A Concise Course in Number Theory

Alan Baker, Cambridge University Press, 1983

Chapter 1: Divisibility

Prime number: a positive integer that cannot be factored into strictly smaller factors. For example, 2, 3, 5, 7.

Every positive integer $n \geq 2$ can be factored into prime numbers: use strong induction on n.

Greatest common divisor of two numbers: maximum common divisor.

Division algorithm: For each pair of integers a and $b \neq 0$ there exists a unique pair of integers q and r so that $a = qb + r$, $0 \leq r < |b|$. Proof: The real number line is partitioned into intervals of the form $[Q|b|, (Q+1)|b|)$ where Q is an integer. Find the one containing a. Then find q so that $qb = Q|b|$ and set $r = a - qb$. A formula for q is $q = \left\lfloor \frac{a}{|b|} \right\rfloor \frac{|b|}{b}$.

Euclid's algorithm for constructing greatest common divisor of a and $b \neq 0$: Form the sequence $a_0, a_1, a_2, \ldots$ with $a_1 > a_2 > \cdots \geq 0$ via $a_0 = a$, $a_1 = b$, and for $k \geq 2$, $a_{k-2} = q_{k-2}a_{k-1} + a_k$ where $0 \leq a_k < a_{k-1}$. The sequence has to terminate with some $a_n = 0$ for some $n \geq 2$, and a_{n-1} is the greatest common divisor. Reason: The recurrence relation can be expressed in the form

$$\begin{bmatrix} a_{k-2} \\ a_{k-1} \end{bmatrix} = \begin{bmatrix} q_{k-2} & 1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} a_{k-1} \\ a_k \end{bmatrix}.$$

This can be used to obtain

$$\begin{bmatrix} q_0 & 1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} q_1 & 1 \\ 1 & 0 \end{bmatrix} \cdots \begin{bmatrix} q_{n-2} & 1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} a_{n-1} \\ 0 \end{bmatrix} = \begin{bmatrix} a_0 \\ a_1 \end{bmatrix}.$$

Simplifying,

$$\begin{bmatrix} x & y \\ z & w \end{bmatrix} \begin{bmatrix} a_{n-1} \\ 0 \end{bmatrix} = \begin{bmatrix} p \\ q \end{bmatrix}.$$

Hence

$$\begin{bmatrix} xa_{n-1} \\ za_{n-1} \end{bmatrix} = \begin{bmatrix} p \\ q \end{bmatrix}.$$

So we can see that a_{n-1} is a common divisor of p and q. Moreover if d is a divisor of both p and q then the recurrence relation can be used to show that

d divides each a_k, including a_{n-1}. Hence $d \leq a_{n-1}$ and a_{n-1} is the greatest common divisor.

Note that the inverse of $\begin{bmatrix} q_k & 1 \\ 1 & 0 \end{bmatrix}$ is $\begin{bmatrix} 0 & 1 \\ 1 & -q_k \end{bmatrix}$. This implies that

$$\begin{bmatrix} a_{n-1} \\ 0 \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ 1 & -q_{n-2} \end{bmatrix} \begin{bmatrix} 0 & 1 \\ 1 & -q_{n-3} \end{bmatrix} \cdots \begin{bmatrix} 0 & 1 \\ 1 & -q_0 \end{bmatrix} \begin{bmatrix} a_0 \\ a_1 \end{bmatrix}.$$

Simplifying,

$$\begin{bmatrix} a_{n-1} \\ 0 \end{bmatrix} = \begin{bmatrix} x' & y' \\ z' & w' \end{bmatrix} \begin{bmatrix} a_1 \\ a_0 \end{bmatrix},$$

$$x'p + y'q = a_{n-1}.$$

In other words, given integers p and q with greatest common divisor d there is always a pair of integers j and k such that $jp + kq = d$. Whenever we have $jp + kq = r$ we must have $d|r$. In particular, when $jp + kq = 1$ we must have $d = 1$.

Example: Let $a = 108$ and $b = 93$. We have

$$108 = 1 \cdot 93 + 15$$

$$93 = 6 \cdot 15 + 3$$

$$15 = 5 \cdot 3 + 0$$

hence $a_0 = 108$, $a_1 = 93$, $a_2 = 15$, $a_3 = 3$, $a_4 = 0$, $q_0 = 1$, $q_1 = 6$, $q_2 = 5$. Therefore $\gcd(108, 93) = 3$. Substituting these values into

$$\begin{bmatrix} a_{n-1} \\ 0 \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ 1 & -q_{n-2} \end{bmatrix} \begin{bmatrix} 0 & 1 \\ 1 & -q_{n-3} \end{bmatrix} \cdots \begin{bmatrix} 0 & 1 \\ 1 & -q_0 \end{bmatrix} \begin{bmatrix} a_0 \\ a_1 \end{bmatrix}$$

yields

$$\begin{bmatrix} 3 \\ 0 \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ 1 & -5 \end{bmatrix} \begin{bmatrix} 0 & 1 \\ 1 & -6 \end{bmatrix} \begin{bmatrix} 0 & 1 \\ 1 & -1 \end{bmatrix} \begin{bmatrix} 108 \\ 93 \end{bmatrix}.$$

Simplifying,

$$\begin{bmatrix} 3 \\ 0 \end{bmatrix} = \begin{bmatrix} -6 & 7 \\ 31 & -36 \end{bmatrix} \begin{bmatrix} 108 \\ 93 \end{bmatrix}.$$

This yields

$$3 = (-6)(108) + 7(93).$$

A useful lemma is that when $(a,b)=1$ and $a|bc$ then $a|c$. Reason: $bc=ak$ and $xa+by=1$ implies $c=cxa+cby=cxa+aky=a(cx+ky)$.

We now prove unique factorization for all integers $n\geq 2$. There is only one factorization of 2 into a weakly descending list of primes. Now assume that every integer ≥ 2 up to n has a unique factorization into a weakly descending list of primes. Suppose $n+1=p_1p_2\cdots p_j=q_1q_2\ldots q_k$ with $p_1\geq p_2\cdots\geq p_j$ and $q_1\geq q_2\geq\cdots\geq q_k$. We will assume wlog that $n+1$ is not prime and that $p_1\geq q_1$If $p_1>q_1$ then $(p_1,q_1)=1$, therefore by the lemma $p_1|q_2\cdots q_k$. If $p_1\neq q_2$ then $(p_1,q_2)=1$ and $p_1|q_3\cdots q_k$. After a finite number of steps we arrive at $p_1=q_i$ for some i, which implies $p_1\leq q_1$. Contradiction. Therefore $p_1=q_1$. Dividing both sides by p_1 we have two factorizations of $(n+1)/p_1\geq 2$ into descending lists of primes, so the factorizations must be the same, so the two factorizations of $n+1$ must be the same.

Note that whenever $p_1,p_2,\ldots,p_n$ are the first n primes then $p_1p_2\cdots p_n+1$ is not divisible by any of these. So it is either prime or has a prime factor not equal to any of these. Hence there are infinitely many primes.

Greatest common divisor and least common multiple construction via prime factorization.

1.8 Exercises, p. 7

(i) Using Euclid's method and matrix calculations, $(x,y)=(191,-42)$.

(ii) Since $(35,55)=5$ there is a solution to $35x+55y=5$. Since $(5,77)=1$ there is a solution to $5p+77q=1$. This yields $35xp+55yp+77q=1$. Given $(x,y)=(-3,2)$ and $(p,q)=(31,-2)$, we obtain $(xp,yp,q)=(-93,62,-2)$.

(iii) Let the primes $\leq n$ be labeled $p_1,p_2,\ldots,p_j$ where $p_1=2$. Let d be the least common multiple of $1,2,\ldots,n$ and set $m=1+\frac{1}{2}+\cdots+\frac{1}{n}$. We will show that dm is an odd integer. This implies that m is not an integer, because if it were then dm would be even since d is even. To construct d we inspect the prime factorization of each of the numbers 1, 2, ..., n, then multiply the highest power of p_1 in these factorizations times the highest power of p_2 in these factorizations times etc. Let 2^a be the largest power of 2 in $\{1,2,\ldots,n\}$. We claim that this is the uniquely highest power of 2 in the prime factorization of these numbers. To see this, let $k\neq 2^a$ be given in this range. Write $k=p_1^{f_1}p_2^{f_2}\cdots p_j^{f_j}$. Then $2^{f_1+f_2+\cdots+f_j}\leq p_1^{f_1}\cdots p_k^{f_j}=k\leq n$, therefore $f_1+\cdots+f_j\leq a$, therefore $f_1<a$. So know we know

$d = p_1^a p_2^{e_2} \cdots p_k^{e_k}$. This implies that $\frac{d}{k}$ is even when $k \neq 2^a$ and that $\frac{d}{2^a}$ is odd, hence dm is odd.

(iv) I assume $(x, y, \dots)$ stands for greatest common divisor and $\{x, y, \dots\}$ stands for least common multiple. To compute these we inspect prime factorizations and pick out lowest or highest powers of primes. The identity boils down to showing

$$\min(\max(a,b), \max(b,c), \max(c,a)) = \max(\min(a,b), \min(b,c), \min(c,a)).$$

This true: both sides are equal to b.

(v) Let the integers in question be $g_1, g_2, \dots, g_k$. $n_0 = a_0 + n_1 g_1$ determines a_0 and n_1 uniquely by the division algorithm. $n_1 = a_1 + n_2 g_2$ determines a_1 and n_2 uniquely by the division algorithm. Keep on going. Now make all the substitutions and solve for n_0.

(vi) Let $p_1, \dots, p_k$ be the complete list of primes of the form $4n+3$. Consider the number $x = 2 + p_1^2 \cdots p_k^2$. It is congruent to 3 mod 4, so its prime factorization cannot include 2 and cannot be comprised exclusively of primes of the form $4n+1$. Hence it must be divisible by some p_i: contradiction. Hence there are infinitely many primes of the form $4n+3$.

(vii) Fermat primes are primes of the form $2^{2^n} + 1$. Now suppose $2^n + 1$ is a prime number. Then the polynomial $x^n + 1$ does not factor. This implies that n does not have any odd divisors, because if $n = pq$ where q is odd then $y^q + 1$ factors (has root -1) hence $(x^p)^q + 1$ factors hence $x^n + 1$ factors. Since n has no odd divisors we must have $n = 2^m$ for some m, which makes $2^n + 1$ a Fermat prime.

(viii) Let d be even. Then $x+1$ divides $x^d - 1$ since the latter has root $x = -1$. So $x^m + 1$ divides $x^{md} - 1$. Now let m and d be powers of 2 to conclude $2^{2^n} + 1$ divides $2^{2^m} - 1$ when $n < m$. Write $2^{2^m} - 1 = k(2^{2^n} + 1)$. Then $1 \cdot (2^{2^m} + 1) - k \cdot (2^{2^n} + 1) = 2$. Any common divisor of $2^{2^m} + 1$ and $2^{2^n} + 1$ is a divisor of 2 and so must be 1 since 2 is ruled out.

(xi) Let $p_1 < p_2 < p_3 < \cdots$ be the prime numbers. Given that $1 + \prod_{i=1}^n p_i$ is not divisible by p_i for any $i \leq n$, it must be divisible by some p_j for some $j > n$. This implies $p_{n+1} \leq p_j \leq 1 + \prod_{i=1}^n p_i$. So we have established that $p_{n+1} \leq 1 + \prod_{i=1}^n p_i$. We have $p_1 \leq 2^{2^0}$. Assume $p_k \leq 2^{2^{k-1}}$ for $1 \leq k \leq n$. Then we know that we can find p_m for some $m \geq n+1$ dividing $p_1 \cdots p_n + 1$, therefore $p_{n+1} \leq p_m \leq p_1 \cdots p_n + 1 \leq 2^{2^0 + \cdots + 2^{n-1}} + 1 = 2^{2^n - 1} + 1 \leq 2^{2^n}$. Hence we have proved $p_n \leq 2^{2^{n-1}}$ for all n. Now let x be an integer in the range $2^{2^{k-1}}+1, 2^{2^k}+2, \ldots, 2^{2^k}$. Given at $p_k \leq 2^{2^{k-1}}$, we have $\pi(x) \geq k$. On the other hand, we have $\log_2 \log_2(x) \leq k$, therefore we have $\pi(x) \geq \log_2 \log_2(x)$ for $x \geq 3$.

Exercise: Show that $\frac{1}{3} + \frac{1}{5} + \cdots + \frac{1}{2n+1}$ is not an integer.

Exercise: Find an upper bound for $q_n = n^{th}$ largest prime of the form $4n+3$.

Exercise: Find a lower bound for $\pi'(x)$ = number of primes of the form $4n+3$ that are $\leq x$.

Chapter 2: Arithmetical Functions

2.1 The function $[x]$:

If $n \leq x < n+1$ then $[x] = n$. Now write $n = qd + r$, $0 \leq r < d$. Then $qd + r \leq x < (q+1)d$, therefore $q + \frac{r}{d} \leq \frac{x}{d} < q+1$, therefore $\left[\frac{x}{d}\right] = q = \left[\frac{n}{d}\right] = \left[\frac{[x]}{d}\right]$.

Write $n = qd + r$, $0 \leq r < d$. Then

$$[n/d] = q = |\{d, 2d, \ldots, qd\}| = \sum_{k=1}^{n} \chi(d|k).$$

Write $m \leq x < m+1$ and $n \leq y < m+1$. Then $m + n \leq x + y$, therefore $[x] + [y] = m + n \leq [x+y]$.

Let p be a prime and $l_p(n)$ the largest power of p in n. Then

$$l_p(n!) = \sum_{k=1}^{n} l_p(k) = \sum_{k=1}^{n} \sum_{j \geq 1} \chi(p^j|k) = \sum_{j \geq 1} \sum_{k=1}^{n} \chi(p^j|k) =$$

$$\sum_{j \geq 1} \left[\frac{n}{p^j}\right] \leq \sum_{j \geq 1} \frac{n}{p^j} = \frac{n}{p} \frac{1}{1 - \frac{1}{p}} = \frac{n}{p-1}.$$

For $n = a + b$,

$$l_p(a!b!) = l_p(a!) + l_p(b!) = \sum_{j \geq 1} \left[\frac{a}{p^j}\right] + \sum_{j \geq 1} \left[\frac{b}{p^j}\right] \leq \sum_{j \geq 1} \left[\frac{n}{p^j}\right] = l_p(n!).$$

This implies $a!b!|n!$. Of course we already knew this because $\frac{n!}{a!b!} = \binom{n}{a} =$ number of a-subsets of $\{1, 2, \ldots, n\}$.

2.2 Multiplicative functions and generating functions: A multiplicative arithmetical function is a function $f : \mathbb{Z}^+ \to \mathbb{R}$ that satisfies $f(ab) = f(a)f(b)$ when $(a, b) = 1$, and more generally

$$f(p_1^{e_1} p_2^{e_2} \cdots) = f(p_1^{e_1}) f(p_2^{e_2}) \cdots .$$

When f is not nontrivial (not identically 0) then $f(1) = 1$.

Generating function of a non-trivial multiplicative function: Let f be a non-trivial multiplicative function and set

$$F_k(t_k) = \sum_{e=0}^{\infty} f(p_k^e) t_k^e.$$

Then

$$f(p_1^{e_1} p_2^{e_2} \cdots) = f(p_1^{e_1}) f(p_2^{e_2}) \cdots = [t_1^{e_1} t_2^{e_2} \cdots] F_1(t_1) F_2(t_2) \cdots .$$

Therefore a generating function for f is $F_f(t) = F_f(t_1, t_2, \ldots) = F_1(t_1)F_2(t_2) \cdots$. Any such product with constant term 1 is the generating function of a multiplicative arithmetic function.

Products of generating functions:

If

$$F(t) = F(t_1, t_2, \ldots) = \sum f(p_1^{e_1} p_2^{e_2} \cdots) t_1^{e_1} t_2^{e_2} \cdots = \sum_{n \geq 1} f(n) t^n$$

and

$$G(t) = G(t_1, t_2, \ldots) = \sum g(p_1^{e_1} p_2^{e_2} \cdots) t_1^{e_1} t_2^{e^2} \cdots = \sum_{n \geq 1} g(n) t^n$$

then

$$F(t)G(t) = \sum f(p_1^{a_1} p_2^{a_2} \cdots) g(p_1^{b_1} p_2^{b_2} \cdots) t_1^{a_1+b_1} t_2^{a_2+b_2} \cdots =$$

$$\sum_{n \geq 1} \sum_{d|n} f(d) g(n/d) t^n.$$

This implies that if a and b are multiplicative functions with generating functions $F_a(t)$ and $F_b(t)$ then the multiplicative function c with generating function $F_a(t)F_b(t)$ is defined by

$$c(n) = \sum_{d|n} a(d)b(n/d) = \sum_{d|n} b(d)a(n/d).$$

Examples:

1. The unit function $u(n) = 1$ has generating function $F_u(t) = \frac{1}{(1-t_1)(1-t_2)\cdots}$. If $f(n)$ is multiplicative then so is

$$g(n) = \sum_{d|n} f(n/d) = \sum_{d|n} f(d)$$

and

$$F_g(t) = F_u(t)F_f(t) = \frac{F_f(t)}{(1-t_1)(1-t_2)\cdots}.$$

2. The identity function $i(n) = n$ has generating function $\frac{1}{(1-p_1t_1)(1-p_2t_2)\cdots}$. If $f(n)$ is multiplicative then so is

$$h(n) = \sum_{d|n} df(n/d) = \sum_{d|n} f(d)\frac{n}{d}$$

and

$$F_h(t) = \frac{F_f(t)}{(1-p_1t_1)(1-p_2t_2)\cdots}.$$

3. The Möbius function $\mu(n)$ defined by

$$\mu(p_1^{e_1}\cdots p_k^{e_k}) = (-1)^k\chi(e_1 = \cdots = e_k = 1)$$

has generating function

$$F_\mu(t) = (1-t_1)(1-t_2)\cdots,$$

hence is multiplicative. If f is a multiplicative function and g is defined by

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

then we have seen by Example 1 above that

$$F_g(t) = \frac{F_f(t)}{(1-t_1)(1-t_2)\cdots} = F_u(t)F_f(t) = \frac{F_f(t)}{F_\mu(t)}.$$

This implies

$$F_f(t) = F_\mu(t)F_g(t),$$

hence

$$f(n) = \sum_{d|n} \mu(d)g(n/d) = \sum_{d|n} g(d)\mu(n/d).$$

In particular,

$$f(p^e) = g(p^e) - g(p^{e-1})$$

when p is prime and $e \geq 1$.

4. The unit characteristic function $\nu(n) = \chi(n = 1)$ has generating function $F_\nu(t) = 1$. Given that $F_\nu(t) = F_u(t)F_\mu(t)$, we have

$$\nu(n) = \sum_{d|n} \mu(d) = \sum_{d|n} \mu(n/d).$$

5. Euler's (totient) function $\phi(n)$: This is defined as the number of natural numbers $\leq n$ that are relatively prime to n. Using the inclusion-exclusion sum formula (see below), we have

$$\phi(n) = \sum_{d|P} \mu(d) \sum_{a\in A_d} 1 = \sum_{d|P} \mu(d)\frac{n}{d} = n\sum_{d|P} \mu d\frac{1}{d} = n(1-\frac{1}{p_1})(1-\frac{1}{p_2})\cdots(1-\frac{1}{p_r}).$$

One can check that ϕ is multiplicative given this formula. Now define

$$g(n) = \sum_{d|n} \phi(d).$$

This is multiplicative. It satisfies

$$g(p^k) = \sum_{i=0}^{k} \phi(p^i) = 1 + (p-1) + (p^2 - p) + \cdots + (p^k - p^{k-1}) = p^k,$$

hence $g(n) = n$ for all n. Therefore

$$\sum_{d|n} \phi(d) = n.$$

To obtain a generating function for $\phi(n)$, note that

$$F_i(t) = F_g(t) = F_u(t)F_\phi,$$

hence

$$F_\phi(t) = F_\mu(t)F_i(t) = \frac{(1-t_1)(1-t_2)\cdots}{(1-p_1t_1)(1-p_2t_2)\cdots}.$$

6. Möbius Inversion: Let $f : \mathbb{R} \to \mathbb{R}$ be given and define

$$g(x) = \sum_{n \le x} f(x/n),$$

summing over positive integers. Then

$$\sum_{n\le x} \mu(n)g(x/n) = \sum_{n\le x} \mu(n) \sum_{m \le x/n} f(x/mn) = \sum_{n \le x} \mu(n) \sum_{mn\le x} f(x/mn) =$$

$$\sum_{l \le x} f(x/l) \sum_{m|l} \mu(l/m) = \sum_{l\le x} f(x/l)\nu(l) = f(x).$$

Conversely, if we define

$$f(x) = \sum_{n\le x} \mu(n)g(x/n)$$

then

$$\sum_{n\le x} f(x/n) = \sum_{n\le x} \sum_{k\le x/n} \mu(n)g(x/kn) = \sum_{n\le x} \sum_{kn\le x} \mu(n)g(x/kn) = \sum_{l\le x} g(x/l) \sum_{m|l} \mu(l/m) =$$

$$\sum_{l\le x} g(x/l)\nu(l) = g(x).$$

When a multiplicative function is used to define the other this way then the second function is also multiplicative, and we obtain

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

if and only if

$$f(n) = \sum_{d|n} \mu(d) g(n/d).$$

We already derived this by the method of generating functions above.

7. Applying Möbius inversion to the functions

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

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

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

we obtain

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

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

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

The identities above also follow from $F_u = F_\tau F_\mu$, $F_i = F_\sigma F_\mu$, $F_\phi = F_i F_\mu$.

8. Summary of generating functions:

$\mu(n)$: $F_\mu = (1 - t_1)(1 - t_2)\cdots$

$\nu(n) = \chi(n = 1) = \sum_{d|n} \mu(n)$: $F_\nu = 1$

$u(n) = 1$: $F_u = \frac{1}{(1-t_1)(1-t_2)\cdots}$

$i(n) = n$: $F_i = \frac{1}{(1-p_1 t_1)(1-p_2 t_2)\cdots}$

$\tau(n) = \sum_{d|n} 1 = \sum_{d|n} u(d)$: $F_\tau = F_u^2 = \frac{1}{(1-t_1)^2(1-t_2)^2\cdots}$

$\sigma(n) = \sum_{d|n} d = \sum_{d|n} i(d)$: $F_\sigma = F_u F_i = \frac{1}{(1-t_1)(1-t_2)\cdots(1-p_1 t_1)(1-p_2 t_2)\cdots}$

$\phi(n)$: $F_\phi = \frac{(1-t_1)(1-t_2)\cdots}{(1-p_1 t_1)(1-p_2 t_2)\cdots} = F_\mu F_i$.

9. The Riemann zeta-function. Take any generating function $F(t) = F(t_1, t_2, \dots) = F_1(t_1)F_2(t_2)\cdots$ for a multiplicative function f. Making the substitution $t_i \mapsto \frac{1}{p_i^s}$ where s is a complex number yields an infinite product. For example, recall that we have

$$F_u(t) = \frac{1}{(1-t_1)(1-t_2)\cdots} = \sum_{e_1,e_2,e_3,\dots\geq 0} t_1^{e_1}t_2^{e_2}t_3^{e_3}\cdots.$$

Hence

$$F_u(s) = F_i(1/p_1^s, 1/p_2^s, \dots) = \sum_{e_1,e_2,e_3,\dots\geq 0} \frac{1}{p_1^{e_1}p_2^{e_2}p_3^{e_3}\cdots} = \sum_{n=1}^{\infty} \frac{1}{n^s}.$$

This is called the Riemann zeta-function $\zeta(s)$. In particular,

$$\zeta(2) = \prod_p \frac{1}{1-(1/p^2)} = \sum_{n=1}^{\infty} \frac{1}{n^2} = \frac{\pi^2}{6}.$$

We will derive this evaluation this shortly.

More generally, if $F_f(t) = \sum_{n=1}^{\infty} f(n)t^n$ then

$$F_f(s) = F_f(1/p_1^s, 1/p_2^s, \dots) = \sum_{n=1}^{\infty} \frac{f(n)}{n^s}.$$

Examples:

1. $F_\mu(s) = \frac{1}{F_u(s)}$. This implies

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

In particular,

$$\sum_{n=1}^{\infty} \frac{\mu(n)}{n^2} = \frac{6}{\pi^2}.$$

2. $F_\tau(s) = F_u(s)^2$. This implies

$$\sum_{n=1}^{\infty} \frac{\mu(n)}{n^s} = \zeta(s)^2.$$

In particular,

$$\sum_{n=1}^{\infty} \frac{\mu(n)}{n^2} = \frac{\pi^4}{36}.$$

3. $F_i(s) = \sum_{n=1}^{\infty} \frac{n}{n^s} = \zeta(s-1)$.
4. $F_\sigma(s) = F_i(s)F_u(s)$. This implies

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

5. $F_\phi(s) = F_\mu(s)F_i(s)$. This implies

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

6. For arbitrary functions $f : \mathbb{Z}^+ \to \mathbb{R}$ and $g : \mathbb{Z}^+ \to \mathbb{R}$ we have

$$\sum_{n=1}^{\infty} \frac{f(n)}{n^s} \sum_{n=1}^{\infty} \frac{g(n)}{n_s} = \sum_{a,b \geq 1} \frac{f(a)g(b)}{(ab)^s} = \sum_{n=1}^{\infty} \frac{1}{n^s} \sum_{d|n} f(d)g(n/d)$$

assuming the expressions converge.

Derivation of $\sum_{n=1}^{\infty} \frac{1}{n^2} = \frac{\pi^2}{6}$:

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

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

$$\int_0^{\frac{\pi}{2}} \frac{\sin^{2n+1} x}{2n+1} \; dx = \frac{2 \cdot 4 \cdots (2n)}{1 \cdot 3 \cdots (2n+1)},$$

$$\frac{\pi^2}{8} = \int_0^{\frac{\pi}{2}} x \; dx = \sum_{n=0}^{\infty} \frac{1}{(2n+1)^2} = \sum_{n=1}^{\infty} \frac{1}{n^2} - \sum_{n=1}^{\infty} \frac{1}{4n^2} = \frac{3}{4} \sum_{n=1}^{\infty} \frac{1}{n^2},$$

$$\frac{\pi^2}{6} = \sum_{n=1}^{\infty} \frac{1}{n^2}.$$

2.5 The functions $\tau(n)$ and $\sigma(n)$: These are the multiplicative functions defined by

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

and

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

We have

$$\tau(p_1^{e_1} p_2^{e_2} \cdots) = (e_1 + 1)(e_2 + 1) \cdots$$

and

$$\sigma(p_1^{e_1} p_2^{e_2} \cdots) = [e_1 + 1]_{p_1} [e_2 + 1]_{p_2} \cdots .$$

Given

$$\log(\tau(p^k)^{\frac{1}{\delta}}) = \frac{1}{\delta} \log(k+1) \leq k \log p$$

for all but a finite number of values of k and p,

$$\tau(p^k) \leq p^{k\delta}$$

and

$$\tau(n) \leq c n^{\delta}$$

where c is large enough to compensate for the exceptions. Also,

$$\sigma(n) = \sum_{d|n} \frac{n}{d} \leq n \sum_{d \leq n} \frac{1}{d} < n(1 + \log n) < 2n \log n,$$

the estimate coming from an integral comparison.

A lower bound for $\phi(n)$: set $f(n) = \sigma(n)\phi(n)/n^2$. This is a multiplicative function, so to evaluate it it suffices to evaluate $f(p^k)$ for a prime p and $k \geq 1$. We have

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

$$\phi(p^k) = p^k - p^{k-1} = p^{k-1}(p-1),$$

$$f(p^k) = \frac{p^{2k} - p^{k-1}}{p^{2k}} = 1 - \frac{1}{p^{k+1}} \geq 1 - \frac{1}{p^2},$$

$$f(n) \geq \prod_{m \geq 2} \left(1 - \frac{1}{m^2}\right) = \frac{1}{2},$$

the latter a limit of finite products, hence

$$\sigma(n)\phi(n)/n^2 \geq \frac{1}{2},$$

$$\phi(n) \geq \frac{n^2}{2\sigma(n)} > \frac{n^2}{4n \log n} = \frac{n}{4 \log n}.$$

The ideas in this proof: (1) To calculate or estimate a multiplicative function, combine them in such a way that things cancel; (2) use properties of inequalities; (3) exploit known formulas such as infinite sums, infinite products, and integrals.

2.6 Average orders: It's time to start using big-O notation. When we say $f(x) = g(x) + O(h(x))$ we mean that

$$g(x) - Ch(x) \leq f(x) \leq g(x) + Ch(x)$$

for some constant $C > 0$ independent of x.

Let x be an integer.

$$\sum_{n \leq x} \tau(n) = \sum_{n \leq x} \sum_{d|n} 1 = \sum_{d \leq x} \sum_{\substack{n \leq x \\ d|n}} 1 = \sum_{d \leq x} |\{d, 2d, 3d, \dots\} \cap [0, x]| =$$

$$\sum_{d \leq x} |\{1, 2, 3, \dots\} \cap [0, \frac{x}{d}]| = \sum_{d=1}^{x} \left[\frac{x}{d}\right] = \sum_{d=1}^{x} \left(\frac{x}{d} + O(1)\right) = x \sum_{d=1}^{x} \frac{1}{d} + O(x) =$$

$$x(\log x + O(1)) + O(x) = x \log x + O(x),$$

hence

$$\frac{1}{x} \sum_{n \leq x} \tau(n) = \log x + O(1).$$

Let x be a positive integer.

$$\sum_{n \leq x} \sigma(n) = \sum_{n \leq x} \sum_{d|n} \frac{n}{d} = \sum_{d \leq x} \sum_{\substack{n \leq x \\ d|n}} \frac{n}{d} =$$

$$\sum_{d\le x}\sum_{n\in\{d,2d,\dots\}\cap[0,x]}\frac{n}{d}=\sum_{d\le x}\sum_{\frac{n}{d}\in\{1,2,\dots\}\cap[0,\left[\frac{x}{d}\right]]}\frac{n}{d}=\sum_{d\le x}\frac{\left[\frac{x}{d}\right]\left(\left[\frac{x}{d}\right]+1\right)}{2}=$$

$$\sum_{d\le x}\frac{1}{2}\left(\left(\frac{x}{d}\right)^2+O(\frac{x}{d})\right)=$$

$$=\frac{x^2}{2}(\sum_{d=1}^{\infty}\frac{1}{d^2}+O(1/x))+O(x\log x)=\frac{\pi^2x^2}{12}+O(x\log x).$$

We have used $\sum_{d=1}^{\infty}\frac{1}{d^2}=\frac{\pi^2}{6}$.

Let x be a positive integer.

$$\sum_{n\le x}\phi(n)=\sum_{n\le x}\sum_{d|n}\mu(d)\frac{n}{d}=\sum_{d\le x}\mu(d)\sum_{\substack{d|n\\ n\le x}}\frac{n}{d}=$$

$$\sum_{d\le x}\mu(d)\left(\frac{1}{2}\left(\frac{x}{d}\right)^2+O(\frac{x}{d})\right)=\frac{x^2}{2}\left(\sum_{d=1}^{\infty}\frac{\mu(d)}{d^2}+O(1/x)\right)+O(x\log x)$$

$$=\frac{3x^2}{\pi^2}+O(x\log x).$$

We have used

$$\sum_{d=1}^{\infty}\frac{\mu(d)}{d^2}=\frac{6}{\pi^2}.$$

We obtain

$$\sum_{n\le x}\phi(n)=\frac{3x^2}{\pi^2}+O(x\log x).$$

Note that the number of integer pairs (p,q) such that $1\le p<q\le x$ and $\gcd(p,q)=1$ is $\sum_{q\le x}\phi(q)$. Given there are $\frac{x^2-x}{2}$ such pairs, the probability that $p<q$ are relatively prime in $\{1,2,\dots,x\}$ is $\sim\frac{6}{\pi^2}$.

2.7: Perfect numbers: Perfect number is a positive integer which is the sum of its proper divisors: $\sigma(n)=2n$. Examples include 6 and 28. More generally, if p is a prime such that 2^p-1 is prime (i.e. a Mersenne prime), then

$$\sigma(2^{p-1}(2^p-1))=\sigma(2^{p-1})\sigma(2^p-1)=(1+2+\dots+2^{p-1})(1+(2^p-1))=(2^p-1)2^p$$

hence $2^{p-1}(2^p - 1)$ is a perfect number. These are the only even examples: Let n be perfect and write $n = 2^a b$ where b is odd and $a \geq 1$. The divisors of n are of the form $2^i d$ where $d|b$. Hence

$$2^{a+1}b = 2n = \sigma(n) = (2^{a+1} - 1)\sigma(b).$$

Unique factorization yields $\sigma(b) = 2^{a+1}b_0$ hence $b = (2^{a+1}-1)b_0$, which yields information about $\sigma(b)$. If $b_0 > 1$ then b and b_0 contribute to $\sigma(b)$, hence

$$2^{a+1}b_0 = \sigma(b) \geq b + b_0 + 1 = (2^{a+1} - 1)b_0 + b_0 + 1 > 2^{a+1}b_0,$$

a contradiction. Therefore $b_0 = 1$, $b = 2^{a+1} - 1$, $n = 2^a(2^{a+1} - 1)$. We now have

$$\begin{aligned} 2^{a+1}(2^{a+1} - 1) = 2n = \sigma(n) &= (2^{a+1} - 1)\sigma(2^{a+1} - 1), \\ 2^{a+1} &= \sigma(2^{a+1} - 1), \\ b + 1 &= \sigma(b), \end{aligned}$$

which implies that b is a prime number. So $2^{a+1} - 1$ is prime. This forces $a + 1$ prime: $x^{hk} - 1 = y^k - 1$ has root $y = 1$, hence

$$x^{hk} - 1 = y^k - 1 = (y - 1)g(y) = (x^h - 1)g(x^h).$$

Given $n = 2^a(2^{a+1} - 1)$ we have $n = 2^{p-1}(2^p - 1)$ where p and $2^p - 1$ are prime.

