6.875: Solutions for Homework 2

Yevgeniy Dodis

Problem 1

Let n be a Blum integer, i.e. $n = p_1 p_2$ where p_1 and p_2 are distinct primes s.t. $p_1 \equiv p_2 \equiv 3 \bmod 4$. Let $QR(n)$ be the subgroup of $\mathbb{Z}_n^*$ of all quadratic residues modulo n. Show that $f(x) = x^2 \bmod n$ is a *bijection* from $QR(n)$ onto $QR(n)$.

Solution: We have to show that each square $a \in QR(n)$ has exactly one of its square roots that is also a square. Let $a = (a_1, a_2)$ (as usual, $a_i \equiv a \bmod p_i$). Let $y = (y_1, y_2)$ be a square root of a with $\left(\frac{y_i}{p_i}\right) = c_i$. We know then that the the 4 square roots of a have the form $(b_1 y_1, b_2 y_2)$, where $b_1, b_2 \in \{\pm 1\}$. Observe that we always have $\left(\frac{1}{p_i}\right) = 1$ but, since $p_i \equiv 3 \bmod 4$, we have that $\left(\frac{-1}{p_i}\right) = (-1)^{(p_i-1)/2} = -1$. Hence, for $b \in \{\pm 1\}$, $\left(\frac{b}{p_i}\right) = b$.

Now, $(b_1 y_1, b_2 y_2)$ is a square if and only if $\left(\frac{b_1 x_1}{p_1}\right) = \left(\frac{b_2 x_2}{p_2}\right) = 1$, i.e. $c_i \left(\frac{b_i}{p_i}\right) = 1$, i.e. $b_i c_i = 1$, $i = 1, 2$. Hence, as b_i's range through ± 1, the pairs of Jacobi symbols of square roots of a w.r.t. p_1 and p_2 range through the set $\{(c_1, c_2), (-c_1, c_2), (c_1, -c_2), (-c_1, -c_2)\}$. No matter what $c_1, c_2 \in \{\pm 1\}$ are, exactly one of these pairs is $(1, 1)$, i.e. exactly one of the square roots is a square.

Problem 2

Let n be an odd integer such that there are $x, y \in \mathbb{Z}_n^*$ satisfying $x^2 = y^2 \bmod n$ but $\left(\frac{x}{n}\right) \neq \left(\frac{y}{n}\right)$ (recall that $\left(\frac{z}{n}\right)$ is the Jacobi symbol of z). Prove that for every quadratic residue in $\mathbb{Z}_n^*$ *exactly* half of its square roots have Jacobi symbol 1.

Solution: Assume we can find z s.t. $z^2 = 1$ but $\left(\frac{z}{n}\right) = -1$. Then I claim that we are done. Indeed, take any quadratic residue w. Let us say that 2 square roots x_1 and x_2 of w are equivalent if $x_1 = z x_2$ (as $z^2 = 1$, this is an equivalence relation where every class has size 2). In other words, we split the square roots of w into disjoint pairs (x_1, x_2) where $x_1 = z x_2$, $x_2 = z x_1$. In each pair, $\left(\frac{x_2}{n}\right) = \left(\frac{z x_1}{n}\right) = -\left(\frac{x_1}{n}\right)$, so one element has Jacobi symbol 1, and the other has Jacobi symbol -1. Since the pairs disjointly split the roots of w, the claim follows. The question is how to construct the needed z.

An easy solution giving little insight is the following. Let $z = xy^{-1}$. As $x^2 = y^2 \bmod n$, we have $z^2 = 1 \bmod n$. On the other hand, $\left(\frac{z}{n}\right) = \left(\frac{x}{n}\right) / \left(\frac{y}{n}\right) = -1$ as the Jacobi symbols of x and y are distinct.

A harder but more insightful solution is the following. Let $n = \prod_{i=1}^{k} p_i^{\alpha_i}$. Set $q_i = \pi^{\alpha_i}$. let $x = (x_1, \ldots, x_k)$, $y = (y_1, \ldots, y_k)$. Since $x^2 = y^2 \bmod n$, we know from homework#1 that each $y_i = b_i x_i$ where $b_i \in \{\pm 1\}$. Then

$$-\left(\frac{x}{n}\right) = \left(\frac{y}{n}\right) = \prod_{i=1}^{k} \left(\frac{b_i}{q_i}\right)\left(\frac{x_i}{q_i}\right) = \left(\frac{x}{n}\right) \prod_{i=1}^{k} \left(\frac{b_i}{q_i}\right)$$

implying that $\prod_{i=1}^{k} \left(\frac{b_i}{q_i}\right) = -1$, so that there is some $1 \le j \le k$ s.t. $\left(\frac{b_j}{q_j}\right) = -1$. Clearly, $b_j = -1$ as $\left(\frac{1}{q_i}\right) = 1$, so that $\left(\frac{-1}{q_j}\right) = -1$. Now define $z = (1, \ldots, 1, -1, 1, \ldots 1)$ where the only -1 is at position j. This z is a square root of 1. On the other hand, $\left(\frac{z}{n}\right) = \left(\frac{-1}{q_j}\right) = -1$.

