Breaking Bad RSA Encryption, Day 1

Michael, Mathcamp 2019

1 Rivest-Shamir-Adleman Encryption

Definition 1 (Euler's Totient Function). For a positive integer n, the **totient** $\phi(n)$ of n is the number of positive integers $a \leq n$ such that $\gcd(a, n) = 1$.

Theorem 2 (Euler's Formula). If $\gcd(a, n) = 1$, then $a^{\phi(n)} \equiv 1 \pmod{n}$.

Proof. Consider the set S of positive $a \leq n$ such that $\gcd(a, n) = 1$. Then, multiplying every element of S by a does not change the product modulo n, proving the claim. □

Definition 3. When $\gcd(a, n) = 1$, the smallest b such that $a^b \equiv 1 \pmod{n}$ is known as the **order** of $a \pmod{n}$.

RSA encryption is a way for Bob to send Alice a secret message that can't be read by anyone but Alice. Alice picks two primes p, q and defines $n = pq$, then finds d, e such that $de \equiv 1 \pmod{\phi(n)}$. Then to send a number $M < n$ such that $\gcd(M, n) = 1$, Alice sends e to Bob, and Bob sends Alice $M^e \pmod{n}$, and to read the message Alice can take $M^{ed} \equiv M \pmod{n}$.

RSA works because prime factorization is hard. Thus the worst pairs p, q are those where p, q are not both prime and those where we have a quick way of factorizing n. We won't talk in depth about specifically how fast and slow things are, but general we can say that we can take the square of a number modulo n and take one number modulo another number relatively quickly.

2 Primitive Roots for Primes

By Euler's formula, we know that every positive integer less than p has order dividing $p-1$. We claim also that one of these, which we call a **primitive root**, has order exactly $p-1$, and thus all of the others can be written modulo p as a power of this number.

Lemma 4. For all positive integers n, we have

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

Proof. Note that $\phi(d)$ is the number of values x at most n such that $\gcd(n, x) = n/d$. Thus this sum counts every number once. □

Lemma 5. For any prime p and polynomial P with degree d such that p does not divide the coefficient of x^d in $P(x)$, there are at most d solutions to P modulo p.

Proof. Say that there are at least d solutions $x_1, \ldots, x_d$ to P modulo p. Then we know that there exists a P_1 such that $P(x) \equiv P_1(x) \times (x - x_1) \pmod{p}$. Similarly, there is a P_2 such that $P_1(x) \equiv P_2(x) \times (x - x_2) \pmod{p}$. Continuing this process, we see that

$$P(x) \equiv c \times (x - x_1)(x - x_2) \ldots (x - x_d) \pmod{p},$$

and thus since p is prime, and c is not a multiple of p due to the constraint on the coefficient of x^d in $P(x)$, there are no solutions other than $x_1, \ldots, x_d$. □

Theorem 6. All primes p have a primitive root.

Proof. Note that any nonzero value modulo p must have an order that divides $p-1$. Note also that by Lemma 5, the number of solutions to $x^d \equiv 1 \pmod{p}$ is at most d, and if one exists then there are d solutions, and thus exactly $\phi(d)$ of them have order d. Since the sum of the phi functions of the divisors of $p-1$ is $p-1$ by Lemma 4, we know that these values must exist, and thus there are $\phi(p-1) \geq 1$ primitive roots. □

2.1 Primitive Root Primality Testing

Say $p-1$ has a known factorization $p-1 = p_1^{a_1} \times \ldots \times p_k^{a_k}$. Now if $b_i^{n-1} \equiv 1 \pmod{n}$, by Euler's Theorem we know that the order of b_i is a factor of $n-1$. If we know also that $b_i^{(n-1)/p_i} \not\equiv 1 \pmod{n}$, then the order of b_i is not a factor of $(n-1)/p_i$, and thus $p_i^{a_i}$ divides the order of b_i, and thus divides $\phi(n)$. If we can find a b_i for all $1 \leq i \leq k$, then $\phi(n) = n-1$, and n is prime.

2.2 Pollard Rho

Consider some arbitrary polynomial $f(x) = x^2+1$. Consider the sequence $x_1 = 1$, $x_{i+1} = f(x_i)$. This sequence is eventually periodic modulo n. Note also that this sequence is eventually periodic modulo any prime, and that this eventual period is likely not the same for every prime. Thus to factorize n, we can take the difference of two values x_i, x_j in the periodic portion of the sequence such that $i-j$ is not a multiple of all of the prime periods but is a multiple of at least one. Then $\gcd(x_i - x_j, n)$ is a nontrivial factor of n. To do this and ensure we find values in the periodic part, we can calculate the values $x_2 - x_1$, $x_4 - x_2$, $x_7 - x_4$, and so on,

This may take some time because the number of x_i we have to calculate is approximately half the square of the smallest period. But we have an idea: find some sequence which has a different period for different primes, and use this to find a value which is not a multiple of n but is a multiple of at least one factor.

