqrt{d}}{T_k},$$

so $T_k > 0$ for such k. Therefore, again for $k \geq N$,

$$0 < T_k \leq T_k T_{k+1} = d - S_{k+1}^2 \leq d,$$

so we deduce simultaneously that $0 < T_k \leq d$ and that $S_{k+1}^2 < d$, the latter, of course, being equivalent to $-\sqrt{d} < S_{k+1} < \sqrt{d}$. We take from all this simply the following:

$$0 < T_k \leq d \quad \text{and} \quad -\sqrt{d} < S_k < \sqrt{d} \quad \text{for all } k > N.$$

There are thus only finitely many possibilities for each of the pairs (S_k, T_k) when $k > N$, so there are integers $j > i$ such that $S_j = S_i$ and $T_j = T_i$. Therefore, because each pair

(S_k, T_k) is determined purely in terms of the previous pair, we must have $S_{j+l} = S_{i+l}$ and $T_{j+l} = T_{i+l}$ for all $l \geq 0$, and the periodicity of the S_k and T_k follows, as does, consequently, the periodicity of the b_k. Now use Proposition 12.3.

Appendix: 13 Purely periodic continued fractions

Lemma 13.1. *Let $k \in \mathbb{Z}_{\geq 1}$. If $a_0, \ldots, a_k$ are positive real numbers, and if $\alpha = (a_0; a_1, \ldots, a_k)$, then*

$$\frac{p_k(\alpha)}{p_{k-1}(\alpha)} = [a_k; a_{k-1}, \ldots, a_1, a_0]$$
$$\frac{q_k(\alpha)}{q_{k-1}(\alpha)} = [a_k; a_{k-1}, \ldots, a_2, a_1]$$

Proof. We prove the assertions by induction on $k \geq 1$. For $k = 1$, observe that

$$\frac{p_1(\alpha)}{p_0(\alpha)} = \frac{a_1 a_0 + 1}{a_0} = a_1 + \frac{1}{a_0} = [a_1; a_0],$$
$$\frac{q_1(\alpha)}{q_0(\alpha)} = \frac{a_1}{1} = a_1.$$

Now let $k \geq 1$, and assume the assertions for all $\alpha = (a_0; a_1, \ldots, a_k)$. Let $\alpha = (a_0; a_1, \ldots, a_{k+1})$. Then

$$\begin{aligned}
\frac{p_{k+1}(\alpha)}{p_k(\alpha)} &= \frac{a_{k+1} p_k(\alpha) + p_{k-1}(\alpha)}{p_k(\alpha)} \\
&= a_{k+1} + \frac{1}{p_k(\alpha)/p_{k-1}(\alpha)} \\
&= a_{k+1} + \frac{1}{p_k(\alpha')/p_{k-1}(\alpha')} \quad \text{where } \alpha' = (a_0; a_1, \ldots, a_k) \\
&= a_{k+1} + \frac{1}{[a_k; a_{k-1}, \ldots, a_0]} \quad \text{by the inductive hypothesis} \\
&= [a_{k+1}; a_k, \ldots, a_0].
\end{aligned}$$

The inductive step for the q's proceeds in identical fashion. □

If x is a quadratic irrational, it is said to be *reduced* if $x > 1$ and $-1 < \overline{x} < 0$, where, as in Appendix Section 12, the bar denotes the conjugate of x.

Proposition 13.2. *A real number x is a reduced quadratic irrational if and only if it has a purely periodic continued-fraction representation, that is, there are positive integers $a_0, a_1, a_2, \ldots, a_{n-1}$ such that $x = [\overline{a_0; a_1, a_2, \ldots, a_{n-1}}]$.*

Proof. Assume that x is a reduced quadratic irrational, and let the numbers a_k, x_k be defined in the usual way, according to the algorithm for the continued-fraction representation of x. Then for all $k \geq 0$, $x_{k+1} = 1/(x_k - a_k)$, so taking conjugates gives

$$\overline{x}_{k+1} = \frac{1}{\overline{x}_k - a_k}.$$

We prove by induction on k that $-1 < \overline{x}_k < 0$ for all $k \geq 0$. The case $k = 0$ is given by assumption, x being reduced. Now assume the inequalities for some $k \geq 0$. Note that $a_0 = \lfloor x \rfloor \geq 1$ because x is reduced, and of course $a_k \geq 1$ if $k \geq 1$. Therefore, $\overline{x}_k - a_k < -1$, so since $\overline{x}_{k+1} = 1/(\overline{x}_k - a_k)$, it follows immediately that $-1 < \overline{x}_{k+1} < 0$, completing the induction.