Remark: According to Davenport (An Introduction to Higher Arithmetic), it is not known if there are infinitely many perfect numbers or if there are any odd perfect numbers.

The ideas in this proof: (1) Compare numbers and use unique factorization; (2) information about how n factors yields information about $\sigma(n)$; (3) exploit inequalities; (4) a number n is prime if $\sigma(n) = n + 1$.

Inclusion-Exclusion Sum and Product

Let $A_1, \ldots, A_n$ be a union of finite sets of natural numbers. Let f be an arithmetic function. For a finite subset A of natural numbers write

$$||A|| = \sum_{a \in A} f(a).$$

Let $s(a)$ the number of sets that a belongs to. Then for each $1 \leq k \leq n$,

$$\sum_{I \in \binom{[n]}{k}} \chi(a \in A_I) f(a) = f(a) \binom{s(a)}{k}.$$

Now sum over each $a \in A_1 \cup \cdots A_n$. We obtain

$$\sum_{I \in \binom{[n]}{k}} ||A_I|| = \sum_{a \in A_1 \cup \cdots \cup A_n} f(a) \binom{s(a)}{k}.$$

Now form the alternating sum

$$\sum_{k=1}^{n} (-1)^{k-1} \sum_{I \in \binom{[n]}{k}} ||A_I|| = \sum_{k=1}^{n} (-1)^{k-1} \sum_{a \in A_1 \cup \cdots \cup A_n} f(a) \binom{s(a)}{k}.$$

The sum on the right-hand side can be reorganized into

$$\sum_{a \in A_1 \cup \cdots \cup A_n} f(a) \sum_{k=1}^{n} (-1)^{k-1} \binom{s(a)}{k}.$$

Each of the expressions $\sum_{k=1}^{n} (-1)^{k-1} \binom{s(a)}{k}$ is equal to 1 by the Binomial Theorem. Hence we obtain

$$\sum_{k=1}^{n} (-1)^{k-1} \sum_{I \in \binom{[n]}{k}} ||A_I|| = \sum_{a \in A_1 \cup \cdots \cup A_n} f(a) = ||A_1 \cup \cdots \cup A_n||.$$

Example: Let f be an arithmetic function. We wish to evaluate

$$\sum_{\substack{1 \leq a \leq n \\ (a,n)=1}} f(a).$$

Let the prime factorization of n be $n = p_1^{e_1} \cdots p_r^{e_r}$. For each $d \leq n$ let $A_d = \{a \leq n : d|a\}$. We have

$$\{a : 1 \leq a \leq n \text{ and } (a,n) > 1\} = A_{p_1} \cup A_{p_2} \cup \cdots \cup A_{p_r}.$$

By the inclusion-exclusion formula,

$$\sum_{\substack{1\le a\le n\\(a,n)>1}} f(a) = -\sum_{d|P}\mu(d)\sum_{a\in A_d} f(a) + \sum_{a=1}^{n} f(a),$$

hence

$$\sum_{\substack{1\le a\le n\\(a,n)=1}} f(a) = \sum_{d|P}\mu(d)\sum_{a\in A_d} f(a)$$

where

$$P = p_1 p_2 \cdots p_r.$$

Example:

$$\prod_{\substack{1\le a\le n\\(a,n)=1}} f(a) = \prod_{d|P}\left(\prod_{a\in A_d} f(a)\right)^{\mu(d)}.$$

Lemma:

$$\sum_{d|P}\mu(d)d^k = (1-p_1^k)(1-p_2^k)\cdots(1-p_r^k).$$

Examples:

1. $f(a) = a$. Then

$$\sum_{a\in A_d} f(a) = n^2/2d + n/2,$$

$$\sum_{\substack{1\le a\le n\\(a,n)=1}} f(a) = \frac{n^2}{2}(1-1/p_1)\cdots(1-1/p_r) = \frac{n\phi(n)}{2}.$$

2. $f(a) = a^3$. Then $\sum_{a\in A_d} f(a) = (dn^2)/4 + n^3/2 + n^4/(4d)$,

$$\sum_{\substack{1\le a\le n\\(a,n)=1}} f(a) = \frac{n^2}{4}(1-p_1)\cdots(1-p_r) + \frac{n^4}{4}(1-1/p_1)\cdots(1-1/p_r) =$$

$$\frac{\phi(n)}{4}((-1)^r p_1\cdots p_r n + n^3).$$

3. $f(a)=a$. Then $\prod_{a\in A_d} f(a) = d^{n/d}(n/d)!$, therefore

$$\prod_{\substack{1\le a\le n\\(a,n)=1}} a = \prod_{d|P}\left(d^{n/d}(n/d)!\right)^{\mu(d)} = \prod_{d|P}\left((n/d)^d d!\right)^{\mu(n/d)} =$$

$$n^{\sum_{d|n} d\mu(n/d)}\prod_{d|n}(d!/d^d)^{\mu(n/d)} = n^{\phi(n)}\prod_{d|n}(d!/d^d)^{\mu(n/d)}.$$

2.10 Exercises:

(i) This is a multiplicative function which evaluates to $-p$ on p^k, so if $n = p_1^{e_1}\cdots p_r^{e_r}$ then we obtain $(-1)^r p_1\cdots p_r$.

(ii) Let $n = p_1^{e_1}\cdots p_r^{e_r}$. $\sum_{d|n}\Lambda(d) = \sum_{i=1}^r\sum_{j=1}^{e_i}\log p_i = \sum_{i=1}^r e_i\log p_i = \log n$. Hence $\sum_{n=1}^\infty \frac{\Lambda(n)}{n^s}\zeta(s) = \sum_{n=1}^\infty \frac{\log n}{n^s}$,

$$\sum_{n=1}^\infty\frac{\Lambda(n)}{n^s} = \zeta(s)^{-1}\sum_{n=1}^\infty\frac{\log n}{n^s} = -\frac{\zeta'(s)}{\zeta(s)}.$$

(iii) See Example 1 above. We have $f = \frac{1}{2}\phi + \frac{1}{2}\nu$. Given $F_\phi F_u = F_i$ and $F_\nu F_u = F_u$, we have $\sum_{d|n} f(d) = \frac{1}{2}(n+1)$.

(iv) See Example 2 above.

(v) See Example 3 above.

(vi) Recall $\{x\} = x-[x]$ and $\left[\frac{x}{d}\right] = \left[\frac{[x]}{d}\right]$. Hence $\left\{\frac{x}{d}\right\} = \frac{x}{n} - \left[\frac{[x]}{d}\right]$. We have

$$\sum_{n\le x}\mu(n)[x/n] = \sum_{n\le x}\mu(n)\sum_{k\le x/n}1 = \sum_{n\le x}\mu(n)\sum_{kn\le x}1 = \sum_{n\le x}\mu(n)\sum_{\substack{m\le x\\ n|m}}1 =$$

$$\sum_{m\le x}\sum_{n|m}\mu(n) = \sum_{m\le x}\nu(m) = 1.$$

Hence

$$\sum_{n\le x}\mu(n)\frac{x}{n} - \sum_{n\le x}\mu(n)\left\{\frac{[x]}{n}\right\} = 1,$$

$$\sum_{n\le x}\mu(n)\frac{1}{n}=\frac{1}{x}\left(1+\sum_{n\le x}\mu(n)\left\{\frac{[x]}{n}\right\}\right),$$

$$\left|\sum_{n\le x}\mu(n)\frac{1}{n}\right|=\frac{1}{x}\left|1+\sum_{n\le [x]}\mu(n)\left\{\frac{[x]}{n}\right\}\right|\le\frac{[x]}{x}\le 1$$

since all the terms in the sum have absolute value ≤ 1 and the last summand is 0.

(vii) If $\mu(n/(m,n))=0$ then every summand is zero and the identity is true. Now consider $\mu(n/(m,n))\ne 0$. Let a be the product of the primes appearing to smaller exponent in m, let b be the product ot the primes appearing to smaller exponent in n, and let c be the product of the primes appearing to the same exponent in m and n. Then we have $(m,n)=abc$, where by convention an empty product is 1. We also have $m=Abc$ and $n=aBc$ where $P(A)=P(a)$ and $P(B)=P(b)$, denoting by $P(k)$ the product of the distinct primes appearing in k. Moreover $n/(m,n)=B/b=P(b)$ since we are assuming $n/(m,n)$ is square free. Any divisor of (m,n) has the form $a'b'c'$ where $a'|a$, $b'|b$, and $c'|c$. Therefore

$$\sum_{d|(m,n)}d\mu(n/d)=\sum_{a'|a,b'|b,c'|c}a'b'c'\mu((a/a')(B/b')(c/c'))=$$

$$\left(\sum_{a'|a}a'\mu(a/a')\right)\left(\sum_{b'|b}b'\mu(B/b')\right)\left(\sum_{c'|c}c'\mu(c/c')\right)=$$

$$\phi(a)\left(\sum_{b'|b}b'\mu(B/b')\right)\phi(c)=$$

$$\phi(a)b\mu(P(b))\phi(c)=\mu(n/(m,n))\frac{\phi(n)}{\phi(n/(m,n))}$$

since the only nonzero contribution by $\mu(B/b')$ is from $b'=b$.

(viii) $\sum_{n=1}^{\infty}\phi(n)\frac{x^n}{1-x^n}=\sum_{n=1}^{\infty}\phi(n)(x^n+x^{2n}+x^{3n}+\cdots)=\sum_{j=1}^{\infty}\left(x^j\sum_{d|j}\phi(d)\right)=\sum_{j=1}^{\infty}jx^j=\frac{x}{(1-x)^2}$.

(ix) Let x be an integer.

$$\sum_{n\le x}\frac{\phi(n)}{n}=\sum_{n\le x}\sum_{d|n}\frac{\mu(d)}{d}=\sum_{d\le x}\frac{\mu(d)}{d}\sum_{\substack{n\le x\\ d|n}}1=\sum_{d\le x}\frac{\mu(d)}{d}\left[\frac{x}{d}\right]=$$

$$\sum_{d\le x}\frac{\mu(d)}{d}\left(\frac{x}{d}+O(1)\right)=x\sum_{d\le x}\frac{\mu(d)}{d^2}+O(1)\sum_{d\le x}\frac{\mu(d)}{d}=$$

$$x((6/\pi^2)+O(1/x))+O(\log x)O(1)=(6/\pi^2)x+O(\log x).$$

Chapter 3: Congruences

Definition. Given a natural number $n\in\{1,2,3,\dots\}$ we say that integers a and b satisfy $a\equiv b \pmod n$ provided $n|(b-a)$. This is an equivalence relation.

Properties: (1) $a\equiv b$ and $a'\equiv b'$ imply $a\pm a'\equiv b\pm b'$ and $aa'\equiv bb'$. (2) $a\equiv r$ where $a=qn+r$ and $0\le r<n$, which implies that there are exactly n different congruence classes mod n. (3) When $(a,n)=1$ then $ab\equiv 1$ mod n has a solution: use x where $xa+ny=1$. b is unique mod n: $ab\equiv 1$ and $ac\equiv 1$ implies $n|a(b-c)$ implies $n|(b-c)$ implies $b\equiv c$ mod n. We write $a^{-1}\equiv b$. (4) $(a,n)=1$ and $a\equiv b$ mod n implies $(b,n)=1$ and $a^{-1}\equiv b^{-1}$. (5) If $(a,n)=1$ and $ax\equiv y$ mod n then $x\equiv a^{-1}y$ mod n. The solution x is unique mod n.

Linear equations: Consider the equation

$$ax\equiv b \pmod n.$$

This is equivalent to $n|(ax-b)$, and if there is a solution then $(a,n)|b$. So this is a necessary condition. If this condition holds then we are attempting to solve $a_0x\equiv b_0 \pmod{n_0}$, where we have divided through by (a,n). We have seen above that has a solution because $(a_0,n_0)=1$. Conclusion: $ax\equiv b \pmod n$ has a solution iff $(a,n)|b$.

Note that there is a unique solution for x mod n_0. So the solutions are all of the form x_0+kn_0 where x_0 is a particular solution. Number of distinct solutions mod n: $x_0+kn_0\equiv x_0+jn_0 \pmod n \iff n|n_0(k-j) \iff (a,n)|(k-j) \iff k\equiv j \pmod{(a,n)}$. So there are (a,n) distinct solutions mod n, and it suffices to use x_0+kn_0 where $0\le k\le (a,n)-1$.

More generally, we now have the means to solve the equation $ax+b \equiv c \bmod n$.

Chinese remainder theorem: Let $n_1, \ldots, n_k$ be natural numbers which are coprime in pairs, meaning $(n_i, n_j) = 1$ when $i \neq j$. Let $c_1, \ldots, c_k$ be integers. Then there is an integer x, unique mod $n_1 n_2 \cdots n_k$, such that

$$(x, x, \ldots, x) \equiv (c_1, c_2, \ldots, c_k) \bmod (n_1, n_2, \ldots, n_k).$$

First proof: repeated substitution and shifting. We are seeking a solution of the form $c_1 + a_1 n_1$ for an appropriate a_1. We require $c_1 + a_1 n_1 \equiv c_2 \bmod n_2$, which has a solution for a_1 because $(n_1, n_2) = 1$. The most general solution is $a_1 + a_2 n_2$. We require $c_1 + (a_1 + a_2 n_2) n_1 \equiv c_3 \bmod n_3$, which has a solution for a_2 because $(n_2 n_1, n_3) = 1$. The most general solution is $a_2 + a_3 n_3$. We require $c_1 + (a_1 + (a_2 + a_3 n_3) n_2) n_1 \equiv c_4 \bmod n_4$, which has a solution for a_3 because $(n_3 n_2 n_1, n_4) = 1$. Keep on going until we have found

$$x = c_1 + a_1 n_1 + a_2 n_2 n_1 + \cdots + a_k n_k n_{k-1} \cdots n_1.$$

The solution is unique mod $n_1 n_2 \cdots n_k$ because if y is another solution then then $n_i | (x - y)$ for $i = 1, \ldots, k$, so $n_1 \cdots n_k | (x - y)$.

Example: Solve $(x, x, x) \equiv (1, 2, 3) \bmod (6, 35, 143)$. Solution: $1 + 6a_1 \equiv 2 \bmod 35$, $a_1 = 6 + 35a_2$, $1 + 6(6 + 35a_2) \equiv 3 \bmod 143$, $a_2 = 1088 + 143a_3$, so our solution is $x = 1 + 6(6 + 35(1088 + 143a_3)) = 228517 + 30030a_3$.

Second proof: decoupling. First find solutions to $x_i \equiv c_i \bmod n_i$ and $x_i \equiv 0 \bmod n_j$ for $j \neq i$, then use the solution $x = x_1 + x_2 + \cdots + x_k$. So we have reduced the problem to solving the simultaneous equations $x_i \equiv c_i \bmod n_i$ and $x_i \equiv 0 \bmod n_1 \cdots \widehat{n_i} \cdots n_k$. Setting $x_i = n_1 \cdots \widehat{n_i} \cdots n_k y_i$ we are seeking a solution to

$$n_1 \cdots \widehat{n_i} \cdots n_k y_i \equiv c_i \pmod{n_i}.$$

There will be a solution for y_i because $(n_1 \cdots \widehat{n_i} \cdots n_k, n_i) = 1$.

With the n_i coprime in pairs, we can now solve the simultaneous equations $a_i x = b_i \pmod{n_i}$: first solve each solution individually, yielding solutions $x_1, \ldots, x_k$, then find x so that $x \equiv x_i \pmod{n_i}$ for $i = 1, \ldots k$.

Alternative proof that ϕ is multiplicative: Let $m \geq 2$, $n \geq 2$ be given such that $(m, n) = 1$. Let $1 \leq m_1 < \cdots < m_r < m$ satisfy $(m_i, m) = 1$, let

$1 \leq n_1 < \cdots < n_s < n$ satisfy $(n_i, n) = 1$, and let $1 \leq x_1 < \cdots < x_t < mn$ satisfy $(x_i, mn) = 1$. Consider the mapping $(m_i, n_j) \mapsto m_i n + n_j m$. If the images are congruent to x_i's and each x_i is uniquely an image modulo mn, then we know that $rs = t$.

We first show $(m_i n + n_j m, mn) = 1$. Let d be a common divisor of these numbers. Then $d|m$ or $d|n$. Without loss of generality $d|m$. Then $(d, n) = 1$ and $d|m_i n$, therefore $d|m_i$. Since $(m, m_i) = 1$, $d = 1$.

The mapping is injective: $m_a n + n_b m = m_c n + n_d m \implies (m_a - m_c)n = (n_d - n_b)m \implies n|(n_d - n_b) \implies n_d = n_b$ and $m_a = m_c$.

The mapping is surjective: We wish to find m_i and n_j such that $m_i n + n_j m \equiv x_k \bmod mn$. We can certainly find integers p and q such that $pm + qn = 1$ since $(m, n) = 1$. This yields $x_k pm + x_k qn = x_k$. The claim is that $(x_k p, n) = 1$ and $(x_k q, m) = 1$. It suffices to prove $(p, n) = 1$ and $(q, m) = 1$, but these are true since $pm + qn = 1$.

Lemma: Let $a_1, a_2, \ldots, a_{\phi(n)}$ a complete list of class representatives in the set $\{a \in \mathbb{Z} : (a, n) = 1\}$. Let $(k, n) = 1$. Then $ka_1, ka_2, \ldots, ka_{\phi(n)}$ is a permutation of $a_1, a_2, \ldots, a_{\phi(n)} \bmod n$.

Proof: The set $\{a \in \mathbb{Z} : (a, n) = 1\}$ is closed with respect to multiplication, so all the elements in the list $ka_1, ka_2, \ldots, ka_{\phi(n)}$ appear in this set. We need only show that the list consists of distinct elements modulo n. The results from $ka_i \equiv ka_j \implies n|k(a_i - a_j) \implies n|(a_i - a_j)$.

Euler's Theorem: Assume $(a, n) = 1$. Then $a^{\phi(n)} \equiv 1 \bmod n$.

Proof: By the previous theorem we know that $aa_1, \ldots, aa_{\phi(n)}$ is a permutation of $a_1, \ldots, a_{\phi(n)} \bmod n$. Forming the product of the lists,

$$a^{\phi(n)} a_1 \cdots a_{\phi(n)} \equiv a_1 \cdots a_{\phi(n)} \bmod n.$$

This implies $a^{\phi(n)} \equiv 1 \bmod n$.

Fermat's Theorem: Let p be prime and assume $(p, n) = 1$. Then $a^{p-1} \equiv 1 \bmod p$.

Proof: This is a corollary of Euler's Theorem with $n = p$ and $\phi(n) = p - 1$.

Wilson's Theorem: Let $p > 2$ be a prime. Then $(p-1)! \equiv -1 \bmod p$. Proof: The class representatives are $1, 2, \ldots, p-1$, and for each class a there is a unique class a' such that $aa' \equiv 1 \bmod p$. Classify the numbers

$1, 2, \ldots, p-1$ into three types: $a < a'$, $a = a'$, $a > a'$. In the $a = a'$ category we have $a^2 \equiv 1$, therefore $p|(a-1)(a+1)$, therefore $p|(a-1)$ or $p|(a+1)$. The only possibilities are $a = 1$ and $a = p-1$. The remaining numbers pair off to form the product $2 \cdot 3 \cdots (p-2) \equiv 1$. This yields $(p-1)! \equiv -1$.

A converse: let $n > 1$ be a natural number that satisfies $(n-1)! \equiv -1 \bmod n$. Then $n|((n-1)!+1)$. So any $d < n$ dividing n divides 1, forcing $d = 1$. Hence n must be prime.

The field $\mathbb{Z}_p$: Define addition and multiplication in $\{0, 1, \ldots, p-1\}$ using modulus class representatives. The set of non-zero elements is closed with respect to multiplication and each element has a unique multiplicative inverse, hence forms a group. We call such a set a field.

Finding $\sqrt{-1}$ in $\mathbb{Z}_p$: Trivial when $p = 2$. For odd p, consider solving the equation $x^2 \equiv -1 \bmod p$. If we can realize x^2 as $(p-1)!$ then we have a solution. Write $p = 2r+1$. We have $-1 \equiv (p-1)! = r!(r+1)(r+2)\cdots(2r) \equiv r!(-r)(-r+1)\cdots(-1) = (-1)^r(r!)^2$. If r is even we get $-1 \equiv x^2$ where $x = r!$. This requires $p \equiv 1 \bmod 4$. Now if $p \equiv 3 \bmod 4$ the equation $x^2 \equiv -1 \bmod p$ implies $x^{p-1} = x^{4n+2} \equiv (-1)^{2n+1} = -1$, which contradicts Fermat's theorem. So there is no $\sqrt{-1}$ in $\mathbb{Z}_p$ when $p \equiv 3 \bmod 4$, but there is when $p \equiv 1 \bmod 4$.

Lagrange's Theorem: Given a polynomial $f(x)$ of degree n and integer coefficients, where p does not divide the leading term, there are at most n solutions to $f(x) \equiv 0 \bmod p$. Proof: Induction argument regarding polynomials in $\mathbb{Z}_p[x]$.

In $\mathbb{Z}_p[x]$ Fermat's theorem implies

$$x^{p-1} - 1 = (x-1)(x-2)\cdots(x-p+1).$$

Therefore $x^{p-1} - 1$ has exactly $p-1$ roots in $\mathbb{Z}_p$.

Now suppose $d|(p-1)$. Write $p-1 = qd$. $(Y-1)|(Y^q-1)$, therefore $(x^d-1)|(x^{p-1}-1)$, so we can write $x^{p-1}-1 = (x^d-1)g(x)$. Since $x^{p-1}-1$ has $p-1$ roots, and $g(x)$ provides at most $p-1-d$ of them by a degree argument, x^d-1 has to provide at least d of them. So x^d-1 has exactly d roots in $\mathbb{Z}_p$.

Second proof of Wilson's theorem: Evaluating at $x = 0$ in the identity above yields

$$-1 \equiv (-1)^{p-1}(p-1)! = (p-1)! \bmod p$$

for an odd prime.

When n is not prime, it is possible for there to be more than $deg(f)$ distinct roots mod n of a monic polynomial $f(x)$. Construction: Let p_1, ..., p_k be distinct primes. Set $f_i(x) = x^{p_i} - x$ for each $i \le k$. Then $f_i(x)$ has exactly p_i roots mod p_i. Now set $f(x) = f_1(x) \cdots f_k(x)$. Choose any vector $(q_1, \ldots, q_k)$ where $0 \le q_i < p_i$ for each i. By the Chinese remainder theorem there is a unique integer x mod $p_1 \cdots p_n$ such that $x \equiv q_i$ mod p_i for each i, and $f(x) \equiv 0$ mod p_1, mod p_2, ..., mod p_k, hence mod $p_1 p_2 \cdots p_n$. So we have constructed a polynomial of degree $p_1 + \cdots + p_k$ with $p_1 \cdots p_k$ roots. For example, $f(x) = (x^3 - x)(x^5 - x)$ has degree 8 and 15 roots mod 15.

Section 3.6: Primitive roots.

Let $(a, n) = 1$. We know that $a^{\phi(n)} \equiv 1$ mod n. The order of a mod n is defined to be $d = o(a)$, the smallest positive integer such that $a^d \equiv 1$ mod n. It is a divisor of $\phi(n)$: write $\phi(n) = pd + r$ where $0 \le r < d$. Raising a to the power of both sides and simplifying, $a^r \equiv 1$. By minimality of d this forces $r = 0$. A primitive root mod n is an integer a such that $(a, n) = 1$ and a has order $\phi(n)$ mod n.

Terminology: when the order of a mod n is d we say that a belongs to d mod n.

Theorem: There are $\phi(p-1)$ primitive roots mod p for a prime p.

Proof: Let $\psi(d)$ be the number of elements a with $1 \le a < p$ and $(a, p) = 1$ and having order d. We wish to prove $\psi(p-1) = \phi(p-1)$. We have

$$\sum_{d|(p-1)} \psi(d) = p - 1.$$

Given that

$$\sum_{d|(p-1)} \phi(d) = p - 1,$$

we will have $\psi(d) = \phi(d)$ for each d dividing $p - 1$ provided we can show $\psi(d) \le \phi(d)$ for each such d. To do this, suppose that $\psi(d) > 0$ for a given d. Let a have order d. Then the numbers $1, a, a^2, \ldots, a^{d-1}$ are distinct mod p and constitute all the roots of $x^d - 1$ in $\mathbb{Z}_p$. We now count the elements of order d mod $p - 1$. Let b be an element with order d. It is a root of $x^d - 1$ mod p, hence it must have the form a^m for a unique m satisfying

$0 \leq m \leq d-1$. We will show $(m,d)=1$. Suppose $(m,d)=k$ and write $m=m_0k$ and $d=d_0k$. Then $b^{d_0}=(a^m)^{d_0}=a^{(m_0d/k)}=(a^d)^{m_0}\equiv 1$, forcing $d_0 \geq d$. But $d \geq d_0$, hence $d=d_0$ and $k=1$. Hence $\psi(d)\leq\phi(d)$, as desired.

Remark: The theorem implies that there is always at least one primitive root mod p for every prime p. It says nothing about how to find them, but in the problems to think about for Chapters 3 and 4 I illustrate some of the techniques. One can always use brute force. In Problem (v) below we will show how all the primitive roots of n are related to any particular one of them.

Remark: Now that we know that there are $\phi(d)\geq 1$ primitive roots of unity of order d mod p when $d|(p-1)$, we can pick any of them, say a, and produce all the roots of x^d-1 via the list $1,a,a^2,\ldots,a^{d-1}$. Hence

$$x^d-1=(x-1)(x-a)\cdots(x-a^{d-1})$$

in $\mathbb{Z}_p[x]$.

Constructing a primitive root mod p^j when p is prime: Let a be a primitive root of p. If the order of a mod p^j is d then $a^d\equiv 1$ mod p, so $\phi(p)|d$, i.e. $(p-1)|d$. On the other hand, $d|\phi(p^j)$, therefore $d|(p-1)p^{j-1}$, therefore $d=(p-1)p^k$ for some $k\leq j-1$. To find a primitive root of p^j we will find a primitive root of p satisfying $k=j-1$.

Lemma: For an odd prime p and an integer z, $(1+pz)^{p^j}=1+p^{j+1}Z$ where $Z\equiv z$ mod p. For any integer x, $(1+2x)^{2^j}=1+2^{j+2}y$ for some integer y for $j\geq 1$.

Proof: We treat the odd prime case by induction on j. The base case $j=0$ is true. Now assume that $(1+pz)^{p^j}=1+p^{j+1}Z$ where $Z\equiv z$ mod p. Raising both sides to the power p we obtain

$$(1+pz)^{p^{j+1}}=1+p^{j+2}Z+\sum_{i=2}^{p}\binom{p}{i}p^{(j+1)i}Z^i.$$

One can check that p^{j+3} is a divisor of all the terms in the sum with a binomial coefficient, using the fact that $p|\binom{p}{i}$ when $0<i<p$ and the fact that $p\geq 3$. Writing $Z'=Z+p^{-j-2}\sum_{i=2}^{p}\binom{p}{i}p^{(j+1)i}Z^i$ we have $(1+pz)^{p^{j+1}}=1+p^{j+2}Z'$ and $Z'\equiv Z$ mod p.

We have $(1+2x)^2 = 1+8\left(\frac{x+x^2}{2}\right)$. Assuming $(1+2x)^{2^j} = 1+2^{j+2}y$, we have

$$(1+2x)^{2^{j+1}} = (1+2^{j+2}y)^2 = 1+2^{j+3}(y+2^{j+1}y^2).//$$

Note that the lemma implies that there are no primitive roots mod 2^j for $j \geq 2$, since $(1+2x)^{2^{j-2}} \equiv 1 \bmod 2^j$.

Now let p be an odd prime and let a be an integer such that $(a,p)=1$ and a is a primitive root of p. Then $a^{p-1} = 1+py$ for some y. For any integer x, $a+px$ is a primitive root of p and

$$(a+px)^{p-1} = a^{p-1} + (p-1)pxa^{p-2} + p^2Z =$$

$$1+py+(p-1)pxa^{p-2}+p^2Z = 1+p(y+(p-1)xa^{p-2}+pZ).$$

Since $((p-1)a^{p-2},p)=1$, we can find x so that $y+(p-1)xa^{p-2}+pZ = Z'$ for some Z' with $Z' \equiv 1 \bmod p$. This implies that $a+px$ is a primitive root mod p^j: suppose that $a+px$ has order d mod p^j. Then $d=(p-1)p^k$ for some $k \leq j-1$. We have

$$(a+px)^{(p-1)p^k} = (1+pZ')^{p^k} = 1+p^{k+1}Z'' \equiv 1 \bmod p^j$$

where $Z'' \equiv Z' \equiv 1 \bmod p$. Since p does not divide Z'', $1+p^{j-1}Z'' \equiv 1 \bmod p^j$ implies $k+1 \geq j$, i.e. $k \geq j-1$. Hence $k=j-1$, $d=(p-1)p^{j-1} = \phi(p^j)$.

Moduli that permit primitive roots: Suppose n factors as $n=n_1n_2$ where $(n_1,n_2)=1$ and $n_1>2$, $n_2>2$. Then $\phi(n)=\phi(n_1)\phi(n_2)$ is even and, for any a with $(a,n)=1$, $a^{\frac{1}{2}\phi(n)} = (a^{\phi(n_1)})^{\frac{1}{2}\phi(n_2)} \equiv 1 \bmod n_2$ and $a^{\frac{1}{2}\phi(n)} = (a^{\phi(n_2)})^{\frac{1}{2}\phi(n_1)} \equiv 1 \bmod n_1$, hence $a^{\frac{1}{2}\phi(n)} \equiv 1 \bmod n$. So there are no primitive roots of n in this case. The lemma implies that there are no primitive roots mod 2^j for $j \geq 3$. This leaves $n = 2, 4, 2p^j$ where p is an odd prime. There are primitive roots in each case: 1 is primitive mod 2, 3 is primitive mod 4. Now let a be primitive mod p^j. Then so is $a+p^j$, and the odd one of these is coprime with $2p^j$. We have $\phi(2p^j)=\phi(p^j)$, so either a or $a+p^j$ is primitive mod $2p^j$.

Section 3.7: Indices.

Let g be a primitive root of n. Then

$$\{1 \leq a \leq n-1 : (a,n)=1\} = \{1, g, \ldots, g^{\phi(n)}-1\} \bmod n.$$

We write $ind(a) = i$ when $(a, n) = 1$ and $a \equiv g^i \bmod n$. More generally, $a \equiv g^l \bmod n$ iff $l \equiv ind(a) \bmod \phi(n)$. Properties include $ind(ab) \equiv ind(a) + ind(b) \bmod \phi(n)$ and, for $n > 2$, $ind(-1) = \frac{1}{2}\phi(n)$.

Example: solve $x^5 \equiv 2 \bmod 7$. The number of distinct solutions to x mod 7 is equal to the number of solutions to $ind(x)$ mod 6. We have $5ind(x) \equiv ind(2) \bmod 6$. Since $(5, 6) = 1$, there is a unique solution for $ind(x)$ mod 6, hence a unique solution for x mod 7: Using the primitive root 3 we have $ind(x) \equiv -ind(2) = -2 \equiv 4$, $x \equiv 3^4 \equiv 4$.

Now consider $n = 2^j$ for $j \geq 3$. We can prove by induction that

$$5^{2^a} = 1 + 2^{a+2}k_a$$

where k_a is odd by induction on $a \geq 0$. This implies that $o(5) = 2^{j-2}$ mod 2^j for $j \geq 3$. Moreover, $5^a \equiv 1 \bmod 4$ and $(-5)^a \equiv 3 \bmod 4$, which implies that $5^a \not\equiv (-5)^b \bmod 2^j$ when $j \geq 3$. So all numbers of the form $(-1)^x 5^y$, $0 \leq x \leq 1$, $0 \leq y \leq 2^{j-2} - 1$, are distinct mod 2^j, and this accounts for all odd residue classes mod 2^j. Hence every odd residue class mod 2^j has a unique expression of the form $(-1)^x 5^y$ mod 2 in x and mod 2^{j-2} in y.

3.9 Exercises:

(i) This equivalent to solving $(x, x, x) \equiv (2, 2, 3) \bmod (3, 5, 7)$. Solution set using repeated substitution: $x = 17 + 105k$.

(ii) Write $ax = q_x n + r_x$ where $0 \leq r_x < n$. As x runs through a set of reduced residues mod n, so does r_x. Moreover $\{ax/n\} = r_x/n$. So for $n \geq 2$ we obtain

$$\frac{1}{n} \sum_{\substack{1 \leq r \leq n \\ (r,n)=1}} r.$$

By Problem (iii) in Chapter 2, this sum is $\frac{\phi(n)}{2}$.