2.3 Pollard p − 1

If $p-1$ has small factors, and in particular is a factor of some arbitrary large factorial 10000!, then $2^{10000!} \equiv 1 \pmod{p}$ for arbitrary 2 relatively prime to p, and so $\gcd(2^{10000!} - 1, n) > 1$, and likely this value is not a multiple of n, since this implies the order of $2 \pmod{p}$ is a factor of every prime power factor of n. This is similar to Pollard rho, in that we have a periodic sequence that has different periods for different primes, which we can exploit to find a nontrivial factor of n.

This algorithm will rarely take a long time (it will take about $O(10000 \log(10000) \log(n)^2)$ time), but it requires us to have that $p-1$ factors into small primes for some prime $p|n$.

3 The Quadratic Sieve

Consider some set B, our **factor base**, containing the first few primes, along with -1. We can take some range of values x near $\sqrt{n}$, and see whether we can write $x^2 - n$ as a product of elements of B. If so, then we can take the parity of each

power and for this value assign a string of bits. If we can find $|B|+1$ such values, we can continue taking the topmost of the leftmost nonzero bits and XORing that bitstring with all of the strings below it which have a 1 in that place until some string consists of zeroes. Then we can take the product of some values of x^2, and this product will be equivalent modulo n to the square product of some small primes (multiplied by an even power of -1). Then we can calculate the square roots of these values, knowing that the sum of the square roots multiplied by the difference is a multiple of n. So we take the gcd of n and the sum, and that of n and the difference, and assuming that the factors are approximately randomly distributed when we get them from this process, then there is about a $1/2$ chance that one of these has a gcd with n between 1 and n, giving us a nontrivial factor of n. Otherwise, if we can find a new x value, we can repeat the process, and continue this pattern until we have a factorization.

3.1 Solving $p^k \mid x^2 - n$ for $p \in B$, when $k > 1$, and various improvements

Factoring $x^2 - n$ may take time depending on the number of values. We know that we need to find $x \pmod p$ to determine whether $x^2 \equiv n \pmod p$, and once we have such an x, we can inductively determine when $x^2 \equiv n \pmod{p^{k+1}}$, by noting that

$$(r + ap^k)^2 \equiv r^2 + 2arp^k \pmod{p^{k+1}}.$$

Now in almost any range of integers, at least half of the numbers have a log within $1 > \log 2$ of the largest number. Thus we can take the log of the largest number, create an array for each value of x in our range, set it to this value, and whenever x is a multiple of a prime, subtract the log of that prime. Then any value which is less than the square of the largest prime in B at the end of our process is prime or can be factored in our factor base.*

Also, we may have some p such that n is not a quadratic residue. If this p is small, we can lose many potential values. In order to solve this problem, we can multiply n by a prime which is not a quadratic residue modulo p, which will make the new product a quadratic residue. This only breaks our above calculation for finding values of x modulo p^2, but we can simply find such a value for the original n. Then, at the end when we have $x+y$ and $x-y$, the odds that the nontrivial factors are in different factors of $x^2 - y^2$ are the same.

*Canonically we divide powers of primes up to about this value, so that we don't divide every power, since this takes time, especially for small primes, and it doesn't give us too many new values.

Breaking Bad RSA Encryption, Day 2

Michael, Mathcamp 2019

4 Quadratic Reciprocity

The $k > 1$ constraint will be removed eventually by use of a new technique, but before we get there, we will build some new prime factorization methods and primality tests, whose mechanics may help us solve this problem.

Definition 7. An integer b is a **quadratic residue** modulo a when there is some t such that $b \equiv t^2 \pmod{a}$.

Definition 8. If p is an odd prime, then the **Legendre symbol** (n/p) equals 0 if $p|n$, 1 if n is a quadratic residue modulo p, and -1 otherwise.

Definition 9. If $n = p_1 \times \ldots \times p_k$ is an odd positive integer and $p_1, \ldots, p_k$ are not necessarily distinct, then the **Jacobi symbol** (a/n) equals the product $(a/p_1) \times \ldots \times (a/p_k)$ of Legendre symbols. Note that this does not determine whether a is a quadratic residue modulo n.

Theorem 10 (Euler's Criterion). For all a and odd prime p, we have $a^{(p-1)/2} \equiv (a/p) \pmod{p}$.

Proof. We can first verify that this is true when $(a/p) = 0$. Since p must have a primitive root r, we can write a as an odd power of r if $(a/p) = -1$, and as an even power otherwise. From this we can verify the result. □

Theorem 11 (Gauss's Criterion). Let a be a positive integer less than an odd prime p. Then if for $1 \leq 2i-1 < p$ we have $(2i-1)a \equiv s_i \pmod{p}$ such that $0 \leq s_i < p$, and t is the number of even values of s_i, we have $(a/p) = (-1)^t$.