Hence, for all $k \geq 0$,

$$-1 < a_k + \frac{1}{\overline{x}_{k+1}} < 0 \quad \text{because } \overline{x}_k = a_k + \frac{1}{\overline{x}_{k+1}},$$

$$\text{i.e.,} \quad -1 - a_k < \frac{1}{\overline{x}_{k+1}} < -a_k,$$

$$\text{i.e.,} \quad a_k < -\frac{1}{\overline{x}_{k+1}} < a_k + 1,$$

so $\lfloor -1/\overline{x}_{k+1} \rfloor = a_k$.

Now, the proof in Section 12 of the periodicity of the continued-fraction representation of a quadratic irrational shows that there are integers $k, l \geq 0$ with $k < l$ such that $x_k = x_l$. We claim that, for all $i \in \{0, \ldots, k\}$, both $a_{k-i} = a_{l-i}$ and $x_{k-i} = x_{l-i}$. We proceed by induction on i. The case $i = 0$ is given: $x_k = x_l$, so also $a_k = \lfloor x_k \rfloor = \lfloor x_l \rfloor = a_l$. Now let $i \in \{0, \ldots, k-1\}$, and assume the equalities for this i. Then

$$a_{k-(i+1)} = \lfloor -1/\overline{x}_{k-i} \rfloor = \lfloor -1/\overline{x}_{l-i} \rfloor = a_{l-(i+1)},$$

and

$$x_{k-(i+1)} = a_{k-(i+1)} + \frac{1}{x_{k-i}} = a_{l-(i+1)} + \frac{1}{x_{l-i}} = x_{l-(i+1)}.$$

The induction is complete.

In particular, in the case $i = k$, we have $x_0 = x_{l-k}$, and then the algorithm for generating the a_j and x_j shows that the sequence $a_0, a_1, a_2, \ldots$ is purely periodic.

Conversely, assume that a real number x has a purely periodic continued-fraction representation, say $x = [\overline{a_0; a_1, \ldots, a_k}]$. Doubling the period if necessary, we may assume that $k \geq 1$. Also, let $x' = [\overline{a_k; a_{k-1}, \ldots, a_0}]$. We show that x and $-1/x'$ are conjugate. Since they are distinct, one being positive and the other negative, it is enough to show that they are both roots of the same quadratic polynomial with rational coefficients.

Let the integers p_k and q_k be those arising from the (p, q)-algorithm for the continued-fraction representation of x, and let p'_k and q'_k be the corresponding integers for x'. Then

$$x = [a_0; a_1, \ldots, a_k, x] = \frac{x p_k + p_{k-1}}{x q_k + q_{k-1}},$$

so

$$q_k x^2 + (q_{k-1} - p_k)x - p_{k-1} = 0.$$

Also,

$$x' = [a_k; a_{k-1}, \ldots, a_0, x'] = \frac{x' p'_k + p'_{k-1}}{x' q'_k + q'_{k-1}},$$

so

$$q'_k (x')^2 + (q'_{k-1} - p'_k)x' - p'_{k-1} = 0,$$

$$\text{i.e.,} \quad p'_{k-1} - (q'_{k-1} - p'_k)x' - q'_k(x')^2 = 0,$$
$$\text{i.e.,} \quad p'_{k-1}\left(-\tfrac{1}{x'}\right)^2 + (q'_{k-1} - p'_k)\left(-\tfrac{1}{x'}\right) - q'_k = 0. \tag{13.1}$$

Now, by Lemma 13.1,

$$\frac{p_k}{p_{k-1}} = [a_k; a_{k-1}, \ldots, a_0] = C_k(x') = \frac{p'_k}{q'_k},$$
$$\frac{q_k}{q_{k-1}} = [a_k; a_{k-1}, \ldots, a_1] = C_{k-1}(x') = \frac{p'_{k-1}}{q'_{k-1}},$$

so because all four fractions involved are in lowest terms and the q's are all positive, we have

$$p'_k = p_k,$$
$$q'_k = p_{k-1},$$
$$p'_{k-1} = q_k,$$
$$q'_{k-1} = q_{k-1}.$$

Using these equations to replace the primed numbers in (13.1) with the unprimed ones, we see that

$$q_k\left(-\tfrac{1}{x'}\right)^2 + (q_{k-1} - p_k)\left(-\tfrac{1}{x'}\right) - p_{k-1} = 0,$$

as desired.