Problem 3

Let $SQRT(a, p)$ be (probabilistic) polynomial time algorithm that given a *prime* p and a *quadratic residue* $a \in Z_p^*$, outputs (w.h.p.) some square root of a mod p. Prove that the following algorithm is a primality test (i.e. that, on input a prime outputs PRIME with probability 1, and that, on input a composite, outputs COMPOSITE with probability at least 1/2). This test is due to Silvio Micali.

1. If $n = 2$, output PRIME.
2. If n is even or of the form $n = p^b$ for some $p, b > 1$, output COMPOSITE.
3. Choose random $r \in \{1 \ldots (n-1)\}$. If $gcd(r, n) > 1$, output COMPOSITE.
4. Compute $x \leftarrow SQRT(r^2 \bmod n, n)$.
5. If $x \equiv r \bmod n$ or $x \equiv -r \bmod n$ then output PRIME, else output COMPOSITE.

Solution: If n is prime, the algorithm is always correct since the only square roots of r^2 are $\pm r$ and we assume that $SQRT$ works correctly when n is a prime.

If n is composite, then we claim that the probability the algorithm outputs COMPOSITE is at least 1/2. Indeed, if n is divisible by fewer than 2 distinct *odd* primes, we will catch it in the second step (with probability 1). Therefore, we only need to examine the case when n is divisible by at least 2 distinct odd primes. If (with a negligible probability) r is not relatively prime with n, we will catch it in step 3. So it remains to show that when r is a random element of $\mathbb{Z}_n^*$, $SQRT$ will return $x = \pm r$ with probability at most 1/2.

From homework#1, as n is divisible by at least 2 distinct odd primes, $x = r^2 \bmod n$ will always have *at least* $2^2 = 4$ distinct square roots. Even though $SQRT$ is not supposed to work when n is composite, we have to assume the worst, i.e. that $SQRT$ might maliciously try to return $\pm r$. However, since r is chosen at random from $\mathbb{Z}_n^*$, it can be any one of at least 4 square roots of $r^2 \bmod n$. Thus, the $SQRT$ will return $x = \pm r$ with probability at most 1/2 indeed.

We remark that, aside from the fact that we assumed that $SQRT$ never makes an error when given correct inputs, the test can indeed be implemented in (probabilistic) polynomial time, since we know how to extract square roots when n is prime. It is true that if SQRT worked correctly even when n was composite, existence of $SQRT$ would imply the existence of the factoring algorithm, so not only we would be able to test primality, but to factor (which is unlikely). However, for the correctness of the test we don't need $SQRT$ to work correctly for composite n. In fact, the worse $SQRT$ performs in this case, the better is the probability of success!

Problem 4

Let us denote by O_n an oracle that, on input a quadratic residue mod n, returns one of its square roots mod n. Show that there is a polynomial time algorithm that, on input n and access to the oracle O_n (but not to any oracle O_m for $m \neq n$), outputs the *complete* factorization of n. Be sure to prove that your algorithm works.

Solution: I had a challenge from one of your fellow students to fit it in one page. Sorry, I failed...
Let $\beta = O(\log\log n)$ with large enough constant (10 will do).

(1) Express $n = 2^t n'$, where n' is odd.

(2) For $j = 1 \ldots \beta$ repeat:
- Pick random $r \in \mathbb{Z}_n^*$. Run $x \leftarrow O_n(r^2)$.
- Let $f_j = gcd(r - x, n')$.

(3) For any of $2^\beta = poly(\log n)$ non-empty subsets $S \subseteq \{1 \ldots \beta\}$ compute the GCD $g(S)$ of the set $\{f_j | j \in S\}$. Test if $g(S) = s^\alpha$, for $s > 1$. If yes, add s to the set P.

(4) Remove from P all the elements that are divisible by some smaller element of P, i.e. leave "minimal" elements only.

(5) For each $p \in P$ extract the largest power $\alpha(p)$ s.t. $p^{\alpha(p)} | n'$.

(6) Output $n = 2^t \prod_{p \in P} p^{\alpha(p)}$.