Proof. Relabel the s_i such that $s_i, \ldots, s_t$ are even, and $s_{t+1}, \ldots, s_{(p-1)/2}$ are odd. Now we claim that $s_1, \ldots, s_t, s_{t+1}, \ldots, s_{(p-1)/2}$ are all distinct. This is because otherwise the sum of two of two values of s_i would be a multiple of p, which is impossible since each can be written as an odd number less than p multiplied by a, and the sum of two positive odd numbers less than p cannot be a multiple of p. Therefore this set of values comprises every odd number less than p, and thus

$$\begin{aligned} & 1 \times 3 \times \ldots \times (p-2) \\ &= (p-s_1) \times \ldots \times (p-s_t) \times s_{t+1} \times \ldots s_{(p-1)/2} \\ &\equiv (-1)^t \, s_1 \times \ldots \times s_{(p-1)/2} \pmod{p} \\ &\equiv (-1)^t b^{(p-1)/2} \, 1 \times 3 \times \ldots \times (p-2) \pmod{p}. \end{aligned}$$

By Euler's Criterion, this is sufficient to prove the claim. □

Note that this is clearly true when $p = 2$ as well. Also, it suffices to consider $p \pmod{4a}$, because if we consider p and $p + 4a$ and both of these are prime, then in the latter, the sequences of consecutive even values of s_i get 2 new even values each, and sequences of consecutive odd values get 2 new values each.

Theorem 12 (Quadratic Reciprocity). Let p and q be odd primes. Then $(p/q) = (q/p)$ if at least one of p, q is equivalent to 1 (mod 4), and $(p/q) = -(q/p)$ otherwise.

Proof. Assume first that $p \neq q$. Let s and t be the values of t from Gauss's Criterion. Note that $(p/q)(q/p) = (-1)^{s+t}$, where s and t are the values of t calculated in Gauss's Criterion. Now consider the set of pairs (a, a') such that a is an odd number at most q and a' is an odd number at most p. Then let $r = ap - a'q$. We will consider the set of values r. First, we see that these values are even and distinct. Note that there is a bijection between even values counted in s and values of r which are positive and at most q, and a bijection between even values counted in t which are negative and at least $-p$. And of course, none of the values are zero. Now note that if (a, a') is such a value then so is $(q-1-a, p-1-a')$, and thus there are an even number unless p and q are both 3 (mod 4), as desired. □

5 Lucas Sequences

Definition 13. Let D be an integer congruent to 0 or 1 (mod 4) which is not a perfect square and let P be an integer with the same parity as D. Then the **Lucas sequences** $\{U_i\}$, $\{V_i\}$ are such that

$$\frac{V_i + U_i\sqrt{D}}{2} = \left(\frac{P + \sqrt{D}}{2}\right)^i,$$

and define Q such that $D = P^2 - 4Q$.

You have probably seen these before. When $P = 1$ and $D = 5$, U_i are the Fibonacci numbers and V_i are the Lucas numbers. The mechanics of the values of U_i and V_i also function similarly to how cosine and sine function in $\cos\theta + i\sin\theta = e^{i\theta}$, in that multiplication between sine and cosine work on a coordinate system in the same way as between U and V, except U values multiply by 5 rather than -1 when multiplied by each other to achieve a V value. They are also similar in that taking multiples of the inputs (the indices or the angles) takes powers of the initial value.

Lemma 14 (Lucas Pythagorean Theorem). For any pair of Lucas sequences, we have

$$V_i^2 - DU_i^2 = 4Q^i$$

Proof. We have that

$$\begin{aligned} V_i^2 - DU_i^2 &= (V_i + U_i\sqrt{D})(V_i - U_i\sqrt{D}) \\ &= 4\left(\frac{P + \sqrt{D}}{2}\right)^i \left(\frac{P - \sqrt{D}}{2}\right)^i \\ &= 4\left(\frac{P^2 - D}{4}\right)^i = 4Q^i \end{aligned}$$

□

Lemma 15 (Lucas Sum and Difference Formulas). For all sequences $\{U_i\}$, $\{V_i\}$ and nonnegative i, j, we have

$$\begin{aligned} 2U_{i+j} &= V_iU_j + U_iV_j \\ 2V_{i+j} &= V_iV_j + DU_iU_j \\ 2Q^jV_{i-j} &= V_iV_j - DU_iD_j \\ 2Q^jU_{i-j} &= U_iV_j - U_jV_i \end{aligned}$$

Proof. For the first two, we have that

$$\begin{aligned} V_{i+j} + U_{i+j}\sqrt{D} &= 2\left(\frac{P+\sqrt{D}}{2}\right)^{i+j} \\ &= 2\left(\frac{P+\sqrt{D}}{2}\right)^{i} \times \left(\frac{P+\sqrt{D}}{2}\right)^{j} \\ &= (V_i + U_i\sqrt{D})(V_j + U_j\sqrt{D})/2 \\ &= (V_iV_j + DU_iU_j)/2 + (V_iU_j + U_iV_j)\sqrt{D}/2 \end{aligned}$$