We have thus succeeded in showing that $\overline{x} = -1/x'$. Since $x' > 1$, it follows that $-1 < \overline{x} < 0$, and of course $x > 1$, so x is reduced. □

Appendix: 14 Proofs concerning Pell's equation

The proof that the solutions to Pell's equation are as claimed in Theorem 8.1 in Section V comprises naturally two parts:

- **Part I:** Show that if (x, y) is a solution, then there is a convergent p/q to $\sqrt{d}$ such that $(x, y) = (p, q)$.
- **Part II:** Among all convergents to $\sqrt{d}$, determine precisely which ones yield solutions.

The books of Barbeau [1], Rosen [6], and Schmidt [7] were all invaluable for the following proofs.

Part I

Lemma 14.1. *Let $n \in \mathbb{Z}_{\geq 2}$, let $a_0, \ldots, a_n \in \mathbb{R}$ with $a_i > 0$ for $i \geq 1$, let $x = [a_0; a_1, a_2, \ldots, a_n]$, and let p_k, q_k for $k \in \{0, \ldots, n\}$ be the numbers appearing in the (p, q)-algorithm for the sequence $(a_0; a_1, a_2, \ldots, a_n)$. Then*

$$q_{n-1}x - p_{n-1} = \frac{(-1)^{n-1}}{a_n q_{n-1} + q_{n-2}}.$$

Proof.

$$\begin{aligned} q_{n-1}x - p_n &= q_{n-1}\,\frac{p_n}{q_n} - p_n \\ &= \frac{p_n q_{n-1} - p_{n-1}q_n}{q_n} \\ &= \frac{(-1)^{n-1}}{q_n} \\ &= \frac{(-1)^{n-1}}{a_n q_{n-1} + q_{n-2}}. \end{aligned}$$

□

Theorem 14.2 (Legendre)**.** *Let $x \in \mathbb{R} \smallsetminus \mathbb{Q}$, and suppose that p and q are integers such that $q \neq 0$ and*

$$\left| x - \frac{p}{q} \right| < \frac{1}{2q^2}.$$

Then p/q is equal to a convergent to x.

Proof. First, we reduce to the case where p and q are coprime and $q > 0$, as follows. Assume the statement to hold in this case, and let p, q be any integers satisfying the assumptions of the theorem. Write $p/q = p'/q'$ where p' and q' are coprime and $q' > 0$. Observe that $|q'| \leq |q|$. Hence,

$$\left| x - \frac{p'}{q'} \right| = \left| x - \frac{p}{q} \right| < \frac{1}{2q^2} \leq \frac{1}{2(q')^2},$$

so p'/q' is a convergent to x, i.e., p/q is a convergent to x.

We now assume that p and q are coprime with $q > 0$ and follow the proof in Schmidt's *Diophantine Approximation*; see [7, Chap. I, Theorem 5C]. By hypothesis, we may write

$$x - \frac{p}{q} = \epsilon\,\frac{y}{q^2}$$

where $\epsilon \in \{1, -1\}$ and $0 < y < 1/2$. Also, we may write

$$\frac{p}{q} = [a_0; a_1, a_2, \ldots, a_{n-1}]$$

for some $n \geq 1$, where $a_i \in \mathbb{Z}$ for all i, $a_i > 0$ for $i \geq 1$, and $(-1)^{n-1} = \epsilon$. (Express p/q in terms of its continued-fraction representation as in Proposition 2.1 in Section V, and if n has the wrong parity, replace a_{n-1} by $(a_{n-1} - 1) + \frac{1}{1}$.)

Right away, we dispense with the case where $n = 1$. In this case, $\epsilon = (-1)^0 = 1$, and $p/q = a_0 \in \mathbb{Z}$, so $q = 1$ and $x - p = y$. Thus, $0 < x - p < 1/2$, so $p = \lfloor x \rfloor = a_0$, and so $p = p/q$ is indeed a convergent to x.

For the remainder of the proof, we assume that $n \geq 2$. Let p_k, q_k, with $k \in \{0, \ldots, n-1\}$, be the numbers appearing in the (p, q)-algorithm for $(a_0; a_1, \ldots, a_{n-1})$. Then $p/q = p_{n-1}/q_{n-1}$, so because p and q are coprime and $q > 0$, we have $p = p_{n-1}$ and $q = q_{n-1}$.