(iii) The contrapositive of this statement is that when n is not prime then either $a^{n-1} \not\equiv 1 \bmod n$ or $a^m \equiv 1 \bmod n$ for some proper divisor of $n - 1$. So assume n is not prime and $a^{n-1} \equiv 1 \bmod n$. We must show $a^m \equiv 1$ mod n. Let the order of a mod n be d. Then $d|(n-1)$ and $d|\phi(n)$. We can set $m = d$ provided we can show $\phi(n) < n - 1$. This follows from $\phi(n) = \phi(n_1)\phi(n_2) \leq (n_1 - 1)(n_2 - 1) \leq n - 3$ where $(n_1, n_2) = 1$, $n_1 n_2 = n$, $n_1, n_2 \geq 2$.

(iv) First suppose that $p > 2$. Using indices,

$$x^{p-1} \equiv 1 \bmod p^j \text{ iff } (p-1)ind(x) \equiv 0 \bmod \phi(p^j) \text{ iff } ind(x) \equiv 0 \bmod p^{j-1}.$$

There are $p-1$ solutions mod p^j, namely the multiples of p^{j-1}. When $p = 2$ there is exactly one solution to $x \equiv 1 \bmod 2^j$.

(v) Let g be a primitive root of n. Then it has order $\phi(n)$. If g^k has order d mod n then $d|\phi(n)$ and $\phi(n)|kd$, and if $(k, \phi(n)) = 1$ then $\phi(n)|d$ and so $d = \phi(n)$. Hence $(k, \phi(n)) = 1$ implies primitive. Moreover, if $(k, \phi(n)) = D > 1$, write $k = k_0 D$, $\phi(n) = \phi_0 D$. Then $(g^k)^{\phi_0} = (g^{k_0})^{\phi(n)} = 1$, hence g^k is not primitive. Hence g^k is primitive iff $(k, \phi(n)) = 1$. This implies $\phi(\phi(n))$ primitive roots mod n, namely

$$\{g^k : 1 \le k \le \phi(n) \text{ and } (k, \phi(n)) = 1\}.$$

(vi) Let g be a primitive root mod p. We have shown in problem (v) that the primitive roots mod p are precisely g^k where $1 \le k \le p-1$ and $(k, p-1) = 1$. So we are evaluating

$$\sum_{\substack{1 \le k \le p-1 \\ (k,p-1)=1}} g^k.$$

By our inclusion-exclusion sum formula of chapter 2 this is equal to

$$\sum_{d|P} \mu(d) \sum_{a \in A_d} g^a$$

where $p_1, \ldots, p_r$ are the primes dividing $p-1$,

$$P = p_1 p_2 \cdots p_r,$$

and

$$A_d = \{a \le p-1 : d|a\}.$$

Given that

$$n_d = \sum_{a \in A_d} g^a = g^d + g^{2d} + \cdots + g^{p-1},$$

we have $g^d n_d \equiv n_d \bmod p$, hence $(g^d - 1)n_d \equiv 0 \bmod p$. For $d < p-1$ this forces $n_d \equiv 0 \bmod p$. This just leaves $\mu(p-1)n_{p-1} = \mu(p-1)$.

(vii) Using 3 as the primitive root of 7, the equation is equivalent to $2Y \equiv 5 + 3X \bmod 6$, where X and Y are the indices of x and y. Hence X must be an odd number. Writing $X = 1 + 2X'$ and substituting, the equation is $2Y \equiv 2 \bmod 6$. Any X' will do, and we must have $Y \equiv 1 \bmod 3$. Hence $y = 3^{1+3a}$, $x = 3^{1+2b}$. In reduced form, $y = 3 \cdot (-1)^a$, $x = 3 \cdot 2^b$.

(viii) Let $m = 1 + \frac{1}{2} + \cdots + \frac{1}{p-1}$. Then $m = n/d$ where $n = (p-1)!m$ and $d = (p-1)!$. Since d is not divisible by p, it suffices to show that n is divisible by p^2, for then any fraction equivalent to m will have numerator divisible by p^2. We have

$$2n = (p-1)![((1/1)+(1/p-1))+((1/2)+(1/p-2))+\cdots+((1/p-1)+(1/1))]$$

$$= p(p-1)!(1/(1(p-1)) + 1/(2(p-2)) + \cdots + (1/((p-1)1))).$$

Write

$$M = (p-1)!(1/(1(p-1)) + 1/(2(p-2)) + \cdots + (1/((p-1)1))).$$

It suffices to show that M is divisible by p. We have

$$M \equiv (p-1)!(1^{-1}(p-1)^{-1} + 2^{-1}(p-2)^{-1} + \cdots + (p-1)^{-1}1^{-1}) \equiv$$

$$1^{-2} + 2^{-2} + \cdots + (p-1)^{-2} \equiv 1^2 + 2^2 + \cdots + (p-1)^2 = \frac{(p-1)p(2p-1)}{6}.$$

Since $(p, 6) = 1$ (we are given $p > 3$), $6|(p-1)(2p-1)$. Hence $M \equiv 0 \bmod p$. NOTE: there is a proof in Hardy and Wright that involves this idea of pairing things off, but this sum of squares business is my idea.

Chapter 4: Quadratic Residues

Sections 4.1 and 4.2: Legendre's Symbol and Euler's Criterion

Solving $ax^2 + bx + c \equiv 0 \bmod n$ requires solving $(2ax + b)^2 \equiv b^2 - 4ac \bmod 4an$. We call a a quadratic residue mod n when there is a solution to $x^2 \equiv a \bmod n$. In other words, $\sqrt{a} \in \mathbb{Z}_n$.

Example: Recall that for an odd prime p, $\sqrt{-1} \in \mathbb{Z}_p$ if and only if $p \equiv 1 \bmod 4$, in which case we have $\sqrt{-1} = \pm(\frac{p-1}{2})!$.

Let p be an odd prime. Each of the numbers $1^2, 2^2, \ldots, (\frac{p-1}{2})^2$ are quadratic residues mod p. They are distinct mod p: given $i \neq j \in \{1, 2, \ldots, \frac{p-1}{2}\}$, the factors $i - j$ and $i + j$ are not divisible by p, hence $i^2 - j^2$ is not divisible by

p. Hence these numbers are the complete set of solutions to $x^{\frac{p-1}{2}} \equiv 1 \bmod p$. Since every $k \in \{1, \ldots, p-1\}$ satisfies $k^{\frac{p-1}{2}} \equiv \pm 1 \bmod p$, k is a non-zero quadratic residue mod p if and only if $k^{\frac{p-1}{2}} \equiv 1 \bmod p$. This gives rise to Euler's Criterion: for an odd prime p and $(a, p) = 1$, a is a quadratic residue mod p if and only if $a^{\frac{p-1}{2}} \equiv 1$. The Legendre symbol is

$$\left(\frac{a}{p}\right) = \left\{\begin{matrix} 1 & a \text{ is a quadratic residue mod } p \\ -1 & a \text{ is not a quadratic residue mod } p \end{matrix}\right\} \equiv a^{\frac{p-1}{2}} \bmod p.$$

Note

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

and since $p \geq 3$ and these symbols are ± 1 this implies $\left(\frac{ab}{p}\right) = \left(\frac{a}{p}\right)\left(\frac{b}{p}\right)$.

Let p be an odd prime and let g be a primitive root of p. A complete set of non-zero quadratic residues mod p is $\{g^2, g^4, \ldots, g^{p-1}\}$.

Section 4.3: Gauss' Lemma

We wish to derive a formula for $\left(\frac{a}{p}\right)$ for an arbitrary odd prime p that does not depend on computing $a^{\frac{p-1}{2}} \bmod p$, which can be difficult when p is large. Toward this end, observe that given an odd prime $p = 2r+1$, every integer k is equivalent to a unique number in $\{-r, -r+1, \ldots, -1, 0, 1, \ldots, r-1\}$. To see this, use the division algorithm to write $k+r = dp+s$ where $0 \leq s \leq 2r$. Then $k \equiv s-r \bmod p$ and $-r \leq s-r < r$. We will say that k has a negative representation mod p if $k \equiv s$ for some $s \in \{-r, -r+1, \ldots, -1\}$ where $p = 2r+1$. The number s is called the numerically least residue of k mod p.

Theorem: Let $p = 2r+1$ be an odd prime and let $(a, p) = 1$. Then

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

where n is the number of integers in the set $\{a, 2a, \ldots, ra\}$ that have a negative representation mod p.

Proof: For each $i \in \{1, 2, \ldots, r\}$ say that $ia \equiv a_i \bmod p$ where $a_i \in \{-r, -r+1, \ldots, r-1\}$. Then $|a_1|, |a_2|, \ldots, |a_r|$ is a rearrangement of $1, 2, \ldots, r$. Hence

$$(r!)a^r = (1a)(2a)\cdots(ra) \equiv a_1 a_2 \cdots a_r = |a_1||a_2|\cdots|a_r|(-1)^n = (r!)(-1)^n \bmod p,$$

$$a^r \equiv (-1)^n \bmod p,$$
$$\left(\frac{a}{p}\right) \equiv (-1)^n \bmod p,$$
$$\left(\frac{a}{p}\right) = (-1)^n.$$

Let's calculate $\left(\frac{2}{p}\right)$ for an odd prime p. Write $p = 2r + 1$. Then

$$\{-1, -2, \ldots, -r\} \equiv \{r+1, r+2, \ldots, 2r-1\}$$

mod p. Using $a = 2$ we must determine

$$n = |\{2, 4, \ldots, 2r\} \cap \{r+1, r+2, \ldots, 2r-1\}|.$$

If $r = 2k$ then

$$n = |\{2, 4, \ldots, 4k\} \cap \{2k+1, 2k+2, \ldots, 4k\}| = |\{2k+2, 2k+4, \ldots, 4k\}| = k$$

and

$$\left(\frac{2}{4k+1}\right) = (-1)^k.$$

If $r = 2k+1$ then

$$n = |\{2, 4, \ldots, 4k+2\} \cap \{2k+2, 2k+3, \ldots, 4k+2\}| = |\{2k+2, 2k+4, \ldots, 4k+2\}| = k+1$$

and

$$\left(\frac{2}{4k+3}\right) = (-1)^{k+1}.$$

Hence

$$\left(\frac{2}{8j+1}\right) = \left(\frac{a}{4(2j)+1}\right) = (-1)^{2j} = 1$$
$$\left(\frac{2}{8j+3}\right) = \left(\frac{a}{4(2j)+3}\right) = (-1)^{2j+1} = -1$$
$$\left(\frac{2}{8j+5}\right) = \left(\frac{a}{4(2j+1)+1}\right) = (-1)^{2j+1} = -1$$
$$\left(\frac{2}{8k+7}\right) = \left(\frac{a}{4(2j+1)+3}\right) = (-1)^{2j+2} = 1.$$

Hence 2 is a quadratic residue mod an odd prime p iff $p \equiv 1, 7 \bmod 8$ and 2 is a non-quadratic residue mod p iff $p \equiv 3, 5 \bmod 8$.

Section 4.4: Law of Quadratic Reciprocity

Let p and q be distinct odd primes. The law of quadratic reciprocity is

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

This formula is useful for deciding whether or not a number is a quadratic residue. For example, we have $\left(\frac{5}{171}\right)\left(\frac{171}{5}\right) = (-1)^{4(170)/4} = 1$, therefore $\left(\frac{5}{171}\right) = \left(\frac{171}{5}\right) \equiv 171^2 \equiv 1 \bmod 5$. Hence 5 is a quadratic residue mod 171.

Proof: For each integer x there is a unique integer $y_{pq}(x)$ such that $xp - y_{pq}(x)q \in (-q/2, q/2]$. We have

$$\left(\frac{p}{q}\right) = (-1)^{|X_{pq}|}$$

where

$$X_{pq} = \{x \in (0, q/2) : xp - y_{pq}(x)q \in (-q/2, 0)\} \cap \mathbf{Z}.$$

Note that for all $x \in X_{pq}$, $y_{pq}(x) \in (0, p/2)$. Setting

$$R_{pq} = \{(x, y) \in (0, q/2) \times (0, p/2) : xp - yq \in (-q/2, 0)\} \cap \mathbf{Z}^2$$

we have

$$\{(x, y_{pq}(x)) : x \in X_{pq}\} = R_{pq}.$$

Therefore

$$\left(\frac{p}{q}\right) = (-1)^{|R_{pq}|}.$$

Similarly, we have

$$\left(\frac{q}{p}\right) = (-1)^{|R_{qp}|}$$

where

$$R_{qp} = \{(x, y) \in (0, p/2) \times (0, q/2) : xq - yp \in (-p/2, 0)\} \cap \mathbf{Z}^2.$$

Reversing the coordinates, this has the same size as the set

$$R'_{qp} = \{(x, y) \in (0, q/2) \times (0, p/2) : xp - yq \in (0, p/2)\} \cap \mathbf{Z}^2.$$

The regions R_{pq} and R'_{qp} correspond to the upper and lower intermediate regions of a rectangular diagram (see Figure 4.1 in the textbook). To determine the relationship between $|R_{pq}|$ and $|R'_{qp}|$, first note that there is a bijection between the integer coordinates in the rectangle $(0, q/2) \times (0, p/2)$ to itself defined by

$$(x, y) \mapsto (x', y') = (\frac{q+1-2x}{2}, \frac{p+1-2y}{2}).$$

Under this mapping we have

$$x'p - y'q = -(xp - yq) + \frac{p-q}{2}.$$

Hence $xp - qy \leq -(q/2)$ if and only if $x'p - y'q \geq p/2$. In other words, there are as many integer coordinates in the top-most region of the associated diagram as there are integer coordinates in the bottom-most region. Hence the total number of coordinates in the top and bottom region is an even number, and the number of coordinates in the entire rectangle is congruent mod 2 to the number of coordinates in the two intermediate regions, namely $|R_{pq}| + |R'_{qp}|$. Hence

$$|R_{pq}| + |R'_{qp}| \equiv \frac{(p-1)(q-1)}{4} \bmod 2,$$

where the latter number is the number of integer coordinates in the rectangle $(0, q/2) \times (0, p/2)$. This implies

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

A second proof: Let p and q be distinct odd primes. Let S be the set of all integers $x \in [1, \frac{pq-1}{2}]$ not divisible by p or q. We have $\frac{pq-1}{2} = p\frac{q-1}{2} + \frac{p-1}{2} = q\frac{p-1}{2} + \frac{q-1}{2}$, therefore

$$\prod_{x \in S} [x]_p = \frac{([1]_p[2]_p \cdots [p-1]_p)^{\frac{q-1}{2}}[1]_p[2]_p \cdots [\frac{p-1}{2}]_p}{[q]_p[2q]_p \cdots [\frac{p-1}{2}q]_p} = [-1]_p^{\frac{q-1}{2}}[\left(\frac{q}{p}\right)]_p$$

and

$$\prod_{x \in S} [x]_q = \frac{([1]_q[2]_q \cdots [q-1]_q)^{\frac{p-1}{2}}[1]_q[2]_q \cdots [\frac{q-1}{2}]_q}{[p]_q[2p]_q \cdots [\frac{q-1}{2}p]_q} = [-1]_q^{\frac{p-1}{2}}[\left(\frac{p}{q}\right)]_q.$$

Write $\theta(x) = ([x]_p, [x]_q)$. Then

$$\prod_{x \in S} \theta(x) = ([-1]_p^{\frac{q-1}{2}} [\left(\frac{q}{p}\right)]_p, [-1]_q^{\frac{p-1}{2}} [\left(\frac{p}{q}\right)]_q).$$

For each $(a, b) \in [p-1] \times [\frac{q-1}{2}]$ there exists a unique $x \in [pq-1]$ such that $\theta(x) = ([a]_p, [b]_q)$ by the Chinese Remainder Theorem. Moreover, exactly one of the two numbers x and $pq - x$ belongs to S. Hence there is a unique $x(a, b) \in S$ and a unique $\epsilon(a, b) \in \{-1, 1\}$ such that $\theta(x(a, b)) = ([\epsilon(a, b)a]_p, [\epsilon(a, b)b]_q)$. This implies

$$S = \{x(a, b) : (a, b) \in [p-1] \times [\frac{q-1}{2}]\}$$

and

$$\prod_{x \in S} \theta(x) = ([\epsilon]_p, [\epsilon]_q) \prod_{b=1}^{\frac{q-1}{2}} \prod_{a=1}^{p-1} ([a]_p, [b]_q)$$

where

$$\epsilon = \prod_{(a,b) \in [p-1] \times [\frac{q-1}{2}]} \epsilon(a, b).$$

We have

$$\prod_{b=1}^{\frac{q-1}{2}} \prod_{a=1}^{p-1} [a]_p = [(p-1)!]_p^{\frac{q-1}{2}} = [-1]_p^{\frac{q-1}{2}}$$

and

$$\prod_{a=1}^{p-1} \prod_{b=1}^{\frac{q-1}{2}} [b]_p = [\left(\frac{q-1}{2}\right)!]_q^{p-1}.$$

We have

$$[-1]_q = [(q-1)!]_q = [(-1)^{\frac{q-1}{2}}]_q [\left(\frac{q-1}{2}\right)!]_q^2,$$

therefore

$$[\left(\frac{q-1}{2}\right)!]_q^{p-1} = [(-1)^{\frac{p-1}{2}}]_q [(-1)^{\frac{(p-1)(q-1)}{4}}]_q.$$

Hence

$$([-1]_p^{\frac{q-1}{2}} [\left(\frac{q}{p}\right)]_p, [-1]_q^{\frac{p-1}{2}} [\left(\frac{p}{q}\right)]_q) = ([\epsilon]_p, [\epsilon]_q)([-1]_p^{\frac{q-1}{2}}, [(-1)^{\frac{p-1}{2}}]_q [(-1)^{\frac{(p-1)(q-1)}{4}}]_q).$$

This implies

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

Comparing the products of the two coordinates, this implies

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

Section 5: Jacobi's Symbol

Let n be a positive odd integer and let $n=p_1\cdots p_k$ be a factorization into primes. Then

$$\left(\frac{a}{n}\right)=\begin{cases}\left(\frac{a}{p_1}\right)\cdots\left(\frac{a}{p_k}\right) & (a,n)=1\\ \\ 0 & (a,n)>1.\end{cases}$$

Properties:

1. $\left(\frac{ab}{n}\right)=\left(\frac{a}{n}\right)\left(\frac{b}{n}\right)$.

2. When $(a,mn)=1$, $\left(\frac{a}{mn}\right)=\left(\frac{a}{m}\right)\left(\frac{a}{n}\right)$. (True even when $(a,mn)>1$.)

3. $\left(\frac{-1}{n}\right)=(-1)^{\frac{n-1}{2}}$.

4. $\left(\frac{2}{n}\right)=(-1)^{\frac{n^2-1}{8}}$.

5. If m and n are odd and $(m,n)=1$, $\left(\frac{m}{n}\right)\left(\frac{n}{m}\right)=(-1)^{\frac{(m-1)(n-1)}{4}}$.

6. If a is a quadratic residue mod n then $\left(\frac{a}{n}\right)=1$. Hence if $\left(\frac{a}{n}\right)=-1$ then a is not a quadratic residue mod n.

Reasons: (1) by the same property of the Legrende symbol, (2) by prime factorization, (3) by induction on $l(n)$, (4) by induction on $l(n)$, (5) by induction on $l(m)+l(n)$, where $l(m)$ is the number of primes in the prime factorization of m, (6) by the lemma and corollary below.

Lemma: Let p be an odd prime. If $(a,p)=1$ and a is a quadratic residue of a mod p then it is a quadratic residue of p^k for all k.

Proof: By induction on k. The base case $k=1$ is trivial. Now assume $x^2\equiv a \bmod p^k$ has a solution. Write $x^2=a+\alpha p^k$. Set $y=x+\beta p^k$. Then

$$y^2=x^2+2x\beta p^k+\beta^2p^{2k}\equiv a+\alpha p^k+2x\beta p^k \bmod p^{k+1}.$$

We wish to find β such that

$$\alpha p^k + 2x\beta p^k \equiv 0 \bmod p^{k+1},$$

or equivalently

$$\alpha + 2x\beta \equiv 0 \bmod p.$$

There is a solution because $(2x, p) = 1$.

Corollary: Let $n = p_1^{e_1} \cdots p_k^{e_k}$ be a product of odd primes. Let $(a, n) = 1$. Then a is a quadratic residue mod n iff a is a quadratic residue mod p_i for $1 \le i \le k$.

Proof: If $x^2 \equiv a \bmod n$ then $x^2 \equiv a \bmod p_i$ for each i. Conversely, suppose that for each $i \le k$ there is an x_i such that $x_i^2 \equiv a \bmod p_i$. By the corollary there is a y_i such that $y_i^2 \equiv a \bmod p_i^{e_i}$ for each $i \le k$. By the Chinese Remainder Theorem, there is a z such that $z \equiv y_i \bmod p_i^{e_i}$ for each i. This implies $z^2 \equiv a \bmod p_i^{e_i}$ for each i, which implies $z^2 \equiv a \bmod n$.

Section 4.7 Exercises:

(i) $\left(\frac{5}{p}\right) = \left(\frac{p}{5}\right) \equiv p^2 \bmod 5$. Hence we need $p \equiv 1, 4 \bmod 5$.

(ii) To decide if 2 is a quadratic residue mod p' we evaluate $2^p \bmod p'$. On the other hand, since $p = 3 + 4k$, $p' = 7 + 8k$, we know that 2 is a quadratic residue mod p'. Hence $2^p \equiv 1 \bmod p'$. Hence $2^p - 1$ is not prime since it has proper divisor p'. We have also proved that 2 is primitive mod p'.

(iii) Let g be a primitive root of p. The quadratic residue product is $P = g^2 g^4 \cdots g^{p-1} = g^{\frac{p^2-1}{4}} = (g^{\frac{p-1}{2}})^{\frac{p+1}{2}} = h^{\frac{p+1}{2}}$ where $h = g^{\frac{p-1}{2}}$. Since $h^2 \equiv 1$, $h \equiv \pm 1$. But $h \not\equiv 1$ since g has order $p-1$, hence $h \equiv -1$ and $P \equiv (-1)^{\frac{p+1}{2}}$.

(iv) We have $\left(\frac{-1}{p}\right) = (-1)^{\frac{p-1}{2}} = 1$ since $p \equiv 1 \bmod 4$. Hence -1 is a quadratic residue. This implies that whenever r is a quadratic residue, so is $-r$. There is a one-to-one correspondence between quadratic residues $\le p/2$ and quadratic residues $\ge p/2$ via $r \leftrightarrow p-r$. There are $\frac{p-1}{2}$ quadratic residues, every pair of which sums to p. Since there are $\frac{p-1}{4}$ pairs, the sum of them all is $\frac{p-1}{4} \cdot p$.

(v) Use the properties repeatedly.

(vi) When $(d, p) > 1$, $d \equiv 0$ and there is exactly one solution, consistent with the formula. When $(d, p) = 1$ and d is not a quadratic residue then there are no solutions, which is consistent with the formula. When $(d, p) = 1$ and d is a quadratic residue, then $x^2 - d$ has a root r, and we can write $x^2 - d \equiv (x - r)(x - s)$ in $F_p[x]$. If $r \equiv s$ then $x^2 - d \equiv x^2 - 2rx + r^2$ hence $r \equiv 0$, $d \equiv 0$: contradiction. Hence there are two solutions, consistent with the formula.

(vii) Since $(a, p) = (2, p) = (4, p) = 1$, we can divide by $a, 2, 4$ in F_p. We have

$$ax^2 + bx + c = (1/a)((ax + b/2)^2 - d/4),$$

hence

$$\left(\frac{f(x)}{p}\right) = \left(\frac{1/a}{p}\right)\left(\frac{(ax + b/2)^2 - d/4}{p}\right) = \left(\frac{a}{p}\right)\left(\frac{(ax + b/2)^2 - d/4}{p}\right),$$

$$\sum_{x=1}^{p}\left(\frac{f(x)}{p}\right) = \left(\frac{a}{p}\right)\sum_{x=1}^{p}\left(\frac{(ax + b/2)^2 - d/4}{p}\right) = \left(\frac{a}{p}\right)\sum_{x=1}^{p}\left(\frac{x^2 - d/4}{p}\right)$$
$$= \left(\frac{a}{p}\right)\sum_{x=1}^{p}\left(\frac{x^2 + d}{p}\right).$$

When $d = 0$ this evaluates to $\left(\frac{a}{p}\right)(p - 1)$. Now consider $d \neq 0$. Since $1 + \left(\frac{k^2+d}{p}\right)$ counts the number of solutions to $x^2 - k^2 = d$,

$$\sum_{k=1}^{p}\left[1 + \left(\frac{k^2 + d}{p}\right)\right]$$

is the size of the set $\{(a, b) \in F_p \times F_p : a^2 - b^2 = d\}$. This is in 1:1 correpondence with the set $\{(u, v) : uv = d/4\}$, and the size of the latter set is $p - 1$. Hence

$$\sum_{k=1}^{p}\left(\frac{k^2 + d}{p}\right) = -1.$$

(viii) To show that 2 is a primitive root mod p we must show that $o(2) = p-1$. We have $p - 1 = 2p'$, which has divisors $1, 2, p', 2p'$. Therefore the order of 2 mod p is one of these numbers. We can rule out 1 and 2 since $p \geq 11$. Moreover if $o(2) = p'$ then $2^{\frac{p-1}{2}} \equiv 1 \bmod p$, which implies that 2 is a quadratic residue mod p, which contradicts the fact that $p \equiv 3 \bmod 8$. Therefore $o(2) = 2p' = p - 1$ and 2 is primitive.

Now $5^{\frac{p-1}{2}} \equiv 1 \bmod p$ if and only if $p \equiv 1, 4 \bmod 5$ by problem (i), hence $5^{\frac{p-1}{2}} \equiv -1 \bmod p$ if and only if $p \equiv 2, 3 \bmod 5$ if and only if $p' \equiv 1, 3 \bmod 5$. For these primes we have $o(5) \in \{1, 2, p-1\} \bmod p$. We can rule out $o(5) = 1, 2$ by a direct inspection of $p = 7$ and $p = 23$. The next smallest p is 47, and the order of 5 in this case is 46.

(ix) If $p = 2$ then we can easily solve $ax + by \equiv c \bmod 2$, hence $ax^2 + by^2 \equiv c \bmod 2$. Now assume that p is an odd prime. We are attempting to show that $\left(\frac{-(a/b)x^2+(c/b)}{p}\right) = 1$ for some x. Now if $-(a/b)x^2 + (c/b) \equiv 0$ for some x then $x^2 \equiv c/a$, hence $(x, 0)$ is a solution to $ax^2 + by^2 \equiv c$. Now assume that $-(a/b)x^2 + (c/b) \not\equiv 0$ for all x. Since the discriminant of this polynomial is $4ac/b^2 \not\equiv 0$, we know by problem (vii) that

$$\sum_{x=1}^{p} \left(\frac{-(a/b)x^2 + (c/b)}{p}\right) = -\left(\frac{-(a/b)}{p}\right) = \pm 1.$$

This implies that some $-(a/b)x^2 + (c/b)$ is a quadratic residue, otherwise the sum would be $-p$.

Generalized Lagrange Theorem

Theorem: Let $f(x_1, x_2, \ldots, x_n) \in \mathbb{Z}_p[x_1, \ldots, x_n]$ be a polynomial with degree $\leq p - 1$ in each variable x_i. Assume that $f(a_1, \ldots, a_n) = 0$ for all $(a_1, \ldots, a_n) \in \mathbb{Z}_p^n$. Then all the coefficients of $f(x_1, \ldots, x_n)$ are equal to zero.

Proof: By induction on n. First consider $n = 1$. Then $f(x_1)$ has roots $0, 1, \ldots, p-1$, therefore $f(x_1) = g(x_1) \prod_{i=0}^{p-1}(x_1 - i)$ for some polynomial $g(x_1)$. If $g(x_1)$ has a non-zero coefficient then we can write $g(x_1) = g_a x_1^a +$ terms of lower degree where $g_a \neq 0$. This implies that $f(x_1)$ has degree $\geq p + a$, contrary to hypothesis. Hence $g(x_1)$ has all zero coefficients, which implies that $f(x_1)$ has all zero coefficients.

Assume that the statement of the theorem is true for some $n \geq 1$. Let $f(x_1, \ldots, x_{n+1})$ be a polynomial that meets the hypothesis of the theorem. We will show that all the coefficients of $f(x_1, \ldots, x_{n+1})$ are equal to zero.

We can write

$$f(x_1, \ldots, x_{n+1}) = \sum_{i=0}^{p-1} f_i(x_1, \ldots, x_n) x_{n+1}^i.$$

Fixing $a_1, \ldots, a_n$, the polynomial $\sum_{i=0}^{p-1} f_i(a_1, \ldots, a_n)x_{n+1}^i$ has p roots, hence each coefficient $f_i(a_1, \ldots, a_n)$ is equal to zero. Now let $a_1, \ldots, a_n$ vary and use the induction hypothesis to show that $f_i(x_1, \ldots, x_n) = 0$ for $0 \leq i \leq p-1$.

(x) Let $f(x_1, \ldots, x_n)$ be a polynomial that vanishes only at $(0, 0, \ldots, 0)$. We will show that the total degree of f is $\geq n$. Note that $1 - f^{p-1}$ vanishes at every non-trivial $(x_1, \ldots, x_n)$ and evaluates to 1 at $(0, \ldots, 0)$. So $1 - f^{p-1}$ is the same function as $h = (1 - x_1^{p-1}) \cdots (1 - x_n^{p-1})$ from $\mathbb{Z}_p^n$ to $\mathbb{Z}_p$. Let g be the polynomial functionally equal to $1 - f^{p-1}$ by repeatedly replacing every instance of x_i^k with $k \geq p$ in $1 - f^{p-1}$ by x_i^{k-p+1}. The is possible because x_i^p and x_i are functionally equal. Then $g - h$ vanishes on $\mathbb{Z}_p^n$ and every variable has exponent at most $p - 1$. By the generalized Lagrange theorem (above), $g - h = 0$, hence $g = h$, hence has total degree $(p-1)n$. This implies that the total degree of $1 - f^{p-1}$ is $\geq (p-1)n$, hence the total degree of f is $\geq n$.

(xi) A special case of (x).

Chapter 5: Quadratic Forms

Section 5.1: Equivalence

Quadratic form:

$$f(x, y) = ax^2 + bxy + cy^2.$$

In matrix form:

$$f(v) = v^T \begin{bmatrix} a & b/2 \\ b/2 & c \end{bmatrix} v.$$

The discriminant of a quadratic form is $d(f) = b^2 - 4ac$. When $f(v) = v^T F v$ we have $d(f) = -4\det(F)$. We will say that forms f and g are equivalent if $f(v) = g(Uv)$ for some 2×2 integer matrix U with determinant 1 (unimodular matrix). In other words, $f = g \circ U$. This is an equivalence relation: $f(v) = f(Iv)$; $f(v) = g(Uv)$ implies $f(U^{-1}v) = g(v)$; $f(v) = g(U_1 v)$ and $g(v) = h(U_2 v)$ implies $f(v) = h(U_2 U_1 v)$. Equivalent forms have the same discriminant: given $f(v) = g(Uv)$ and $g(v) = v^T G v$ we have $f(v) = v^T(U^T G U)v$, hence $F = U^T G U$, hence $\det(F) = \det(G)$. Equivalent forms produce the same set of output values. Note also that $4af(x, y) = (2ax + by)^2 - dy^2$, hence when $d < 0$ the output values of f are all ≤ 0 when $a < 0$ and all ≥ 0 when $a > 0$. Moreover the only zero output occurs when $x = y = 0$.

More facts about forms:

1. If $f = ax^2 + bxy + cy^2$ then $(f \circ U) = f(p,r)x^2 + b'xy + f(q,s)y^2$ where $U = \begin{bmatrix} p & q \\ r & s \end{bmatrix}$.

2. Let $U_k = \begin{bmatrix} 1 & k \\ 0 & 1 \end{bmatrix}$ and $V = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}$. If $f = ax^2 + bxy + cy^2$ then $f \circ U_k = ax^2 + (b + 2ak)xy + (ak^2 + bk + c)y^2$ and $f \circ V = cx^2 - bxy + ay^2$.

3. Let f be a form with $a > 0$ and $d < 0$ and having minimum positive output a_{min} using integer inputs. Then we can produce an equivalent form $g = a_{min}x^2 + \cdots$. To find a_{min}, proceed as follows (following Niven and Zuckerman): Given any output m, check all (x, y) such that $f(x, y) < m$, i.e. solve

$$\frac{(2ax + by)^2 - dy^2}{4a} < m.$$

This requires

$$y^2 < \frac{4am}{-d},$$

which has a finite number of solutions in y. Given y in this range, we hunt for integers x such that

$$(2ax + by)^2 < 4am + dy^2,$$

and there are finitely many values of x to check for each y. Having found (p, r) such that $f(p, r) = a_{min}$, the fact that f is homogeneous implies that $\gcd(p, r) = 1$. Hence there exist integers q, s such that $U = \begin{bmatrix} p & q \\ r & s \end{bmatrix}$ has determinant 1. We set $g = f \circ U$.