For the last two, note by the Lucas Pythagorean Theorem that

$$(V_j + U_j\sqrt{D})^{-1} = \frac{V_j - U_j\sqrt{D}}{4Q^j},$$

so we know that

$$\begin{aligned} V_{i-j} + U_{i-j}\sqrt{D} &= 2\left(\frac{P+\sqrt{D}}{2}\right)^{i-j} \\ &= 2(V_i + U_i\sqrt{D}) \times (V_j + U_j\sqrt{D})^{-1} \\ &= \frac{V_iV_j - DU_iU_j + (U_iV_j - V_iU_j)\sqrt{D}}{2Q^j}. \end{aligned}$$

□

Breaking Bad RSA Encryption, Day 3

Michael, Mathcamp 2019

Lemma 16. If an odd prime p divides U_i and j is a multiple of i, then p divides U_j. If p divides V_i and j is an odd multiple of i, then p divides V_j.

Proof. For the first part, take a look at the U-sum property. For the second, by induction we can show that for all j, if V_j is a multiple of p then V_{j+2i} is a multiple of p. We can do this using the V-sum and U-difference properties. □

Let's prove one of the converses:

Lemma 17. Let m be relatively prime to $2Q$, and let e be the smallest value such that m divides U_e (e is called the **rank** of m). Then m divides U_k if and only if $e|k$.

Proof. We know one direction already, so we'll just prove the other. By the U-difference formula, if $i = qe + r$ where $r < e$, then

$$2Q^{qe}U_r = U_iV_{qe} - U_{qe}V_i.$$

Since the left-hand side is not a multiple of m, and the right-most term is, the middle term cannot be, so m does not divide U_i. □

Theorem 18. If p is a prime coprime with $2Q$, then the rank of p divides $p-(D/p)$. Furthermore, if p does not divide $2QD$ then

$$V_{p-(D/p)} \equiv 2Q^{(1-(D/p))/2} \pmod{p}.$$

Proof. Remember that

$$V_p + U_p\sqrt{D} = 2^{1-p}(P + \sqrt{D})^p,$$

and 2^{1-p} must be equivalent to $1 \pmod{p}$, so by the binomial theorem

$$V_p \equiv P^p \equiv P \pmod{p}$$
$$U_p \equiv D^{(p-1)/2} \equiv (D/p) \pmod{p}$$

Thus if $(D/p) = 0$, then the rank must divide p. Otherwise, this follows from the above two equations, along with the Lucas Sum and Difference formulas. □

5.1 An improvement on the Quadratic Sieve

Lemma 19. For all P, Q and odd prime p, we have that

$$V_{2i} = V_i^2 - 2Q^i$$
$$V_{2i+1} = V_iV_{i+1} - PQ^i$$

Proof. By adding the V-sum and V-difference formulas when $i = j$, we have that

$$2V_{2i} + 4Q^i = 2V_i^2$$
$$V_{2i} = V_i^2 - 2Q^i$$

When $i = j + 1$, this becomes

$$2V_{2j+1} + 2PQ^j = 2V_jV_{j+1}$$
$$V_{2j+1} = V_jV_{j+1} - PQ^j$$

□

Given V_i and V_{i+1}, we can calculate V_{2i}, V_{2i+1}, and V_{2i+2}. Thus we can calculate V_i in $O(\log i)$ calculations.

Theorem 20. If we have an odd prime p and a D such that $(D/p) = -1$ and $(Q/p) = 1$, then

$$\left(\frac{p+1}{2} \times V_{(p+1)/2}\right)^2 \equiv Q \pmod{p}$$

Proof. Since $(D/p) \neq 0$, we know that p is coprime with $2QD$. Then by Lemma 19, Theorem 18 and Euler's Criterion we have that

$$V_{p+1} \equiv 2Q \pmod{p}$$
$$V_{(p+1)/2}^2 \equiv 2Q + 2Q^{(p+1)/2} \equiv 4Q \pmod{p},$$

which is sufficient. □

By setting $Q = n$ we can solve for which values of x will when squared be equivalent to $n \pmod{p}$, provided that $(n/p) = 1$. Finding values of D is not difficult.

6 Factorization and Primality Testing using Lucas Sequences

From Theorem 18, it seems feasible to use Lucas sequences as a sequence which has pattern which is cyclic within the context of prime factors, allowing us to make a primality test as before.