Now, if $z \in \mathbb{R}$, then

$$\begin{aligned} &[a_0; a_1, a_2, \ldots, a_{n-1}, z] = x \\ &\quad \iff \frac{zp_{n-1} + p_{n-2}}{zq_{n-1} + q_{n-2}} = x \\ &\quad \iff (p_{n-1} - xq_{n-1})z = xq_{n-2} - p_{n-2}. \end{aligned} \tag{14.1}$$

Since x is irrational, ensuring that $p_{n-1} - xq_{n-1} \neq 0$, this last equation has a unique solution z, and z is also necessarily irrational by (14.1).

Our immediate task is to show that $z > 1$. To that end, observe that

$$\begin{aligned} \epsilon \frac{y}{q^2} &= x - \frac{p}{q} \\ &= \frac{1}{q_{n-1}}(q_{n-1}x - p_{n-1}) \\ &= \frac{1}{q_{n-1}} \frac{(-1)^{n-1}}{zq_{n-1} + q_{n-2}} \quad \text{by Lemma 14.1} \\ &= \frac{1}{q_{n-1}} \frac{\epsilon}{zq_{n-1} + q_{n-2}}. \end{aligned}$$

Therefore, because $q = q_{n-1}$, we have

$$y = \frac{q_{n-1}}{zq_{n-1} + q_{n-2}},$$

and rearranging this gives

$$z = \frac{1}{y} - \frac{q_{n-2}}{q_{n-1}}.$$

Hence, because $0 < y < 1/2$, so that $1/y > 2$, and because $q_{n-2} < q_{n-1}$, it follows that $z > 2 - 1 = 1$, as desired.

If the continued-fraction representation of z is

$$z = [a_n; a_{n+1}, a_{n+2}, \ldots],$$

where of course $a_{n+1}, a_{n+2}, \ldots$ are positive, then because $z > 1$, a_n is positive as well. Therefore,

$$x = [a_0; a_1, a_2, \ldots, a_{n-1}, z] = [a_0; a_1, a_2, \ldots] \quad \text{by Proposition 11.2.}$$

As $p/q = [a_0; a_1, a_2, \ldots, a_{n-1}]$, this shows that p/q is a convergent to x. □

Corollary 14.3. *Let d be a positive integer that is not a square, and n an integer such that $|n| < \sqrt{d}$. If x and y are positive integers such that $x^2 - dy^2 = n$, then x/y is a convergent to $\sqrt{d}$.*

Proof. Observe that because $x, y > 0$ and d is not a square, n must be non-zero. Assume first that $n > 0$. Then $0 < n = (x + y\sqrt{d})(x - y\sqrt{d})$, so $x - y\sqrt{d} > 0$, i.e., $x > y\sqrt{d}$ and $\frac{x}{y} > \sqrt{d}$. Hence,

$$0 < \frac{x}{y} - \sqrt{d} = \frac{x - y\sqrt{d}}{y} = \frac{x^2 - dy^2}{y(x + y\sqrt{d})} = \frac{n}{y(x + y\sqrt{d})} < \frac{n}{y \cdot 2y\sqrt{d}} < \frac{\sqrt{d}}{2y^2\sqrt{d}} = \frac{1}{2y^2},$$

so by the theorem, x/y is a convergent to $\sqrt{d}$.

Now assume that $n < 0$. Then

$$0 < -\frac{n}{d} = y^2 - \frac{1}{d}\,x^2 = \left(y + x\,\frac{1}{\sqrt{d}}\right)\left(y - x\,\frac{1}{\sqrt{d}}\right),$$

so $y > x\,\frac{1}{\sqrt{d}}$, i.e., $\frac{y}{x} > \frac{1}{\sqrt{d}}$, and in a similar way to the above, we find that

$$0 < \frac{y}{x} - \frac{1}{\sqrt{d}} = \frac{-n/d}{x\left(y + x\frac{1}{\sqrt{d}}\right)} < \frac{-n/d}{x \cdot 2x\frac{1}{\sqrt{d}}} < \frac{1/\sqrt{d}}{2x^2\frac{1}{\sqrt{d}}} = \frac{1}{2x^2},$$

so y/x is a convergent to $1/\sqrt{d}$. This implies that x/y is a convergent to $\sqrt{d}$, for if

