**The Vigenere Cipher**: Let $m$ be a positive integer. Define $P = C = K = (\mathbb{Z}_{26})^m$. For a key $K = (k_1, k_2, \ldots, k_m)$, we define

$$e_K(x_1, x_2, \ldots, x_m) = (x_1 + k_1, x_2 + k_2, \ldots, x_m + k_m)$$

and $d_K(y_1, y_2, \ldots y_m) = (y_1 - k_1, y_2 - k_2, \ldots, y_m - k_m)$, where all operations are performed in $\mathbb{Z}_{26}$.

**Example**: Suppose $m = 5$ and the keyword is PLAIN. This corresponds to the numerical equivalent $K = (15, 11, 0, 8, 13)$. Suppose the plaintext is the string: "wewillmeetatmidnight". We convert the plaintext elements to residues mod 26, and write them in groups of 5:

22 4 22 8 11 | 11 12 4 4 19

0 19 12 8 3 | 13 8 6 7 19

We add the key $K = (15, 11, 0, 8, 13)$ mod 26 to each group:

11 15 22 16 24 | 0 23 4 12 6

15 4 12 16 16 | 2 19 6 15 6

The corresponding ciphertext is:

LPWQYAXEMGPEMQQCTGPG

In a Vigenere Cipher having keyword length $m$, an alphabetic character can be mapped to one of $m$ possible alphabetic characters (assuming that the keyword contains $m$ distinct characters). Such a cryptosystem is called a **polyalphabetic cryptosystem**

**The Permutation Cipher (Transposition Cipher):** Let $m$ be a positive integer. Let $\mathcal{P} = \mathcal{C} = (\mathbb{Z}_{26})^m$ and let $\mathcal{K}$ consist of all permutations of $\{1, 2, \ldots, m\}$. For a key (i.e. a permutation) $\pi$, we define

$$e_\pi(x_1, x_2, \ldots, x_m) = (x_{\pi(1)}, x_{\pi(2)}, \ldots, x_{\pi(m)}),$$

and $d_\pi(y_1, y_2, \ldots, y_m) = (y_{\pi^{-1}(1)}, y_{\pi^{-1}(2)}, \ldots, y_{\pi^{-1}(m)})$,

where $\pi^{-1}$ is the inverse permutation to $\pi$.

Example: Suppose $m = 5$ and the key is the following permutation $\pi$: $\pi(1) = 3$, $\pi(2) = 5$, $\pi(3) = 1$, $\pi(4) = 2$, $\pi(5) = 4$. Suppose the plaintext is "wewillmeetatmidnight". We have divided the plaintext in groups of $m = 5$. Now we will permute each group according to $\pi$:

WLWEI ETLME MDATI GTNIH

The inverse permutation $\pi^{-1}$ is:

$\pi^{-1}(1) = 3$, $\pi^{-1}(2) = 4$, $\pi^{-1}(3) = 1$, $\pi^{-1}(4) = 5$, $\pi^{-1}(5) = 2$.

We can apply $\pi^{-1}$ to the ciphertext to get back the original plaintext.

Cryptanalysis of Classical Ciphers: We make an assumption that the opponent, Oscar, knows the cryptosystem being used. This is called Kerckhoffs' principle. We will consider the weakest attack model called the ciphertext only attack in which the opponent possesses a string of ciphertext, $y$ and is trying to decrypt it.
The simplest type of attack is the brute-force attack. We consider all possible keys one by one, and decrypt the ciphertext $y$ using the selected key. If we are able to get a meaningful English message, then we have found the key. So for a cryptosystem to be secure, first requirement is that it should have a large keyspace. Now we will analyze the size of keyspace for the classical cryptosystems:

① Shift Cipher: $|K_1| = 26$

② Affine Cipher: $|K_2| = \phi(26) \times 26 = 312$

③ Substitution Cipher: $|K_3| = 26! = 4.03 \times 10^{26}$

④ Permutation Cipher: $|K_4| = m!$ where $m$ is the key size

⑤ Vigenere Cipher: $|K_5| = 26^m$ where $m$ is the key size.

For security against brute-force attack, we can arrange the above cryptosystems in the following order:
$1 < 2 < 3 < 5 < 4$.

Example brute-force attack on Shift Cipher: given the ciphertext string "JBCRCLQRWCRVNBJENBWRWN," we successively try the decryption keys 0, 1, 2, ... etc. The following is obtained:

| | | | | | | | | | | | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| j | b | c | r | c | l | q | r | w | c | r | v | n | b | j | e | n | b | w | r | w | n | | | | |
| i | a | b | q | b | k | p | q | v | b | q | u | m | a | i | d | m | a | v | q | u | m | | | | |
| h | z | a | p | a | j | o | p | u | a | p | t | l | z | h | c | l | z | u | p | u | l | | | | |
| g | y | z | o | z | i | n | o | t | z | o | s | k | y | g | b | k | y | t | o | t | k | | | | |
| f | x | y | n | y | h | m | n | s | y | n | r | j | x | f | a | j | x | s | n | s | j | | | | |
| e | w | x | m | x | g | l | m | r | x | m | q | i | w | e | z | i | w | r | m | r | i | | | | |
| d | v | w | l | w | f | k | l | q | w | l | p | h | v | d | y | h | v | q | l | q | h | | | | |
| c | u | v | k | v | e | j | k | p | v | k | o | g | u | c | x | g | u | p | k | p | g | | | | |
| b | t | u | j | u | d | i | j | o | u | j | n | f | t | b | w | f | t | o | j | o | f | | | | |
| a | s | t | i | t | c | h | i | n | t | i | m | e | s | a | v | e | s | n | i | n | e | | | | |

The plaintext is "a stich in time saves nine".

A more efficient attack than the brute-force attack uses frequency analysis of English alphabet. We can group the 26 letters based on the probability of occurrence as follows:

(1) E has highest probability about 0.120.
(2) T, A, O, I, N, S, H, R, each having probability between 0.06 and 0.09.
(3) D, L, each having probability around 0.04.
(4) C, U, M, W, F, G, Y, P, B, each having probability between 0.015 and 0.028.
(5) V, K, J, X, Q, Z, each having probability less than 0.01.

It is also useful to consider sequences of two or three consecutive letters, called digrams and trigrams, respectively. The 30 most common digrams are (in decreasing order):

TH, HE, IN, ER, AN, RE, ED, ON, ES, ST, EN, AT, TO, NT, HA, ND, OU, EA, NG, AS, OR, TI, IS, ET, IT, AR, TE, SE, HI, OF.

The 12 most common trigrams are:
THE, ING, AND, HER, ERE, ENT, THA, NTH, WAS, ETH, FOR, DTH.

Given a ciphertext, we find the frequency of each letter and try to match it with English alphabet. We can also consider mapping digrams and trigrams. Frequency analysis can be applied for breaking the monoalphabetic ciphers: Shift cipher, Affine Cipher, and Substitution Cipher.