Definition 21. Let $n = p_1^{a_1} \times \ldots \times p_k^{a_k}$. Then given a D such that n is relatively prime to D, we define

$$\psi(n) = 2\left(\frac{p_1 - (D/p_1)}{2}\right) \times p_1^{a_1 - 1} \times \ldots \times \left(\frac{p_k - (D/p_k)}{2}\right) \times p_k^{a_k - 1}.$$

Lemma 22. If n is relatively prime to $2QD$, then the rank* of n divides $\psi(n)$.

Proof. **Case 1.** Say first that $n = p^i$ for an odd prime p. We note that when $i = 1$ this is equivalent to Theorem 18, and we proceed by induction. It suffices to show that if p^i divides U_m then p^{i+1} divides U_{pm}. We can do this by expanding $V_{pm} + U_{pm}\sqrt{D} = (V_m + U_m\sqrt{D})^p$.

*This is the smallest e such that n divides U_e.

Case 2. Say that n has $k > 1$ unique prime factors. From Case 1, we can induct on k, by noting if m and n are relatively prime to each other and to $2QD$, and the both satisfy the condition in the problem, then because they are relatively prime to D the values of ψ are both even, and thus $\psi(m)$ and $\psi(n)$ are factors of $\psi(mn) = \psi(m)\psi(n)/2$, and thus the rank of their product divides $\psi(mn)$ as desired. □

Theorem 23. Let n be relatively prime to $2QD$ and let (D/n) be the Jacobi symbol. If the rank of n is $n-(D/n)$ then n is prime.

Proof. **Case 1.** Say that $n = p^i$. Then the rank must divide $\psi(n) = (p-(D/p))p^{i-1}$, but if the rank is $n-(D/n)$ then it must be relatively prime to p, so if n is not prime then the rank must divide $(p-(D/p)) < n-1$, contradiction.

Case 2. Say that n has at least 2 distinct prime factors. Then

$$\begin{aligned}\psi(n) &= 2\left(\frac{p_1-(D/p_1)}{2}\right)\times p_1^{a_1-1}\times\ldots\times\left(\frac{p_k-(D/p_k)}{2}\right)\times p_k^{a_k-1}\\ &\leq 2n\left(\frac{p_1+1}{2p_1}\right)\times\ldots\times\left(\frac{p_k+1}{2p_k}\right)\leq 2n\times\frac{4}{6}\times\frac{6}{10} < n-1\end{aligned}$$

□

This is useful given that sometimes the rank of n is indeed $n-(D/n)$, since then we can take $U_{(n+1)/d}$ for $d|n+1$. We will look at this later.

Breaking Bad RSA Encryption, Day 4

Michael, Mathcamp 2019

For a primality test, we look at the other portion of Theorem 18, which considers V_i, which are easier to calculate for large values, like our friend 10000!. Luckily, it can also be used to extract prime factors similarly to Pollard $p-1$.

Lemma 24. If we define $U_i(P)$ and $V_i(P)$ as the U_i and V_i generated from P and $Q=1$, then $V_{mk}(P) = V_m(V_k(P))$.

Proof. Define $P' = V_k(P)$ and $D' = P'^2 - 4$. Then by the Lucas Pythagorean Theorem we have that $D' = DU_i(P)^2$, and so

$$\begin{aligned} V_{mk}(P) + U_{mk}(P)\sqrt{D} &= 2\left(\frac{P+\sqrt{D}}{2}\right)^{mk} \\ &= 2\left(\frac{P'+\sqrt{D'}}{2}\right)^{m} \\ &= V_m(P') + U_m(P')U_k(P)\sqrt{D}. \end{aligned}$$

□

Theorem 25. Say we have a set of Lucas sequences where $Q=1$ and $(D/p) = -1$ where p is an odd prime. Then

$$V_{m(p+1)} \equiv 2 \pmod{p}.$$

Proof. We have that

$$\begin{aligned} V_{m(p+1)} + U_{m(p+1)}\sqrt{D} &= 2^{1-m}\left(2\left(\frac{P+\sqrt{D}}{2}\right)^{p+1}\right)^{m} \\ &= 2^{1-m}\left(V_{p+1} + U_{p+1}\sqrt{D}\right)^{m}, \end{aligned}$$

so by Theorem 18, binomial expansion gives us that $V_{m(p+1)} \equiv 2^{1-m} \times 2^m \equiv 1 \pmod{p}$. □

Thus we can calculate $V_{10000!} - 2$ quickly, and if we are lucky, and we have picked a D which is not a quadratic residue modulo the unknown prime, then we will have successfully factored n.

7 Groups

Definition 26. A **group** is a set of elements combined with a closed **binary operation** $\times$ such that:

- There is an **identity** e such that for all a, $a \times e = e \times a = a$,
- For all a, there is an **inverse** a^{-1} such that $a \times a^{-1} = a^{-1} \times a = e$,
- The operation $\times$ is **associative**, so $a \times (b \times c) = (a \times b) \times c$.

Definition 27. The **order** of an element a is the smallest o such that $a^o = e$.

Theorem 28. If x is in some finite G, the order of x divides the order of G.

