

Математична криптологія

Основні поняття.
Алгоритм Евкліда.



Основні поняття

Алфавітом називають довільну скінченну множину:

$\{a, б, в, г, \dots, ь, ю, я\}$ – український алфавіт,

$\{ _ , a, б, в, г, \dots, ь, ю, я\}$ – український алфавіт з пропуском між словами,

$\{ . , , , ? , a, б, в, г, \dots, ь, ю, я\}$ – український алфавіт з розділовими знаками,

$\{0,1\}^6$ – всі 0–1 послідовності довжини 6,

$Z_{33} = \{0, 1, 2, \dots, 32\}$ – кільце лишків за модулем 33.

Алфавіт позначатимемо рукописними літерами A, B, \dots

Елементи алфавіту називатимемо символами або буквами.

Слово в алфавіті A – це скінчення послідовність букв цього алфавіту.

A^* – множина всіх слів у алфавіті A .

Множина слів у алфавіті A довжини l – A^l .

Довжина слова w – $|w|$.

Для слів v та u через vu позначатимо результат їх злиття.

Слово v називатимемо префіксом слова w у випадку, коли $v=w$ або $w=vi$ для деякого слова i .

Основні поняття

Коли говорять про криптосистему або шифр, мають на увазі такі об'єкти:

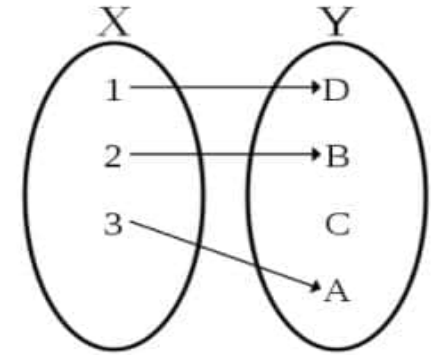
- Алфавіт \mathcal{A} , в якому записуються повідомлення (відкриті тексти). Повідомлення M є словом в алфавіті \mathcal{A} (яке може складатися з багатьох слів у звичному лінгвістичному розумінні), тобто $M \in \mathcal{A}^*$. Множина \mathcal{A}^* називається простором повідомлень або відкритих текстів.
- Алфавіт \mathcal{B} , в якому записуються криптотексти. Множина \mathcal{B}^* називається простором криптотекстів. Часто $\mathcal{A} = \mathcal{B}$.
- Простір ключів \mathcal{K} , який складається із слів у деякому алфавіті, що називаються ключами.
- Шифруюче відображення $E : \mathcal{K} \times \mathcal{A}^* \rightarrow \mathcal{B}^*$
- Дешифруюче відображення $D : \mathcal{K} \times \mathcal{B}^* \rightarrow \mathcal{A}^*$. Відображення E і D повинні мати таку властивість:

$$D(K, E(K, M)) = M \text{ для всіх } M \in \mathcal{A}^* \text{ і } K \in \mathcal{K}. \quad (1)$$

Фіксація ключа K визначає відображення $E_K : \mathcal{A}^* \rightarrow \mathcal{B}^*$ за формулою $E_K(M) = E(K, M)$ для кожного $M \in \mathcal{A}^*$. Подібним чином задається й відображення $D_K : \mathcal{B}^* \rightarrow \mathcal{A}^*$.

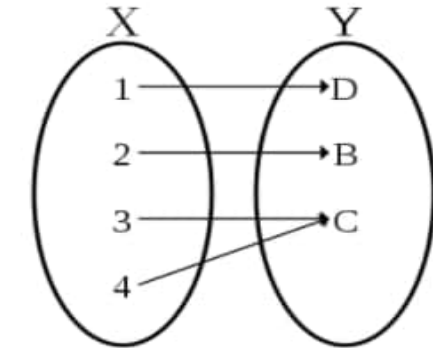
Відображення

Відображення $f: X \rightarrow Y$ називається **ін'єктивним**, якщо воно різним аргументам співставляє різні значення: $f(x_1) \neq f(x_2)$ для $x_1 \neq x_2$.



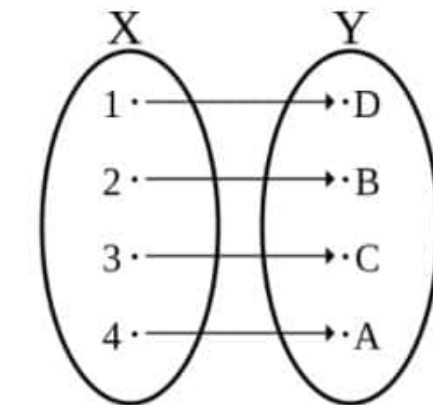
Ін'єктивне, але не сюр'єктивне відображення

Відображення $f: X \rightarrow Y$ називається **сюр'єктивним**, якщо кожен елемент y з множини Y має прообраз – такий елемент $x \in X$, що $f(x) = y$.