First, the running time is polynomial in 2^β, i.e. polynomial in $\log n$ which is what we can afford. Assume $n' = \prod_{i=1}^k q_i$, where $q_i = p_i^{\alpha_i}$ and p_i are odd. Since t is found correctly, we only need to worry about odd factors. Also, even though k is unknown, we know that $k < \log n$, so $\beta = \Omega(\log k)$ with large enough constant in front.

Let's analyze Step (2) first. We know from homework#1 that there are at least 2^k square root of r^2 modulo n. Each of them comes from a square root modulo 2^t (at least one, $r \bmod 2^t$, exists) and one of exactly 2^k square roots modulo n'. We also know that if $r' = r \bmod n' = (r_1, \ldots, r_k)$, then each one of these 2^k square roots is of the form $x_{\vec{b}} = (b_1 r_1, \ldots, b_k r_k)$, $b_i \in \{\pm 1\}$. Given any set $T \subseteq \{1 \ldots k\}$, we let $x(T) = x_{\vec{b}}$, where $b_i = 1$ if $i \in T$, and $b = -1$ otherwise. We observe that

$$gcd(r' - x(T), n') = \prod_{i \in T} q_i \tag{1}$$

Indeed, let $z = r' - x(T)$. Then $z \equiv (1 - b_i) r_i \bmod q_i$ and since $r_i \in \mathbb{Z}_{q_i}^*$ and p_i is odd, $gcd(z, q_i) = q_i$ if $b_1 = 1$ (i.e. $i \in T$) and is 1 otherwise, justifying (1). Let $x' = x \bmod n'$. We clearly have that $f_j = gcd(r - x, n') = gcd(r' - x', n')$. But since r was random in $\mathbb{Z}_n^*$, $r' = r \bmod n'$ is random in $Z_{n'}^*$, so that no matter how O_n computes x we have that $x = x(T)$ for a *random* subset T of $\{1 \ldots k\}$. Let us denote by $n(T) = \prod_{i \in T} q_i$. Hence, Step (2) is equivalent to *choosing β random subsets $T_1, \ldots, T_\beta$ of $\{1 \ldots k\}$ and letting $f_j = n(T_j)$ be the product of q_i for $i \in T_j$.*

We remark that picking a set S in step (3) and letting $g(S)$ be the gcd of $\{f_j | j \in S\} = \{n(T_j) | j \in S\}$ means that

$$g(S) = n\Big(\bigcap_{j \in S} T_j\Big) = \prod_{i \in \cap_{i \in S} T_j} q_i$$

Hence, a trivial inspection of Steps (3)-(6) tells us that our algorithm will correctly find the complete factorization of n if with high probability *each "singleton" set $\{i\}$, $i = 1 \ldots k$, is equal to the intersection of some of the β* **randomly chosen** *subsets $T_1, \ldots, T_\beta$ of $\{1 \ldots k\}$*. Indeed, in this case for each $i = 1 \ldots k$, some $g(S) = q_i = p_i^{\alpha_i}$, so P contains all p_i's yielding a complete factorization of n.

Take any $1 \leq i \leq k$. Define (a random variable) set $S_i = \{j | i \in T_j\}$, $D_i = \cap_{j \in S_i} T_j$, so that $i \in D_i$ by construction. In other words, S_i has all the indices j s.t. i was chosen in T_j.

Claim 1

$$\Pr(\text{all } |S_i| > 3 \log\log n) \geq 1 - \frac{1}{\log n}$$

Proof: Each T_j contains i with probability $1/2$, so we expect half of T_j to contain i. By Chernoff bound and our choice of β, with probability at least $1 - 1/\log^2 n$, S_i has size $\beta/3 > 3 \operatorname{loglog} n$, so with probability at least $1 - 1/\log n$ (by union bound) all S_i have size at least $3 \operatorname{loglog} n$. $\square$

Claim 2

$$\Pr(\text{all } D_i = \{i\} \mid \text{all } |S_i| > 3 \operatorname{loglog} n) \geq 1 - \frac{1}{\log n}$$

Proof: We know by definition that $i \in D_i$. Take any i and any $i' \neq i$. Then $i' \in D_i$ iff $i' \in T_j$ for all $j \in S_i$. Since $i' \neq i$ and each T_j contains i' with probability $1/2$, this happens with probability $(1/2)^{|S_i|} < 1/\log^3 n$, since $|S_i| > 3 \operatorname{loglog} n$. Taking the union bound over i and i' ($k^2 < \log^2 n$ choices) yields the desired result. $\square$

The above two claims show that

$$\Pr(\text{all } D_i = \{i\}) \quad \geq \quad \Pr(\text{all } |S_i| > 3 \operatorname{loglog} n) \cdot \Pr(\text{all } D_i = \{i\} \mid \text{all } |S_i| > 3 \operatorname{loglog} n) \geq \left(1 - \frac{1}{\log n}\right)^2$$

Thus, with probability at least $(1 - 1/\log n)^2 > 1 - 2/\log n$ we have that $D_i = \{i\}$ for all $i = 1 \ldots k$, so indeed every singleton set is an intersection of some of T_j's (for $j \in S_i$), so that w.h.p. our algorithm obtains all prime factors of n.