Proof. Consider the powers of x. If we consider any other value times these powers, then they all must be distinct from each other, and from the powers of x. If we make multiple of these, they must all be pairwise distinct for the same reason, and we can partition G into these groups. □

Note that we've used the **multiplicative group** and the group whose elements are pairs (U_i, V_i) for positive integers n and the group which we call $L(D, n)$, where $Q = 1$ and n is coprime with $2D$, in which the elements are pairs a, b of values modulo n such that $a^2 - Db^2 \equiv 4 \pmod{n}$, giving us pairs (U_i, V_i). In this group, the inverse of (a, b) is $(a, -b)$.

Definition 29. A **group modulo n** is one in which every element can be described as a vector of values modulo n, and whose binary operation can be described in terms of arithmetic operations modulo n on these values. Furthermore, if G is a group modulo n, the **restricted group modulo d**, denoted $G|d$, is the same group in which the values within the vectors are taken modulo d and the arithmetic operations are similarly restricted.

Theorem 30 (Primality Testing with Group Theory). Consider a group G modulo n. If for any $p|n$ the order of any element of $G|p$ is at most $f(p)$ for some f which increases on primes, then if there is some x such that $x^m = e$ where $m > f(\lfloor \sqrt{n} \rfloor)$ and for all $q|m$ we have that some coordinate of $x^{(m/q)}$ is relatively prime to e then n is prime.

Proof. If a coordinate of $x^{(m/q)}$ is relatively prime to e, then that element is not the identity in any restricted group. Thus, in each restricted group the order of this element is m. Thus if n is composite, there is a prime at most $\lfloor \sqrt{n} \rfloor$ which has an element with an order greater than $f(\lfloor \sqrt{n} \rfloor)$, contradiction. □

Theorem 31 (Pocklington's Theorem). Let $n - 1 = FR$ where $F = p_1^{a_1} \times \ldots \times p_k^{a_k}$, F and R are relatively prime and $R < \sqrt{n}$. If for all $1 \leq i \leq k$ we have some x_i such that $x_i^{n-1} \equiv 1 \pmod{n}$ and $\gcd(x^{(n-1)/p_i}, n) = 1$, then n is prime.

Proof. In this case, $G(n)$ is the multiplicative group modulo n and $f(p) = p - 1$ for primes. From this we know from each x_i that $p_i^{a_i}$ divides $p - 1$, and thus F divides $p - 1$ and thus by Theorem 30 we are done. □

Theorem 32. Let $n + 1 = FR$ where $F = p_1^{a_1} \times \ldots \times p_k^{a_k}$, F and R are relatively prime and $F > \sqrt{n} + 1$. If we have a D such that $(D/n) = -1$ and $\gcd(n, 2D) = 1$ and set $Q = 1$, if for all $1 \leq i \leq k$ we have some x_i such that the rank of x_i divides $(n + 1)/p_i$ and $\gcd(U_{(n-1)/p_i}, n) = 1$, then n is prime.

Proof. This uses the same logic as Theorem 30. Note from the ideas in Theorem 23 that $f(n) = n + 1$. □

Theorem 33 (Prime Factorization with Group Theory). Consider a group G modulo n such that there is an x such that some prime factor of n has the order of x dividing 10000! and another has it not dividing 10000!. Then $\gcd(x^{10000!} - e, n)$ is a nontrivial divisor of n.

8 Elliptic Curves

Definition 34. An **elliptic curve** is an equation of the form $y^2 = x^3 + ax + b$, where $4a^3 + 27b \neq 0$, implying that $x^3 + ax + b$ has 3 distinct complex roots.

Importantly, if we count multiplicity then any non-vertical line intersecting two points intersects a third. We will omit the proof of this fact from this discussion, in order to arrive at the power of elliptic curves, rather than their underlying mechanics.

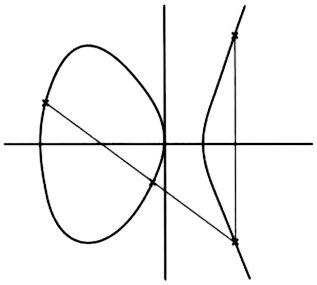

Figure 1: An elliptic curve and an example of the group action.

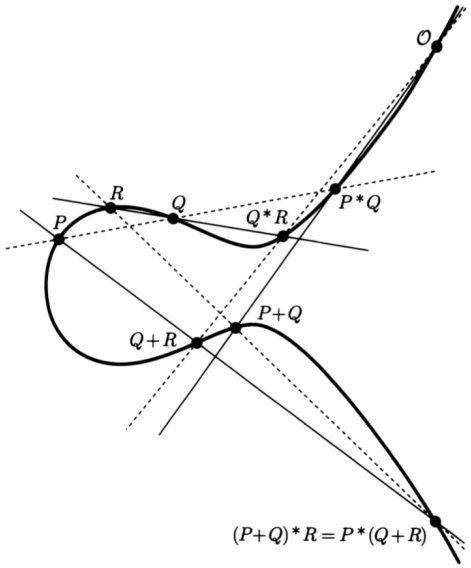

Figure 2: Multiplying the linear equations on the solid lines gives a cubic, as does doing so with the dotted lines.

