FINAL EXAM

1. (a) Describe the XOR cipher. **(2 points)**

   (b) Prove that the number of those keys in the 23-bit XOR cipher which contain at least 7 and at most 16 zeros is divisble by 23. **(4 points)**

   Solution. (a) In the XOR cipher, we fix a positive integer t, and then

   $$\mathcal{M} = \mathcal{C} = \mathcal{K} = \{0-1 \text{ sequences of length } t\}.$$

   We define the $\oplus$ operation as the bitwise addition, i.e. if $a = \sum_{j=0}^{t-1} a_j 2^j$, $b = \sum_{j=0}^{t-1} b_j 2^j$ (where a_j, b_j's are binary digits, 0 or 1 each), then let

   $$a \oplus b = \sum_{j=0}^{t-1} c_j 2^j,$$

   where $c_j = 0$ if $a_j = b_j$, and $c_j = 1$ if $a_j \neq b_j$.

   Given m and k, $e_k(m) = m \oplus k$. The decryption function is the same: $d_k = e_k$, i.e. $d_k(c) = c \oplus k$.

   (b) From elementary enumeration, we know that the number of keys containing k zeros is

   $$\binom{23}{k} = \frac{23!}{k!(23-k)!}.$$

   Clearly, when $0 < k < 23$, then this is divisible by 23, since 23 is a prime. Then

   $$\binom{23}{7} + \ldots + \binom{23}{16}$$

   is divisible by 23, since each term is divisible by 23.

2. (a) Describe the RSA cryptosystem. **(2 points)**

   (b) Assume Eve has a machine which, for any input (a, b, N) (with positive integers a, b, N), returns in polynomial time

$$\begin{cases} 1, \text{ if there exists } d \mid N \text{ such that } a < d < b, \\ 0, \text{ if there is no } d \mid N \text{ satisfying } a < d < b. \end{cases}$$

   Prove that using this machine, Eve can break the RSA in polynomial time. **(4 points)**

Solution. (a) Alice takes two (large) prime numbers p, q, then computes their product N. She also computes $\varphi(N) = (p-1)(q-1)$. Then she takes an exponent $e \in \mathbf{N}$ coprime to $\varphi(N)$, and computes its inverse d modulo $\varphi(N)$. She publishes N, e and keeps $p, q, \varphi(N), d$ in secret.

Now anyone (say, Bob) can send her a message m (a residue class modulo N) using the following protocol. Bob raises the message to power e modulo N and sends $c \equiv m^e \bmod N$ to Alice.

Now Alice raises the incoming cipher c to power d modulo N. With high probability, m is coprime to N, and then, by Euler-Fermat,

$$c^d \equiv (m^e)^d \equiv m^{\varphi(N)u+1} \equiv (m^{\varphi(N)})^e \cdot m \equiv 1 \cdot m \equiv m \bmod N,$$

which is the original message.

(b) Let N be as in RSA. The machine combined with binary search captures a divisor of N in polynomial time. Indeed, set $a_0 = 1, b_0 = N$, of course $M(a_0, b_0, N) = 1$ (where M is the result of the machine). In each step, we take $c_n = \lfloor (a_n + b_n)/2 \rfloor$, and if $M(a_n, c, N) = 1$, then we set $a_{n+1} = a_n$, $b_{n+1} = c_n$, while if $M(a_n, c, N) = 0$, then we set $a_{n+1} = c_n - 1$, $b_{n+1} = b_n$. Clearly (a_n, b_n) will always contain a divisor of N, and $b_{n+1} - a_{n+1} \leqslant (b_n - a_n)/2 + 2$, i.e. in, say, $O(\log N)$ steps (with the machine), the interval containing a divisor of N gets smaller than 10. In an interval of length 10, we can easily find a divisor in polynomial time. Obtaining hence p or q, we get the factorization of N, and can play the role of Alice then.

3. (a) Define entropy. **(2 points)**
   (b) Alice and Bob use an XOR cipher on t bits, and they choose the message and the key independently and uniformly (i.e. for each t-bit sequences m and k, $P(M = m) = 2^{-t}$, $P(K = k) = 2^{-t}$, $P(M = m, K = k) = 2^{-2t}$). Compute the key equivocation $H(K \mid C)$. **(4 points)**

Solution. (a) The entropy function H is defined on finite sets of positive numbers summing up to 1, i.e. on tuples $(p_1, \dots, p_n) \in \mathbf{R}_+^n$ if $p_1 + \ldots + p_n = 1$, for any $n \in \mathbf{N}$. For such a tuple,

$$H(p_1, \dots, p_n) = -\sum_{j=1}^{n} p_j \log_2 p_j.$$

(b) We proved in the lecture that if M and K are independent, then

$$H(K \mid C) = H(M) + H(K) - H(C).$$

We know that M, K are uniform distributions on 2^t elements. It is easy to see that this also holds for C as well: each single cipher is obtained 2^t ways out of the 2^{-2t} choices for (m, k), hence each cipher is obtained with probability 2^{-t}. For uniform distributions on 2^t elements, the entropy is $\log_2(2^t) = t$. Then

$$H(K \mid C) = H(M) + H(K) - H(C) = t + t - t = t.$$

4. (a) Describe the elliptic curve ElGamal cryptosystem. **(2 points)**

   (b) Let E be the elliptic curve given by the equation $y^2 = x^3 + x + 1$ over the field $\mathbf{F}_5$. Show that the points $P = (4, 2)$ and $Q = (3, 4)$ lie on E, and solve the elliptic curve discrete logarithm problem $nP = Q$. (It is enough to give one such n, you don't have to compute all of them.) **(4 points)**

Solution. (a) Alice chooses a prime number $p > 3$, an elliptic curve E over the prime field $\mathbf{F}_p$, and a point P on the elliptic curve. She further chooses a positive integer n_A, and computes the point

$$Q = n_A P = \underbrace{P + \ldots + P}_{n_A \text{ many}}.$$

Now she publishes p, E, P, Q and keeps n_A in secret.

Anyone (say, Bob) can send her a message M (a point on the elliptic curve) using the following protocol. Bob chooses an ephemeral key $k \in \mathbf{N}$, and computes

$$C_1 = kP, \qquad C_2 = M + kQ.$$

Then he sends the pair (C_1, C_2) to Alice.

Now Alice computes $C_2 - n_A C_1$, obtaining

$$C_2 - n_A C_1 = M + kQ - n_A kP = M + kn_A P - n_A kP = M,$$

which is the original message.

(b) Clearly $2^2 \equiv 4^3 + 4 + 1 \bmod 5$, $4^2 \equiv 3^3 + 3 + 1 \bmod 5$ hold, hence P, Q are indeed on E.

We compute $2P$. From the lecture, we know that the slope of the tangent line at $P = (x_P, y_P)$ is $(3x_P^2 + 1)/(2y_P)$, which is 1 in our case, therefore the tangent line is $y = x + 3$. We need hence the third solution of the system $y^2 = x^3 + x + 1$ and $y = x + 3$. Writing $y = x + 3$ into the cubic one,

$$\begin{aligned} (x+3)^2 &= x^3 + x + 1, \\ 0 &= x^3 - x^2 + 2, \\ 0 &= (x+1)^2(x+2). \end{aligned}$$

Then the third intersection point is $(3, 1)$, hence $2P = (3, 4)$. This is just Q, so $n = 2$ is a solution.