Сюр'єктивне, але не ін'єктивне відображення

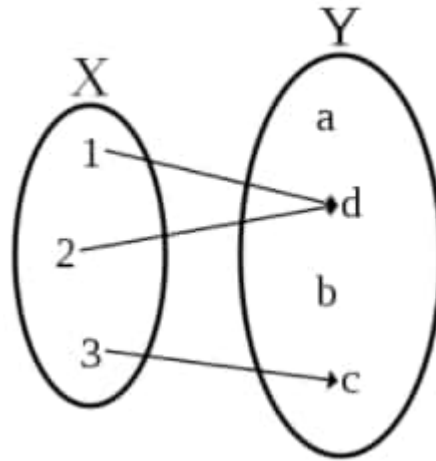
Бієктивним є відображення, яке **ін'єктивне** та **сюр'єктивне** одночасно.



Бієктивне відображення (сюр'єктивне та ін'єктивне)

Відображення

Яким є наступне відображення?



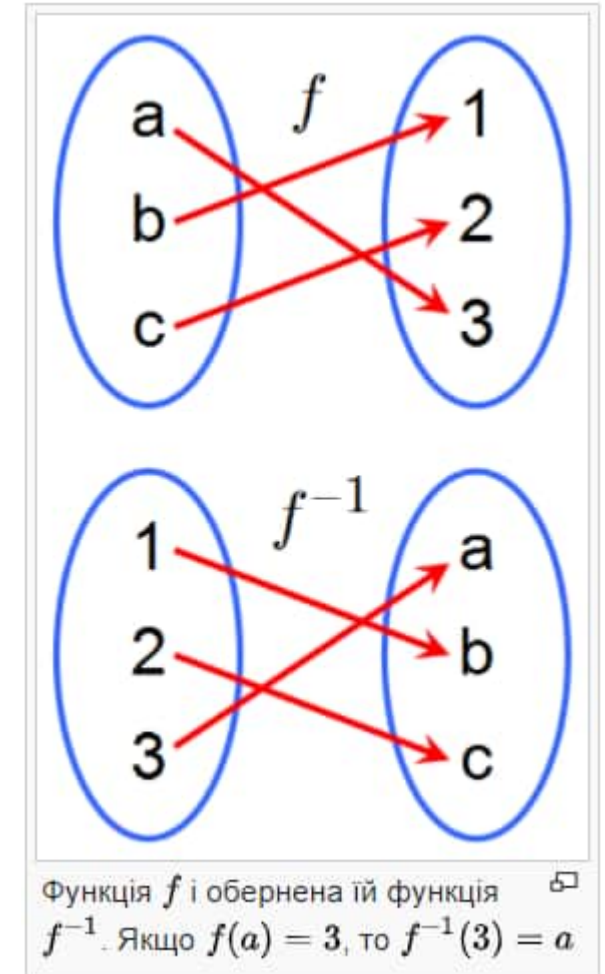
Для двох відображень $f: X \rightarrow Y$ та $g: Y \rightarrow Z$ їх **композиція** задається співвідношенням $g \circ f(x) = g(f(x))$ для $x \in X$.

Тотожне відображення $id_x: X \rightarrow X$ залишає елемент множини X на місці: $id_x(x) = x$.

Відображення $g: Y \rightarrow X$ вважається **лівим оберненим** до відображення $f: X \rightarrow Y$ за умови що їх композиція рівна $g \circ f = id_x$.

Відображення $g: Y \rightarrow X$ вважається **правим оберненим** до відображення $f: X \rightarrow Y$ за умови що їх композиція рівна $f \circ g = id_y$.

Відображення $g: Y \rightarrow X$ називається **оберненим** до відображення f , якщо воно є одночасно і лівим, і правим оберненим до f .



