YALE UNIVERSITY
DEPARTMENT OF COMPUTER SCIENCE

CPSC 467a: Cryptography and Computer Security

Notes 14 (rev. 2)
October 20, 2005

Professor M. J. Fischer

Lecture Notes 14

1 Jacobi Symbol

The *Jacobi symbol* extends the Legendre symbol to the case where the "denominator" is an arbitrary odd positive number n.

1.1 Definition

Let n be an odd positive integer with prime factorization $\prod_{i=1}^{k} {p_i}^{e_i}$. We define the *Jacobi symbol* by

$$\left(\frac{a}{n}\right) = \prod_{i=1}^{k} \left(\frac{a}{p_i}\right)^{e_i}, \tag{1}$$

where the symbol on the left is the Jacobi symbol, and the symbol on the right is the Legendre symbol. (By convention, this product is 1 when $k = 0$, so $\left(\frac{a}{1}\right) = 1$.) Clearly, when $n = p$ is an odd prime, the Jacobi symbol and Legendre symbols agree, so the Jacobi symbol is a true extension of our earlier notion.

What does the Jacobi symbol mean when n is not prime? If $\left(\frac{a}{n}\right) = -1$ then a is definitely not a quadratic residue modulo n, but if $\left(\frac{a}{n}\right) = 1$, a might or might not be a quadratic residue. Consider the important case of $n = pq$ for p, q distinct odd primes. Then

$$\left(\frac{a}{n}\right) = \left(\frac{a}{p}\right)\left(\frac{a}{q}\right)$$

so there are two cases that result in $\left(\frac{a}{n}\right) = 1$: either $\left(\frac{a}{p}\right) = \left(\frac{a}{q}\right) = +1$ or $\left(\frac{a}{p}\right) = \left(\frac{a}{q}\right) = -1$. In the first case, a is a quadratic residue modulo both p and q, so a is a quadratic residue modulo n. In the second case, a is not a quadratic residue modulo either p or q, so it is *not* a quadratic residue modulo n, either. Such numbers a are sometimes called "pseudo-squares" since they have Jacobi symbol 1 but are not quadratic residues.

1.2 Identities

The Jacobi symbol is easily computed if the factorization of n is known using Equation 1 above and Theorem 1 of lecture 13, section 3. Similarly, $\gcd(u, v)$ is easily computed given the factorizations of u and v, without resort to the Euclidean algorithm. The remarkable fact about the Euclidean algorithm is that it lets us compute $\gcd(u, v)$ efficiently even without knowing the factors of u and v. A similar algorithm allows the Jacobi symbol $\left(\frac{a}{n}\right)$ to be computed efficiently without knowing the factorization of a or n.

The algorithm is based on identities satisfied by the Jacobi symbol:

1. $\left(\frac{0}{1}\right) = 1$; $\left(\frac{0}{n}\right) = 0$ for $n \neq 1$;
2. $\left(\frac{2}{n}\right) = 1$ if $n \equiv \pm 1 \pmod 8$; $\left(\frac{2}{n}\right) = -1$ if $n \equiv \pm 3 \pmod 8$;

3. $\left(\frac{a_1}{n}\right) = \left(\frac{a_2}{n}\right)$ if $a_1 \equiv a_2 \pmod{n}$;
4. $\left(\frac{2a}{n}\right) = \left(\frac{2}{n}\right)\left(\frac{a}{n}\right)$;
5. $\left(\frac{a}{n}\right) = -\left(\frac{n}{a}\right)$ if $a \equiv n \equiv 3 \pmod{4}$.
6. $\left(\frac{a}{n}\right) = \left(\frac{n}{a}\right)$ if $a \equiv 1 \pmod{4}$ or ($a \equiv 3 \pmod{4}$ and $n \equiv 1 \pmod{4}$);

There are many ways to turn these identities into an algorithm. Below is a straightforward recursive approach. Slightly more efficient iterative implementations are also possible.

```
int jacobi(int a, int n)
/* Precondition: a, n >= 0; n is odd */
{
  if (a == 0)                        /* identity 1 */
    return (n==1) ? 1 : 0;
  if (a == 2) {                      /* identity 2 */
    switch (n%8) {
    case 1:
    case 7:
      return 1;
    case 3:
    case 5:
      return -1;
    }
  }
  if ( a >= n )                      /* identity 3 */
    return jacobi(a%n, n);
  if (a%2 == 0)                      /* identity 4 */
    return jacobi(2,n)*jacobi(a/2, n);
  /* a is odd */                     /* identities 5 and 6 */
  return (a%4 == 3 && n%4 == 3) ? -jacobi(n,a) : jacobi(n,a);
}
```