$$\sqrt{d} = [a_0; a_1, a_2, \ldots],$$

then

$$\frac{1}{\sqrt{d}} = [0; a_0, a_1, a_2, \ldots],$$

so $y/x = [0; a_0, a_1, \ldots, a_k]$ for some $k \geq 0$, and then

$$\frac{x}{y} = [a_0; a_1, \ldots, a_k],$$

a convergent to $\sqrt{d}$. □

Taking $n = 1$ and $n = -1$ in the corollary shows that any solution to $x^2 - dy^2 = 1$ or $x^2 - dy^2 = -1$ with $x, y \in \mathbb{Z}_{\geq 1}$ must be such that x/y is a convergent to $\sqrt{d}$, say $x/y = p_k/q_k$ for some $k \geq 0$. Further, because the equation $x^2 - dy^2 = \pm 1$ forces x and y to be coprime, we must in fact have $x = p_k$ and $y = q_k$.

Part II

Lemma 14.4. *Let $x \in \mathbb{R} \smallsetminus \mathbb{Q}$, and for $k \geq 0$ let a_k, x_k be the usual numbers appearing in the construction of the continued-fraction representation of x, so that $x = [a_0; a_1, a_2, \ldots]$. Also, let l and n be integers with $l \geq 0$ and $n \geq 1$. Then the following are equivalent:*

(i) $x_{k+n} = x_k$ for all $k \geq l$.

(ii) $a_{k+n} = a_k$ for all $k \geq l$.

Proof. That (i) implies (ii) is obvious: Just take the floor of both sides of the equation and use the fact that $a_k = \lfloor x_k \rfloor$.

Conversely, assume (ii). We saw in the existence part of Section 10 that

$$x = [a_0; a_1, a_2, \ldots, a_{k-1}, x_k]$$

for all $k \geq 0$, so by Proposition 11.2,

$$x_k = [a_k; a_{k+1}, a_{k+2}, \ldots].$$

Hence, if $k \geq l$, then

$$\begin{aligned} x_{k+n} &= [a_{k+n}; a_{k+n+1}, a_{k+n+2}, \ldots] \\ &= [a_k; a_{k+1}, a_{k+2}, \ldots] \quad \text{by the periodicity assumption on the } a_i \\ &= x_k. \end{aligned}$$

□

If $x \in \mathbb{R} \smallsetminus \mathbb{Q}$ and the numbers a_k, x_k are as usual, then define $\tau_k(x) = x_k$ for each $k \geq 0$. We saw in the proof of Lemma 14.4 that

$$\tau_k(x) = [a_k; a_{k+1}, a_{k+2}, \ldots] \tag{14.2}$$

for all $k \geq 0$.

Lemma 14.5. *If $x \in \mathbb{R} \smallsetminus \mathbb{Q}$ and $c \in \mathbb{Z}$, then $\tau_k(x + c) = \tau_k(x)$ for all $k \geq 1$.*

Proof. The case $k = 1$ holds because

$$\tau_1(x + c) = \frac{1}{(x + c) - \lfloor x + c \rfloor} = \frac{1}{x + c - \lfloor x \rfloor - c} = \frac{1}{x - \lfloor x \rfloor} = \tau_1(x).$$

Hence, the desired equality holds for all $k \geq 1$ by induction:

$$\tau_{k+1}(x + c) = \frac{1}{\tau_k(x + c) - \lfloor \tau_k(x + c) \rfloor} = \frac{1}{\tau_k(x) - \lfloor \tau_k(x) \rfloor} = \tau_{k+1}(x).$$

□

Lemma 14.6. *If $x \in \mathbb{R} \smallsetminus \mathbb{Q}$ and the integers p_k, q_k are those appearing in the (p, q)-algorithm applied to x, then for all $k \geq 1$,*

$$x = \frac{\tau_{k+1}(x) p_k + p_{k-1}}{\tau_{k+1}(x) q_k + q_{k-1}}.$$

Proof. Write $x = [a_0; a_1, a_2, \ldots]$. Then using once again the observation that $x = [a_0; a_1, a_2, \ldots, a_k, x_{k+1}]$ where $x_{k+1} = \tau_{k+1}(x)$, we deduce the claimed expression from Theorem 3.1 in Section V. □

In the remainder, d is a positive integer that is not a square, and $\xi = \lfloor \sqrt{d} \rfloor + \sqrt{d}$. We let y_k, b_k, S_k, T_k be the numbers appearing in the (S, T)-algorithm of Section 12 applied to $x = \xi$ with $S_0 = \lfloor \sqrt{d} \rfloor$ and $T_0 = 1$. Note, in particular, that $\tau_k(\xi) = y_k = (S_k + \sqrt{d})/T_k$ for all $k \geq 0$.