Функція f і обернена їй функція f^{-1} . Якщо $f(a) = 3$, то $f^{-1}(3) = a$ \square

Теорема про обернене відображення

- 1) Відображення $f: X \rightarrow Y$ ін'єктивне тоді і лише тоді, коли до нього існує ліве обернене.
- 2) Відображення $f: X \rightarrow Y$ сюр'єктивне тоді і лише тоді, коли до нього існує праве обернене.
- 3) Якщо відображення бієктивне, то його ліве обернене збігається із правим оберненим.
- 4) У випадку, коли множини X та Y скінченні та містять однакову кількість елементів, відображення $f: X \rightarrow Y$ має ліве обернене тоді і лише тоді, коли воно має праве обернене.

$$D(K, E(K, M)) = M \text{ для всіх } M \in \mathcal{A}^* \text{ і } K \in \mathcal{K}. \quad (1)$$

Для довільного K дешифруєче відображення $D_K: B^* \rightarrow A^*$ є оберненим зліва до шифруючого відображення $E_K: A^* \rightarrow B^*$. З теореми про обернене відображення випливає що E_K має бути ін'єкцією. Ця умова рівнозначна можливості дешифрування.

Будь-яка родина ін'єктивних відображень $\{E_k: A^* \rightarrow B^*\}_{k \in K}$ може розглядатись як шифруюча в тому плані, що для них існують обернені зліва відображення $\{D_k: B^* \rightarrow A^*\}_{k \in K}$, які можуть вважатись дешифруючими.

Досить часто шифруюче відображення крім ін'єктивності володіє також сюр'єктивністю. За теоремою про обернене відображення, дешифруєче відображення D_k є оберненим до E_k не тільки зліва, а й справа.

$$D_K(E_K(M)) = E_K(D_K(M)) = M.$$

Криптосистему називають ефективною, якщо шифруюче і дешифруюче відображення реалізуються швидким алгоритмом.

Блоковий шифр. Дешифрування ітераціями.

Шифр називається **блоковим** з періодом l , якщо шифруюче відображення задається на словах довжини l :

$$K \times A^l \rightarrow B^l$$

та поширюється на слова довільної довжини $M = M_1 M_2 \dots M_t$: $E(K, M) = E(K, M_1) E(K, M_2) \dots E(K, M_t)$.

Часто у блоковому шифрі шифруюче відображення зберігає довжину: $E: K \times A^l \rightarrow B^l$, окрім того A^l та B^l можуть містити однакову кількість символів, наприклад $A=B$.

Щоб перевірити чи відображення такого вигляду можна використовувати як шифруєче, достатньо довести одну з чотирьох умов теореми про обернене відображення:

1) ін'єктивність, 2) сюр'єктивність, 3) існування оберненого зліва, 4) існування оберненого справа відображення.

Дешифрування ітераціями: Розглянемо блоковий шифр вигляду $E: K \times A^l \rightarrow A^l$. Для кожного ключа K шифруюче відображення є елементом групи $\text{Sym } A^l$. Шифр утворює групу, якщо множина $\{E_k\}_{k \in K}$ є в $\text{Sym } A^l$ підгрупою.

Шифруюче відображення $E_k: A^l \rightarrow A^l$ можна застосовувати декілька разів: $E_k^1 = E_k$; $E_k^i = E_k \circ E_k^{i-1}$, $i > 1$.

Методом ітерацій суперник може скористатись у разі, коли він має доступ до шифруючого відображення з фіксованим ключем. Підслухавши криптотекст $C = E_k(M)$, суперник може спробувати знайти повідомлення M , обчислюючи послідовність $E_k^1(C), E_k^2(C), E_k^3(C), \dots$ доки не виявиться що $E_k^m(C) = C$. Це означає, що відкрите повідомлення було отримане на попередньому кроці: $M = E_k^{m-1}(C)$.

Групи

Б.3. Групи. Групою називається множина G , наділена бінарною операцією \star з такими властивостями:

- 1) $(x \star y) \star z = x \star (y \star z)$ для будь-яких елементів $x, y, z \in G$ (асоціативність);
- 2) в G існує нейтральний елемент e такий, що $x \star e = e \star x = x$ для всіх $x \in G$;
- 3) для кожного елемента $x \in G$ в G є обернений елемент x' такий, що $x \star x' = x' \star x = e$.

Якщо підмножина H множини G утворює групу відносно тієї ж операції \star , то вона називається *підгрупою* групи G . Так, множина раціональних чисел \mathbb{Q} утворює групу за додаванням ($e = 0, x' = -x$), а множина цілих чисел \mathbb{Z} є її підгрупою. Множина додатніх раціональних чисел \mathbb{Q}_+ утворює групу за множенням ($e = 1, x' = 1/x$). Якщо групову операцію називаємо множенням, то саму групу називаємо *мультиплікативною*, а її нейтральний елемент — одиницею. Якщо ж операцію називаємо додаванням, то групу називаємо *адитивною*, нейтральний елемент нулем, а обернений елемент — протилежним елементом.