Problem 5

Assume that there exists an algorithm A that, on input a Blum integer n (See Problem 1) and a pair of numbers $a, b \in \mathbb{Z}_n^*$, where $\left(\frac{a}{n}\right) = \left(\frac{b}{n}\right) = 1$ and $a \equiv -b \bmod n$, with probability $> .51$ outputs the element of the pair which is a quadratic residue mod n. (For each n, this probability is taken only over the uniform choices of the input pairs and the random coins of A).

Design another probabilistic algorithm B that, given any number $z \in \mathbb{Z}_n^*$ with $\left(\frac{z}{n}\right) = 1$, determines whether or not z is a quadratic residue with probability $> .99$. (Here, for each n and z, the probability is taken only over the random coins of B.)

Solution:

Lemma 1 *If n is a Blum integer, then -1 is a quadratic non-residue with Jacobi symbol 1. Moreover, for any $a \in \mathbb{Z}_n^*$, if $\left(\frac{a}{n}\right) = 1$, exactly one of a and $-a$ is a quadratic residue.*

Proof: We saw that $\left(\frac{-1}{p_1}\right) = \left(\frac{-1}{p_2}\right) = -1$, so -1 is not a quadratic residue, but $\left(\frac{-1}{n}\right) = (-1)(-1) = 1$. Take any a with $\left(\frac{a}{n}\right) = 1$. Then $1 = \left(\frac{a}{n}\right) = \left(\frac{a}{p_1}\right)\left(\frac{a}{p_2}\right)$, so that $\left(\frac{a}{p_1}\right) = \left(\frac{a}{p_2}\right) = \alpha$, where $\alpha \in \{\pm 1\}$. As n is Blum integer, $\left(\frac{-a}{p_1}\right) = \left(\frac{-a}{p_2}\right) = -\alpha$. If $\alpha = 1$ then a is a square but $-a$ is not. If $\alpha = -1$, then $-a$ is a square and a is not. $\square$

This lemma also justifies why A makes sense. Indeed, *exactly* one of a, $b = -a$ will be a quadratic residue so A indeed has a non-trivial choice to make.

We first construct B' from A s.t. for any $z \in \mathbb{Z}_n^*$, B' determines if $z \in QR(n)$ with probability at least 0.51, where the probability is only over the random choices of B'. Once we have B', the rest is trivial. Simply repeat B' independently enough (say 10000) times and take the majority vote. As each run is independent, the resulting B is what we are looking for.

To construct B' we want to exploit our lemma, i.e. the fact that we *know* a quadratic non-residue with Jacobi symbol 1. Recall that we had a lot of complications for the general case when we did not know such a quadratic non-residue (we had to estimate certain probabilities by sampling, etc.). Now everything will be very clean; here is B' on input z.

1. Pick a random $x \in \mathbb{Z}_n^*$, let $a' = zx^2 \bmod n$, $b' = -a' = -zx^2 \bmod n$.

2. Flip a coin u. If $u = 0$ then let $c = A(a', b')$, else let $c = A(b', a')$.

3. If $c = a'$ return "yes", else return "no".

We claim that B' is correct for *any* z with exactly the same probability that A is for a *random* pair (a, b), i.e. at least 0.51. Indeed, $a' = zx^2$ is a random element of $\mathbb{Z}_n^*$ with Jacobi symbol 1 and the same quadratic residuosity as z. $b' = -zx^2$ is a random element of $\mathbb{Z}_n^*$ with Jacobi symbol 1 and the opposite quadratic residuosity with z. Since the first argument to A is either a' or b' with probability 1/2, we feed A with a random pair (a, b) where $a = -b$ and Jacobi symbol of a is 1. Thus, A returns the correct answer with probability at least 0.51. Since we ought to output "yes" if and only if a' is indeed a quadratic residue, we are correct w/pr at least 0.51 indeed. We note that it is crucial to flip a coin, i.e. to feed a' and b' in random order. Otherwise, the "men/men" vs. "women/women" paradox is possible: B' could be correct on quadratic residues w/pr .53, but on quadratic non-residues – with probability 0.49, even though the average of the two is 0.51.