Example: Consider $f(x, y) = 37x^2 + 59xy + 25y^2$. The discriminant is -219. Some outputs are $\{37, 37 - 59 + 25, 25\}$, the least of which is $m = 3$. Solving $y^2 < 148/219$ we have $y = 0$. Number of non-zero solutions to $4ax^2 < 4am$: none. So the smallest output is $3 = f(1, -1)$. We set

$$g = f \circ \begin{bmatrix} 1 & 3 \\ -1 & -2 \end{bmatrix} = 3x^2 + 27xy + 79y^2.$$

Section 5.2: Reduction

From here on out we are going to consider $f(x, y) = ax^2 + bxy + cy^2$ where $a > 0$ and $d < 0$. This forces $c > 0$ since $b^2 < 4ac$. We will show that every such form is equivalent to $Ax^2 + Bxy + Cy^2$ where $-A < B \leq A < C$ or $0 \leq B \leq A = C$. We call such forms reduced, and we will show that distinct reduced forms are inequivalent.

Given $f(x, y) = ax^2 + bxy + cy^2$, let $g(x, y) = a_0x^2 + b_0xy + c_0y^2$ be any form equivalent to $f(x, y)$ where a_0 is the minimum positive output of f. Now choose k so that $-a_0 < b_0 + 2ka_0 \leq a_0$ (division algorithm applied to $b + a - 1$ and $2a$). Then we obtain the equivalent form $h(x, y) = a_0x^2 + (b_0 + 2ka_0)xy + (a_0k^2 + b_0k + c_0)y^2$. Since $a_0k^2 + b_0k + c_0$ is an output of h, it is an output of f, hence $a_0 \leq a_0k^2 + b_0k + c_0$. This is reduced if $a_0 < a_0k^2 + b_0k + c_0$, and if $a_0 = a_0k^2 + b_0k + c_0$ and $b_0 + 2ka_0 < 0$ then the equivalent form $k(x, y) = a_0x^2 - (b_0 + 2ka_0)xy + a_0y^2$ is reduced.

Example: We have already shown that $f(x, y) = 37x^2 + 59xy + 25y^2$ has minimum output 3 and is equivalent to $3x^2 + 27xy + 79y^2$. In order to subtract 24 from 3 we will compose with U_{-4}. This yields

$$h(x, y) = (3x^2 + 27xy + 79y^2) \circ \begin{bmatrix} 1 & -8 \\ 0 & 1 \end{bmatrix} = 3x^2 + 3xy + 19y^2.$$

We now show that distinct reduced forms are inequivalent. Let $f(x, y) = ax^2 + bxy + cy^2$ and $g(x, y) = Ax^2 + Bxy + Cy^2$ be reduced quadratic forms with $a, A > 0$ and the same discriminant $d < 0$. We will show that if they are equivalent then $a = A$, $b = B$, and $c = C$. Since they are equivalent, they have the same outputs. For $x^2 \geq y^2 > 0$ we have

$$f(x, y) \geq ax^2 - |b|x^2 + cy^2 = (a - |b|)x^2 + cy^2 \geq c$$

and for $y^2 \geq x^2 > 0$ we have

$$f(x, y) \geq ax^2 - |b|y^2 + cy^2 = ax^2 + (c - |b|)y^2 \geq a + c - |b| \geq c.$$

We also have

$$f(x, 0) = ax^2 \geq a$$

and

$$f(0, y) = cy^2 \geq c$$

for $x, y \neq 0$. Therefore the five smallest outputs of f are $0, a, a, c, c$ and similarly the five smallest outputs of g are $0, A, A, C, C$. This implies $a = A$ and $c = C$, which implies $b = \pm B$ since the discriminants are equal. We must prove $B = -b$ implies $b = B = 0$. This is clear if $a = c$ since the forms are reduced, so we can assume $a < c$.

Suppose $f(x, y) = ax^2 + bxy + cy^2$, $g(x, y) = ax^2 - bxy + cy^2$, $b \geq 0$, $a < c$, and $f(v) = g(Uv)$ where $U = \begin{bmatrix} p & q \\ r & s \end{bmatrix}$. We wish to show $U = I$ and $b = 0$. We must have $a > b$ given $-a < |b| \leq a$. If $pr \neq 0$ then we have $c > a = f(1, 0) = g(p, r) \geq C = c$, which is impossible. Therefore $p = 0$ or $r = 0$. If $qs \neq 0$ then we have $c = f(0, 1) = g(q, s) > c$, which is impossible. Given $F = U^T G U$ and

$$F = \begin{bmatrix} a & b/2 \\ b/2 & c \end{bmatrix}$$

and

$$G = \begin{bmatrix} a & -b/2 \\ -b/2 & c \end{bmatrix},$$

we obtain

$$F \in \left\{ \begin{bmatrix} cr^2 & -(b/2)qr \\ -(b/2)qr & aq^2 \end{bmatrix}, \begin{bmatrix} ap^2 & (-b/2)ps \\ (-b/2)ps & cs^2 \end{bmatrix} \right\}.$$

The first possibility contradicts $a = c$. The second possibility implies that $b = -b$ since $ps = 1$, hence $b = 0$.

There are finitely many reduced forms $h(d)$ with $a > 0$ and $d < 0$: We have $-a \leq b \leq a \leq c$, hence $b^2 \leq ac$, hence $-d = 4ac - b^2 \geq 3ac$, $|b| \leq a \leq c \leq -d/3a \leq -d/3$, hence there are finitely many choices for a, b, and c.

Example: $x^2 + y^2$ is reduced and satisfies $d = -4$. The equivalent reduced forms $ax^2 + bxy + cy^2$ satisfy $|b| \leq a \leq c \leq 4/3$, which forces $a = c = 1$ and $b = 0$. In other words, $x^2 + y^2$ is the unique reduced form with discriminant -4.

Note that $c = \frac{b^2 - d}{4a}$. Reducing $|b|$ reduces c when a is unchanged. This leads to another algorithm for finding reducing a form and for finding a_{min} given an arbitrary form with $a > 0$ and $d < 0$: If the form is reduced then $a_{min} = a$. If the form is not reduced, then either (1) $a > c$ or (2) $a = c$ and $b > a$ or (3) $a = c$ and $b < 0$ or (4) $a < c$ and $b > a$ or (5) $a < c$ and $b \leq -a$. The following actions either lower $[x^2]$ or identify a reduced form:

(1) $a > c :$ Apply V, lowering $[x^2]$.

(2) $a = c$ and $b > a$: We have $b > b - 2a > -a > -b$. Find the largest $k \geq 1$ such that $b - 2ka > -b$, such that $|b - 2ka| < |b|$, then apply U_k, then apply V, lowering $[x^2]$.

(3) $a = c$ and $b < 0$: If $b \geq -a$ then applying V produces a reduced form. But if $b < -a$ then $b < b + 2a < a < -b$, and there is a largest $k \geq 1$ such that $b < b + 2ka < -b$. Apply U_k, then V. Summarizing: Find the largest $k \geq 0$ such that $|b + 2ka| < |b|$, then apply U_k, then apply V, either producing a reduced form or lowering $[x^2]$.

(4) $a < c$ and $b > a$: Applying U_{-1} leads to $b > b - 2a > -a > -b$. Find the largest value of $k \geq 1$ such that $|b - 2ka| < |b|$, then apply U_{-k}, then V, lowering $[x^2]$.

(5) $a < c$ and $b \leq -a$: Applying U_1 leads to $b + 2a \leq a \leq -b$. Find the largest $k \geq 1$ such that $|b + 2ka| \leq |b|$, then apply U_k, then V, either producing a reduced form or lowering $[x^2]$.

Section 5.3: Representations by Binary Forms

Definition: Let a be a natural number. We say that a is properly represented by the binary form f iff $a = f(p, r)$ for some coprime pair p and r. For example, if $f(x, y) = 37x^2 + 59xy + 25y^2$ then $f(4, -3) = 109$ hence 109 is properly represented by f.

Theorem: A necessary and sufficient condition that a be properly represented by a binary form with discriminant d is that $b^2 \equiv d \bmod 4a$ has a solution. In other words, d is a quadratic residue mod $4a$.

Proof: Suppose $b^2 \equiv d \bmod 4a$. Then $b^2 - d = 4ac$ for some c, therefore $b^2 - 4ac = d$. Setting $f(x, y) = ax^2 + bxy + cy^2$ we have $d(f) = d$ and $a = f(1, 0)$.

Conversely, suppose $a = f(p, r)$ where $(p, r) = 1$. Then $a = g(1, 0)$ where $g(v) = f(Uv)$ and $U = \begin{bmatrix} p & q \\ r & s \end{bmatrix}$. We know that $g(x, y) = ax^2 + bxy + cy^2$ for some b and c, therefore $d = d(f) = d(g) = b^2 - 4ac \equiv b^2 \bmod 4a$.

Example: we will determine the primes representable as a sum of two squares. We have $2 = 1^2 + 1^2$. The odd primes are of the form $4n + 1$ and $4n + 3$. No prime of the form $4n + 3$ can be represented as a sum of 2 squares, because the latter is congruent to 0, 1, or 2 mod 4. When $p \equiv 1 \bmod 4$, -1 is a quadratic residue mod p, hence -4 is a quadratic residue mod $4p$, hence p

can be properly represented by a form with discriminant -4, hence by a reduced form with discriminant -4. Since x^2+y^2 is the unique reduced form with discriminant -4, p can be properly represented by x^2+y^2.

Example: Primes of the form $4n+3$ are 3, 7, 11, 19, ... and primes of the form $4n+1$ are of the form 5, 13, 17, ... Setting

$$f(n) = \{\sqrt{n^2-x^2} : 0 \le x \le \sqrt{n/2}\}$$

we have

$$f(3) = \{1.73205, 1.41421\},$$
$$f(7) = \{2.64575, 2.44949\},$$
$$f(11) = \{3.31662, 3.16228, 2.64575\},$$
$$f(19) = \{4.3589, 4.24264, 3.87298, 3.16228\},$$
$$f(5) = \{2.23607, 2.\},$$
$$f(13) = \{3.60555, 3.4641, 3.\},$$
$$f(17) = \{4.12311, 4., 3.60555\}.$$

Section 5.4: Sums of Two Squares

Necessary Conditions: Suppose x^2+y^2 is divisible by an odd prime p. Then $x^2 \equiv -y^2 \bmod p$, hence $(y,p)=1$ implies $(x/y)^2 \equiv -1 \bmod p$ implies $p \equiv 1 \bmod 4$. So if $p \equiv 3 \bmod 4$ then $p|y$, which implies $p|x$, which $p^2|(x^2+y^2)$. Hence in the prime factorization of x^2+y^2, primes $\equiv 3 \bmod 4$ occur to even exponent.

Sufficient Conditions: Suppose n is any arbitrary number with this property. Write $n = n_0^2 p_1 p_2 \cdots p_k$, the p_i distinct primes. Then $p_i \equiv 1 \bmod 4$ (or $p_i = 2$) for each i, hence $x_i^2 \equiv -1 \bmod p_i$ has a solution for each i, hence by the Chinese remainder theorem $b^2 \equiv -1 \bmod p_1 \cdots p_k$ has a solution, hence $(2b) \equiv -4 \bmod 4p_1 \cdots p_k$ has a solution, hence $p_1 \cdots p_k$ is representable by a binary quadratic form with discriminant -4, hence by a reduced form with this discriminant, which can only be x^2+y^2. So we have $p_1 \cdots p_k = x^2+y^2$, $n = n_0^2 p_1 \cdots p_k = n_0^2(x^2+y^2) = (n_0x)^2 + (n_0y)^2$.

A second proof: Write each p_i in the form $x_i^2 + y_i^2$ and use the fact that the set of sums of squares is closed with respect to multiplication:

$$(x_1^2+y_1^2)(x_2^2+y_2^2) = |x_1+y_1 i|^2 |x_2+y_2 i|^2 = |(x_1+y_1 i)(x_2+y_2 i)|^2 =$$

$$|(x_1x_2 - y_1y_2) + (x_1y_2 + x_2y_1)i|^2 = (x_1x_2 - y_1y_2)^2 + (x_1y_2 + x_2y_1)^2.$$

Example:

$$1485154 = 2 \cdot 11^2 \cdot 17 \cdot 19^2 = (1^2 + 1^2)(4^2 + 1)11^2 19^2 = (3^2 + 5^2)11^2 19^2 =$$
$$627^2 + 1045^2.$$

Section 5.5: Sums of Four Squares

The set of sums of four square integers is closed with respect to multiplication: using quaternions,

$$(x_1^2+x_2^2+x_3^2+x_4^2)(y_1^2+y_2^2+y_3^2+y_4^2) = |x_1-x_2i-x_3j-x_4k|^2|y_1+y_2i+y_3j+y_4k|^2 =$$
$$|(x_1y_1 + x_2y_2 + x_3y_3 + x_4y_4) + (x_1y_2 - x_2y_1 - x_3y_4 + x_4y_3)i+$$
$$(x_1y_3 + x_2y_4 - x_3y_1 - x_4y_2)j + (x_1y_4 - x_2y_3 + x_3y_2 - x_4y_1)k|^2 =$$
$$(x_1y_1 + x_2y_2 + x_3y_3 + x_4y_4)^2 + (x_1y_2 - x_2y_1 - x_3y_4 + x_4y_3)^2+$$
$$(x_1y_3 + x_2y_4 - x_3y_1 - x_4y_2)^2 + (x_1y_4 - x_2y_3 + x_3y_2 - x_4y_1)^2.$$

Another derivation: Let z and w be complex numbers. Then

$$\begin{bmatrix} z & w \\ \overline{w} & \overline{z} \end{bmatrix}$$

has a determinant which is a sum of four squares. The product of two such matrices has a similar form, and $\det(AB) = \det(A)\det(B)$, hence a product of two sums of four squares is a sum of four squares.

The numbers 1 and 2 can be expressed as the sum of four squares. If we can show that every odd prime can be expressed as the sum of four squares, then every natural number can be.

Observation 1: If n is a sum of four squares and n is even then $\frac{n}{2}$ is a sum of four squares. This follows from the identity

$$\frac{a^2 + b^2 + c^2 + d^2}{2} = \left(\frac{a+b}{2}\right)^2 + \left(\frac{a-b}{2}\right)^2 + \left(\frac{c+d}{2}\right)^2 + \left(\frac{c-d}{2}\right)^2,$$

grouping together numbers of equal parity.

Observation 2: If

$$n =$$

$$(x_1y_1 + x_2y_2 + x_3y_3 + x_4y_4)^2 + (x_1y_2 - x_2y_1 - x_3y_4 + x_4y_3)^2 +$$
$$(x_1y_3 + x_2y_4 - x_3y_1 - x_4y_2)^2 + (x_1y_4 - x_2y_3 + x_3y_2 - x_4y_1)^2$$

and $m^2|n$ and $x_i \equiv y_i \bmod m$ for $i = 1, 2, 3, 4$ then each of the squares in the sum is divisible by m^2.

Observation 3: If m is odd and mN is an odd sum of four squares and $m > 1$ then $m'N$ is a sum of four squares for some m' satisfying $1 \leq m' < m$. Proof: write $mN = a^2 + b^2 + c^2 + d^2$ and choose $a_0, b_0, c_0, d_0 \in (-m/2, m/2]$ such that $(a, b, c, d) \equiv (a_0, b_0, c_0, d_0) \bmod m$. Set $n = a_0^2 + b_0^2 + c_0^2 + d_0^2$. Then $n < 4\frac{m^2}{4} = m^2$. We have $n \equiv mN \equiv 0 \bmod m$, hence $n = m_1m$ for some m_1 satisfying $1 \leq m_1 < m$. The product $n(mN) = (m_1m)(mN) = m_1Nm^2$ is a sum of four squares, and by Observation 2 each of the squares is divisible by m^2. Hence m_1N is a sum of four squares.

Observation 4: If mN is a sum of four squares for some m then N is a sum of four squares. Proof: construct the sequence $m = m_0 > m_1 > m_2 > \cdots$ where each m_iN is a sum of four squares, setting $m_{i+1} = \frac{m_i}{2}$ when m_i is even as in Observation 1 and constructing m_{i+1} from m_i when m_i is odd as in Observation 3. At some point we must have $m_i = 1$.

To prove that an odd prime p is a sum of four squares it suffices to show that mp is a sum of four squares for some m. A fancy proof: setting $f(x, y, z) = x^2 + y^2 + z^2$ there is a non-trivial solution to $f(x, y, z) \equiv 0 \bmod p$ by Exercise (x), Chapter 4, so there are integers x, y, z, not all congruent 0 mod p, such that $x^2 + y^2 + z^2 = mp$ for some m. A plain proof (which a computer can find): Let $p = 2r + 1$ be given. The numbers $0^2, 1^2, \ldots, r^2$ are distinct mod p, as are the numbers $-1 - 0^2, -1 - 1^2, \ldots, -1 - r^2$: If $0 \leq i < j \leq r$ then $r \geq j - i \geq 1$ and $1 \leq i + j \leq 2r - 1$, hence $i - j \not\equiv 0$ and $i + j \not\equiv 0 \bmod p$, hence $i^2 - j^2 \not\equiv 0 \bmod p$, hence $i^2 \not\equiv j^2$. Since there are only $2r + 1$ residue classes, there has to be some overlap in the list: $x^2 \equiv -1 - y^2 \bmod p$ where $0 \leq x, y \leq r$. This yields $x^2 + y^2 + 0^2 + 1^2 = mp$ for some m.

Chapter 5 Exercises:

(i) Using Mathematica, the unique reduced forms with discriminant in the range $-1, -2, \ldots, -200$ are:

$d = -3$: $x^2 + xy + y^2$

$d = -4$: $x^2 + y^2$

$d = -7$: $x^2 + xy + 2y^2$

$d = -8$: $x^2 + 2y^2$

$d = -11$: $x^2 + xy + 3y^2$

$d = -19$: $x^2 + xy + 5y^2$

$d = -43$: $x^2 + xy + 11y^2$

$d = -67$: $x^2 + xy + 17y^2$

$d = -167$: $x^2 + xy + 41y^2$.

(ii) By Problem (i) the form $x^2 + xy + 5y^2$ is the unique reduced form with discriminant $d = -19$. It suffices to determine the odd primes properly represented by a form with discriminant -19 (the input must be coprime since the output will be prime). We must determine the odd primes p such that $b^2 \equiv -19 \bmod 4p$ has a solution. Now $b^2 \equiv -19 \bmod 4p$ if and only if $b^2 \equiv -19 \bmod 4$ and $b^2 \equiv -19 \bmod p$. The first congruence always has a solution: any odd number b. The equation $b^2 \equiv -19 \bmod p$ has a solution iff $\left(\frac{-19}{p}\right) = 1$. Any even solution b yields on odd solution $b + 19$. Using the Jacobi symbol and quadratic reciprocity, we have $\left(\frac{-19}{p}\right) = \left(\frac{-1}{p}\right)\left(\frac{19}{p}\right) = (-1)^{\frac{p-1}{2}}\left(\frac{p}{19}\right)(-1)^{\frac{(p-1)(19-1)}{4}} \equiv p^9 \bmod 19$. We have $p^9 \equiv 1 \bmod 19$ when p is equivalent to 1, 4, 5, 6, 7, 9, 11, 16, 17 mod 19.

The least primes congruent to one of these are $191 = f(6,5), 23 = f(1,2), 43 = f(1,-3), 101 = f(3,4), 197 = f(9,4), 47 = f(6,1), 163 = f(11,2), 73 = f(4,3), 131 = f(1,5)$.

(iii) Let $n = x^2 + 2y^2$. Let $p|n$ be an odd prime. If $(y,p) = 1$ then $0 \equiv x^2 + 2y^2 \bmod p$ implies $(xy^{-1})^2 \equiv -2 \bmod p$, hence $\left(\frac{-2}{p}\right) = 1$, $1 = \left(\frac{-1}{p}\right)\left(\frac{2}{p}\right) = (-1)^{\frac{p-1}{2}}(-1)^{\frac{p^2-1}{8}} = (-1)^{\frac{(p-1)(p+5)}{8}}$, hence $p \equiv 1$ or $p \equiv 3 \bmod 8$. Contrapositive: $p|n$ is congruent to 5 or 7 mod 8 then $p|y$, hence $p|x$, hence $p^2|n$. I will conjecture that the set of integers that can be expressed in the form $x^2 + 2y^2$ are those in which its prime divisors congruent to 5 or 7 mod 8 appear with an even exponent. Evidence of this: $1 = 1^2 + 2(0^2)$,

$2 = 0^2 + 2(1^2)$, $3 = 1^2 + 2(1^2)$, $4 = 2^2 + 2(0^2)$, 5 cannot be expressed in the form $x^2 + 2y^2$, $6 = 2^2 + 2(1^2)$, 7 cannot be expressed in the form $x^2 + 2y^2$.

Let $P(n)$ be the statement that if $n = x^2 + 2y^2$ then prime divisors of n congruent to 5 or 7 mod 8 appear with even exponent. Then $P(1)$ is true. Assume $P(1)$ through $P(n-1)$ are true. Now suppose $n = x^2+2y^2$ is possible and let $p|n$ where $p \equiv 5$ or $p \equiv 7$ mod 8. We have seen that $p|x$ and $p|y$, hence $p^2|n$ and we can write $n = p^2 n_0$ where $n_0 = x_0^2 + 2y_0^2$. Since $P(n_0)$ is true, so is $P(n)$. Hence $P(n)$ is true for all $n \geq 1$.

Conversely, let n be such that prime divisors congruent to 5 or 7 mod 8 appear with even exponent. Write $n = ms^2$ where m is square-free. Then m is a product of distinct primes not congruent to 5 or 7 mod 8. It will suffice that all such primes p are representable in the form $p = x_p^2 + 2y_p^2$, because the set of integers of the form $x^2 + 2y^2$ is closed with respect to multiplication:

$$(a^2 + 2b^2)(A^2 + 2B^2) = (aA + 2bB)^2 + 2(aB - Ab)^2.$$

This can be derived as follows: Set

$$M_k(x, y) = \begin{bmatrix} x & y \\ ky & x \end{bmatrix}.$$

Then $\det M_k(x, y) = x^2 - ky^2$. Given that we have

$$M_k(x_1, y_1) M_k(x_2, y_2) = M_k(x_1x_2 + ky_1y_2, x_1y_2 + y_1x_2),$$

after taking determinants we obtain

$$(x_1^2 - ky_1^2)(x_2^2 - ky_2^2) = (x_1x_2 + ky_1y_2)^2 - k(x_1y_2 + y_1x_2)^2.$$

The discriminant of $x^2 + 2y^2$ is -8, and there is just one form with this discriminant up to equivalence. The prime 2 can be expressed in this form. Given a prime p congruent to 1 or 3 mod 8, we have seen that $x^2 \equiv -2$ mod p has a solution. The same solution yields $(2x)^2 \equiv -8$ mod $4p$. So p can be expressed in the form $x_p^2 + 2y_p^2$. Note: we can use an infinite descent algorithm to express primes equal to 2 or congruent to 1,3 mod 8 in the form $x^2 + 2y^2$, adapting Problem 7 in the Problems to Think About for Chapter 5.

(iv) Integers congruent to $0, 1, 3$ mod 4 can be represented this way: $4n = (n+1)^2 - (n-1)^2$, $4n+1 = (2n+1)^2 - (2n)^2$, $4n+3 = (2n+2)^2 - (2n+1)^2$.

Integers congruent to 2 mod 4 cannot be presented this way: $(2a)^2 - (2b)^2 = 4(a^2 - b^2) \equiv 0$ and $(2a+1)^2 - (2b+1)^2 = 4(a^2 + a - b^2 - b) \equiv 0$.

(v) According to Mathematica, there are two reduced forms with discriminant -20: $x^2 + 5y^2$ and $2x^2 + 2xy + 3y^2$. Now let p be an odd prime not equal to 5. It is representable by one of these forms iff there is a solution to $x^2 \equiv -20$ mod $4p$, i.e. an even solution to $x^2 \equiv -20$ mod p, i.e. any solution at all (if an odd one exists, add p to obtain an even one), iff $\left(\frac{-20}{p}\right) = 1$, iff $(-1)^{\frac{p-1}{2}} p^2 \equiv 1$ mod 5 using quadratic reciprocity. Setting $p = 4k+1$ yields $p \equiv 1, 9$ mod 20. Setting $p = 4k+3$ yields $p \equiv 3, 7$ mod 20. We must show primes of the form $1 + 20k$ and $9 + 20k$ are representable in the form $x^2 + 5y^2$ and primes of the form $3 + 20k$ and $7 + 20k$ are not. A brute-force calculation shows that the only values taken on by $2x^2 + 2xy + 3y^2$ mod 20 are 0, 2, 3, 7, 8, 10, 12, 15, 18. So primes congruent to 1 or 9 mod 20 cannot be represented in this form and must be representable by $x^2 + 5y^2$. Another brute-force calculation shows that the only values taken on by $x^2 + 5y^2$ mod 20 are 0, 1, 4, 5, 6, 9, 10, 14, 16, so primes congruent to 3 or 7 mod 20 cannot be represented in this form and must be representable by $2x^2 + 2xy + 3y^2$.

Examples: The first primes congruent to 1 or 9 mod 20 are $41 = 6^2 + 5(1^2)$ and $29 = 3^2 + 5(2^2)$. The first primes after 101 that are congruent to 3 or 7 mod 20 are $103 = 2(-7)^2 + 2(-7)(5) + 3(5^2)$ and $107 = 2(-8)^2 + 2(-8)(3) + 3(3^2)$.

(vi) Mathematica: reduced forms with $d = -31$ are

$$x^2 + xy + 8y^2, 2x^2 - xy + 4y^2, 2x^2 + xy + 4y^2.$$

Hence $h(-31) = 3$.

(vii) It suffices to find the reduced form and calculate a. Applying U_k with $k = -2$ we can convert $4x^2 + 17xy + 20y^2$ to $4x^2 + xy + 2y^2$. Applying V we can further convert this to this $2x^2 - xy + 4y^2$. (Check: this is indeed one of the three reduced forms with discriminant -31.) The smallest output is 2.

(viii) Assuming that the number of representations of a by a form with discriminant d is equal to the number of disinct solutions to $b^2 \equiv d$ mod $4a$ in $[0, 2a)$ times the number of automorphs of discriminant d forms (not proved in the book), it suffices to show that the number of solutions to $x^2 \equiv -4$ mod $4n$ is the same as the number of solutions to $x^2 \equiv -4$ mod $8n$, since the discriminant of $x^2 + y^2$ is -4 and there is one such form up to equivalence. Since each solution to $x^2 \equiv -4$ mod $8n$ is a solution to

$x^2 \equiv -4 \bmod 4n$, it suffices to show that each solution to $x^2 \equiv -4 \bmod 4n$ is also a solution to $x^2 \equiv -4 \bmod 8n$. Suppose $x^2 \equiv -4 \bmod 4n$. Then $n = x^2 + y^2$ is possible. Using the closure-under-multiplication formula, $2n = (1^2 + 1^2)(x^2 + y^2) = (x - y)^2 + (x + y)^2$, therefore $x^2 \equiv -4 \bmod 8n$ is possible.

This solution suggest a bijection: $\phi(x, y) = (x-y, x+y)$. This maps solutions to $n = x^2+y^2$ injectively into solutions to $2n = x^2+y^2$. Moreover, if an even number m satisfies $m^2 = x^2+y^2$ then x and y have the same parity, therefore $p = \frac{x+y}{2}$ and $q = -\frac{x-y}{2}$ are integers, and $p^2 + q^2 = \frac{x^2}{4} + \frac{y^2}{4} = \frac{m^2}{4} = (m/2)^2$. Moreover $\phi(p, q) = (p - q, p + q) = (x, y)$, so the mapping is surjective.

(ix) Given $n = 3^k - 1 = x_1^k + x_2^k + \cdots + x_s^k$, each $x_i \in \{1, 2\}$. So there is some non-negative solution to $a + b = s$ where $a + b2^k = n$. The larger b is, the smaller s is. We need to find the maximum value of b such that $n - b2^k \geq 0$. This yields $b = [n/2^k]$, $a = n - b2^k = n - \lfloor \frac{n}{2^k} \rfloor 2^k$, $s = n - [n/2^k]2^k + [n/2^k]$.

Chapter 6: Diophantine Approximation

Introduction: Numbers can be classified as natural, integer, rational, real, complex. The rationals can be listed out uniquely in a sequence. The complex numbers cannot, because the reals in $[0, 1]$ cannot by Cantor's argument. So there exist irrational numbers. In fact, we can prove θ is irrational. Algebraic numbers are complex numbers that are roots to polynomials with integer coefficients. For example, $\sqrt{2}$ is algebraic. Since we can list out integer-coefficient polynomials sequentially, and each has a finite number of roots, we can list out their roots sequentially. So there exist transcendental (non-algebraic) numbers. We will prove in this chapter that e is transcendental and show how to construct other transcendental numbers.

Section 6.1: Dirichlet's Theorem

Theorem: Let θ be a real number. Then θ is irrational if and only if there exist an infinite number of reduced fractions $\frac{p}{q}$ such that $|\theta - \frac{p}{q}| < \frac{1}{q^2}$.

Proof: Consider a rational number $\theta = \frac{a}{b}$. When $\frac{p}{q} \neq \theta$ we have $|\theta - \frac{p}{q}| = \frac{|aq-bp|}{bq} \geq \frac{1}{bq}$, and $q \geq b \implies \frac{1}{bq} \geq \frac{1}{q^2}$. So $|\theta - \frac{p}{q}| < \frac{1}{q^2}$ can only be achieved for a finite number of values of q, namely $q < b$, which limits p to a finite number of values.

Consider an irrational number θ. To illustrate the construction we will find coprime integers p and q, $1 \leq q < 10$, such that $|q\sqrt{2} - p| \leq \frac{1}{10}$. This yields

$|\sqrt{2}-\frac{p}{q}| \leq \frac{1}{10q} < q^2$. Write $k\sqrt{2} = a_k+b_k$ for $0 \leq k \leq 9$, where a_k is an integer and $0 \leq b_k < 1$. Set $a_{10} = 0$ and $b_{10} = 1$. Now write $I_k = [k/10, (k+1)/10)$ for $0 \leq k \leq 8$ and $I_9 = [9/10, 1]$. Then the numbers $b_0, b_1, \ldots, b_{10}$ lie in the disjoint intervals $I_0, I_1, \ldots, I_9$, and two of these numbers lie in the same interval. Say that $b_i, b_j \in I_k$ where $0 \leq i < j < 10$. Then we have $|b_j - b_i| \leq \frac{1}{10}$. If $j < 10$ then we can write $|(j-i)\sqrt{2} - (a_j - a_i)| \leq \frac{1}{10}$. If $j = 10$ then we can write $|i\sqrt{2} - (1 + a_i)| \leq \frac{1}{10}$. Mathematica yields

k	b[k]	a[k]
0	0.	0
1	0.414214	1
2	0.828427	2
3	0.242641	4
4	0.656854	5
5	0.0710678	7
6	0.485281	8
7	0.899495	9
8	0.313708	11
9	0.727922	12
10	1	0.

We can choose $i = 1$, $j = 6$, which yields $|5\sqrt{2} - 7| \leq \frac{1}{10}$. Check: $5\sqrt{2} - 7 \approx 0.0710678$. So we can use $q = 5$, $p = 7$.

More generally, given $\theta \in \mathbb{R}$ and $1 < Q \in \mathbb{Z}$ there exist a pair of integers p, q with $1 \leq q < Q$ such that $|q\theta - p| \leq \frac{1}{Q}$. We can assume that p and q are coprime, dividing through if necessary by (p, q). Hence $|\theta - \frac{p}{q}| \leq \frac{1}{qQ} < \frac{1}{q^2}$. Having found $\frac{p}{q}$ satisfying this condition, choose any integer $Q' > \frac{1}{|\theta - \frac{p}{q}|}$. If $|q'\theta - p'| \leq \frac{1}{Q'}$ then

$$|\theta - \frac{p'}{q'}| < \frac{1}{q'}|\theta - \frac{p}{q}| \leq |\theta - \frac{p}{q}|,$$

hence $\frac{p'}{q'} \neq \frac{p}{q}$. So there are infinitely many such $\frac{p}{q}$.

Section 6.2: Continued Fractions

Let $x_0, x_1, x_2, \ldots$ be a sequence of real numbers with $x_i > 0$ for $i \geq 1$. The associated sequence of continued fractions is $[x_0], [x_0, x_1], [x_0, x_1, x_2], \ldots$

defined by the recurrence relation $[x_0] = x_0$ and

$$[x_0, x_1, \dots, x_n] = x_0 + 1/[x_1, x_2, \dots, x_n]$$

for $n \geq 1$. The first few terms in the sequence are

$$x_0, x_0 + \frac{1}{x_1}, x_0 + \frac{1}{x_1 + \frac{1}{x_2}}, x_0 + \frac{1}{x_1 + \frac{1}{x_2 + \frac{1}{x_3}}}, \dots.$$

Any rational number $\frac{a}{b}$ where $b > 0$ can be expressed in the form $[x_0, x_1, \dots, x_n]$ for some choice of integers $x_0, x_1, \dots, x_n$: by strong induction on b. For $b = 1$ we have $\frac{a}{b} = [x_0]$ where $x_0 = a$. Now assume for $1 \leq b \leq n$ that $\frac{a}{b} = [x_0, x_1, \dots, x_k]$ for some choice of integers $x_0, x_1, \dots, x_k$. Given $b = n+1$, write $a = qb+r$ where $0 \leq r \leq n$. If $r = 0$ then $\frac{a}{b} = [q]$, but if $1 \leq r < b$ then $\frac{b}{r} = [x_0, x_1, \dots, x_k]$ and $\frac{a}{b} = q + \frac{r}{b} = b + 1/[x_0, x_1, \dots, x_k] = [b, x_0, \dots, x_k]$.