2 Strassen-Solovay Test of Compositeness

Recall that a test of compositeness for n is a set of predicates $\{\tau_a(n)\}_{a\in\mathbf{Z}_n^*}$ such that if $\tau(n)$ succeeds (is true), then n is composite. The Strassen-Solovay Test is the set of predicates $\{\nu_a(n)\}_{a\in\mathbf{Z}_n^*}$, where

$$\nu_a(n) = \textsf{true} \text{ iff } \left(\frac{a}{n}\right) \not\equiv a^{(n-1)/2} \pmod{n}.$$

If n is prime, the test always fails by Theorem 1 of lecture 13, section 3. Equivalently, if some $\nu_a(n)$ succeeds, then n must be composite. Hence, the test is a valid- test of compositeness.

Let $b = a^{(n-1)/2}$. There are two possible reasons why the test might succeed. One possibility is that $b^2 \equiv a^{n-1} \not\equiv 1 \pmod{n}$ in which case $b \not\equiv \pm 1 \pmod{n}$. This is just the Fermat test $\zeta_a(n)$ from section 1.4 of lecture notes 11. A second possibility is that $a^{n-1} \equiv 1 \pmod{n}$ but nevertheless, $b \not\equiv \left(\frac{a}{n}\right) \pmod{n}$. In this case, b is a square root of $1 \pmod{n}$, but it might have the opposite sign from $\left(\frac{a}{n}\right)$, or it might not even be ± 1 since 1 has additional square roots when n is composite. We claim without proof that for some constant $c > 0$ and all composite numbers n, the probability that $\nu_a(n)$ succeeds for a randomly-chosen $a \in \mathbf{Z}_n^*$ is at least c. I believe that $c \geq 1/4$, but this fact must be checked.

3 Miller-Rabin Test of Compositeness

The Miller-Rabin Test is more complicated to describe than the Solovay-Strassen Test, but the probability of error (that is, the probability that it fails when n is composite) seems to be lower than for Solovay-Strassen, so that the same degree of confidence can be achieved using fewer iterations of the test. This makes it faster when incorporated into a primality-testing algorithm. It is also closely related to the algorithm presented in lecture notes 12, section 1.3 for factoring an RSA modulus given the encryption and decryption keys.

3.1 The test

The test $\mu_a(n)$ is based on computing a sequence $b_0, b_1, \ldots, b_k$ of integers in $\mathbf{Z}_n^*$. If n is prime, this sequence ends in 1, and the last non-1 element, if any, is $n-1$ ($\equiv -1 \pmod{n}$). If the observed sequence is *not* of this form, then n is composite, and the Miller-Rabin Test succeeds. Otherwise, the test fails.

The sequence is computed as follows:

1. Write $n - 1 = 2^k m$, where m is an odd positive integer. Computationally, k is the number of 0's at the right (low-order) end of the binary expansion of n, and m is the number that results from n when the k low-order 0's are removed.

2. Let $b_0 = a^m \bmod n$.

3. For $i = 1, 2, \ldots, k$, let $b_i = (b_{i-1})^2 \bmod n$.

An easy inductive proof shows that $b_i = a^{2^i m} \bmod n$ for all i, $0 \leq i \leq k$. In particular, $b_k \equiv a^{2^k m} = a^{n-1} \pmod{n}$.

3.2 Validity

To see that the test is valid, we must show that $\mu_a(p)$ fails for all $a \in \mathbf{Z}_p^*$ when p is prime. By Euler's theorem[1], $a^{p-1} \equiv 1 \pmod{p}$, so we see that $b_k = 1$. Since 1 has only two square roots, 1 and -1, modulo p, and b_{i-1} is a square root of b_i modulo p, the last non-1 element in the sequence (if any) must be $-1 \bmod p$. This is exactly the condition for which the Miller-Rabin test fails. Hence, it fails whenever n is prime, so if it succeeds, n is indeed composite.