Lemma 35. Let (x_1, y_1), (x_2, y_2) be two points on the curve described above, which do not lie on a vertical line ($x_1 = x_2 \Rightarrow y_1 \neq -y_2$). We then calculate the third point $(x_3, y_3) = (x_1, y_1)\partial(x_2, y_2)$ of intersection in the following manner: first, set $\lambda = (y_2 - y_1)/(x_2 - x_1)$ if $x_1 \neq x_2$, and $\lambda = (3x_1^2 + a)/2y_1$. Then

$$x_3 = \lambda^2 - x_1 - x_2$$
$$y_3 = \lambda(x_3 - x_1) + y_1$$

In order to make a group action on this set, we can move our understanding into the **projective plane**.

Definition 36. The **projective plane** is the union of the plane $\mathbb{R}^2$ with one point for each possible line acting as the single endpoint of that line. We can represent these points as triples (x, y, z), such that $(xa, ya, za) = (x, y, z)$ for all real $a \neq 0$.

With this, we have a third point on our curve, $(0, 1, 0)$, which we can denote as ∞, for any vertical line to intersect. We now claim that this is a group action. It can be verified that there is an identity ∞ and an inverse $(x, -y, z)$, it must be shown that this action is associative.

Theorem 37. The elliptic action ∂ is associative.

Proof. We can do this with any value as the identity, as long as we replace the "take the negative of the y-value" operation with "take the third point with the existing point and the identity. We will only prove the generic case where these points are distinct, but the remaining cases can be verified more easily. This gives us Figure 2, which defines ∂ as $+$ and the third point on a line as $*$. If we can show that the two points S, T it claims are equal are equal, then we are done. Note that in projective space, we can write lines in homogenous coordinates, replacing $ax + by - c = 0$ with $ax + by - cz = 0$. Thus the product of the dotted lines and the product of the solid lines can both be written as homogenous cubics with 8 common points, and the points that we want to show are equal are both shared between one of these and the original curve. We can calculate the location of the 8 points on the vector space whose basis includes cubic monomials. Since we are proving the generic case, these are linearly independent, so by rank nullity the dimension of the set of polynomials which vanishes at these points has dimension 2. Then our elliptic curve $y^2z = x^3 + axz^2 + bz^3$ can be written as the sum $f = ag + bh$ for functions g, h such that $g(S) = 0, h(T) = 0$ and constant a, b. Note that $f(S) = f(T) = 0$ and $g(T), h(S) \neq 0$, and thus $a = b = 0$, contradicting that f is not the zero polynomial. □

Breaking Bad RSA Encryption, Day 5

A View from the Top

Michael, Mathcamp 2019

Note that if we write our points in projective space as triples of integers, and then take these values modulo n, then we still have closure, identity, and inverses, so we know that we can form a group with them. We will call this the **elliptic group modulo n**, and when we take the polynomial $y^2 = x^3 + ax + b$, we will call this group $E(a,b)/n$.

Theorem 38 (Hasse, 1934). For prime p, the order of $E(a,b)/p$ lies in the interval $I(p) = (p+1-2\sqrt{p}, p+1+2\sqrt{p})$.

Theorem 39 (Waterhouse, 1969). Given a prime p larger than 3 and any $n \in I(p)$, there exist a, b such that $|E(a,b)/p| = n$. Furthermore, the orders of these curves are fairly uniformly distributed.

8.1 Primality testing with Elliptic Curves

Theorem 40. If there exists a $q > \sqrt[4]{n}+1$ such that there is some m, prime $q|m$, and element P of our elliptic group E_n such that $P^m = e$ and $P^{(m/q)} \neq e$, then n is prime.

Proof. Assume the contrary. Then there is some prime $p \leq \sqrt{n}$ that divides n. From Waterhouse, we know that the order m_p of p is less than q, and thus is relatively prime to q. Therefore there is some u such that $uq \equiv 1 \pmod{m_p}$, and thus in the restricted group modulo p we have $P^{(m_p/q)} = P^{uq(m_p/q)} = e$, contradiction. □

9 Calculating the Order of E(a,b)/p

By calculating the order of $E(a,b)/p$ for prime p, we can find the m used in Theorem 40.

9.1 Quadratic Forms

Say we want to solve in integers an equation such as $a^2 + b^2 = n$ or $a^2 + ab + b^2 = n$. We can define a set of operations which we will say produces an **equivalent** quadratic form:

- Replace x or y with $x + ky$ for integer k
- Replace x and $-x$.
- Switch x and y.