Example: Consider the sequence of Fibonacci numbers $F_0, F_1, F_2, F_3, F_4, \dots = 1, 1, 2, 3, 5 \dots$. For $n \geq 2$ we have

$$\frac{F_n}{F_{n-1}} = \frac{F_{n-1} + F_{n-2}}{F_{n-1}} = 1 + \frac{F_{n-2}}{F_{n-1}} = [1, \frac{F_{n-1}}{F_{n-2}}],$$

hence

$$\frac{1}{1} = [1],\ \frac{2}{1} = [1, 1],\ \frac{3}{2} = [1, 1, 1],\ \frac{5}{3} = [1, 1, 1, 1], \dots.$$

Given an irrational number θ, we have $\theta_0 = \theta = a_0 + 1/\theta_1$ where $\theta_1 > 1$, $\theta_1 = a_1 + 1/\theta_2$ there $\theta_2 > 1$, $\theta_2 = a_2 + 1/\theta_3$ where $\theta_3 > 1$, etc via $a_n = [\theta_n]$ and $\theta_{n+1} = 1/\{\theta_n\}$ for all n. This gives rise to the continued fractions

$$\theta = a_0 + \frac{1}{\theta_1}$$

$$\theta = a_0 + \frac{1}{a_1 + \frac{1}{\theta_2}}$$

$$\theta = a_0 + \frac{1}{a_1 + \frac{1}{a_2 + \frac{1}{\theta_3}}}$$

etc. Hence for all n we have

$$\theta = [a_0, a_1, \dots, a_{n-1}, \theta_n].$$

Example: we have

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

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

hence

$$\sqrt{2} = [1, 2, 2, \ldots, 2, \theta_n]$$

for all $n \geq 1$.

Terminology: The numbers $a_0, a_1, \ldots$ are the partial quotients of θ, the numbers $\theta_1, \theta_2, \ldots$ are the complete quotients of θ, and the numbers $[a_0, a_1, \ldots, a_n]$ are the convergents of θ.

Theorem: $\lim_{n\to\infty} [a_0, a_1, \ldots, a_n] = \theta$ when θ is irrational.

Proof: For each $n \geq 0$ let (p_n, q_n) be the coprime pair with $q_n > 0$ that satisfies

$$\frac{p_n}{q_n} = [a_0, a_1, \ldots, a_n].$$

For each $n \geq 1$ let (p'_n, q'_n) be the coprime pair with $q'_n > 0$ that satisfies

$$\frac{p'_n}{q'_n} = [a_1, a_2, \ldots, a_{n+1}].$$

We can check directly that $p_0 = a_0$, $q_0 = 1$, $p_1 = a_0 a_1 + 1$, $q_1 = a_1$. Hence

$$\begin{bmatrix} p_1 & p_0 \\ q_1 & q_0 \end{bmatrix} = \begin{bmatrix} a_0 & 1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} a_1 & 1 \\ 1 & 0 \end{bmatrix}.$$

Given

$$\frac{p_n}{q_n} = a_0 + \frac{q'_{n-1}}{p'_{n-1}} = \frac{a_0 p'_{n-1} + q'_{n-1}}{p'_{n-1}}$$

for $n \geq 2$, we also have

$$\begin{bmatrix} p_n & p_{n-1} \\ q_n & q_{n-1} \end{bmatrix} = \begin{bmatrix} a_0 & 1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} p'_{n-1} & p'_{n-2} \\ q'_{n-1} & q'_{n-2} \end{bmatrix}.$$

Hence we can prove by induction that

$$\begin{bmatrix} p_n & p_{n-1} \\ q_n & q_{n-1} \end{bmatrix} = \begin{bmatrix} a_0 & 1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} a_1 & 1 \\ 1 & 0 \end{bmatrix} \cdots \begin{bmatrix} a_n & 1 \\ 1 & 0 \end{bmatrix}.$$

Taking determinants, this yields

$$\begin{vmatrix} p_n & p_{n-1} \\ q_n & q_{n-1} \end{vmatrix} = (-1)^{n+1}.$$

This implies

$$\left| \frac{p_n}{q_n} - \frac{p_{n-1}}{q_{n-1}} \right| = \frac{1}{q_{n-1}q_n}.$$

The matrix identity implies

$$\begin{bmatrix} p_n & p_{n-1} \\ q_n & q_{n-1} \end{bmatrix} = \begin{bmatrix} p_{n-1} & p_{n-2} \\ q_{n-1} & q_{n-2} \end{bmatrix} \begin{bmatrix} a_n & 1 \\ 1 & 0 \end{bmatrix},$$

hence

$$p_n = a_n p_{n-1} + p_{n-2}$$

and

$$q_n = a_n q_{n-1} + q_{n-2}.$$

Since $q_n \to \infty$, this implies

$$\lim_{n\to\infty} \left| \frac{p_n}{q_n} - \frac{p_{n-1}}{q_{n-1}} \right| = 0.$$

If we can show that

$$\frac{p_{2n+1}}{q_{2n+1}} > \theta > \frac{p_{2n}}{q_{2n}}$$

for all n, then this limit implies that $\frac{p_n}{q_n} \to \theta$.

Let $x_0, x_1, x_2, \ldots$ with $x_i > 0$ for all $i \geq 1$. Let x_k^+ denote a quantity larger than x_k. We can prove by induction on n that $[x_0, \ldots, x_{2n}^+] > [x_0, \ldots, x_{2n}]$ and $[x_0, \ldots, x_{2n+1}^+] < [x_0, \ldots, x_{2n+1}]$. Given that $\theta_n > a_n$ for all n, we have

$$\frac{p_{2n+1}}{q_{2n+1}} = [a_0, \ldots, a_{2n+1}] > [a_0, \ldots, \theta_{2n+1}] = \theta = [a_0, \ldots, \theta_{2n}] > [a_0, \ldots, a_{2n}] = \frac{p_{2n}}{q_{2n}}.$$

Example: Applying this to $\sqrt{2} = [1, 2, 2, \ldots]$ we have $p_n = 2p_{n-1} + p_{n-2}$ and $q_n = 2q_{n-1} + q_{n-2}$ for $n \geq 2$ with $p_0 = 1$, $p_1 = 3$, $q_0 = 1$, $q_1 = 2$. This yields the sequence of fractions

$$1, \frac{3}{2}, \frac{7}{5}, \frac{17}{12}, \frac{41}{29}, \frac{99}{70}, \frac{239}{169}, \frac{577}{408}, \frac{1393}{985}, \frac{3363}{2378}, \frac{8119}{5741} \cdots \longrightarrow \sqrt{2}.$$

Example: Setting $a_i = 1$ for all i yields $p_i = F_{i+1}$ and $q_i = F_i$ for all $i \geq 0$. Given that

$$\frac{p_{n+1}}{q_{n+1}} = 1 + \frac{q_n}{p_n}$$

for all n, in the limit we obtain $\theta = 1 + 1/\theta$. This implies $\theta = \frac{1+\sqrt{5}}{2}$. Therefore

$$\lim_{n\to\infty} \frac{F_{n+1}}{F_n} = \frac{1+\sqrt{5}}{2}.$$

Section 6.3: Rational Approximations

We can extract a lot of information from the proof above about the convergents $\frac{p_n}{q_n} = [a_0, \ldots, a_n]$ to an irrational θ:

1. The sequence of differences $|\theta - \frac{p_n}{q_n}|$ is strictly decreasing. To see this, note that the recurrence relation

$$P_{n+1} = x_{n+1}P_n + P_{n-1}$$

and

$$Q_{n+1} = x_{n+1}Q_n + Q_{n-1}$$

holds for any arbitrary sequence $x_0, x_1, x_2, \ldots$ satisfying $x_i > 0$ for $i \geq 1$ and $P_i/Q_i = [x_0, x_1, \ldots x_i]$ for all i with $P_0 = x_0$ and $Q_0 = 1$. Setting $x_i = a_i$ for $0 \leq i \leq n$ and $x_{n+1} = \theta_{n+1}$ we obtain

$$\theta = [a_0, a_1, \ldots, a_n, \theta_{n+1}] = \frac{P_{n+1}}{Q_{n+1}} = \frac{\theta_{n+1}p_n + p_{n-1}}{\theta_{n+1}q_n + q_{n-1}}.$$

Substituting this into $q_n\theta - p_n$ and simplifying the numerator using

$$|p_nq_{n-1} - p_{n-1}q_n| = 1$$

we obtain

$$|q_n\theta - p_n| = \frac{1}{\theta_{n+1}q_n + q_{n-1}}.$$

Given that

$$\theta_{n+1}q_n + q_{n-1} > q_n + q_{n-1} = (a_n+1)q_{n-1} + q_{n-2} > \theta_n q_{n-1} + q_{n-2},$$

this implies

$$|q_{n+1}\theta - p_{n+1}| < |q_n\theta - p_n|,$$

hence

$$\left|\theta - \frac{p_{n+1}}{q_{n+1}}\right| < \frac{q_n}{q_{n+1}}\left|\theta - \frac{p_n}{q_n}\right|.$$

2. Given $a_{n+1} = [\theta_{n+1}]$ and $q_{n-1} < q_n$, we have

$$a_{n+1}q_n < \theta_{n+1}q_n + q_{n-1} < (a_{n+1}+1)q_n + q_n = (a_{n+1}+2)q_n.$$

This yields

$$\frac{1}{(a_{n+1}+2)q_n^2} < \left|\theta - \frac{p_n}{q_n}\right| < \frac{1}{a_{n+1}q_n^2},$$

which provides an alternative proof of Dirichlet's theorem.

3. Infinitely many convergents $\frac{p_n}{q_n}$ satisfy $|\theta - \frac{p_n}{q_n}| < \frac{1}{2q_n^2}$. To see this, use the fact that $\frac{p_{2n+1}}{q_{2n+1}} > \theta > \frac{p_{2n}}{q_{2n}}$ and $\left|\frac{p_n}{q_n} - \frac{p_{n-1}}{q_{n-1}}\right| = \frac{1}{q_{n-1}q_n}$ to obtain

$$\left|\theta - \frac{p_{2n}}{q_{2n}}\right| + \left|\theta - \frac{p_{2n+1}}{q_{2n+1}}\right| = \left|\frac{p_{2n+1}}{q_{2n+1}} - \frac{p_{2n}}{q_{2n}}\right| = \frac{1}{q_{2n}q_{2n+1}} < \frac{1}{2q_{2n}^2} + \frac{1}{2q_{2n+1}^2}.$$

So either

$$\left|\theta - \frac{p_{2n}}{q_{2n}}\right| < \frac{1}{2q_{2n}^2}$$

or

$$\left|\theta - \frac{p_{2n+1}}{q_{2n+1}}\right| < \frac{1}{2q_{2n+1}^2}.$$

4. Every rational number p/q satisfying $|\theta - p/q| < 1/2q^2$ is a convergent to θ: We must have $q_n \le q < q_{n+1}$ for some n. Given that the matrix $\begin{bmatrix} p_{n+1} & p_n \\ q_{n+1} & q_n \end{bmatrix}$ has determinant $(-1)^n$, we can find integers u and v such that

$$\begin{bmatrix} p \\ q \end{bmatrix} = \begin{bmatrix} p_{n+1} & p_n \\ q_{n+1} & q_n \end{bmatrix} \begin{bmatrix} u \\ v \end{bmatrix}.$$

Since

$$q_n \le uq_{n+1} + vq_n < q_{n+1},$$

$v \ne 0$ and u and v have opposite signs. Given that $q_{n+1}\theta - p_{n+1}$ and $q_n\theta - p_n$ are also of opposite signs, we obtain

$$|q\theta - p| = |u(q_{n+1}\theta - p_{n+1}) + v(q_n\theta - p_n)| = |u||q_{n+1}\theta - p_{n+1}| + |v||q_n\theta - p_n| \ge |q_n\theta - p_n|.$$

Therefore

$$\left|\frac{pq_n - qp_n}{qq_n}\right| = \left|\frac{p}{q} - \frac{p_n}{q_n}\right| \le \left|\theta - \frac{p}{q}\right| + \left|\theta - \frac{p_n}{q_n}\right| =$$

$$\frac{1}{q}|q\theta - p| + \frac{1}{q_n}|q_n\theta - p_n| \le \left(\frac{1}{q} + \frac{1}{q_n}\right)|q\theta - p| < \frac{q + q_n}{qq_n}\frac{1}{2q} \le \frac{1}{qq_n} \le 1.$$

This forces $pq_n - qp_n = 0$, $p/q = p_n/q_n$.

5. We can actually show that infinitely many convergents satisfy $|\theta - p_n/q_n| < 1/\sqrt{5}q^2$. Suppose that there are three consecutive convergents

$$p_n/q_n,\ p_{n+1}/q_{n+1},\ p_{n+2}/q_{n+2}$$

that satisfy $|\theta - p/q| \ge c/q^2$. We will show that $c < 1/\sqrt{5}$. Adding them in pairs as #3 above we obtain

$$\frac{c}{q_n^2} + \frac{c}{q_{n+1}^2} \le \left|\theta - \frac{p_n}{q_n}\right| + \left|\theta - \frac{p_{n+1}}{q_{n+1}}\right| \le \frac{1}{q_nq_{n+1}}$$

and

$$\frac{c}{q_{n+1}^2} + \frac{c}{q_{n+2}^2} \le \left|\theta - \frac{p_{n+1}}{q_{n+1}}\right| + \left|\theta - \frac{p_{n+2}}{q_{n+2}}\right| \le \frac{1}{q_{n+1}q_{n+2}}.$$

Rearranging, we obtain

$$\lambda + \frac{1}{\lambda} \le \frac{1}{c}$$

and

$$\mu + \frac{1}{\mu} \le \frac{1}{c}$$

where $\lambda = q_{n+1}/q_n$ and $\mu = q_{n+2}/q_{n+1}$. This implies

$$\lambda \le x, \qquad \mu \le x,$$

where

$$x = \frac{1 + \sqrt{1 - 4c^2}}{2c}.$$

Given that $q_{n+2} = a_{n+2}q_{n+1} + q_n$ we have $\mu = a_{n+2} + 1/\lambda \geq 1 + 1/\lambda$. This yields

$$1 + \frac{1}{x} \leq 1 + \frac{1}{\lambda} \leq \mu \leq x,$$
$$x + 1 \leq x^2.$$

This forces

$$x \geq \frac{1+\sqrt{5}}{2}.$$

In fact, the inequality is strict because equality implies μ is irrational. Given that x satisfies $cx^2 - x + c = 0$, we have

$$c = \frac{x}{x^2+1}.$$

This is a decreasing function for $x \geq 1$, so

$$c < \frac{\frac{1+\sqrt{5}}{2}}{\left(\frac{1+\sqrt{5}}{2}\right)^2 + 1} = \frac{1}{\sqrt{5}}.$$

In summary, when $c \geq \frac{1}{\sqrt{5}}$, at least one of three consecutive convergents always satisfies $|\theta - p/q| < c/q^2$, hence infinitely many of them do. This is best possible: when $c < \frac{1}{\sqrt{5}}$ and $\theta = \frac{1+\sqrt{5}}{2}$, only finitely many convergents p_k/q_k satisfy $|\theta - p/q| < c/q^2$ and none of the non-convergents do. This is Hurwitz's Theorem, proved via Liouville's Theorem (Section 6.5).

Section 6.4: Quadratic Irrationals

A quadratic irrational is an irrational solution to $ax^2 + bx + c = 0$ where a, b, c are integers.

Theorem: Let θ be an irrational number, and let $a_0, a_1, \ldots$ be the corresponding sequence of partial quotients. Then θ is a quadratic irrational if and only if its partial quotients are ultimately periodic, i.e. there exists $m \geq 1$ and N such that $n \geq N$ implies $a_n = a_{n+m} = a_{n+2m} = \cdots$.

Proof: First suppose that the partial quotients are purely periodic, i.e. $a_n = a_{n+m} = a_{n+2m} = \cdots$ for all $N \geq 0$. We can assume without loss of generality that $m \geq 2$. Given that $\theta_m = a_0 + 1/\theta_{m+1}$, $\theta_{m+1} = a_1 + 1/\theta_1$, ..., θ_m has the same convergents as θ, hence is equal to θ. This implies

$$\theta = [a_0, a_1, \ldots, a_{m-1}, \theta],$$

which implies

$$\theta = \frac{\theta p_{m-1} + p_{m-2}}{\theta q_{m-1} + q_{m-2}},$$

hence θ is the root of a quadratic equation.

More generally, assume that $a_n = a_{n+m} = \cdots$ for $n \geq N$. We can assume $N \geq 2$. Then

$$\theta = [a_0, \ldots, a_{N-1}, \psi]$$

where ψ is purely periodic. Then

$$\theta = \frac{\psi p_{N-1} + p_{N-2}}{\psi q_{N-1} + q_{N-2}},$$

and since ψ has the form $r + \sqrt{s}$ where $r, s \in \mathbb{Q}$, so does θ.

Conversely, let θ be an irrational solution to $ax^2 + bx + c = 0$. Define $f(x, y) = ax^2 + bxy + cy^2$ and, for $n \geq 1$, the equivalent binary form $f_n(x, y) = a_n x^2 + b_n xy + c_n y^2$ where

$$f_n(v) = f(V_n v)$$

and

$$V_n = \begin{bmatrix} p_n & p_{n-1} \\ q_n & q_{n-1} \end{bmatrix}.$$

Given

$$\theta = \frac{\theta_{n+1} p_n + p_{n-1}}{\theta_{n+1} q_n + q_{n-1}}$$

and

$$f_n(x, y) = f(p_n x + p_{n-1} y, q_n x + q_{n-1} y),$$

we have

$$f_n(\theta_{n+1}, 1) = f(\theta(q_n \theta_{n+1} + q_{n-1}), \theta_{n+1} q_n + q_{n-1}) =$$
$$(\theta_{n+1} q_n + q_{n-1})^2 f(\theta, 1) = 0.$$

Hence θ_{n+1} is a root of $a_n x^2 + b_n x + c_n = 0$. If we can show that there finitely many triples (a_n, b_n, c_n) then there must be a finite number of possibilities for θ_n. So at some point we have $\theta_{N+m} = \theta_N$, which implies ultimate periodicity in the sequence $a_0, a_1, a_2, \ldots$ by virtue of its definition.

We have $a_n = f_n(1, 0) = f(p_n, q_n)$, $c_n = f_n(0, 1) = f(p_{n-1}, q_{n-1}) = a_{n-1}$, and $f(\theta, 1) = 0$. We also have

$$a_n = f(p_n, q_n) = q_n^2 f(p_n/q_n, 1) = q_n^2 (f(p_n/q_n, 1) - f(\theta, 1)) =$$

$$q_n^2(a((p_n/q_n)^2 - \theta^2) + b((p_n/q_n) - \theta)).$$

The inequality

$$|\theta - p_n/q_n| < 1/q_n^2$$

implies

$$q_n^2|\theta^2 - p_n^2/q_n^2| = |q_n\theta - p_n||q_n\theta + p_n| < |\theta + p_n/q_n| \le 3|\theta|$$

for sufficiently large n. Hence $|a_n| \le M$ for some M. In other words, there are finitely many values for a_n. Since c_n and b_n are determined by a_n and the common discriminant d, there are a finite number of triples (a_n, b_n, c_n).

We now characterize the purely periodic quadratic irrationals in terms of continued fractions. If θ is purely periodic then it satisfies

$$\theta = [a_0, a_1, \ldots, a_{m-1}, \theta]$$

for some m, therefore

$$\theta = \frac{\theta p_{m-1} + p_{m-2}}{\theta q_{m-1} + q_{m-2}},$$

therefore θ is a root of $f(x) = q_{m-1}x^2 + (q_{m-2} - p_{m-1})x - p_{m-2}$. We have $\theta = a_0 + 1/\theta_1 > 1$. The other root θ' lies between -1 and 0 by the intermediate value theorem since $f(-1) = q_{m-1} - q_{m-2} + p_{m-1} - p_{m-2} > 0$ and $f(0) = -p_{m-2} < 0$.

Conversely, let θ be a quadratic irrational that satisfies $\theta > 1$ and $-1 < \theta' < 0$, where θ' denotes the other root of the quadratic that θ satisfies. Setting $\theta_0 = \theta$ we have $-1 < (\theta_0)' < 0$. Now assume $-1 < (\theta_n)' < 0$. Then $\theta_n = a_n + 1/\theta_{n+1}$, hence $(\theta_n)' = a_n + 1/(\theta_{n+1})'$, hence $-1 < a_n + 1/(\theta_{n+1})' < 0$, hence

$$-\frac{1}{a_n} < (\theta_{n+1})' < -\frac{1}{1 + a_n}.$$

Therefore $-1 < (\theta_n)' < 0$ for all n. We also have

$$a_n - (\theta_n)' = -\frac{1}{(\theta_{n+1})'},$$

hence, after computing the floor of each expression,

$$a_n = \left[-\frac{1}{\theta'_{n+1}}\right]$$

for each n. Since θ is ultimately periodic with some period $m \geq 1$, there is a minimum value of n such that $a_k = a_{k+m}$ for all $k \geq n$. If $n \geq 1$ we have $\theta_n = \theta_{n+m}$, therefore $(\theta_n)' = (\theta_{n+p})'$, therefore $a_{n-1} = a_{n+p-1}$. Contradiction. Therefore $n = 0$ and θ is purely periodic.

Now consider $\theta = \sqrt{d} + [\sqrt{d}]$ where d is not a perfect square. Then $\theta' = -\sqrt{d} + [\sqrt{d}]$, therefore $-1 < \theta < 0$, therefore θ is purely periodic. If

$$\sqrt{d} + [\sqrt{d}] = [\overline{a_0, a_1, \ldots, a_{p-1}}]$$

then

$$\sqrt{d} = [a_0 - [\sqrt{d}], \overline{a_1, \ldots, a_p}].$$

For example, if $\theta = \sqrt{2} + 1$ then

$$\theta = 2 + \frac{1}{\theta_1}$$

$$\theta_1 = \frac{1}{\sqrt{2} - 1} = \sqrt{2} + 1 = \theta$$

hence

$$\sqrt{2} + 1 = [2, 2, 2, \ldots]$$

and

$$\sqrt{2} = [1, 2, 2 \ldots].$$

Section 6.5: Liouville's Theorem

We know that when θ is irrational and $|\theta - p/q| < 1/2q^2$ for some rational number p/q, $p/q = p_n/q_n$ for some n. Hence if p/q is not a convergent then $|\theta - p/q| \geq 1/2q^2$. Moreover, we proved earlier that

$$\left|\theta - \frac{p_n}{q_n}\right| > \frac{1}{(a_n + 2)q_n^2}$$

for all n. When θ is a quadratic irrational the sequence of partial quotients $a_0, a_1, \ldots$ is bounded, so a number $c > 0$ can be found so that $|\theta - p_n/q_n| \geq c/q_n^2$ for all convergents p_n/q_n. In summary, a quadratic irrational θ satisfies $|\theta - p/q| > c/q^2$ for all rational p/q and some $c > 0$. Quadratic irrational numbers fall into a class of numbers called algebraic numbers, and Liouville's theorem states that for any algebraic number α with minimal polynomial of degree $n > 1$ there exists a number sufficiently small real number $c > 0$ such that $|\alpha - p/q| > c/q^n$ for all rationals p/q. We will define algebraic number carefully, then prove Liouville's theorem.

A real or complex number α is said to be algebraic if it is a zero of a non-zero polynomial $P(x)$ with integer coefficients. We can always assume that the coefficients of $P(x)$ do not have a common divisor. If $P(x)$ and $Q(x)$ are reduced polynomials of least degree n such that $P(\alpha) = Q(\alpha) = 0$ then for any integers a and b we see that α is a root of $aP(x) - bQ(x)$. We can choose a coprime pair a and b so that $aP(x) - bQ(x)$ has smaller degree than n, which forces $aP(x) = bQ(x)$. If p is a prime dividing b then, since p cannot divide all the coefficients of $P(x)$, p divides a. Since $(a, b) = 1$, this forces $|b| = 1$, and similarly $|a| = 1$. If we further assume that $P(x)$ and $Q(x)$ have positive leading coefficient then we must have $P(x) = Q(x)$. In other words, there is a unique reduced polynomial $P(x)$ of minimal degree and positive leading coefficient such that $P(\alpha) = 0$, and we call $P(x)$ the minimal polynomial of α. For example, the quadratic irrational rational $\sqrt{2}$ has minimal polynomial $P(x) = x^2 - 2$ and we can see that $\sqrt{2}$ is an algebraic number. Note also that rational numbers p/q are algebraic with minimal polynomial $qx - p$, but the degree of the minimal polynomial in this case is 1.

Minimal polynomials are irreducible over the rationals: If $P(x)$ is the minimal polynomial of α then $P(x) = f(x)g(x)$ implies $f(\alpha)$ or $g(\alpha) = 0$. We can multiply $f(x)$ and $g(x)$ by a suitable integers to obtain reduced polynomials $F(x)$ and $G(x)$ with positive leading term satisfying $F(\alpha) = 0$ or $G(\alpha) = 0$, and minimality of $P(x)$ implies that $P(x) = F(x)$ or $P(x) = G(x)$. Hence $f(x)$ or $g(x)$ is a scalar multiple of $P(x)$, which implies $P(x)$ is irreducible.

For the purposes of presenting this material rapidly in a lecture, we can say that α is algebraic if and only if it is the root of a non-zero polynomial with rational coefficients. The minimal polynomial $P(x)$ is the unique polynomial of minimal degree and leading coefficient 1 and must be irreducible in $\mathbb{Q}[x]$.

To prove Liouville's theorem, let α be a real algebraic number with minimal polynomial $P(x)$ of degree $n > 1$. Let $r = p/q$ be given where $q > 0$. Then by the mean-value theorem,

$$P(\alpha) - P(r) = (\alpha - r)P'(\xi_r)$$

for some $\xi_r \in (\alpha, r)$. We have $P(\alpha) = 0$ and, since $P(x)$ is irreducible of degree ≥ 2 over the rationals, $P(r) \neq 0$. This implies $P'(\xi_r) \neq 0$. We have

$$|\alpha - r| = \left| \frac{P(r)}{P'(\xi_r)} \right|.$$

Choosing a positive integer M so that $MP(x)$ has integer coefficients, $q^n MP(p/q)$ is a non-zero integer, hence $|P(r)| \geq 1/Mq^n$. Hence

$$|\alpha - p/q| \geq \frac{1}{Mq^n |P'(\xi_r)|}.$$

For $|\alpha - p/q| \leq 1$ have $|\xi_r| \leq |\xi_r - \alpha| + |\alpha| \leq |r - \alpha| + |\alpha| \leq 1 + |\alpha|$, hence we can find $C \geq 1$ such that $|P'(\xi_r)| \leq C$, which implies

$$|\alpha - p/q| \geq \frac{1}{MCq^n}.$$

The latter inequality is also satisfied when $|\alpha - p/q| \geq 1$.

For example, consider $\alpha = \frac{1+\sqrt{5}}{2}$. Its minimal polynomial is $P(x) = x^2 - x - 1$. We have

$$|\alpha - p/q| \geq \frac{1}{q^2 |2\xi_r - 1|}$$

for some ξ between α and p/q. As $p/q \to \alpha$, $\xi_r \to \alpha$, hence $|2\xi_r - 1| \to \sqrt{5}$. When $|\alpha - p/q| < c'/q^2$ for some $c' < 1/\sqrt{5}$ we know that $p/q = p_k/q_k$ for some k, which implies

$$c'/q_k^2 > |\alpha - p_k/q_k| \geq \frac{1}{q_k^2 |2\xi_k - 1|},$$

which implies

$$c' > \frac{1}{|2\xi_k - 1|},$$

which implies

$$|2\xi_k - 1| > 1/c' > \sqrt{5},$$

which can only happen for a finite number of k. So for all rational numbers except a finite number of convergents of the form p_k/q_k, $|\alpha - p/q| \geq c'/q^2$ when $c' < 1/\sqrt{5}$. This proves Hurwitz's theorem.

Liouville's theorem says that algebraic numbers α of degree $d \geq 2$ are separated from rational numbers in the sense that $q^d|\alpha - p/q| \geq c > 0$ for every rational p/q. So if α is a real number for which there exists a sequence $p_1/q_1, p_2/q_2, \ldots$ such that $q_n^n|\alpha - p_n/q_n| \to 0$ as $n \to \infty$ then for any $d \geq 2$ we have (eventually)

$$q_n^d|\alpha - p_n/q_n| \leq q_n^n|\alpha - p_n/q_n| \to 0,$$

hence α is not algebraic of any degree d (i.e. transcendental). We can replace the expression q_n^n by $f(n)$ where $q_n^d \leq f(n)$ for any $d \geq 2$ and sufficiently large n, for example $f(n) = q_n^{g(n)}$ where $g(n) \to \infty$.

Example: Let $r \in (0, 1)$ be a rational number. Set $\theta = \sum_{k=0}^{\infty} r^{\psi(k)}$. We will choose r and $\psi(k) \in \mathbb{Z}$ so that θ is a transcendental convergent infinite series. The partial sums $s_n = \sum_{k=0}^{n} r^{\psi(k)}$ are rational and we set $p_n/q_n = s_n$. We have

$$\theta - p_n/q_n = \sum_{k=n+1}^{\infty} r^{\psi(k)} = r^{\psi(n+1)} \sum_{k=n+1}^{\infty} r^{\psi(k)-\psi(n+1)}.$$

Assuming $\psi(n+i) - \psi(n+1) \geq i$ for all $i \geq 1$ we have

$$\left(\frac{1}{r}\right)^{\psi(n+1)} |\theta - p_n/q_n| = \sum_{k=n+1}^{\infty} r^{\psi(k)-\psi(n+1)} \to 0.$$

For example, if $r = 1/2$ and $\psi(n+1) = (n+1)!$ then we can set $q_n = 2^{n!}$ and $p_n = q_n s_n$ (verify that p_n is an integer), and we have $q_n^n < 2^{\psi(n+1)}$, therefore

$$q_n^n|\theta - p_n/q_n| \to 0.$$

This yields the transcendental number

$$\theta = \sum_{k=0}^{\infty} \frac{1}{2^{k!}}.$$

A proof that e is transcendental: The proof begins with the observation that for any differentiable function f, if we set $I(t, f) = e^t \int_0^t e^{-x} f(x)\, dx$, then integration by parts yields

$$I(t, f) = e^t f(0) - f(t) + I(t, f').$$

When $f(x)$ is a polynomial, this yields

$$I(t,f) = e^t \sum_{j\geq 0} f^{(j)}(0) - \sum_{j\geq 0} f^{(j)}(t)$$

where the index j is bounded above by the degree of $f(x)$. Now suppose e is algebraic. Then there exist integers $a_0, a_1, \ldots, a_n$, coefficients of the minimal polynomial of e, that satisfy

$$a_0 + a_1 e + \cdots + a_n e_n = 0.$$

This yields

$$a_0 I(0,f) + a_1 I(1,f) + \cdots + a_n I(n,f) = -\sum_{k=0}^{n} \sum_{j\geq 0} a_k f^{(j)}(k).$$

The right-hand side can be evaluated given information about the coefficients of $f(x)$, and the left-hand side can be approximated using properties of the definite integral. The idea is to choose $f(x)$ to yield a contradiction.

Details: consider the polynomial

$$f(x) = x^{p-1}(x-1)^p \cdots (x-n)^p$$

where $p > n$ is prime. We claim that all the expressions $f^{(j)}(k)$ are divisible by $p!$ for $j \geq 0$ and $0 \leq k \leq n$ except $f^{(p-1)}(0)$, and the latter is divisible by $(p-1)!$ and not p. To see this, note that for $1 \leq k \leq n$ the polynomial $f(x+k)$ is divisible by x^p. Since the coefficient of x^j in $f(x+k)$ is $\frac{f^{(j)}(k)}{j!}$, $f^{(j)}(k) = 0$ for $j < p$ and $f^{(j)}(k)$ is a multiple of $j!$ for $j \geq p$. Since the coefficient of x^j in $f(x)$ is $\frac{f^{(j)}(0)}{j!}$, $f^{(j)}(0) = 0$ for $j < p-1$ and $f^{(j)}(0)$ is a multiple of $j!$ for $j > p-1$. The coefficient of x^{p-1} in $f(x)$ is $(-1)^{np}(n!)^p$, hence $f^{(p-1)}(0) = (p-1)!(-1)^{np}(n!)^p$ is a multiple of $(p-1)!$ that is not divisible by p.

Let $\overline{f}(x)$ denote the polynomial

$$\overline{f}(x) = x^{p-1}(x+1)^p \cdots (x+n)^p.$$

Then $\overline{f}(k) \leq (2n)^{2np}$ for each $0 \leq k \leq n$.

