Homework 2

1. **Defining Multiplication over $\mathbb{Z}_{27}^*$.** In the class, we had considered the group $(\mathbb{Z}_{26}, +)$ to construct a one-time pad for one alphabet messages. A few students were interested to define a group with 26 elements using a "multiplication"-like operation. This problem will assist you to define the $(\mathbb{Z}_{27}^*, \times)$ group.

   Interpret $\mathbb{Z}_{27}^*$ as the set of all triplets (a_0, a_1, a_2) such that $a_0, a_1, a_2 \in \mathbb{Z}_3$ and at least one of them is non-zero (you can think of the triplets as the ternary representation of the elements in $\mathbb{Z}_{27}^*$). We shall equivalently interpret the element (a_0, a_1, a_2) as the polynomial $a_0 + a_1X + a_2X^2$. So, every element in $\mathbb{Z}_{27}^*$ has an associated non-zero polynomial of degree at most 2, and every non-zero polynomial of degree at most 2 has an element in $\mathbb{Z}_{27}^*$ associated with it.

   The multiplication ($\times$ operator) of the element (a_0, a_1, a_2) with the element (b_0, b_1, b_2) is defined as the element corresponding to the polynomial

   $$(a_0 + a_1X + a_2X^2) \times (b_0 + b_1X + b_2X^2) \mod X^3 + 2X + 2$$

   According to this definition of the $\times$ operator, find

   - (10 points) $(1, 2, 1) \times (2, 2, 1)$, and
   - (15 points) the inverse of $(1, 2, 1)$.

2. **One-time Pad for 3-Alphabet Words.** We interpret $a, b, \ldots, z$ as $0, 1, \ldots, 25$. We will work over the group $(\mathbb{Z}_{26}^3, +)$, where $+$ is coordinate-wise integer-sum $\mod 26$. For example, $abx + acd = ada$.

   Now, consider the one-time pad encryption scheme over the group $(\mathbb{Z}_{26}^3, +)$.

   - (12.5 points) What is the probability that the encryption of the message *cat* is the cipher text *cat*?
   - (12.5 points) What is the probability that the encryption of the message *cat* is the cipher text *dog*?

3. **Left Identity and Left Inverse.** Recall that when we defined a group $(G, \circ)$, we stated that there exists an element e such that for all $x \in G$ we have $x \circ e = x$. Note that e is "applied on x from the right."

   Similarly, for every $x \in G$, we are guaranteed that there exists $\mathsf{inv}(x) \in G$ such that $x \circ \mathsf{inv}(x) = e$. Note that $\mathsf{inv}(x)$ is again "applied to x from the right."

   Intuitively, we shall explore the following questions: (a) Is there an "identity from the left?," and (b) Is there an "inverse from the left?"

   We shall formalize and prove these results in this question.

   - (10 points) Prove that $e \circ x = x$, for all $x \in G$.
   - (10 points) Prove that if there exists an element $\alpha \in G$ such that for all $x \in G$ we have $\alpha \circ x = x$, then $\alpha = e$.

   Note that these two steps prove that the "left identity" is identical to the right identity e.

   - (10 points) Prove that $\mathsf{inv}(x) \circ x = e$.
   - (10 points) Prove that if there exists an element $\alpha \in G$ and $x \in G$ such that $\alpha \circ x = e$, then $\alpha = \mathsf{inv}(x)$.

   Note that these two steps prove that the "left inverse of x" is identical to the left inverse $\mathsf{inv}(x)$.

   Finally, we can prove the following result crucial to the proof of security of one-time pad over the group $(G, \circ)$.

   - (10 points) Suppose $m \in G$ is a message and $c \in G$ is a cipher text. Prove that there exists a unique $\mathsf{sk} \in G$ such that $m \circ \mathsf{sk} = c$.

4. **One-time Pad with non-uniform secret key.** (25 points) Consider the one-time pad encryption scheme over a group $(G, +)$. Suppose the a priori distribution of messages is the uniform distribution over the set G. Suppose the generation algorithm samples the secret-key sk according to the distribution $\mathcal{D}$ over the sample space G such that $\mathcal{D}$ is *not* the uniform distribution over G. Is this encryption scheme secure?

   (*Remark:* To prove that the scheme is secure, provide a proof that the a priori distribution of messages is same as the a posteriori distribution. To prove that the scheme is insecure, provide a proof that the a priori distribution of messages is different from the a posteriori distribution.)

5. **Designing Encryption Scheme.** We shall work over the field $(\mathbb{Z}_{11}, +, \times)$. Assume that there are ten people $\{1, 2, \ldots, 10\}$. Design a private-key encryption scheme for the following scenario.

   Alice meets the ten people $\{1, 2, \ldots, 10\}$ today. She can provide each of them information $\{s_1, s_2, \ldots, s_{10}\}$.

   Tomorrow, Alice shall encrypt a message $m \in \mathbb{Z}_{11}$. The encryption has to ensure that decryption should be possible if and only if two people among $\{1, \ldots, 5\}$ and three people among $\{6, \ldots, 10\}$ get together.

   - (15 points) Provide the $(\mathsf{Gen}, \mathsf{Enc}, \mathsf{Dec})$ algorithms.
   - (15 points) Proof of security of this scheme.

6. **A property of 2-wise Independence.** Let $\mathcal{H}$ be a hash function family from the domain $\mathcal{D}$ to the range $\mathcal{R}$.

   - (20 points) Similar to the proof in the lectures for universal hash function family, prove the following. There exists distinct $x_1^*, x_2^* \in \mathcal{D}$ and $y_1^*, y_2^* \in \mathcal{R}$ such that

     $$\mathbb{P}\left[h(x_1^*) = y_1^*, h(x_2^*) = y_2^* \colon h \xleftarrow{\$} \mathcal{H}\right] \geqslant \frac{1}{|\mathcal{R}|^2}$$

     (*Remark:* Note that this result does not depend on whether $|\mathcal{R}| < |\mathcal{D}|$ or not.)

   - (25 points) Now, suppose that $|\mathcal{R}| < |\mathcal{D}|$. Suppose that for all distinct $x_1, x_2 \in \mathcal{D}$ the following holds.

     $$\mathbb{P}\left[h(x_1) = h(x_2) \colon h \xleftarrow{\$} \mathcal{H}\right] < \frac{1}{|\mathcal{R}|}$$

     Prove that there exists distinct $x_1^*, x_2^* \in \mathcal{D}$ and $y_1^*, y_2^* \in \mathcal{R}$ such that

     $$\mathbb{P}\left[h(x_1^*) = y_1^*, h(x_2^*) = y_2^* \colon h \xleftarrow{\$} \mathcal{H}\right] > \frac{1}{|\mathcal{R}|^2}$$

   This result proves that if a universal hash-function family has collision probability $< \frac{1}{|\mathcal{R}|}$ then it is not pairwise independent.

7. **Extra Credit.** Suppose $\mathcal{D} = \{0,1\}^n$ and $\mathcal{R} = \{0,1\}^{n-1}$. Construct a hash function family such that for all distinct $x_1, x_2 \in \mathcal{D}$ we have

$$\mathbb{P}\left[h(x_1) = h(x_2) \colon h \xleftarrow{\$} \mathcal{H}\right] = \frac{1}{M} \cdot \left(\frac{N-M}{N-1}\right),$$

where $N = 2^n$ and $M = 2^{n-1}$. Try to construct a hash function family such that each hash function can be efficiently evaluated.