Для елемента x групи G через x^i позначається його i -тий степінь — елемент $x \star x \star \dots \star x$, де операція виконана $i - 1$ раз. (В адитивній формі те ж саме записується як $ix = x + \dots + x$.) Нескладною вправою є доведення того, що для кожного елемента x скінченної групи для деякого показника m виконується рівність $x^m = e$. Найменше з таких m називається *порядком елемента x* у групі G .

Порядок групи. Теорема Лагранжа. Гомоморфізм

Порядком скінченної групи називається кількість її елементів. Легко зауважити, що всі степені елемента групи утворюють у ній підгрупу, порядок якої дорівнює порядку елемента.

ТЕОРЕМА ЛАГРАНЖА (1771).

- 1) *Порядок підгрупи в дільником порядку групи.*
- 2) *Порядок елемента в дільником порядку групи.*

Нехай маємо дві групи G і G' з операціями \star і \circ відповідно. Кажемо, що відображення $f: G \rightarrow G'$ зберігає операцію, якщо $f(x \star y) = f(x) \circ f(y)$ для всіх $x, y \in G$. Таке відображення $f: G \rightarrow G'$ називається *гомоморфізмом* з групи G у групу G' . *Ядром* гомоморфізму $f: G \rightarrow G'$ є множина всіх тих елементів групи G , які f відображає в нейтральний елемент групи G' . Наприклад, відображення, яке кожному цілому числу ставить у відповідність його остачу від ділення на натуральне n , є гомоморфізмом із адитивної групи \mathbb{Z} в адитивну групу \mathbb{Z}_n , ядро якого утворюють цілі числа, кратні n . Легко показати, що ядро гомоморфізму $f: G \rightarrow G'$ утворює в G підгрупу.

Гомоморфізм $f: G \rightarrow G'$ є ін'єктивним тоді і лише тоді, коли його ядро складається тільки із нейтрального елемента групи G . Таке ядро називається *тривіальним*.

Гомоморфізм, який є бієктивним відображенням, називається *ізоморфізмом*. Дві групи *ізоморфні*, якщо існує ізоморфізм з однієї з них на іншу. Наприклад, двійковий логарифм $\log_2: \mathbb{R}_+ \rightarrow \mathbb{R}$ задає ізоморфізм із мультиплікативної групи невід'ємних дійсних чисел \mathbb{R}_+ в адитивну групу всіх дійсних чисел \mathbb{R} . Ізоморфні групи мають тотожні алгебраїчні властивості.

Група перестановок

Б.4. Група перестановок. Іншим прикладом групи є множина $\text{Sym } X$ бієктивних відображень множини X на себе з операцією композиції. Нейтральним елементом є тотожне відображення id_X , а оберненим елементом до $f \in \text{Sym } X$ є відображення f' , обернене до f . Бієкцію множини на себе називають ще *перестановкою* цієї множини. Відповідно, $\text{Sym } X$ називається *групою перестановок* множини X . Якщо множина X налічує n елементів, то очевидним чином група $\text{Sym } X$ ізоморфна групі $\text{Sym}\{1, 2, \dots, n\}$. Остання називається *симетричною групою степеня n* і позначається через S_n .

Перестановку σ із S_n звично записують у вигляді таблицьки

$$\begin{pmatrix} 1 & 2 & \dots & n \\ \sigma(1) & \sigma(2) & \dots & \sigma(n) \end{pmatrix}$$

Цикл $(i_1 i_2 \dots i_l)$, де i_1, i_2, \dots, i_l різні числа від 1 до n , — це перестановка, яка елемент i_j , $j < l$, відображає в i_{j+1} , i_l в i_1 , а всі інші елементи самі в себе. l називається *довжиною* циклу. Цикли $(i_1 i_2 \dots i_l)$ та $(i'_1 i'_2 \dots i'_l')$ *незалежні*, якщо множини елементів $\{i_1, i_2, \dots, i_l\}$ та $\{i'_1, i'_2, \dots, i'_l'\}$ не перетинаються. Кожна перестановка є композицією (або добутком) попарно незалежних циклів. Наприклад, перестановка

$$\begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ 5 & 2 & 6 & 3 & 1 & 4 \end{pmatrix}$$

дорівнює добутку $(15)(364)$. Якщо перестановка σ розкладається в добуток m незалежних циклів з довжинами l_1, l_2, \dots, l_m , то її знак $\varepsilon(\sigma)$ обчислюється за формулою

$$\varepsilon(\sigma) = (-1)^{\sum_{j=1}^m (l_j - 1)}$$

Знак тотожної перестановки приймається рівним 1.