We return to the identity

$$a_0 I(0,f) + a_1 I(1,f) + \cdots + a_n I(n,f) = -\sum_{k=0}^{n} \sum_{j \geq 0} a_k f^{(j)}(k).$$

Since $a_0 \neq 0$, and given the information we have about $f(x)$ above, the right-hand side in this identity is a non-zero multiple of $(p-1)!$. This yields

$$|a_0||I(0,f)| + |a_1||I(1,f)| + \cdots + |a_n||I(n,f)| \geq (p-1)!.$$

On the other hand,

$$|I(t,f)| = \left| e^t \int_0^t e^{-x} f(x)\ dx \right| \leq e^t \int_0^t e^{-x} \overline{f}(x)\ dx.$$

Since

$$\int_0^t e^{-x} x^j\ dx \leq \int_0^t x^j \leq t^{j+1},$$

we have

$$|I(t,f)| \leq t e^t \overline{f}(t).$$

Combined with the inequalities above this implies

$$(|a_1|e + 2|a_2|e^2 + \cdots + n|a_n|e^n)(2n)^{2np} \geq (p-1)!.$$

So there exist integers $a, k > 0$ such that

$$ak^p \geq (p-1)!$$

for all primes $p > n$. This is impossible: for $p \geq k+2$,

$$(p-1)! = k!(k+1)\cdots(p-1) \geq k!(k+1)^{p-k-1},$$

$$a \geq \frac{(p-1)!}{k^p} \geq \frac{k!}{(k+1)^{k+1}} \left(\frac{k+1}{k} \right)^p,$$

and $\left(\frac{k+1}{k}\right)^p \to \infty$ as $p \to \infty$.

Section 6.7: Minkowski's Theorem

Given a bounded region $X \subseteq \mathbb{R}^n$ we define

$$\mathrm{vol}(X) = \int_X 1\ dx_1 dx_2 \cdots dx_n.$$

Given an $n \times n$ matrix A, we have

$$\mathrm{vol}(AX) = |\det(A)|\mathrm{vol}(X)$$

by the change-of-variables theorem.

Let $a_1, a_2, \ldots, a_n$ be a basis for $\mathbb{R}^n$. These vectors define an integer lattice Λ via

$$\Lambda = \mathrm{span}_{\mathbb{Z}}(a_1, a_2, \ldots, a_n).$$

In other words, Λ consists of all vectors of the form Au where A is the matrix whose columns are $a_1, a_2, \ldots, a_n$ and u is a column vector with integer coordinates. The parameter $d(\Lambda)$ denotes the determinant of A.

A symmetric convex body $S \subseteq \mathbb{R}^n$ is an open, bounded set that satisfies three properties: $0 \in S$, $x \in S$ implies $-x \in S$, and $x, y \in S$ and $0 < t < 1$ implies $tx + (1-t)y \in S$.

Minkowski's theorem states that if $\mathrm{vol}(S) > 2^n d(\Lambda)$ then S contains a non-zero point in Λ. For the proof, set $S_0 = A^{-1}S$. Then S_0 is a convex body with $\mathrm{vol}(S_0) > 2^n$, and it suffices to find a non-zero point integral point in S_0. Let $S_1 = \frac{1}{2}S_0$. Then S_1 is a symmetric convex body with $\mathrm{vol}(S_1) > 1$. A partition of S_1 is $\bigcup_u S_1(u)$, where

$$S_1(u) = \{x \in S_1 : u_i \leq x_i < u_{i+1} \text{ for } 1 \leq i \leq n\}.$$

Note that $S_1(u) - u$ has the same volume as $S_1(u)$ for each u and $S_1(u) - u \subseteq [0,1]^n$. Since

$$\sum_u \mathrm{vol}(S_1(u) - u) > 1$$

and

$$\mathrm{vol}([0,1]^n) = 1,$$

there must be distinct integral points u, v such that

$$(S_1(u) - u) \cap (S_1(v) - v) \neq \emptyset.$$

Let z be an element in the intersection. Then $z = x - u$ for some $x \in S_1(u)$ and $z = y - v$ for some $y \in S_1(v)$. So we have $z = \frac{1}{2}X - u$ and $z = \frac{1}{2}Y - v$ for some $X, Y \in S_0$. Writing $Y = -Y'$ we have $Y' \in S_0$. So now we have

$$\frac{1}{2}X - u = z = -\frac{1}{2}Y' - v,$$

$$\frac{1}{2}X + \frac{1}{2}Y' = u - v.$$

By convexity, $\frac{1}{2}X + \frac{1}{2}Y' \in S_0$, and this is a non-trivial integral point in $\mathbb{R}^n$.

In summary, if a convex body S is symmetric about the origin and has volume greater than 2^n times the determinant defining a lattice, then it contains a non-trivial point in the lattice. In particular, if $\lambda_1 \cdots \lambda_n > \det(A)$ then for $m = 1, 2, 3, \ldots$ there is a non-zero $u_m \in \mathbb{Z}^n$ such that

$$Au \in (-\lambda_1, \lambda_1) \times \cdots \times (-\lambda_n - 1/m, \lambda_n + 1/m).$$

Inspecting the sequence $u_1, u_2, u_3, \ldots$ we see that there are only a finite number of distinct terms, so one of them, call it u, satisfies

$$Au \in (-\lambda_1, \lambda_1) \times \cdots \times [-\lambda_n, \lambda_n].$$

Example 1: Let $\theta_1, \ldots, \theta_n$ be real numbers and define

$$A = \begin{bmatrix} 1 & 0 & \cdots & 0 & 0 \\ 0 & 1 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \cdots & 1 & 0 \\ \theta_1 & \theta_2 & \cdots & \theta_n & -1 \end{bmatrix},$$

$\lambda_1 = Q, \ldots, \lambda_n = Q, \lambda_{n+1} = Q^{-n}$. Then there exist integers $q_1, q_2, \ldots, q_n, p$, not all zero (the coordinates of u), such that $|q_1|, \ldots, |q_n| < Q$ and

$$|q_1\theta_1 + \cdots + q_n\theta_n - p| \leq Q^{-n}.$$

Example 2: Let $\theta_1, \ldots, \theta_n$ be real numbers and define

$$A = \begin{bmatrix} -1 & 0 & \cdots & 0 & \theta_1 \\ 0 & -1 & \cdots & 0 & \theta_2 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \cdots & -1 & \theta_n \\ 0 & 0 & \cdots & 0 & 1 \end{bmatrix},$$

$\lambda_1 = 1/Q, \ldots, \lambda_n = 1/Q, \lambda_{n+1} = Q^n$. Then there exist integers $p_1, p_2, \ldots, p_n, q$, not all zero (the coordinates of u), such that

$$|q\theta_1 - p_1|, |q\theta_2 - p_2|, \ldots, |q\theta_n - p_n| < 1/Q$$

and $|q| \leq Q^n$. For further results, see Cassels' *An Introduction to the Geometry of Numbers*.

Chapter 6 Exercises

1. Write $\theta = [1, 2, 3, \overline{1, 4}]$. Then θ is a quadratic irrational. We first determine $\phi = [\overline{1, 4}]$. We have

$$\phi = [1, 4, \phi].$$

Writing the convergents to $[a_0, a_1, a_2]$ we have

$$\begin{bmatrix} p_2 & p_1 \\ q_2 & q_1 \end{bmatrix} = \begin{bmatrix} a_0 & 1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} a_1 & 1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} a_2 & 1 \\ 1 & 0 \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} 4 & 1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} \phi & 1 \\ 1 & 0 \end{bmatrix} = \begin{bmatrix} 1+5\phi & 5 \\ 1+4\phi & 4 \end{bmatrix}.$$

Hence

$$\phi = \frac{p_2}{q_2} = \frac{1+5\phi}{1+4\phi},$$

$$\phi \in \{\frac{1}{2}\left(1 - \sqrt{2}\right), \frac{1}{2}\left(1 + \sqrt{2}\right)\}.$$

Since $a_0 = 1$ we must have

$$\phi = \frac{1}{2}\left(1 + \sqrt{2}\right).$$

Now we can write $\theta = [1, 2, 3, \phi]$. Writing the convergents to $[a_0, a_1, a_2, a_3]$ we have

$$\begin{bmatrix} p_3 & p_2 \\ q_3 & q_2 \end{bmatrix} = \begin{bmatrix} a_0 & 1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} a_1 & 1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} a_2 & 1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} a_3 & 1 \\ 1 & 0 \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} 2 & 1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} 3 & 1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} \phi & 1 \\ 1 & 0 \end{bmatrix} =$$

$$\begin{bmatrix} 3+10\phi & 10 \\ 2+7\phi & 7 \end{bmatrix}.$$

Hence

$$\theta = \frac{p_3}{q_3} = \frac{3+10\phi}{2+7\phi} = -\frac{2}{23}\left(-18 + \sqrt{2}\right).$$

Mathematica yields

$$\text{ContinuedFraction}\left[-\frac{2}{23}\left(-18+\sqrt{2}\right),10\right] = \{1,2,3,1,4,1,4,1,4,1\}.$$

(ii) Using the program myContinuedFraction which I wrote in Mathematica, the continued fraction representing 3.1415926 is $[3,7,15,1,243,\cdots]$. Applying myConvergents to this yields the list $\{3, 22/7, 333/106, 355/113, 86598/27565, \ldots\}$. So $|3.1415926 - 355/113| < \frac{1}{243(113)^2}$. This implies

$$|\pi - 355/113| \leq |\pi - 3.1415926| + |3.1415926 - 355/113| < 10^{-7} + \frac{1}{243(113)^2} < 10^{-6}.$$

(iii) We have $\theta = \frac{a+\sqrt{a^2+4}}{2} = [a, a, \ldots]$, $\theta' = \frac{a-\sqrt{a^2+4}}{2}$. This yields the recurrence relation

$$q_0 = 1,\ q_1 = a,\ q_n = aq_{n-1} + q_{n-2}\ (n \geq 2),$$

the solution to which is

$$q_n = \alpha(\theta)^{n+1} + \beta(\theta')^{n+1}$$

for a suitable α and β. Using the initial conditions we obtain

$$\alpha = \frac{1}{\sqrt{a^2+4}} = \frac{1}{\theta-\theta'}, \beta = \frac{-1}{\sqrt{a^2+4}} = \frac{-1}{\theta-\theta'}.$$

We obtain the Fibonacci sequence when $a = 1$.

(iv) The recurrence relation in (iii) suggests that a floor for q_n is given by the Fibonacci numbers $F_0, F_1, F_2, \ldots$. So we must argue $F_n \geq \left(\frac{1+\sqrt{5}}{2}\right)^{n-1}$. This is true for $n = 0$ and $n = 1$. Assuming it true for F_0 through F_n, we have

$$F_{n+1} \geq \left(\frac{1+\sqrt{5}}{2}\right)^{n-1} + \left(\frac{1+\sqrt{5}}{2}\right)^{n-2} = \left(\frac{1+\sqrt{5}}{2}\right)^{n}$$

since the number $x = \frac{1+\sqrt{5}}{2}$ satisfies $x^{-1} + x^{-2} = 1$. On the other hand, if $a_n \leq a$ for all n then the recurrence relation in (iii) says that $q_n \leq Q_n$ where

Q_n is the solution to the recurrence relation in (iii). The latter sequence satisfies $Q_n \leq \left(\frac{a+\sqrt{a^2+4}}{2}\right)^n$ for $n = 0$ and $n = 1$. Assuming this is true for Q_0 through Q_n, we have

$$Q_{n+1} = aQ_n + Q_{n-1} \leq a\left(\frac{a+\sqrt{a^2+4}}{2}\right)^n + \left(\frac{a+\sqrt{a^2+4}}{2}\right)^{n-1} = \left(\frac{a+\sqrt{a^2+4}}{2}\right)^{n+1}$$

since $x = \frac{a+\sqrt{a^2+4}}{2}$ satisfies $ax + 1 = x^2$.

(v) We will first look at convergents. We have $|e - \frac{p_n}{q_n}| > \frac{1}{(a_n+2)q_n^2}$. We want to show $\frac{1}{a_n+2} > \frac{c}{\log q_n}$ for an appropriate $c > 0$. In other words, $q_n > e^{c(a_n+2)}$. This is clear because q_n grows exponentially by (iv) and a_n is bounded by a linear function. Now if some $|e - p/q| < c/q^2 \log q$ then it must be a convergent (choosing c sufficiently small), and this is not possible.

(vi) Thue-Siegel-Roth says that algebraic numbers α of degree $d \geq 2$ are separated from rational numbers in the sense that $q^\kappa|\alpha - p/q| \geq c(\alpha, \kappa) > 0$ for every rational p/q for any given $\kappa > 2$. So if α is a real number for which there exists a sequence $p_1/q_1, p_2/q_2, \ldots$ such that $f(n)|\alpha - p_n/q_n| \to 0$ as $n \to \infty$ where $\frac{q_n^\kappa}{f(n)} = O(1)$ for some $\kappa > 2$ then we have

$$q_n^\kappa|\alpha - p_n/q_n| = \frac{q_n^\kappa}{f(n)} f(n)|\alpha - p_n/q_n| \to 0,$$

hence α is transcendental.

Now set $\alpha = \sum_{k=1}^\infty \frac{1}{a^{b^k}}$ where $a \geq 2$ and $b \geq 3$ are integers. This is convergent by comparison with the geometric series. Set

$$\frac{p_n}{q_n} = \sum_{k=1}^n \frac{1}{a^{b^k}}.$$

We have

$$\alpha - \frac{p_n}{q_n} = \sum_{k=n+1}^\infty \frac{1}{a^{b^k}} = \frac{1}{a^{b^{n+1}}} \sum_{k=n+1}^\infty \frac{1}{a^{b^k - b^{n+1}}},$$

$$a^{b^{n+1}}|\alpha - \frac{p_n}{q_n}| = \sum_{k=n+1}^\infty \frac{1}{a^{b^k - b^{n+1}}} \to 0.$$

(The sum on the right approaches 0 as $n \to \infty$ by comparison with the tails of the geometric series.) Setting $q_n = a^{b^n}$ and $f(n) = a^{b^{n+1}}$ we have

$$\frac{q_n^b}{f(n)} = 1.$$

Hence α is transcendental using $\kappa = b$.

(vii) Minkoski's theorem says that if $\mathrm{vol}(S) > 4|\Delta|$ then S will contain a point of the lattice. Setting $S = \{(x, y) : |x| + |y| \leq \sqrt{2|\Delta|}\}$ we obtain a square with vertices at $(0, \pm\sqrt{2|\Delta|})$ and $(\pm\sqrt{2|\Delta|}, 0)$ with area $4|\Delta|$, so L and M can be found. The point $(|L|, |M|)$ lives in a rectangle inside the region bounded by the x-axis, the y-axis, and the line $y = \sqrt{2|\Delta|} - x$, hence $|LM|$ is the bounded above by the maximum inscribed area. The latter is generated by the square corresponding to the point $(\frac{\sqrt{2|\Delta|}}{2}, \frac{\sqrt{2|\Delta|}}{2})$, which has area $\frac{|\Delta|}{2}$.

(viii) Construct a counterexample using the parameters given.

(ix) I'd like to see a proof of Kronecker's Theorem first!

Chapter 7: Quadratic Fields

Vector space and field: Let d be a square-free integer other than 1. Then

$$\mathbb{Q}(\sqrt{d}) = \{u + v\sqrt{d} : u, v \in \mathbb{Q}\}.$$

This is a vector space over $\mathbb{Q}$ with basis $\{1, \sqrt{d}\}$. It is also a field: one can check closure with respect to addition and multiplication and the existence of additive inverses. Moreover, since $\sqrt{d}$ is irrational, $u + v\sqrt{d} \in \mathbb{Q}(\sqrt{d})^*$ implies $u^2 - v^2 d \neq 0$ implies $(u + v\sqrt{d})^{-1} = \frac{u}{u^2 - v^2 d} - \frac{v}{u^2 - v^2 d}\sqrt{d} \in \mathbb{Q}(\sqrt{d})$. Since $\mathbb{Q}(\sqrt{d})$ has dimension 2 over the rationals, every $\alpha \in \mathbb{Q}(\sqrt{d})$ its the root of a non-zero rational polynomial of degree 2, hence is algebraic.

Linear Operator and norm: For each $\alpha \in \mathbb{Q}(\sqrt{d})$ we obtain a linear operator $L_\alpha : \mathbb{Q}(\sqrt{d}) \to \mathbb{Q}(\sqrt{d})$ via $L_\alpha(\beta) = \alpha\beta$. Now set $\alpha = u + v\sqrt{d}$. Then $L_{u+v\sqrt{d}}(1) = u + v\sqrt{d}$ and $L_{u+v\sqrt{d}}(\sqrt{d}) = dv + u\sqrt{d}$, hence $L_{u+v\sqrt{d}}$ has matrix representation $\begin{bmatrix} u & dv \\ v & u \end{bmatrix}$. The norm of $u + v\sqrt{d}$ is the determinant of $L_{u+v\sqrt{d}}$, hence $N(u + v\sqrt{d}) = u^2 - dv^2$. Since $L_\alpha L_\beta = L_{\alpha\beta}$, $N(\alpha)N(\beta) = N(\alpha\beta)$. This yields the identity

$$(u_1^2 - dv_1^2)(u_2^2 - dv_2^2) = (u_1u_2 + v_1v_2d)^2 - d(u_1v_2 + u_2v_1)^2.$$

Note also that $N(\alpha) = \alpha\overline{\alpha}$ where $\overline{\alpha}$ is the conjugate of α.

Algebraic integer: A number whose minimal polynomial (rational polynomial of least degree with leading coefficient 1) has integer coefficients.

Degree 1 algebraic integer: minimal polynomial $x - k$ for some $k \in \mathbb{Z}$, hence ordinary integers.

Degree 2 algebraic integers: Let $x = u + v\sqrt{d}$ be an algebraic integer with $v \neq 0$. We can find a rational polynomial satisfied by x as follows: Using L_x as above we have

$$\begin{bmatrix} x & 0 \\ 0 & x \end{bmatrix} \begin{bmatrix} 1 \\ \sqrt{d} \end{bmatrix} = \begin{bmatrix} u & dv \\ v & u \end{bmatrix} \begin{bmatrix} 1 \\ \sqrt{d} \end{bmatrix}$$

$$\begin{bmatrix} x-u & -dv \\ -v & x-u \end{bmatrix} \begin{bmatrix} 1 \\ \sqrt{d} \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}.$$

Hence

$$x^2 - 2xu + (u^2 - dv^2) = \det \begin{bmatrix} x-u & -dv \\ -v & x-u \end{bmatrix} = 0.$$

Hence the minimal polynomial is $x^2 - 2ux + (u^2 - dv^2)$. Since x is algebraic, $2u \in \mathbb{Z}$ and $N(x) = u^2 - dv^2 \in \mathbb{Z}$. Conversely, any x having this minimal polynomial is an algebraic integer.

Characterization of degree 2 algebraic integers:

Write $2u = m$, $u^2 - dv^2 = n$, $v = p/q$ where $(p, q) = 1$. Then

$$u^2 - dv^2 = n$$

$$4u^2 - 4dv^2 = 4n$$
$$m^2 - 4dv^2 = 4n$$
$$q^2m^2 - 4dp^2 = 4nq^2$$
$$q^2(m^2 - 4n) = (2p)^2d.$$

Since d is square-free, we must have $q^2|(2p)^2$, hence $q|2p$, hence $q \in \{1, 2\}$. We will write $2v = k$. We now have $m^2 - dk^2 = 4n$, hence $m^2 \equiv dk^2 \bmod 4$. Bearing in mind that $m^2, k^2 \equiv 0, 1 \bmod 4$, consider the cases:

$d \equiv 0 \bmod 4$: Not possible since d is square-free.

$d \equiv 1 \bmod 4$: $m^2 \equiv k^2 \bmod 4$, hence m and k have the same parity. Writing $m = k + 2p$, we have

$$x = \frac{k+2p}{2} + \frac{k}{2}\sqrt{d} = p + k\frac{1+\sqrt{d}}{2}.$$

$d \equiv 2 \bmod 4$: $m^2 \equiv 2k^2 \bmod 4$, therefore m and k are even, u and v are integers, and

$$x = u + v\sqrt{d}.$$

$d \equiv 3 \bmod 4$: $m^2 \equiv 3k^2 \bmod 4$, therefore m and k are even, u and v are integers, and

$$x = u + v\sqrt{d}.$$

The ring R_d of algebraic integers in $\mathbb{Q}(\sqrt{d})$: We have found necessary conditions above for x to be an algebraic integer, but one can check that they are also sufficient, given the polynomial satisfied by x. Setting R_d equal to the set of algebraic integers $\mathbb{Q}(\sqrt{d})$, we have

$$R_d = \begin{cases} \mathbb{Z}[1, \frac{1+\sqrt{d}}{2}] & d \equiv 1 \bmod 4 \\ \mathbb{Z}[1, \sqrt{d}] & d \equiv 2, 3 \bmod 4. \end{cases}$$

In particular, R_d is a ring and is closed with respect to conjugation.

Every algebraic integer has an integer norm. To see this, let $\alpha \in R_d$ and write $\alpha = x + y\omega$ where

$$\omega = \begin{cases} \sqrt{d} & d \equiv 1 \bmod 4 \\ \frac{1+\sqrt{d}}{2} & d \equiv 2, 3 \bmod 4. \end{cases}$$

Then

$$N(\alpha) = \alpha\overline{\alpha} = \begin{cases} x^2 + xy + \frac{1-d}{4}y^2 & d \equiv 1 \bmod 4 \\ x^2 - dy^2 & d \equiv 2, 3 \bmod 4. \end{cases}$$

We can interpret $N(x + y\omega)$ as a binary quadratic form $f(x, y)$ with discriminant

$$d(f) = \begin{cases} d & d \equiv 1 \bmod 4 \\ 4d & d \equiv 2, 3 \bmod 4. \end{cases}$$

Units, primes, and irreducibles in R_d: We will define the units in R_d to be the set U_d of invertible elements in R_d. If $\epsilon \in U_d$ then $\epsilon\epsilon' = 1$ for some $\epsilon' \in U_d$, hence $1 = N(1) = N(\epsilon\epsilon') = N(\epsilon)N(\epsilon')$, hence $N(\epsilon) = \pm 1$. Conversely, if $\alpha \in R_d$ and $N(\alpha) = \pm 1$ then $\alpha\overline{\alpha} = \pm 1$ hence $\alpha^{-1} = \pm\overline{\alpha} \in R_d$. So we have proved

$$U_d = \{\alpha \in R_d : N(\alpha) = \pm 1\}.$$

In other words, the units in R_d are the numbers in R_d norm a unit in $\mathbb{Z}$.

We say that $a|b$ in R_d if $a = bc$ for some $c \in R_d$. Primes $\pi \in R_d$ are non-zero non-units that satisfy $\pi|ab \implies \pi|a$ or $\pi|b$. Irreducibles $\pi \in R_d$ are non-zero non-units that satisfy $\pi = ab \implies a$ or b is a unit.

Primes are irreducible: Let π be prime and suppose $\pi = ab$. Then $\pi|a$ or $\pi|b$. If $\pi|a$, write write $a = \pi a_0$. Then $1 = a_0 b$, therefore b is a unit. But if $\pi \not| a$ then $\pi|b$, therefore a is a unit.

Not all irreducibles are primes: Note that $a = bc$ implies $N(a) = N(b)N(c)$ and $a|b$ implies $N(a)|N(b)$. We can use these properties to show that 2 is irreducible but not prime in R_{-5}. Since $-5 \equiv 3 \bmod 4$ we have $R_{-5} = \mathbb{Z}[1, \sqrt{-5}]$. Suppose

$$2 = (x_1 + y_1\sqrt{-5})(x_2 + y_2\sqrt{-5}).$$

Take norms,

$$4 = (x_1^2 + 5y_1^2)(x_2^2 + 5y_2^2).$$

Since $x^2 + 5y^2 = 2$ has no solution, one of the two norms on the right is 1, hence one of the two factors is a unit. Therefore 2 is irreducible. On the other hand, $2|(1+\sqrt{-5})(1-\sqrt{-5})$ yet 2 is a divisor of neither factor since there is no algebraic integer $x + y\sqrt{-5} \in R_{-5}$ that satisfies $1 \pm \sqrt{-5} = 2(x + y\sqrt{2})$. Hence 2 is not prime.

Every non-zero non-unit α in R_d can be factored into irreducibles: by induction on $|N(\alpha)| \geq 2$. If $|N(\alpha)| = 2$, write $\alpha = \beta\gamma$. Taking norms, $N(\alpha) = N(\beta)N(\gamma)$ so $|N(\beta)| = 1$ or $|N(\gamma)| = 1$, therefore β or γ is a unit. Now consider $|N(\alpha)| > 2$. If α is not irreducible then there must be a way to factor it in the form $\alpha = \beta\gamma$ where neither factor is a unit. Taking norms we see that $1 < |N(\beta)|, |N(\gamma)| < |N(\alpha)|$, hence β and γ are products of irrreducibles, hence α is a product of irrreducibles.

Unique factorization into irreducibles in R_d: When all irreducibles are primes in R_d, we have unique factorization of non-units in R_d into irreducibles in the following sense: When $x_1 \cdots x_m = y_1 \cdots y_n$ in R_d with each x_i and y_j irreducible, then $m = n$ and there are units $u_1, \ldots, u_n$ and a permutation σ such that $y_i = u_i x_{\sigma(i)}$ for each i. We will prove this by induction on n. When $n = 1$ we have $x_1 \cdots x_m = y_1$. Since y_1 is irreducible, this forces $m = 1$ and $x_1 = y_1$. Assume the statement is true for a given n. Suppose $x_1 \cdots x_m = y_1 \cdots y_{n+1}$. Then $m > 1$. The product on the right is divisible by x_1, and since x_1 is prime it has to be a divisor of some y_j. Reordering if necessary, $j = 1$ and $y_1 = u_1 x_1$ for some unit u_1. Cancelling off x_1 we obtain $x_2 \cdots x_m = u_1 y_2 \cdots y_{n+1}$. Equivalently, $(v_1 x_2) \cdots x_m = y_2 \cdots y_{n+1}$ where $u_1 v_1 = 1$. We can use the induction hypothesis provided $v_1 x_2$ is irreducible. It is: If $v_1 x_2 = \alpha\beta$ then $x_2 = (u_1\alpha)\beta$, therefore $u_1\alpha$ is a unit or β is a unit, hence α is a unit or β is a unit.

By consideration of norms we can prove that $2, 3, 1 + \sqrt{-5}, 1 - \sqrt{-5}$ are irreducible in R_{-5} and neither number in $\{2, 3\}$ is an associate of either of the numbers in $\{1 + \sqrt{-5}, 1 - \sqrt{-5}\}$. On the other hand,

$$(2)(3) = (1 + \sqrt{-5})(1 - \sqrt{-5}).$$

Hence unique factorization into irreducibles fails in R_{-5}.

Units in R_d where $d < 0$: For $d \equiv 1 \bmod 4$, i.e. $d = 1 - 4k$, we have

$$N(x + y\omega) = x^2 + xy + ky^2.$$

This is a reduced binary quadratic form. When $k \geq 2$ the output 1 occurs only with $(x, y) = (\pm 1, 0)$, so the units in R_d are ± 1. When $k = 1$ the output 1 occurs only with $(x, y) = (\pm 1, 0), (0, \pm 1), \pm(1, -1)$, so the units in R_{-3} are ± 1 and $\pm\frac{1+\sqrt{-3}}{2}$ and $\pm\frac{1-\sqrt{-3}}{2}$.

For $d \equiv 2 \bmod 4$, i.e. $d = 2 - 4k$, we have

$$N(x + y\omega) = x^2 + (4k - 2)y^2.$$

This form is reduced for all $k \geq 1$ and the only units are ± 1. For $d \equiv 3 \bmod 4$, i.e. $d = 3 - 4k$, we have

$$N(x + y\omega) = x^2 + (4k - 3)y^2.$$

This form is reduced for all $k \geq 1$. For $k \geq 2$ the only units are ± 1. When $k = 1$ the units in R_{-1} are $\pm 1, \pm i$.

Summary:

R_{-1} has units $\pm 1, \pm i$, roots of x^2+1.

R_{-2} has units ± 1, roots of x^2-1.

R_{-3} has units $\pm 1, \pm\frac{1\pm\sqrt{3}i}{2}$, roots of x^6-1.

R_{-k} has units ± 1 for $k \geq 5$ and k square-free, roots of x^2-1.

Units in R_d where $d > 0$: Consider a square-free integer $d \geq 2$. (Actually, all we require is $\sqrt{d}$ irrational.) We can construct infinitely many units in $\mathbb{Q}[\sqrt{d}]$ as follows: $\sqrt{d}$ is irrational and its convergents p_n/q_n satisfy

$$|\sqrt{d} - p_n/q_n| < 1/q_n^2.$$

Hence

$$p_n - q_n\sqrt{d} = \cos(\theta_n)/q_n$$
$$p_n + q_n\sqrt{d} = \cos(\theta_n)/q_n + 2q_n\sqrt{d}$$
$$|N(p_n - q_n\sqrt{d})| = |\cos(\theta_n)^2/q_n^2 + 2\cos(\theta_n)\sqrt{d}| \leq 1 + 2\sqrt{d}.$$

Since the sequence of norms $N(p_n - q_n\sqrt{d})$ is bounded, there is an infinite subsequence of constant norm N. We can finitely partition this subsequence according to $([p_n]_N, [q_n]_N)$, congruence classes mod N, and one of the parts of this partition is infinite. Hence there exist infinitely many pairs $m < n$ such that $p_m \equiv p_n \bmod N$ and $q_m \equiv q_n \bmod N$ and $p_m^2 - dq_m^2 = p_n^2 - dq_n^2 = N$. Setting

$$\eta = \frac{p_m - q_m\sqrt{d}}{p_n - q_n\sqrt{d}} = \frac{p_m p_n - dq_m q_n}{N} + \frac{p_m q_n - p_n q_m}{N}\sqrt{d},$$

we have

$$p_m p_n - dq_m q_n \equiv p_n^2 - dq_n^2 \equiv 0 \bmod N$$

and

$$p_m q_n - p_n q_m \equiv 0 \bmod N.$$

Hence $\eta \in R_d$ and has norm 1. Note that $\eta = x - y\sqrt{d}$ satisfies $x^2 - dy^2 = 1$, so it is a solution to the Diophantine equation known as Pell's equation. This argument shows that there are an infinite number of solutions to this equation.

We can characterize the set of units U_d in R_d as follows: we first claim that there is a unit $\mu = x + y\sqrt{d} > 1$. If $\eta > 1$, use η. If $0 < \eta < 1$, use $1/\eta$. If $-1 < \eta < 0$, use $-1/\eta$. If $\eta < -1$, use $-\eta$. Secondly, we claim that any unit $x + y\sqrt{d} > 1$ satisfies $x, y > 0$. For if $N(x + y\sqrt{d}) = 1$ then $x + y\sqrt{d} > (x + y\sqrt{d})^{-1} = x - y\sqrt{d} > 0$, and if $N(x + y\sqrt{d}) = -1$ then $x + y\sqrt{d} > (x + y\sqrt{d})^{-1} = -x + y\sqrt{d} > 0$, and both statements yield $x, y > 0$. Therefore $U_d \cap (1, \infty)$ has a minimum element ϵ: let x_0 be minimum such that there exists at unit $x_0 + y_0\sqrt{d} > 1$. If $x_1 + y_1\sqrt{d} < x_0 + y_0\sqrt{d}$ in $U_d \cap (1, \infty)$ then $0 \le x_1 - x_0 < (y_0 - y_1)\sqrt{d}$, hence $y_1 < y_0$. This is satisfied by only finitely many values of y_1, hence by only finitely many values of $x_1 + y_1\sqrt{d}$ since the value of y_1 determines the value of x_1 uniquely. We can identify ϵ as the minimum element of $\{x_i + y_i\sqrt{d} : i \ge 0\}$. Every other unit can be expressed in terms of ϵ: for any other unit δ with $\delta > 1$ we have $\epsilon^n \le \delta < \epsilon^{n+1}$ for some n, hence $1 \le \delta/\epsilon^n < \epsilon$. Since δ/ϵ^n is a unit and ϵ is the smallest unit > 1, we must have $\delta/\epsilon^n = 1$, i.e. $\delta = \epsilon^n$. So the set of all units > 1 is $\{\epsilon^k : k \ge 1\}$, which implies that the set of all units is $\{\pm\epsilon^k : k \in \mathbb{Z}\}$ by the argument at the beginning of the pargraph.

Euclidean Fields: Certain quadratic fields, called Euclidean, are endowed with an analogue of the division algorithm: for each $\alpha, \beta \in R_d$ with $\beta \ne 0$ there exist $\delta, \rho \in R_d$ such that

$$\alpha = \delta\beta + \rho$$

with $|N(\rho)| < |N(\delta)|$. This gives rise to an analogue of Euclid's algorithm for constructing the greatest common divisor of α and $\beta \ne 0$: Form the sequence $\alpha_0, \alpha_1, \alpha_2, \dots$ with $|N(\alpha_1)| > |N(\alpha_2)| > \cdots \ge 0$ via $\alpha_0 = \alpha$, $\alpha_1 = \beta$, and for $k \ge 2$, $\alpha_{k-2} = \delta_{k-2}\alpha_{k-1} + \alpha_k$ where $0 \le |N(\alpha_k)| < |N(\alpha_{k-1})|$. The sequence has to terminate with some $\alpha_n = 0$ for some $n \ge 2$, and α_{n-1} is a greatest common divisor in the sense that $\alpha_{n-1}|\alpha$ and $\alpha_{n-1}|\beta$, and whenever $x|\alpha$ and $x|\beta$ we must have $x|\alpha_{n-1}$. All greatest common divisors divide each other, hence are associates of each other. To see that α_{n-1} is a greatest common divisor, observe that the recurrence relation can be expressed in the form

$$\begin{bmatrix} \alpha_{k-2} \\ \alpha_{k-1} \end{bmatrix} = \begin{bmatrix} \delta_{k-2} & 1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} \alpha_{k-1} \\ \alpha_k \end{bmatrix}.$$

This can be used to obtain

$$\begin{bmatrix} \delta_0 & 1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} \delta_1 & 1 \\ 1 & 0 \end{bmatrix} \cdots \begin{bmatrix} \delta_{n-2} & 1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} \alpha_{n-1} \\ 0 \end{bmatrix} = \begin{bmatrix} \alpha_0 \\ \alpha_1 \end{bmatrix}.$$

Simplifying,

$$\begin{bmatrix} x & y \\ z & w \end{bmatrix} \begin{bmatrix} \alpha_{n-1} \\ 0 \end{bmatrix} = \begin{bmatrix} \alpha \\ \beta \end{bmatrix}.$$

Hence

$$\begin{bmatrix} xa_{n-1} \\ za_{n-1} \end{bmatrix} = \begin{bmatrix} \alpha \\ \beta \end{bmatrix}.$$

So we can see that α_{n-1} is a common divisor of α and β. Moreover if δ is a divisor of both α and β then the recurrence relation can be used to show that δ divides each α_k, including α_{n-1}. Hence $\delta|\alpha_{n-1}$ and α_{n-1} is a greatest common divisor. All greatest common divisors are associates of each other.