Proposition 14.7. *Let notation be as above, and let p_k, q_k be the integers appearing in the (p, q)-algorithm applied to $\sqrt{d}$. Then $p_k^2 - dq_k^2 = (-1)^{k+1} T_{k+1}$ for all $k \geq 0$.*

Proof. Apply the (S, T)-algorithm to $\sqrt{d}$ with the initial S and T equal to 0 and 1 respectively, and let the S's and T's in this case be denoted S'_k and T'_k. Then $\tau_k(\sqrt{d}) =$

$(S'_k + \sqrt{d})/T'_k$, but by Lemma 14.5, $\tau_k(\sqrt{d}) = \tau_k(\xi)$ for all $k \geq 1$, so in particular, $S'_k = S_k$ and $T'_k = T_k$ for all $k \geq 1$. Hence, for all $k \geq 1$,

$$\begin{aligned}\sqrt{d} &= \frac{\tau_{k+1}(\sqrt{d})p_k + p_{k-1}}{\tau_{k+1}(\sqrt{d})q_k + q_{k-1}} \quad \text{by Lemma 14.6} \\ &= \frac{(S_{k+1} + \sqrt{d})p_k + T_{k+1}p_{k-1}}{(S_{k+1} + \sqrt{d})q_k + T_{k+1}q_{k-1}}.\end{aligned}$$

Rearranging this and using the linear independence of 1 and $\sqrt{d}$ over $\mathbb{Q}$, we obtain

$$\begin{aligned}S_{k+1}q_k + T_{k+1}q_{k-1} &= p_k, \\ S_{k+1}p_k + T_{k+1}p_{k-1} &= dq_k.\end{aligned}$$

Multiply the first of these two equations by p_k and the second by q_k, and then perform the obvious subtraction of equations:

$$p_k^2 - dq_k^2 = T_{k+1}(p_kq_{k-1} - p_{k-1}q_k).$$

Recall from Proposition 4.1 in Section V that

$$p_kq_{k-1} - p_{k-1}q_k = (-1)^{k-1} = (-1)^{k+1}.$$

Thus, $p_k^2 - dq_k^2 = (-1)^{k+1}T_{k+1}$ for all $k \geq 1$. In fact, this equality holds for $k = 0$ as well. Indeed, by the formulas for S_1 and T_1 (for the real number $\xi = \lfloor\sqrt{d}\rfloor + \sqrt{d}$), we have

$$\begin{aligned}T_1 &= \frac{d - S_1^2}{T_0} \\ &= d - (b_0T_0 - S_0)^2 \\ &= d - (2\lfloor\sqrt{d}\rfloor - \lfloor\sqrt{d}\rfloor)^2 \\ &= d - \lfloor\sqrt{d}\rfloor^2,\end{aligned}$$

while $p_0^2 - dq_0^2 = \lfloor\sqrt{d}\rfloor^2 - d$. In summary, then, we have

$$p_k^2 - dq_k^2 = (-1)^{k+1}T_{k+1} \quad \text{for all } k \geq 0.$$

□

We recall that our only assumption on d is that it be a positive integer that is not a square.

Theorem 14.8. *Let $\epsilon \in \{1, -1\}$, let n be the period length of the continued-fraction representation of $\sqrt{d}$, and let p_k, q_k be the numbers appearing in the (p, q)-algorithm for the continued-fraction representation of $\sqrt{d}$. Then the solutions (x, y) to the equation $x^2 - dy^2 = \epsilon$ are the pairs (p_{rn-1}, q_{rn-1}) for which the positive integer r satisfies $(-1)^{rn} = \epsilon$.*

Proof. By the discussion following Corollary 14.3, any solution to $x^2 - dy^2 = \epsilon$ has to be (p_k, q_k) for some $k \geq 0$. Therefore, in light of Proposition 14.7, we will be done if we can show all of the following concerning the numbers T_k:

(i) For all $k \geq 0$, $T_k \neq -1$.

(ii) If $T_{k+1} = 1$, then n divides $k + 1$.

(iii) $T_{rn} = 1$ for all $r \geq 1$.