3.3 Accuracy

How likely is it to succeed when n is composite? It succeeds whenever $a^{n-1} \not\equiv 1 \pmod{n}$, so it succeeds whenever the Fermat test $\zeta_a(n)$ would succeed. (See lecture notes 11, section 1.4.) But even when $a^{n-1} \equiv 1 \pmod{n}$ and the Fermat test fails, the Miller-Rabin test will succeed if the last non-1 element in the sequence of b's is one of the square roots of 1 other than ± 1. It can be proved that $\mu_a(n)$ succeeds for at least $3/4$ of the possible values of a. Empirically, the test almost always succeeds when n is composite, and one has to work to find a such that $\mu_a(n)$ fails.

[1]This is also called Fermat's little theorem.

3.4 Example

For example, take $n = 561 = 3 \cdot 11 \cdot 17$. This number is interesting because it is the first Carmichael number. A *Carmichael number* is an odd composite number n that satisfies $a^{n-1} \equiv 1 \pmod{n}$ for all $a \in \mathbf{Z}_n^*$. (See `http://mathworld.wolfram.com/CarmichaelNumber.html`.) These are the numbers that I have been calling "pseudoprimes". Let's go through the steps of computing $\mu_{37}(561)$.

We begin by finding m and k. 561 in binary is 1000110001 (a palindrome!). Then $n - 1 = 560 = (1000110000)_2$, so $k = 4$ and $m = (100011)_2 = 35$. We compute $b_0 = a^m = 37^{35} \bmod 561 = 265$ with the help of the computer. We now compute the sequence of b's, also with the help of the computer. The results are shown in the table below:

i	b_i
0	265
1	100
2	463
3	67
4	1

This sequence ends in 1, but the last non-1 element $b_3 \not\equiv -1 \pmod{561}$, so the test $\mu_{37}(561)$ succeeds. In fact, the test succeeds for every $a \in \mathbf{Z}_{561}^*$ except for $a = 1, 103, 256, 460, 511$. For each of those values, $b_0 = a^m \equiv 1 \pmod{561}$.

3.5 Optimization

In practice, one only wants to compute as many of the b's as necessary to determine whether or not the test succeeds. In particular, one can stop after computing b_i if $b_i \equiv \pm 1 \pmod{n}$. If $b_i \equiv -1 \pmod{n}$ and $i < k$, the test fails. If $b_i \equiv 1 \pmod{n}$ and $i \geq 1$, the test succeeds. This is because we know in this case that $b_{i-1} \not\equiv -1 \pmod{n}$, for if it were, the algorithm would have stopped after computing b_{i-1}.

4 Digital Signatures

4.1 Definition

A *digital signature* is a string attached to a message that is used to guarantee the integrity and authenticity of the message. It is very much like the message authentication codes (MACs) discussed in lecture notes 7, section 1. Recall that Alice can protect a message m (encrypted or not) by attaching a MAC $\xi = C_k(m)$ to the message m. The pair (m, ξ) is an *authenticated message*. To produce a MAC requires possession of the secret key k. To verify the integrity and authenticity of m, Bob, who also must know k, checks a received pair (m', ξ') by verifying that $\xi = C_k(m)$. Assuming Alice and Bob are the only parties who share k, then Bob knows that either he or Alice must have sent the message.

A digital signature can be viewed as a 2-key MAC, just as a public key cryptosystem is a 2-key version of a classical cryptosystem. The basic idea is the same. Let $\mathcal{M}$ be a *message space* and $\mathcal{S}$ a *signature space*. A *signature scheme* consists of a private *signing key* d, a public *verification key* e, a *signature function* $S_d : \mathcal{M} \to \mathcal{S}$, and a *verification predicate* $V_e \subseteq \mathcal{M} \times \mathcal{S}$.[2] A *signed message*

[2] As with RSA, we denote the private component of the key pair by the letter d and the public component by the letter e, although they no longer have same same mnemonic significance.

is a pair $(m, s) \in \mathcal{M} \times \mathcal{S}$. A signed message is *valid* if $V_e(m, s)$ holds, and we say that (m, s) is *signed with* e.

The basic property of a signature scheme is that the signing function always produces valid signatures, that is,

$$V_e(m, S_d(m)) \tag{2}$$

always holds. Assuming d is Alice's private signing key, and only Alice knows d, then a valid message signed with Alice's key d identifies her with m (possibly erroneously, as we shall see).