Note that the inverse of $\begin{bmatrix} \delta_k & 1 \\ 1 & 0 \end{bmatrix}$ is $\begin{bmatrix} 0 & 1 \\ 1 & -\delta_k \end{bmatrix}$. This implies that

$$\begin{bmatrix} \alpha_{n-1} \\ 0 \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ 1 & -\delta_{n-2} \end{bmatrix} \begin{bmatrix} 0 & 1 \\ 1 & -\delta_{n-3} \end{bmatrix} \cdots \begin{bmatrix} 0 & 1 \\ 1 & -\delta_0 \end{bmatrix} \begin{bmatrix} \alpha_0 \\ \alpha_1 \end{bmatrix}.$$

Hence given algebraic integers α and β with greatest common divisor δ there is a pair of algebraic μ and ν such that $\mu\alpha + \nu\beta = \delta$. Whenever we have $\mu\alpha + \nu\beta = \rho$ we must have $\delta|\rho$. In particular, when $\mu\alpha + \nu\beta = 1$ we must have $\delta|1$, hence δ is a unit. Conversely when the greatest common divisor of α and β is a unit, i.e. α and β are coprime, there exist μ and ν such that $\mu\alpha + \nu\beta = 1$.

Euclidean quadratic fields have unique factorization: let π be an irreducible algebraic integer in a Euclidean quadratic field $\mathbb{Q}(\sqrt{d})$ and suppose $\pi|\alpha\beta$. We will show that $\pi|\alpha$ or $\pi|\beta$. Assume $\pi \not| \alpha$. We claim that π and α are coprime. To see this, suppose $\delta|\pi$ and $\delta|\alpha$. Write $\pi = \pi_0\delta$ and $\alpha = \alpha_0\delta$. If δ is not a unit then π_0 is a unit and $\alpha = \alpha_0\pi_0^{-1}\pi$ hence $\pi|\alpha$: contradiction. Therefore δ must be a unit, hence a divisor of 1. Given that π and α are coprime, there exist μ and ν such that $\mu\pi + \nu\alpha = 1$, which yields $\mu\pi\beta + \nu\alpha\beta = \beta$, and since $\pi|\alpha\beta$, $\pi|\beta$. Hence π is prime. We have shown that all irreducibles are primes, so there is unique factorization in R_d when it is Euclidean.

Our task now is to identify d such that R_d is Euclidean.

Necessary conditions for R_d to be Euclidean:

First consider $d \equiv 2, 3 \bmod 4$. Then $R_d = \mathbb{Z}[1, \sqrt{d}]$ and there exist integers x_1, y_1, x_2, y_2 such that

$$\sqrt{d} = 2(x_1 + y_1\sqrt{d}) + (x_2 + y_2\sqrt{d})$$

where $|x_2^2 - dy_2^2| < 4$. When $d \leq -5$ we have $x_2^2 + 5y_2^2 < 4$, hence $y_2 = 0$ and

$$\sqrt{d} = (2x_1 + x_2) + 2y_1\sqrt{d},$$

which is not possible because $2y_1 \neq 1$. Hence $d \geq -2$ is necessary.

Next consider $d \equiv 1 \bmod 4$. Then $R_d = \mathbb{Z}[1, \frac{1+\sqrt{d}}{2}]$, and there exist integers x_1, y_1, x_2, y_2 such that

$$\frac{1+\sqrt{d}}{2} = 2\left(x_1 + y_1\frac{1+\sqrt{d}}{2}\right) + \left(x_2 + y_2\frac{1+\sqrt{d}}{2}\right)$$

where $|(x_2 + y_2)^2 - dy_2^2/4| < 4$. When $d \leq -15$, $y_2 = 0$. This yields

$$\frac{1+\sqrt{d}}{2} = (2x_1 + x_2) + 2y_1\frac{1+\sqrt{d}}{2},$$

which forces $2y_1 = 1$, which is not possible. Hence $d \geq -11$ is necessary.

Sufficient conditions for R_d to be Euclidean:

Let $\alpha, \beta \in R_d$ with $\beta \neq 0$. Write $\alpha/\beta = u + v\sqrt{d}$ where $u, v \in \mathbb{Q}$. Consider the cases.

$d \equiv 2, 3 \bmod 4$: Algebraic integers are of the form $x + y\sqrt{d}$ where $x, y \in \mathbb{Z}$. Choosing x closest to u and y closest to v we have

$$\alpha = (x + y\sqrt{d})\beta + (r + s\sqrt{d})\beta$$

where $|r|, |s| \leq \frac{1}{2}$. Restricting d to $|d| \leq 3$ we have $d \in \{-2, -1, 2, 3\}$. When $|d| \leq 2$ we have

$$|r^2 - ds^2| \leq r^2 + 2s^2 \leq \frac{3}{4}.$$

When $d = 3$ we have

$$-3/4 \leq -3s^2 \leq r^2 - 3s^2 \leq r^2 \leq \frac{1}{4}.$$

In all cases

$$|N((r + s\sqrt{d})\beta)| \leq \frac{3}{4}|N(\beta)| < |N(\beta)|.$$

Conclusion:

$$\mathbb{Q}(\sqrt{-2}), \mathbb{Q}(\sqrt{-1}), \mathbb{Q}(\sqrt{2}), \mathbb{Q}(\sqrt{3})$$

are Euclidean.

$d \equiv 1 \bmod 4$: Algebraic integers are of the form $x + \frac{y}{2} + \frac{y}{2}\sqrt{d}$ where $x, y \in \mathbb{Z}$. Choose y so that $\frac{y}{2}$ is closest to v, then choose x so that $x + \frac{y}{2}$ is closest to u, i.e. x is closest to $u - \frac{y}{2}$. This yields

$$\alpha = (x + y\frac{1+\sqrt{d}}{2})\beta + (r + s\sqrt{d})\beta$$

where $|r| \leq \frac{1}{2}$ and $|s| \leq \frac{1}{4}$. Restricting d to $|d| \leq 13$ we have $d \in \{-11, -7, -3, 5, 13\}$. When $|d| \leq 11$,

$$|r^2 - s^2 d| \leq r^2 + 11s^2 \leq \frac{15}{16}.$$

When $d = 13$,

$$\frac{-13}{16} \leq -13s^2 \leq r^2 - 13s^2 \leq r^2 \leq \frac{1}{4}.$$

In all cases

$$|N((r + s\sqrt{d})\beta)| \leq \frac{15}{16}|N(\beta)| < |N(\beta)|.$$

Conclusion:

$$\mathbb{Q}(\sqrt{-11}), \mathbb{Q}(\sqrt{-7}), \mathbb{Q}(\sqrt{-3}), \mathbb{Q}(\sqrt{5}), \mathbb{Q}(\sqrt{13})$$

are Euclidean.

Section 7.6: The Gaussian Field

The Gaussian Field $\mathbb{Q}(i)$ is Euclidean and has ring of algebraic integers $R = \mathbb{Z}[1, i]$. We will give a complete description of the units and primes in R.

Units: As determined above, $1, -1, i, -i$.

Primes: Any irreducible π divides its norm, therefore divides a prime number, which must be unique ($\pi | p$ and $\pi | q$ where $p \neq q$ implies $\pi | (px + qy)$ and in particular $\pi | 1$: contradiction). We will characterize irreducibles according to the prime numbers they divide.

Let p be a prime number. Write $p = \pi_1 \cdots \pi_n$ (factored into irreducibles). Taking norms, $p^2 = N(\pi_1) \cdots N(\pi_n)$, hence $n \leq 2$. There are two cases to consider.

Case 1: -1 is a quadratic residue mod p. Then we have $x^2 + 1$ divisible by p, which rules out p irreducible (lest $p | (x + \sqrt{-1})$ or $p | (x - \sqrt{-1})$, which impossible). Hence $p = (a_1 + b_1\sqrt{-1})(a_2 + b_2\sqrt{-1})$, which forces $p = (a + b\sqrt{-1})(a - b\sqrt{-1}) = a^2 + b^2$. The corresponding irreducibles are associates of $a + b\sqrt{-1}$ where $a^2 + b^2 = p$.

Case 2: When -1 is not a quadratic residue mod p we cannot have $a^2+b^2 = p$, so p is irreducible.

The prime numbers p for which -1 is a quadratic residue mod p are 2 and odd primes of the form $p \equiv 1 \bmod 4$. So the irreducibles in R_{-1} are associates of $a+b\sqrt{-1}$ whenever $a^2+b^2 = p$ for some prime p, which occurs when $p = 2$ and $p \equiv 3 \bmod 4$, and all primes $p \equiv 3 \bmod 4$.

Chapter 7 Exercises

(i) $N(1 + \sqrt{2}) = -1$, hence $1 + \sqrt{2}$ is a unit in $\mathbb{Q}(\sqrt{2})$. It is certainly the smallest unit greater than 1 of the form $x + y\sqrt{d}$ where $x, y > 0$ belong to $\mathbb{Z}$, hence it generates all the units as described in Section 7.3 above. Also, $N(2 + \sqrt{3}) = 1$ and $N(1 + \sqrt{3}) = -2$ and $1 + k\sqrt{3} > 2 + \sqrt{3}$ when $k \geq 2$, hence the units in $\mathbb{Q}(\sqrt{3})$ are $\pm(2 + \sqrt{3})^n$, $n \in \mathbb{Z}$. Just out of curiosity, we have $(2 + \sqrt{3})^5 = 362 + 209\sqrt{3}$, and $362^2 - 3(209)^2 = 1$.

(ii) By construction, $\alpha = \frac{1+n\sqrt{d}}{1-n\sqrt{d}}$ satisfies $N(\alpha) = 1$. We just have to verify that it can expressed in terms of the appropriate integral basis, $\{1, \sqrt{d}\}$ if $d \equiv 2, 3 \bmod 4$ or $\{1, \frac{1+\sqrt{d}}{2}\}$ if $d \equiv 1 \bmod 4$. Checking cases yields a finite number of values of n and d. Examples are $\alpha = 1, -3 - 2\sqrt{2}, -2 - \sqrt{3}, \frac{-3-\sqrt{5}}{2}, i$.

(iii) Let p be a prime number. Then p is divisible by at least one irreducible π_p and we can write $p = \alpha_p \pi_p$. If p and q are distinct primes with $\pi_p = \pi_q$ then, choosing integers r and s such that $rp + sq = 1$, we have $r\alpha_p\pi_p + s\alpha_q\pi_p = 1$, hence $\pi_p(r\alpha_p + s\alpha_q) = 1$, hence π_p is a unit. Contradiction. So the π_p are distinct.

(iv) $2 = (-1 + \sqrt{3})(1 + \sqrt{3})$. Factors have norm -2, hence are not units. Therefore 2 is not irreducible.

(v) 2 is irreducible: If $2 = \alpha\beta$ then $4 = N(\alpha)N(\beta)$. If neither α nor β is a unit then $\pm 2 = N(\alpha) = x^2 + 6y^2$, which is impossible. Similarly 3 is irreducible. $\sqrt{-6}$ is irreducible: suppose $\sqrt{-6} = \alpha\beta$. Then $6 = N(\alpha)N(\beta)$. We have seen that neither norm in the product is 2 or 3, which just leaves 1 or 6. The numbers 2 and 3 are not associates of the numbers $\sqrt{-6}$ by comparison of norms. So R_{-6} does not have unique factorization.

(vi) Write $(x_1 + y_1\sqrt{-17})(x_2 + y_2\sqrt{-17}) = 1 + \sqrt{-17}$. Taking norms, $(x_1^2 + 17y_1^2)(x_2^2 + 17y_2^2) = 18$. The only possibilities for $x_1^2 + 17x_2^2$ are divisors of 18, namely 1, 2, 3, 6, 9, 12 or 18. The divisors 2, 3, 6, 9 and 12 are not possible, hence one of the two factors is a unit. This implies $1 + \sqrt{-17}$ is irreducible. Similarly, $1 - \sqrt{-17}$ is irreducible. On the other hand, we have $2 \cdot 3^2 = (1 + \sqrt{-17})(1 - \sqrt{-17})$. The number 2 is irreducible: $2 = \alpha\beta$ implies $N(\alpha)|4$, the only possibilities being $N(\alpha) \in \{1, 4\}$. By comparison of norms, 2 is not an associate of $1 \pm \sqrt{-17}$. Hence we have non-unique factorization into irreducibles.

(vii) $-2 \cdot 5 = \sqrt{-10}\sqrt{-10}$ and no integer solution to $x^2 + 10y^2 = 2, 5$. $2 \cdot 7 = (1 + \sqrt{-13})(1 - \sqrt{-13})$ and no integer solution to $x^2 + 13y^2 = 2, 7$. $-2 \cdot 7 = \sqrt{-14}\sqrt{-14}$ and no integer solution to $x^2 + 14y^2 = 2, 7$. $-3 \cdot 5 = \sqrt{-15}\sqrt{-15}$ and no integer solution to $(x + y/2)^2 + 15y^2/4 = 3, 5$ because no integer solution to $(2x + y)^2 + 15y^2 = 12, 20$.

(viii) We have $N(x + y\sqrt{10}) = x^2 - 10y^2 \equiv x^2 \equiv 0, 1, 4 \bmod 5$, hence there is no solution to $N(x + y\sqrt{10}) = \pm 2, \pm 3$. This implies that $4 + \sqrt{10}$ is irreducible: $(x_1 + y_1\sqrt{10})(x_2 + y_2\sqrt{10}) = 4 + \sqrt{10}$ implies, after computing norms, $(x_1^2 - 10y_1^2)(x_2^2 - 10y_2^2) = 6$. Hence each norm is a divisor of 6, hence ± 1 or ± 6, hence one of the factors is a unit. Given that $2 \cdot 3 = (4 + \sqrt{10})(4 - \sqrt{10})$ and that none of the irreducible divisors of 2 is an associate of $4 + \sqrt{10}$, we have non-unique factorization into irreducibles.

(ix) Applying Euclid's Algorithm we have

$$5 + 4\sqrt{3} = (2 + 0\sqrt{3})(1 + 2\sqrt{3}) + (3 + 0\sqrt{3})$$

$$1 + 2\sqrt{3} = (0 + \sqrt{3})(3 + 0\sqrt{3}) + (1 - \sqrt{3})$$

$$2 + 0\sqrt{3} = (-2 - 2\sqrt{3})(1 - \sqrt{3}) - 1$$

$$1 - \sqrt{3} = (-1 + \sqrt{3})(-1) + 0.$$

This yields

$$\begin{bmatrix} -1 \\ 0 \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ 1 & 1 - \sqrt{3} \end{bmatrix} \begin{bmatrix} 0 & 1 \\ 1 & 2 + 2\sqrt{3} \end{bmatrix} \begin{bmatrix} 0 & 1 \\ 1 & 0 - \sqrt{3} \end{bmatrix} \begin{bmatrix} 0 & 1 \\ 1 & -2 + 0\sqrt{3} \end{bmatrix} \begin{bmatrix} 5 + 4\sqrt{3} \\ 1 + 2\sqrt{3} \end{bmatrix},$$

$$(5 + 2\sqrt{3})(5 + 4\sqrt{3}) + (-12 - 6\sqrt{3})(1 + 2\sqrt{3}) = 1.$$

(x) Any irreducible divides its norm, therefore divides a prime number, which must be unique because if π divides distinct primes then it divides their greatest common divisor 1. We will characterize irreducibles according to the prime numbers they divide. Let p be a prime. Write $p = \pi_1 \cdots \pi_n$. Taking norms, $p^2 = N(\pi_1) \cdots N(\pi_n)$, hence $n \leq 2$. When 2 is a quadratic residue mod p we have $x^2 - 2$ divisible by p, which rules out p irreducible (lest $p|(x+\sqrt{2})$ or $p|(x-\sqrt{2})$, which impossible given that algebraic integers are of the form $a + b\sqrt{2}$ for integers a and b). Hence $p = (a_1 + b_1\sqrt{2})(a_2 + b_2\sqrt{2})$, which forces $p = (a + b\sqrt{2})(a - b\sqrt{2}) = a^2 - 2b^2$. The corresponding irreducibles are associates of $a+b\sqrt{2}$ where $a^2-2b^2 = p$. In short, expressions of the form $(1+\sqrt{2})^k(a+b\sqrt{2})$ where $a^2-2b^2 = p$. When 2 is not a quadratic residue mod p we cannot have $a^2 - 2b^2 = p$, so p is irreducible. The prime numbers p for which 2 is a quadratic residue mod p are 2 and odd primes of the form $p \equiv \pm 1 \bmod 8$.

(xi) The problem statement is not quite right. Let π be a Gaussian prime. If $\pi = p$, a prime number, then writing $\alpha = (px + r) + (py + s)i$, we have $\alpha \equiv r + is \bmod p$, where $0 \leq r < p$, $0 \leq s < p$. The representatives $r + is$ are distinct mod π. Now suppose $(\alpha, p) = 1$. Then $\alpha(r_1 + s_1 i) \equiv \alpha(r_2 + s_2 i)$ mod π implies $p|\alpha((r_2 - r_1) + (s_2 - s_1)i)$ implies $p|((r_2 - r_1) + (s_2 - s_1)i)$ implies $r_1 + s_1 i = r_2 + s_2 i$. Hence multiplication by α permutes the non-zero $r + si$ mod π. Forming the product of all $p^2 - 1$ expressions of the form $\alpha(r + si)$ where $r + si \neq 0$ yields the product of all $p^2 - 1$ of the non-zero class representatives, hence $\alpha^{N(\pi)-1} = \alpha^{p^2-1} \equiv 1 \bmod \pi$.

Next, consider $\pi = a + bi$ where $a^2 + b^2 = p$, a prime number. Then $b \not\equiv 0$ mod p. We claim that every Gaussian integer is equivalent to some element in $\{0, 1, \ldots, p-1\}$ mod π. Given $\alpha = x + yi$ we have $\alpha \equiv \alpha - z\pi = (x - az) + (y - bz)i \bmod \pi$ for any integer z. There is a solution to $y - bz \equiv 0$ mod p, hence mod π. Using this value of z we obtain $\alpha \equiv x - az \bmod \pi$. We also have $x - az \equiv r \bmod p$, hence mod π, for some $0 \leq r < p$. The representatives $0, 1, \ldots, p-1$ are distinct mod π: $\pi|(r-s)$ implies $r - s = \kappa\pi$ implies $(r-s)^2 = N(\kappa)p$ implies $p|(r-s)^2$ implies $p|(r-s)$ implies $r = s$.

Now consider $(\alpha, \pi) = 1$. Then $\{\pi, 2\pi, \ldots, (p-1)\pi\}$ is a permutation of $\{1, 2, \ldots, p-1\}$ mod π: $\alpha r \equiv \alpha s \bmod \pi$ implies $\pi|\alpha(r-s)$ implies $\pi|(r-s)$ implies $r \equiv s \bmod \pi$. Hence $(1\alpha)(2\alpha)\cdots((p-1)\alpha) \equiv (p-1)! \bmod \pi$, which implies $\pi|(\alpha^{p-1} - 1)(p-1)!$. Since $((p-1)!, p) = 1$, $a(p-1)! + b\pi\overline{\pi} = 1$ for integers some pair of integers a, b, hence $((p-1)!, \pi) = 1$. Therefore $\pi|(\alpha^{p-1} - 1)$. Hence $\alpha^{N(\pi)-1} = \alpha^{p-1} \equiv 1 \bmod \pi$.

Chapter 8: Diophantine Equations

Section 8.1: The Pell Equation

The Pell equation is $x^2 - dy^2 = 1$ where d is a non-square positive integer. We will consider more generally solutions to $x^2 - dy^2 = \pm 1$ where x and y are positive integers. Since we have already found all units in $Q(\sqrt{2})$ and $Q(\sqrt{3})$ in Exercise 7.1, we will assume without loss of generality that $d \geq 5$ when considering the equation $x^2 - dy^2 = -1$.

When $x^2 - dy^2 = 1$ then we have we have

$$x - y\sqrt{d} = 1/(x + y\sqrt{d}) > 0,$$

hence $x > y\sqrt{d}$ and $x/y > \sqrt{d}$. Substituting this into

$$x - y\sqrt{d} = 1/(x + y\sqrt{d})$$

yields

$$|x - y\sqrt{d}| < 1/2y\sqrt{d}.$$

This implies $x/y = p_n/q_n$, one of the convergents to $\sqrt{d}$. Since $(x, y) = 1$ and $(p_n, q_n) = 1$, this forces $x = p_n$ and $y = q_n$. Since $p_n/q_n > \sqrt{d}$, n must be an odd number.

When $x^2 - dy^2 = -1$ we have $(2 - \sqrt{d})y < 0 < x$ hence $2y < x + y\sqrt{d}$ hence $|x - y\sqrt{d}| = \frac{1}{x+y\sqrt{d}} < \frac{1}{2y}$ hence $|\sqrt{d} - \frac{x}{y}| < \frac{1}{2y^2}$ hence x/y is a convergent of the form p_n/q_n. We also have $x - y\sqrt{d} = \frac{-1}{x+y\sqrt{d}} < 0$, hence $\sqrt{d} > \frac{p_n}{q_n}$, hence n must be even.

We next consider which convergents p_n/q_n satisfy $p_n^2 - dq_n^2 = \pm 1$. Let $\theta = \sqrt{d} + [\sqrt{d}]$. Then $\theta' = -\sqrt{d} + [\sqrt{d}]$. In other words, $\theta > 1$ and $-1 < \theta' < 0$. Therefore θ is purely perodic and we have

$$\theta = [\overline{b_0, \cdots, b_{m-1}}]$$

for some minimal value of $m \geq 1$. This implies

$$\sqrt{d} = [b_0 - [\sqrt{d}], \overline{b_1, \ldots, b_m}] = [a_0, \overline{a_1, \ldots, a_m}].$$

In other words, $a_k = a_{k+m} = a_{k+2m} = \cdots$ for all $k \geq 1$. When $p_n^2 - dq_n^2 = 1$ we have n odd and

$$a_{n+1} + \frac{1}{\theta_{n+2}} = \theta_{n+1} = \frac{p_{n-1} - q_{n-1}\sqrt{d}}{q_n\sqrt{d} - p_n} =$$

$$(-p_{n-1} + q_{n-1}\sqrt{d})(p_n + q_n\sqrt{d}) =$$
$$-p_{n-1}p_n + dq_{n-1}q_n + (p_n q_{n-1} - p_{n-1}q_n)\sqrt{d} =$$
$$-p_{n-1}p_n + dq_{n-1}q_n + (-1)^{n+1}\sqrt{d} =$$
$$-p_{n-1}p_n + dq_{n-1}q_n + a_0 + \frac{1}{\theta_1}.$$

Comparing the expressions that are less than 1, $1/\theta_{n+2} = 1/\theta_1$, hence $\theta_{n+2} = \theta_1$. This implies a period of $n+1$ in the sequence $a_1, a_2, a_3, \ldots$, which forces $m|(n+1)$. Hence we have $n = km - 1$ for some k.

When $p_n^2 - dq_n^2 = -1$ we have n even and

$$a_{n+1} + \frac{1}{\theta_{n+2}} = \theta_{n+1} = \frac{p_{n-1} - q_{n-1}\sqrt{d}}{q_n\sqrt{d} - p_n} =$$
$$(p_{n-1} - q_{n-1}\sqrt{d})(p_n + q_n\sqrt{d}) =$$
$$p_{n-1}p_n - dq_{n-1}q_n - (p_n q_{n-1} - p_{n-1}q_n)\sqrt{d} =$$
$$-p_{n-1}p_n + dq_{n-1}q_n + (-1)^{n+2}\sqrt{d} =$$
$$-p_{n-1}p_n + dq_{n-1}q_n + a_0 + \frac{1}{\theta_1}.$$

Comparing the expressions that are less than 1, $1/\theta_{n+2} = 1/\theta_1$, hence $\theta_{n+2} = \theta_1$. This implies a period of $n+1$ in the sequence $a_1, a_2, a_3, \ldots$, which forces $m|(n+1)$. Hence we have $n = km - 1$ for some k. Since n must be even, m must be odd, so there is no solution to $x^2 - dy^2$ when the period of the $\sqrt{d}$ is even.

We next show that every $n = km - 1$ of the right parity is a solution to $x^2 - y^2 = \pm 1$. By periodicity we have $\theta_{n+2} = \theta_1$, hence

$$\sqrt{d} = \frac{\theta_{n+2}p_{n+1} + p_n}{\theta_{n+2}q_{n+1} + q_n} = \frac{\theta_1 p_{n+1} + p_n}{\theta_1 q_{n+1} + q_n}.$$

Substituting $1/\theta_1 = \sqrt{d} - a_0$ we obtain

$$\sqrt{d} = \frac{p_{n+1} + p_n(\sqrt{d} - a_0)}{q_{n+1} + q_n(\sqrt{d} - a_0)},$$

and rearranging we obtain

$$\sqrt{d}(q_{n+1} - q_n a_0 - p_n) = p_{n+1} - p_n a_0 - q_n d.$$

Hence both sides are zero. Equating the two resulting expressions for a_0 yields

$$\frac{q_{n+1} - p_n}{q_n} = \frac{p_{n+1} - q_n d}{p_n}.$$

Rearranging this yields

$$p_n^2 - q_n^2 d = -(p_{n+1}q_n - q_{n+1}p_n) = -(-1)^{n+2} = (-1)^{n+1}.$$

Theorem: $p_{mk-1} + q_{mk-1}\sqrt{d} = (p_{m-1} + q_{m-1}\sqrt{d})^k$ for all $k \geq 1$.

Proof: We have $U_d \cap (1, \infty) = \{\eta^k : k \geq 1\}$. Let k_0 be the least positive integer such that η^{k_0} has integer coefficients of 1 and $\sqrt{d}$. Then for $k \geq 1$, η^k has integer coefficients if and only $k_0|k$. This is clearly sufficient. To prove necessity, suppose η^k has integer coefficients and write $k = qk_0 + r$ with $0 \leq r < k_0$. Then $\eta^r = \eta^k \eta^{-qk_0}$. Since $\eta^{qk_0} = (\eta^{k_0})^q$, η^{qk_0} has integer coefficients, and since it has norm ± 1, η^{-qk_0} also has integer coefficients. Hence η^r has integer coefficients, which forces $r = 0$. Hence the set of units in $(1, \infty)$ with integer coefficients is $\{\mu^k : k \geq 1\}$ where $\mu = \eta^{k_0}$. So the solutions to $|x^2 - dy^2| = 1$ are embedded as the coefficients in the list $\mu < \mu^2 < \mu^3 < \cdots$. But this list is equal to $p_{m-1} + q_{m-1}\sqrt{d} < p_{2m-1} + q_{2m-1}\sqrt{d} < p_{3m-1} + q_{3m-1}\sqrt{d} \cdots$, hence the theorem is true.

An example is given in the textbook to the fundamental solutions to $x^2 - 97y^2 = -1$ and $x^2 - 97y^2 = 1$. It is hard enough to find the fundamental solution to the first equation by looking at convergents. It would be much harder to find the fundamental solution to the second equation by convergents (at least by hand), but having found the first solution we just square it to produce the second solution.

When $p \equiv 1 \bmod 4$ is a prime number, then $x^2 - py^2 = -1$ always has a solution. Reason: First choose a solution to $x^2 - py^2 = 1$ with $x, y > 0$. Then $x^2 - y^2 \equiv 1 \bmod 4$, which forces y even and x odd. Write $x = 2k+1$, $y = 2j$. Then $x^2 - 1 = py^2$ yields $k(k+1) = pj^2$. There are two cases to consider.

Case 1: $p|k$. Write $k = k_0 p$. Then we have $k_0(k+1) = j^2$. Since k and $k+1$ are coprime, k_0 and $k+1$ are coprime, and we have $k_0 = Y^2$ and

$k+1 = X^2$. This yields $X^2 = k+1 = pY^2+1$, $X^2 - pY^2 = 1$. Note that $X \le X^2 = k+1 = \frac{x+1}{2} < x$ since $x > 1$.

Case 2. $p|(k+1)$. Write $k+1 = k_0 p$. Then $kk_0 = j^2$, and since k and $k+1$ are coprime, k and k_0 are coprime, and we have $k = X^2$, $k_0 = Y^2$, $X^2+1 = k+1 = pk_0 = pY^2$, $X^2 - pY^2 = -1$.

So a solution to $x^2 - py^2 = 1$ yields either another solution $X^2 - pY^2 = 1$ with $x < X$ or to a solution $X^2 - pY^2 = -1$. We cannot fall into Case 1 indefinitely, so eventually we will arrive in Case 2 and produce the desired solution.

Section 8.3: The Mordell Equation $y^3 = x^2 + k$

1. Chords and tangents in projective space.

Homogenous polynomials $F(x,y,z)$ satisfy $F(\lambda x, \lambda y, \lambda z) = \lambda^n F(x,y,z)$ where n is the total degree of $F(x,y,z)$. Solutions to a polynomial equation $f(x,y) = 0$ can be embedded in the set of solutions to a polynomial equation $F(x,y,z) = 0$ where $F(x,y,z)$ is homogeneous and $F(x,y,1) = f(x,y)$. Solutions to $F(x,y,z) = 0$ with $z = 0$ are called points at infinity. Any rational solution (x,y,z) to $F(x,y,z) = 0$ with $z \neq 0$ gives rise to the rational solution $(x/z, y/z)$ to $f(x,y) = 0$ since $f(x/z, y/z) = F(x/z, y/z, 1) = z^{-n}F(x,y,z) = 0$.

On page 80 it is stated that for any non-zero integer k the curve $y^2 = x^3 + k$ has the property that the chord joining any two rational points on the curve $y^2 z = x^3 + kz^3$ intersects the curve again at a rational point. For example, the curve $y^2 = x^3 + 17$ is associated with the homogeneous equation $y^2 z = x^3 + 17z^3$ and two of its rational points are $(0,1,0)$ and $(-2,-3,1)$. The chord between them has coordinates $(1-t)(0,1,0) + t(-2,-3,1) = (-2t, 1-4t, t)$. Solutions to the homogeneous equation on this chord satisfy $(1-4t)^2 t = (-2t)^3 + 17t^3$, i.e. $t(t-1)(7t-1) = 0$. There are three solutions, $t = 0, 1, \frac{1}{7}$. Solutions $t = 0$ and $t = 1$ correspond to $(0,1,0)$ and $(-2,-3,1)$ while $t = \frac{1}{7}$ yields $(\frac{-2}{7}, \frac{3}{7}, \frac{1}{7})$, which yields the solution $(x,y) = (-2,3)$. Of course, this example is trivial because when (a,b) is a solution, so is $(a,-b)$. To take another example, two rational solutions are $(2,5,1)$ and $(-2,3,1)$. The chord between them is $(2-4t, 5-2t, 1)$. Solutions to the homogeneous equation satisfy $t(t-1)(16t-7) = 0$ and $t = \frac{7}{16}$ yields the solution $(\frac{1}{4}, \frac{33}{8}, 1)$. Hence we obtain the solution $(1/4, 33/8)$. Another way to generate new solutions from old is to find the point of intersection between an existing solution and the tangent to the curve at that solution. Apparently the rational solutions on the Mordell curve $y^2 = x^3 + k$ satisfy a kind of group law.

Solutions to a homogeneous equation live on lines through the origin. The set of all non-trivial lines forms projective space and one can study solutions to polynomial equations in this context.

2. Quadratic residues.

Rearranging $y^2 = x^3 + k$ into

$$y^2 = x^3 - a^3 + a^3 + k = (x - a)(x^2 + ax + a^2) + (a^3 + k)$$

and reducing by a prime that divides $x - a$ or $x^2 + ax + a^2$ yields $y^2 \equiv a^3 + k \mod p$. Hence $a^3 + k$ is a quadratic residue mod p and we can ask if such a thing is possible.

For example, suppose there is a solution to $y^2 = x^3 + 11$. As x and y range through $0, 1, 2, 3 \mod 4$, y^2 ranges through $0, 1, 0, 1 \mod 4$ and $x^3 + 11$ ranges through $3, 0, 3, 2 \mod 4$, hence we must have $y \equiv 0, 2 \mod 4$ and $x \equiv 1 \mod 4$. Choosing $a = -3$ we obtain $a^3 + k = -16$ and $x^2 + ax + a^2 \equiv 3 \mod 4$. Therefore $x^2 + ax + a^2$ has a prime divisor $p \equiv 3 \mod 4$. Reducing mod p we find that -16 is a quadratic residue mod p, which implies that -1 is a quadratic residue mod p, which is not possible given $p \equiv 3 \mod 4$. So there is no solution.