Theorem 41. All equivalent quadratic forms have the same discriminant.

Proof. This can be proven by consulting the three quadratic form operations. □

Definition 42. Every quadratic form where $D < 0 < a$ is equivalent to exactly one quadratic form $ax^2 + bxy + cy^2 = n$ such that $0 \leq b \leq a \leq c$.

Thus, given a quadratic form equivalent to $x^2 + y^2$, we can solve it and then have a solution to this equation.

Theorem 43. Given a reduced quadratic form, we have $a \leq \sqrt{|D|/3}$.

Proof. Note that we have

$$a \leq c = \frac{b^2 - D}{4a} \leq \frac{a^2 - D}{4a}.$$

□

From this, note that if our discriminant is -4 or -3, as our examples are, then our reduced form must be one of $x^2 + y^2$ or $x^2 + xy + y^2$.

Theorem 44. If D is congruent to 0 or 1 modulo p and $b^2 \equiv D \pmod{p}$, then $(b^2 - D)/4p$ is an integer and

$$px^2 + bxy + \left(\frac{b^2 - D}{4p}\right)y^2$$

is a quadratic form with discriminant D and which equals p when $x = 1, y = 0$.

Note that we can use our algorithm from our Lucas Sequence improvement of the Quadratic Sieve in order to find b.

9.2 Power Residues

To test primality, it is good to know what the order of a prime number of a given value would be.

Definition 45. A **prime number** is a number that does not divide 1 or 0 such that if it is a factor of ab, then it is a factor of either a or b. We will call a number which is prime in the integers an **ordinary prime**, but some ordinary primes will not be prime in other contexts.

Definition 46. The **Gaussian integers** $\mathbb{Z}[i]$ are values $a + bi$ for integer a, b, and a **Gaussian prime** is a number which is prime in $\mathbb{Z}[i]$.

Definition 47. The **conjugate** of $a + bi$ is $\overline{a + bi} = a - bi$, and the **norm** $N(a + bi)$ of $a + bi$ is $a^2 + b^2$.

Theorem 48. Norms are multiplicative.

Theorem 49. All primes $p = 4k + 1$ are expressible as the sum of two squares $a^2 + b^2$.

Proof. Consider the tautology

$$x^{p-1} - 1 \equiv (x^{(p-1)/2} - 1)(x^{(p-1)/2} + 1) \pmod{p}.$$

Since the left hand side has $p - 1$ roots, and the left factor has exactly $(p - 1)/2$ roots, the right factor must have at least $(p - 1)/2$ roots, thus we have $(x^k)^2 \equiv -1 \pmod{p}$ for some x.

Now note that if p is a Gaussian prime, then since $p|(x^k + i)(x^k - i)$ we have that $p|(x^k + 1)$ or $p|(x^k - 1)$, so $p(m + ni) = x^k + i$ for some m, n, so $1|pn$, contradiction. Thus p is composite in the Gaussian integers. So $(a + bi)(c + di) = p$, so since norms are multiplicative and p is an ordinary prime, so one of these has norm p, so $a^2 + b^2 = p$ or $c^2 + d^2 = p$. □

Therefore, there is some complex prime $a + bi$ which divides n.

Theorem 50. For any Gaussian prime p, exactly one of $p, ip, -p, -ip$ is 1 more than a multiple of $2 + 2i$.

Given this system of modular arithmetic, there is an equivalent to the Legendre symbols for fourth powers:

Definition 51. Consider some Gaussian prime p which does not divide 2 or $n = 4k + 1$. Then the **fourth power symbol** $(n/p)_4$ is the unique integer j such that

$$n^{(N(p)-1)/4} \equiv i^j \pmod{p},$$

which is to say that the value on the left is i^j more than a multiple of p.

Theorem 52 (Weil, 1952). Say that n is an ordinary prime in the integers which is equivalent to 1 modulo 4 and p divides n in the Gaussian integers and is congruent to 1 modulo $2 + 2i$. Then if D is not a multiple of n then

$$|E(D, 0)/n| = n + 1 - \overline{(D/p)_4} \times p - (D/p)_4 \times \overline{p}$$

We can create a similar definition for $\omega = e^{2\pi/3}$, where our new number system has norm $N(a - b\omega) = a^2 + ab + b^2$.

Theorem 53. Let n be an ordinary prime congruent to 1 modulo 3, and let p be a **cubic prime** that divides n and is congruent to 2 modulo 3. If D is not divisible by n then the order of $E(0, D)$ is

$$n + 1 + \overline{(4D/p)_6} \times p + (4D/p)_6 \times \overline{p}.$$

Acknowledgements

Thanks to J-Lo, Misha, and AnLin for making this class possible! This class was adapted from Bressoud's *Factorization and Primality Testing*, and takes its proof of the Cayley-Bacharach Theorem from MIT's 18.783 lecture notes.