Let us note, before continuing, that since $\xi > 1$ and $-1 < \overline{\xi} < 0$, i.e., ξ is reduced, Proposition 13.2 tells us that the continued-fraction representation of ξ is purely periodic. But ξ and $\sqrt{d}$ differ only by an integer, so their minimum periods are equal, and so in fact ξ is purely periodic with minimum period n. Now let us continue to the proofs of (i)–(iii).

(i) Assume that $T_k = -1$. Then $y_k = -S_k - \sqrt{d}$. But by (14.2), $y_k = \tau_k(\xi)$ has a purely periodic continued-fraction representation because ξ does, so by Proposition 13.2, $-S_k - \sqrt{d}$ is reduced. In particular, $-S_k - \sqrt{d} > 1$ and $-S_k + \sqrt{d} < 0$, so

$$\sqrt{d} < S_k < -1 - \sqrt{d},$$

giving $2\sqrt{d} < -1$, a contradiction.

(ii) Suppose that $T_{k+1} = 1$. Then $y_{k+1} = S_{k+1} + \sqrt{d}$, but $y_{k+1} = \tau_{k+1}(\xi)$ has a purely periodic continued-fraction representation, so $S_{k+1} + \sqrt{d}$ is reduced, and so in particular,

$$-1 < S_{k+1} - \sqrt{d} < 0,$$

or, to put it another way,

$$0 < \sqrt{d} - S_{k+1} < 1.$$

Thus, $\lfloor \sqrt{d} \rfloor = S_{k+1}$, so

$$y_{k+1} = \frac{S_{k+1} + \sqrt{d}}{T_{k+1}} = \lfloor \sqrt{d} \rfloor + \sqrt{d} = y_0.$$

The sequence $y_0, y_1, y_2, \ldots$ is determined by the first-order recurrence relation $y_{j+1} = 1/(y_j - \lfloor y_j \rfloor)$, so the equality $y_{k+1} = y_0$ implies that $k + 1$ is a period of the sequence $(y_0, y_1, y_2, \ldots)$, and so the period length, i.e., minimum period, of this sequence divides $k + 1$. (See the exercise on periodic functions in Section I–1.) But by Lemma 14.4, this period length is equal to the period length n of the continued-fraction representation of ξ, so n divides $k + 1$, as desired.

(iii) Recall that $\tau_k(\xi) = y_k = (S_k + \sqrt{d})/T_k$ for all $k \geq 0$. Now, by Lemma 14.4, $\tau_{k+n}(\xi) = \tau_k(\xi)$ for all $k \geq 0$, so in particular, $T_{k+n} = T_k$. Because $T_0 = 1$, we therefore have $T_{rn} = 1$ for all $r \geq 1$. □

References

[1] Edward J. Barbeau. *Pell's equation.* Problem Books in Mathematics. Springer-Verlag, New York, 2003.

[2] Frazer Jarvis. *Algebraic number theory.* Springer Undergraduate Mathematics Series. Springer, Cham, 2014.

[3] J. L. Lagrange. Solution d'un problème d'arithmétique. *Miscellanea Taurinensia*, IV:671–731, 1768. In *Œuvres de Lagrange.*

[4] Serge Lang. *Algebraic number theory*, volume 110 of *Graduate Texts in Mathematics.* Springer-Verlag, New York, second edition, 1994.

[5] Jürgen Neukirch. *Algebraic number theory*, volume 322 of *Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences].* Springer-Verlag, Berlin, 1999. Translated from the 1992 German original and with a note by Norbert Schappacher, With a foreword by G. Harder.

[6] Kenneth H. Rosen. *Elementary number theory and its applications.* Pearson, fifth edition, 2005.

[7] Wolfgang M. Schmidt. *Diophantine approximation.* Springer, Berlin, 1980.

[8] J.-P. Serre. *A course in arithmetic*, volume No. 7 of *Graduate Texts in Mathematics.* Springer-Verlag, New York-Heidelberg, 1973. Translated from the French.

[9] Joseph H. Silverman. *The arithmetic of elliptic curves*, volume 106 of *Graduate Texts in Mathematics.* Springer, Dordrecht, second edition, 2009.

[10] John Stillwell. *Mathematics and its history.* Undergraduate Texts in Mathematics. Springer, Cham, concise edition, [2020] ©2020.

[11] Lawrence C. Washington. *Introduction to cyclotomic fields*, volume 83 of *Graduate Texts in Mathematics.* Springer-Verlag, New York, second edition, 1997.