3. Factorization in $\mathbb{Q}[\sqrt{k}]$.

Consider the equation $y^2 = x^3 - 11$. We will establish some necessary conditions on x and y, assuming that there is a solution. We have

$$(y - \sqrt{-11})(y + \sqrt{-11}) = x^3.$$

We will show that the two factors $y + \sqrt{-11}$ and $y - \sqrt{-11}$ are coprime, then exploit unique factorization to determine x and y.

By unique factorization in $\mathbb{Q}(\sqrt{-11})$, any common irreducible divisor π of $y - \sqrt{-11}$ and $y + \sqrt{-11}$ is a divisor of x^3, hence of x by primality. Since π must be a divisor of $(y + \sqrt{-11}) - (y - \sqrt{-11}) = 2\sqrt{-11}$, we must have $N(\pi)|44$. Hence $N(\pi) \in \{1, 2, 4, 11, 22, 44\}$. We also have $N(\pi)|x^2$, which implies that any prime divisor of $N(\pi)$ is a divisor of x. We can rule out some of the prime divisors of $N(\pi)$ by considering the prime divisors of x. If $2|x$ then $y^2 \equiv -11 \equiv 5 \mod 8$, which is not possible. Also, if $11|x$ then $11|y$

and the equation implies $11y_0^2 = 11^2x_0^3 - 1$, which is also not possible. So in fact $N(\pi) = 1$ and $y - \sqrt{-11}$ and $y + \sqrt{-11}$ must be coprime.

Irreducible factorizations yield $y - \sqrt{-11} = \alpha_1 \cdots \alpha_j$, $y + \sqrt{-11} = \beta_1, \ldots \beta_k$, $x = \gamma_1 \cdots \gamma_l$. This yields

$$\alpha_1 \cdots \alpha_j \beta_1, \ldots \beta_k = \gamma_1^3 \cdots \gamma_l^3.$$

By unique factorization, any irreducible appears a multiple of 3 times (up to associates) in $y + \sqrt{-11}$, hence $y + \sqrt{-11}$ is a perfect cube times a unit. Given that the units in $\mathbb{Q}(\sqrt{-11})$ are ± 1, this implies

$$y + \sqrt{-11} = (a + \frac{b}{2}(1 + \sqrt{-11}))^3 =$$

$$a^3 + (3a^2b)/2 - (15ab^2)/2 - 4b^3 + (3/2a^2b + 3/2ab^2 - b^3)\sqrt{-11}$$

for some pair of integers a and b. Since a basis for $\mathbb{Q}(\sqrt{-11})$ is $\{1, \sqrt{-11}\}$, comparing coefficients we obtain

$$1 = 3/2a^2b + 3/2ab^2 - b^3$$

and

$$y = a^3 + (3a^2b)/2 - (15ab^2)/2 - 4b^3.$$

Taking norms in $y + \sqrt{-11} = (a + \frac{b}{2}(1 + \sqrt{-11}))^3$ we obtain

$$x^3 = (a^2 + ab + 3b^2)^3,$$

hence

$$x = a^2 + ab + 3b^2.$$

Multiplying the first equation by 2 and substituting the expression for x yields

$$2 = 3bx - 11b^3 = b(3x - 11b^2),$$

therefore $b \in \{-2, -1, 1, 2\}$. Chasing through the possibilities, $x = 3, 15$. This yields $y^2 = 3^3 - 11 = 4^2$, $y^2 = 15^3 - 11 = 58^2$. We have proved that if (x, y) is a solution to $y^2 = x^3 - 11$ then it must satisfy $(x, y) \in \{(3, \pm 4), (15, \pm 58)\}$. Conversely, one can check that these are all in fact solutions.

Section 8.4: The Fermat Equation

The equation $x^n + y^n = z^n$ has no non-trivial integer solutions for an integer $n \geq 3$: Conjectured by Fermat in 1637, proved by Wiles in 1995.

1. $x^2 + y^2 = z^2$.

A modulus 4 argument shows that x and y cannot both be odd. Given a solution (x, y, z) with $(x, y) = d > 1$, we can obtain another solution (x_0, y_0, z_0) after division by d^2. So it suffices to characterize primitive solutions (x, y, z) where $x, y, z > 0$, $(x, y) = 1$, and x is odd and y is even and z is odd. A solution satisfies $(z + x)(z - x) = y^2 = 4y_0^2$. The factors $z - x$ and $z + x$ are even. Any prime divisor p of $z + x$ and $z - x$ must be a divisor of $(z + x) - (z - x) = 2x$. It cannot divide x, otherwise it divides z and therefore y. Hence $p = 2$. Writing $z + x = 2u$ and $z - x = 2v$ we have $uv = y_0^2$. This yields $u = a^2$, $v = b^2$, hence $(x, y, z) = (a^2 - b^2, 2ab, a^2 + b^2)$ where $a > b$ have opposite parity and are coprime. Conversely, every such triple is a reduced solution: If p is a common prime divisor of $a^2 - b^2$ and $2ab$ then it must divide a or b, but if so then it divides both a and b: contradiction. Therefore $a^2 - b^2$ and $2ab$ are coprime.

REMARK: The first element x in a primitive Pythagorean triple (x, y, z) is a difference of squares $a^2 - b^2$ where a and b are coprime and of opposite parity. When x itself is a perfect square X^2 we obtain another Pythagorean triple (X, b, a). It is primitive: X and b are coprime since a and b are, and $X^2 = a^2 - b^2$ forces b to be even.

2. $x^4 + y^4 = z^4$.

The method of infinite descent can be used to show that $x^4 + y^4 = z^2$ has no non-trivial solutions. If there is a solution then (x^2, y^2, z) is a Pythagorean triple. Choose a primitive solution in which x^2 is odd, y^2 is even, and x^2, y^2, z are coprime in pairs. Then there exists a coprime pair a, b of opposite parity such that $x^2 = a^2 - b^2$, $y^2 = 2ab$, and $z = a^2 + b^2$. By the remark above, (x, b, a) is a primitive Pythagorean triple and there is a coprime pair A, B of opposite parity such that $x = A^2 - B^2$, $b = 2AB$, $a = A^2 + B^2$. This yields $y^2 = 2ab = 4AB(A^2 + B^2)$. Since A, B, and $A^2 + B^2$ are coprime in pairs, $A = u^2$, $B = v^2$, $A^2 + B^2 = w^2$. That is, $u^4 + v^4 = w^2$. This yields another primitive solution with $w < z$ since $w^2 = A^2 + B^2 = a < a^2 + b^2 = z$. This can't continue forever, so there were no solutions to begin with.

3. $x^3 + y^3 = z^3$.

Suppose there is a positive integer solution to $x^3 + y^3 = z^3$. Then $x^3 + y^3 \equiv z^3$ mod 9, which is only possible if one of x, y, z is divisible by 9 in $\mathbb{Z}$. Let

$\lambda = \frac{3-\sqrt{-3}}{2} \in R_{-3}$. Since $N(\lambda) = 3$, λ is irreducible in R_{-3}. Moreover $\lambda^4 = -\frac{9}{2} - \frac{9i\sqrt{3}}{2} = 9\omega$ where $\omega = \frac{-1-\sqrt{-3}}{2}$ is a unit. We have found non-zero α, β, γ in R_{-3} such that $\alpha^3 + \beta^3 + \gamma^3 = 0$ where $\lambda^4 | \gamma$. In other words, $\alpha^3 + \beta^3 + \lambda^{12}\gamma_1^3 = 0$. This justifies Baker's statement that a positive integer solution to $x^3 + y^3 = z^3$ gives rise to a non-zero solution to $\alpha^3 + \beta^3 + \eta\lambda^{3n}\gamma^3 = 0$ in R_{-3} where η is a unit and $n \geq 2$ and γ is not divisible by λ. We can assume further that α and β have no common factors. To complete the proof, we should be able to derive another such solution with n replaced by $n-1$ with $n-1 \geq 2$. Iterating this yields a contradiction.

Details: ω is a primitive 3^{rd} root of unity. Hence

$$(\alpha + \beta)(\alpha + \omega\beta)(\alpha + \omega^2\beta) = -\eta\lambda^{3n}\gamma^3.$$

Hence λ divides one of the factors on the left hand side. We claim that λ divides all three factors and that $\frac{\alpha+\beta}{\lambda}, \frac{\alpha+\omega\beta}{\lambda}, \frac{\alpha+\omega^2\beta}{\lambda}$ have no common factors in R_{-3}. To see this, suppose that λ divides $\alpha + \omega^i\beta$. Then

$$(\alpha + \omega^i\beta) - (\alpha + \omega^{i+1}\beta) = \omega^i(1-\omega)\beta = -\omega^{i+1}\lambda\beta,$$

hence λ divides $\alpha + \omega^{i+1}\beta$. This implies λ divides $\alpha + \omega^{i+2}\beta$. So λ divides all three factors. Now suppose that an irreducible π divides $\alpha + \omega^i\beta$ and $\alpha + \omega^{i+1}\beta$. Then it divides their difference $-\omega^{i+1}\lambda\beta$. If π is not an associate of λ then it must divide β, so it also divides α, a contradiction. Hence the three algebraic integers $\frac{\alpha+\beta}{\lambda}, \frac{\alpha+\omega\beta}{\lambda}, \frac{\alpha+\omega^2\beta}{\lambda}$ have no common factors in R_{-3}. So now we can write

$$\frac{\alpha + \omega^i\beta}{\lambda} = \eta_1\alpha_1^3$$

$$\frac{\alpha + \omega^{i+1}\beta}{\lambda} = \eta_2\beta_1^3$$

$$\frac{\alpha + \omega^{i+2}\beta}{\lambda} = \eta_3\gamma_1^3\lambda^{3n-3}$$

where η_1, η_2, η_3 are units in R_{-3} and λ does not divide γ_1. Since $1+\omega+\omega^2 = 0$, we have

$$\eta_1\alpha_1^3 + \omega\eta_2\beta_1^3 + \omega^2\eta_3\gamma_1^3\lambda^{3n-3} = 0.$$

Rescaling,

$$\alpha_1^3 + \eta_1\beta_1^3 + \eta_2\gamma_1^3\lambda^{3n-3} = 0.$$

We have $3n-3 \geq 3$. Reducing by $\lambda^3 = 3\sqrt{-3}$ we obtain

$$\alpha_1^3 + \eta_1\beta_1^3 \equiv 0 \bmod 3\sqrt{-3}.$$

Hence

$$\alpha_1^3 + \eta_1\beta_1^3 \equiv 0 \bmod 9.$$

Lemma 1: The distinct congruence class representatives mod $\sqrt{-3}$ in R_{-3} are $-1, 0, 1$.

Proof: Given that $\omega - 1 = (-\frac{1}{2} + \frac{i\sqrt{3}}{2})\sqrt{-3} \equiv 0 \bmod \sqrt{-3}$, we have $\omega \equiv 1 \bmod \sqrt{-3}$. Hence $x + y\omega \equiv x + y \bmod \sqrt{-3}$. Since every integer is in the class of $-1/0/1 \bmod 3$, it is in these clases mod $\sqrt{-3}$. These classes are distinct mod $\sqrt{-3}$. So every element in $R_{-3} = \mathbb{Z}[\omega]$ falls into one of these classes.

Lemma 2: When $\sigma \equiv 1 \bmod \sqrt{-3}$, $\sigma^3 \equiv 1 \bmod 9$.

Proof: $\sigma^3 - 1 = (\sigma - 1)(\sigma^2 + \sigma + 1)$. But $\sigma - 1 \equiv 0 \bmod \sqrt{-3}$ and $\sigma^2 + \sigma + 1 \equiv 1 + 1 + 1 \equiv 0 \bmod \sqrt{-3}$, so the result follows.

Using the lemmas, we obtain $1 \pm \eta_1 \equiv 0 \bmod 9$. The only units in R_{-3} satisfying this are $\eta_1 = \pm 1$. So we can rescale to

$$\alpha_1^3 + \beta_1^3 + \eta_2\gamma_1^3\lambda^{3n-3} = 0.$$

To complete the proof we must show that $n - 1 \geq 2$. If this is not true then we have

$$1 + (\pm 1) + \eta_2 3\sqrt{-3} \equiv 0 \bmod 9.$$

This is not possible in R_{-3}, as can be checked by brute force, checking $\eta_2 = \pm 1, \pm\omega, \pm\omega^2$.

Section 8.5: The Catalan Equation.

It is conjectured that $x^p - y^q = 1$ has only one solution in positive integers, namely $3^2 - 2^3 = 1$. For example, consider the equation $x^5 - y^2 = 1$. Suppose there is a solution (x, y). As x ranges through $0, 1, 2, 3 \bmod 4$, x^5 ranges through $0, 1, 0, 3 \bmod 4$. As y ranges through $0, 1, 2, 3 \bmod 4$, y^2 ranges through $0, 1, 0, 1 \bmod 4$. Therefore x and y have the same parity and must both be odd. We have $x^5 = y^2 + 1 = (y + i)(y - i)$. We claim that $y + i$ and $y - i$ are coprime: Suppose π is a common irredicuble divisor of $y + i$ and

$y-i$. Then π divides their difference $2i$, hence $N(\pi)|4$. On the other hand, π is a divisor of x^5, hence of x by primality, given that $\mathbb{Q}(i)$ is Euclidean. This implies $N(\pi)|x^2$. Since $(4,x^2)=1$, $N(\pi)=1$: contradiction. Given that $y+i$ and $y-i$ are coprime, $y+i=uz^5$ for some unit u, and wlog $y+i=z^5$ after absorbing the unit into the fifth power. Writing $z=a+bi$ and comparing coefficients in $y+i=(a+bi)^5$ we obtain

$$1=5a^4b-10a^2b^3+b^5=b(5a^4-10a^2b^2+b^4).$$

Hence $b=\pm 1$ and $5a^4-10a^2+1=\pm 1$. Since there is no integer solution to the latter equation, there was no solution to $x^5-y^2=1$ to begin with.

REMARK: What we seem to be doing here is exploiting unique factorization in a quadratic field, where we can factor things further than we can in $\mathbb{Z}$, hence imposing more conditions on a potential solution.

Chapter 8 Exercises:

(i) The positive solutions are of the form $x_n+y_n\sqrt{d}=(a+b\sqrt{d})^n$. Hence $x_{n+1}+y_{n+1}\sqrt{d}=(a+b\sqrt{d})(x_n+y_n\sqrt{d})$. This yields

$$x_{n+1}=ax_n+bdy_n,$$

$$y_{n+1}=bx_n+ay_n.$$

Using just the first recurrence relation and $a^2-b^2d=1$ yields

$$x_{n+1}=ax_n+bd(bx_{n-1}+ay_{n-1})=ax_n+b^2dx_{n-1}+a(x_n-ax_{n-1})=$$

$$2ax_n+(b^2d-a^2)x_{n-1}=2ax_n-x_{n-1},$$

$$x_{n+1}-2ax_n+x_{n-1}=0.$$

Similarly,

$$y_{n+1}-2ay_n+y_{n-1}=0.$$

Given $\sqrt{7}=[2,\overline{1,1,1,4}]$, we have $m=4$, hence the convergent $p_3/q_3=8/3$ yields $a+b\sqrt{7}=8+3\sqrt{7}$. Hence $a=8$.

(ii) $\sqrt{31}=[5,\overline{1,1,3,5,3,1,1,10}]$ has period $m=8$, hence there is no solution to $x^2-31y^2=-1$. Another solution is merely that $31\equiv 3 \bmod 4$, hence 31 is a Gaussian prime and cannot be a divisor of $x^2+1=(x+i)(x-i)$ because it divides neither factor.

(iii) We should assume $p \neq q$. We will use an infinite-descent argument. First find a solution to

$$x^2 - pqy^2 = 1$$

where $x > 0, y > 0$. This yields $x^2 - y^2 \equiv 1 \bmod 4$, hence x is odd and y is even. Write $x = 2k + 1$, $y = 2j$. Then

$$k(k + 1) = pqj^2.$$

There are four cases to consider.

Case 1: $pq|k$. Write $k = pqk_0$. Then we have $k_0(k + 1) = j^2$, therefore $k_0 = u^2$, $k + 1 = v^2$, $v^2 - pqu^2 = 1$.

Case 2: $p|k$ and $q|(k + 1)$. Write $k = pk_0$ and $k + 1 = qh$. Then we have $k_0h = j^2$, therefore $k_0 = u^2$, $k + 1 = qv^2$, $pu^2 - qv^2 = -1$.

Case 3: $q|k$ and $p|(k + 1)$. As in Case 2 this yields $qu^2 - pv^2 = -1$, hence $pv^2 - qu^2 = 1$.

Case 4: $pq|(k + 1)$. Write $k + 1 = pqh$. Then we have $kh = j^2$, $k = u^2$, $k + 1 = pqv^2$, $u^2 - pqv^2 = -1$. This yields $u^2 + 1 = pqv^2$, $p|(u + i)(u - i)$, which is not possible because p is a Gaussian prime and divides neither factor.

Conclusion: finding a solution to $x^2 - pqy^2 = 1$ yields another solution with smaller $x > 0, y > 0$ or to a solution to $px^2 - qy^2 = \pm 1$. Eventually we arrive at a solution to the latter.

Example: Let $p = 3$, $q = 7$. Then $55^2 - 21(12^2) = 1$ using the continued fraction expansion of $\sqrt{21}$ and $m = 6$. Writing $55 = 2k + 1$ yields $k = 27$, $k + 1 = 28$. This is Case 2 and yields $u = 3$, $v = 2$, $3(3^2) - 7(2^2) = -1$.

Example: Let $p = 31$, $q = 41$. Then $32799^2 - (31)(41)(920^2) = 1$ using the continued fraction expansion of $\sqrt{31 \cdot 41}$ and $m = 10$. Writing $32799 = 2k+1$ yields $k = 16399$, $k + 1 = 16400$. This is Case 2 and yields $u = 23$, $v = 20$, $31(23^2) - 41(20^2) = -1$.

A generalization (with Russell Jahn): Let $d > 1$ be square-free and congruent to 1 mod 4 and divisible by at least one prime p congruent to 3 mod 4. Then there is an integer solution to

$$ax^2 - by^2 = 1$$

for some $a > 1, b > 1$ satisfying $ab = d$.

Proof: Start with an integer solution to $x^2 - dy^2 = 1$ with $x > 0$, $y > 0$. By a mod 4 argument, x is odd and y is even. Write $x = 2k+1$, $y = 2j$. Then

$$k(k+1) = dj^2.$$

There are three cases to consider.

Case 1: $d|k$. Write $k = dk_0$. Then we have $k_0(k+1) = j^2$, therefore $k_0 = u^2$, $k+1 = v^2$, $v^2 - du^2 = 1$. We have $0 < v < x$.

Case 2: $d|(k+1)$. Write $k+1 = dh$. Then we have $kh = j^2$, $k = u^2$, $k+1 = dv^2$, $u^2 - dv^2 = -1$. This yields $u^2+1 = dv^2$. Therefore $p|(u+i)(u-i)$, which is not possible because p is a Gaussian prime and divides neither factor. So Case 2 can't happen.

Case 3: $d = d_1d_2$ where $d_1, d_2 > 1$ and $d_1|k$ and $d_2|k+1$. Write $k = d_1k_1$ and $k+1 = d_2k_2$. Then we have $k_1k_2 = j^2$, so we can write $k_1 = u_1^2$ and $k_2 = u_2^2$. Therefore $k = d_1u_1^2$ and $k+1 = d_2u_2^2$, and we have $d_2u_2^2 - d_1u_1^2 = 1$.

An infinite descent argument shows that we must eventually arrive in Case 3.

(iv) Suppose there is a rational solution to $x^4 - ay^4 = c$. Then there is an integer solution to $x^4 = ay^4 + cz^4$ with at least one of x, y, z odd. There are only two possible congruence classes for n^4 mod 16: 0 or 1. This yields $0/1 \equiv 0/a + 0/c$ mod 16, which can only be realized with $0 \equiv a + c$ mod 16. Hence x is even, y is odd, z is odd, $a + c = 16$. One possible solution is $x = 2$ and $y = z = 1$ and $a = c = 8$. A good trick question.

(v) The only solution to $a^3 + 2b^3 \equiv 0$ mod 7 is $a, b \equiv 0$ mod 7. So if $x^3 + 2y^3 = 7(z^3 + 2w^3)$ then each term is divisible by 7, and an infinite descent argument shows there is no non-trivial solution. Another good trick question.

(vi) $(2t-1)^4 + (t^2-1)^4 + (t^2-2t)^4 = 2(t^2-t+1)^4$.

(vii) Necessary conditions for $y^2 = x^3 - 17$: Reducing mod 4, $y^2 \equiv x^3 - 1$. Since $y^2 \equiv 0, 1$ and $x^3 - 1 \equiv 0, 3$, y must be even and x must be odd. Now write $y^2 = (x^3+8) - 25 = (x+2)(x^2+2x+4) - 25$. When $x \equiv 1$ mod 4, $x + 2 \equiv 3$, hence $x^3 + 8$ has a prime divisor $p \equiv 3$ mod 4. This implies $y^2 \equiv -25$ mod p, which contradicts the fact that -1 is not a quadratic residue mod p. When $x \equiv 3$, $x^3+8 \equiv 3 \bmod 4$, so again there is no solution. Hence this equation has no integer solutions.

(viii) The field $\mathbb{Q}(\sqrt{-2})$ is Euclidean. If there is any solution to $y^2 = x^3 - 2$ then y must be odd and x must be congruent to 3 mod 4 since $y^2 \equiv 0, 1, 0, 1$ and $x^3 - 2 \equiv 2, 3, 2, 1 \bmod 4$. Rearranging $y^2 = x^3 - 2$ to $y^2 + 2 = x^3$, we obtain

$$(y - \sqrt{-2})(y + \sqrt{-2}) = x^3.$$

Now let π be a Gaussian prime divisor of both $y - \sqrt{-2}$ and $y + \sqrt{-2}$. Then it is a divisor of their difference $2\sqrt{-2}$, hence $N(\pi)|8$, hence $|N(\pi)| \in \{1, 2, 4, 8\}$. Since $\pi|x^3$, $N(\pi)|x^6$, forcing $N(\pi)$ to be odd. Hence $|N(\pi)| = 1$. Therefore $y - \sqrt{-2}$ and $y + \sqrt{-2}$ are coprime and by unique factorization $y + \sqrt{-2} = \mu\omega^3$ for some unit μ. Since $\mu = \pm 1$ in $\mathbb{Q}(\sqrt{-2})$, we can write $y + \sqrt{-2} = \gamma^3 = (a + b\sqrt{-2})^3$. Comparing coefficients of $\sqrt{-2}$ yields

$$1 = 3a^2b - 2b^3 = b(3a^2 - 2b^2).$$

This yields $(a, b) = (\pm 1, \pm 1)$. We also have

$$x^3 = (a - b\sqrt{-2})^3(a + b\sqrt{-2})^3 = a^6 + 6a^4b^2 + 12a^2b^4 + 8b^6 = 27.$$

So $y^2 = x^3 - 2$ is solvable with $x = 3$, $y = \pm 5$.

(ix) Let S be the set of all coprime and positive (x, y, z) satisfying $x^4 - y^4 = z^2$. We will show that $(x, y, z) \in S \implies (x', y', z') \in S$ with $x' < x$. Hence S must be empty. We will let P represent the set of primitive Pythagorean triples.

Suppose $(x, y, z) \in S$. Then x must be odd and y and z must be of opposite parity.

Case 1: y is odd and z is even. Then $(y^2, z, x^2) \in P$, hence $(y^2, z, x^2) = (a^2 - b^2, 2ab, a^2 + b^2)$, hence

$$a^4 - b^4 = (a^2 - b^2)(a^2 + b^2) = (yx)^2,$$

hence $(a, b, yx) \in S$ with $a < x$.

Case 2: y is even and z is odd. Then $(z, y^2, x^2) \in P$, hence $(z, y^2, x^2) = (a^2 - b^2, 2ab, a^2 + b^2)$.

Case 2.1: a is even and b is odd. Then $a = 2a_1^2$ and $b = b_1^2$ and $(b_1^2, 2a_1^2, x) \in P$, hence $(b_1^2, 2a_1^2, x) = (a_2^2 - b_2^2, 2a_2b_2, a_2^2 + b_2^2)$, hence $a_2 = a_3^2$ and $b_2 = b_3^2$, hence $b_1^2 = a_3^4 - b_3^4$, hence $(a_3, b_3, b_1) \in S$ with $a_3 \le a_2 < x$.

Case 2.2: a is odd and b is even. Then $a = a_1^2$ and $b = 2b_1^2$ and $(a_1^2, 2b_1^2, x) \in P$, hence $(a_1^2, 2b_1^2, x) = (a_2^2 - b_2^2, 2a_2b_2, a_2^2 + b_2^2)$, hence $a_2 = a_3^2$ and $b_2 = b_3^2$, hence $a_1^2 = a_3^4 - b_3^4$, hence $(a_3, b_3, a_1) \in S$ with $a_3 \leq a_2 < x$.

(x) If $x^4 + y^4 = z^3$ has a primitive solution then z is odd. We have

$$(x^2 + iy^2)(x^2 - iy^2) = z^3.$$

If π is a common divisor of $x^2 + iy^2$ and $x^2 - iy^2$ then it divides $2x^2$ and $2iy^2$ and z^3. Hence $N(\pi)$ divides $4x^4$ and $4y^2$ and z^6. If p is a prime divisor of $N(\pi)$ then it divides $4x^4$ and $4y^2$ and z^6. Since z is odd, p is odd and must divide x and y. Contradiction. Hence $N(\pi) = 1$ and $x^2 + iy^2$ and $x^2 - iy^2$ are coprime. Hence $x^2 + iy^2 = u(a + bi)^3$ for some $u \in \{1, i, -1, -i\}$ by unique factorization. Multiplying through by u^{-1} and absorbing the minus sign into the cube we can write $X^2 + iY^2 = (a + bi)^3$. We have $X^2 = a(a^2 - 3b^2)$ and $Y^2 = b(3a^2 - b^2)$. Suppose p is a prime dividing a and $a^2 - 3b^2$. Then $p = 3$ and in fact $(a, a^2 - 3b^2) = 3$. Similarly, if q is a prime dividing b and $3a^2 - b^2$ then $q = 3$ and $(b, 3a^2 - b^2) = 3$. Therefore $(a, a^2 - 3b^2) = 1$ or $(b, 3a^2 - b^2) = 1$. If both pairs are coprime then, as we argued before, we can derive a positive solution to $X^4 - 3Y^4 = z^2$. Now suppose $(a, a^2 - 3b^2) = 3$ and $(b, 3a^2 - b^2) = 1$. Then $a = 3a_1$ and $|b| = b_1^2$ and we can write $X^2 = 3a_1(9a_1^2 - 3b_1^4)$. Writing $X = 3X_1$ we obtain $X_1^2 = a_1(3a_1^2 - b_1^4)$. Hence we have $|a_1| = a_2^2$ and $|3a_1^2 - b_1^4| = c^2$ and we have $|3a_2^4 - b_1^4| = c^2$, which by a mod 4 argument implies $b_1^4 - 3a_2^4 = c^2$. The other case yields the same conclusion. Hence there is a primitive solution to $x^4 - 3y^4 = z^2$. We will show using computations in R_3 that this gives rise to $X^4 - 3Y^4 = Z^4$ with $X < x$.

Given a primitive solution to $x^4 - 3y^4 = z^2$, a mod 4 argument shows that z must be odd and that x and y must have opposite parity. We have

$$(x^2 - \sqrt{3}y^2)(x^2 + \sqrt{3}y^2) = z^2.$$

The two factors are coprime in R_3, hence

$$x^2 + \sqrt{3}y^2 = u(a + b\sqrt{3})^2$$

for some positive unit u. Since the positive units in R_3 are integer powers of $2 + \sqrt{3}$, by creative grouping we can write

$$x^2 + \sqrt{3}y^2 = (A + B\sqrt{3})^2$$

or

$$x^2 + \sqrt{3}y^2 = (2 + \sqrt{3})(A + B\sqrt{3})^2.$$

The latter equation can be ruled out using congruences mod 4 as follows: it implies

$$x^2 = 2A^2 + 6AB + 6B^2$$

and

$$y^2 = A^2 + 4AB + 3B^2.$$

Hence x is even, y is odd. Since $y^2 \equiv A^2 + 3B^2 \bmod 4$, A is odd and B is even. Writing $x = 2x_0$, we have $2x_0^2 = A^2 + 3AB + 3B^2$, which is impossible. Given that $x^2 + \sqrt{3}y^2 = (A + B\sqrt{3})^2$, we have

$$x^2 = A^2 + 3B^2$$

and

$$y^2 = 2AB.$$

Hence y is even, x is odd, A is odd, B is even, and all four numbers are coprime in pairs. Writing $A = A_1^2$ and $B = 2B_1^2$ we obtain

$$x^2 = A_1^4 + 12B_1^4,$$

$$y^2 = 4A_1^2B_1^2.$$

Factoring,

$$(x - 2\sqrt{3}B_1^2)(x + 2\sqrt{3}B_1^2) = A_1^4,$$

and repeating the argument above, we arrive at

$$x = A_2^2 + 3B_2^2,$$

$$2B_1^2 = 2A_2B_2,$$

hence $A_2 = A_3^2$, $B_2 = B_3^2$, $B_1^2 = A_3^2B_3^2$,

$$x = A_3^4 + 3B_3^4,$$

$$A_1^4 = x^2 - 12B_1^4 = (A_3^4 + 3B_3^4)^2 - 12(A_3^2B_3^2) = (A_3^4 - 3B_3^4)^2,$$

hence

$$A_3^4 - 3B_3^4 = \pm A_1^2.$$

A modulo 4 argument shows that in fact we have

$$A_3^4 - 3B_3^4 = A_1^2.$$

Given that $A_3 < x$, we have infinite descent. So there is no solution to $x^4 - 3y^4 = z^2$ in positive x, y, z.

(xi) For any integer n, $n^3 \equiv n \bmod 6$, as can be checked directly using $0 \leq n \leq 5$. Therefore

$$n = (n - n^3) + n^3 = 6k + n^3 = (k+1)^3 + (k-1)^3 + (-k)^3 + (-k)^3 + n^3.$$

(xii) Suppose $x^2 + 7 = 2^{3k+2}$ has a solution. Then x must be odd. Writing $x = 2X + 1$ we have $4X^2 + 4X + 8 = 2^{3k+2}$, which implies $X^2 + X + 2 = 2^{3k}$. We will find all solutions to $x^2 + x + 2 = y^3$.

Assume $x^2 + x + 2 = y^3$. A modulus 4 argument shows that x and y must both be even. Factoring, we obtain

$$(x + \frac{1}{2} + \frac{\sqrt{-7}}{2})(x + \frac{1}{2} - \frac{\sqrt{-7}}{2}) = y^3.$$

Any common divisor δ of $x + \frac{1}{2} + \frac{\sqrt{-7}}{2}$ and $x + \frac{1}{2} - \frac{\sqrt{-7}}{2}$ is a common divisor of $2x + 1$ and $2\sqrt{-7}$ and y^3. Hence $N(\delta)$ is a common divisor of $(2x+1)^2$ and 28 and y^6. If p is a prime divisor of $N(\delta)$ then $p \in \{2, 7\}$. But 2 does not divide $(2x+1)^2$ and 7 does not divide y^6. So in fact $|N(\delta)| = 1$ and δ is a unit and $x + \frac{1}{2} + \frac{\sqrt{-7}}{2}$ and $x + \frac{1}{2} - \frac{\sqrt{-7}}{2}$ are coprime. By unique factorization in R_{-7}, this forces

$$x + \frac{1}{2} + \frac{\sqrt{-7}}{2} = u(a + b\frac{1 + \sqrt{-7}}{2})^3$$

for some unit u, and since the units in R_{-7} are ± 1, we can assume without loss of generality that $u = 1$. Given that

$$(a+b\frac{1+\sqrt{-7}}{2})^3 = a^3 + \frac{3a^2b}{2} + \frac{3}{2}\sqrt{-7}a^2b - \frac{9ab^2}{2} + \frac{3}{2}\sqrt{-7}ab^2 - \frac{5b^3}{2} - \frac{1}{2}\sqrt{-7}b^3,$$

comparing coefficients of $\sqrt{-7}$ we obtain

$$\frac{1}{2} = \frac{3}{2}a^2b + \frac{3}{2}ab^2 - \frac{1}{2}b^3,$$

$$1 = b(3a^2 + 3ab - b^2).$$

The only integer solutions to this are $(a, b) = (0, -1)$ and $(a, b) = (1, -1)$. Given that

$$x + \frac{1}{2} = a^3 + \frac{3a^2b}{2} - \frac{9ab^2}{2} - \frac{5b^3}{2},$$
$$x \in \{-3, 2\}.$$

Hence we obtain solutions $(x, y) = (-3, 2)$ and $(x, y) = (3, 2)$.

In summary, the only integer solution to $x^2 + x + 2 = 2^{3k}$ is $x \in \{-3, 2\}$ and $2^k = 2$, which forces $k = 1$.