Дешифрування ітераціями

ТВЕРДЖЕННЯ 1.1. Для шифруючого відображення $E: K \times A^l \rightarrow A^l$ метод ітерацій через деяку кількість кроків приводить до успіху. Точніше, якщо алфавіт складається з n букв, то для будь-якого ключа K існує число $m \leq (n^l)!$ таке, що $E_K^m(C) = C$ для всіх $C \in A^l$

Доведення. В якості m можна взяти порядок відображення E_K як елемента групи $\text{Sym } A^l$, який за теоремою Лагранжа не перевищує $(n^l)!$, порядку цієї групи (див. додаток Б). Тоді для будь-якого $C \in A^l$ матимемо $E_K^m(C) = \text{id}_{A^l}(C) = C$. ■

Вправа для самостійного виконання:

Довести, що для розкриття шифру Віженера над n -символьним алфавітом достатньо n ітерацій.



Алгоритм Евкліда

Для будь-якого цілого a та натурального b однозначно визначені цілі числа q та r такі, що:

$$a = bq + r \text{ та } 0 \leq r < b;$$

$$r = a \bmod b.$$

Якщо $r=0$, то кажуть, що a ділиться на b націло або без остачі: $b \mid a$.

Алгоритм Евкліда:

$\text{НСД}(a, b) = \text{НСД}(a, a \bmod b)$ для $a \geq b$, а $\text{НСД}(a, 0) = a$, і складається з кроків:

1) $r_0 = a, r_1 = b, i = 1$;

2) i -й крок: $r_{i-1} = r_i q_i + r_{i+1}$ (ділення з остачею);

3) якщо $r_{i+1} > 0$, то $i++$ та перехід на крок 2; якщо $r_{i+1} = 0$ то $\text{НСД}(a, b) = r_i$.

$$211 = 79 \cdot 2 + 53$$

$$79 = 53 \cdot 1 + 26$$

$$53 = 26 \cdot 2 + 1$$

$$26 = 1 \cdot 26 + 0$$

$$\text{Маємо } \text{НСД}(211, 79) = \text{НСД}(79, 53) = \text{НСД}(53, 26) = \text{НСД}(26, 1) = \text{НСД}(1, 0) = 1.$$

Ефективність алгоритму Евкліда

Під ефективністю алгоритму Евкліда розуміють кількість кроків m .

Із співвідношення $0 < r_{i+1} \leq r_i$ випливає $m \leq b$. Але ця оцінка є незадовільною.

Покажемо що:

$m \leq 2 \log_2 b + 1$. Для цього використаємо нерівність $r_{i+1} < r_{i-1}/2$. Вона доводиться розглядом двох випадків: якщо $r_i < r_{i-1}/2$, то використовуємо нерівність $r_{i+1} < r_i$; якщо ж $r_i \geq r_{i-1}/2$, то використовуємо рівність $r_{i+1} = r_{i-1} \bmod r_i$. Таким чином, кожні два кроки зменшують r_{i+1} принаймні вдвічі, і не пізніше, ніж за $2 \log_2 b + 1$ кроків ми прийдемо до $r_{i+1} = 0$.

Наслідок з алгоритму Евкліда.

ТВЕРДЖЕННЯ 2.2. Для кожної пари взаємно простих чисел a і b можна знайти такі цілі u і v , що $ua + vb = 1$.

Доведення. За умовою $\text{НСД}(a, b) = 1$. Тому на передостанньому кроці алгоритму Евкліда матимемо $r_{m-2} = r_{m-1}q_{m-1} + 1$. Звідси $1 = 1 \cdot r_{m-2} + (-q_{m-1})r_{m-1}$. Використавши це як базу для зворотної індукції за i від $m-1$ до 1, доведемо, що $1 = u_i r_{i-1} + v_i r_i$ для деяких цілих u_i та v_i . Справді, якщо $1 = u_{i+1} r_i + v_{i+1} r_{i+1}$, то після підстановки замість r_{i+1} його значення, визначеного з (3), маємо $1 = v_{i+1} r_{i-1} + (u_{i+1} - q_i v_{i+1}) r_i$.

При $i = 1$ отримуємо потрібне твердження. ■

Зазначимо, що алгоритм Евкліда дає ефективний спосіб знаходження коефіцієнтів u і v для заданої пари a, b .

Приклад 2.3. Нехай $a = 211$, $b = 79$. Протокол роботи алгоритму Евкліда виписаний у прикладі 2.1. Рухаючись знизу вгору, отримуємо

$$\begin{aligned} 1 &= 1 \cdot 53 + (-2) \cdot 26 & 26 &= 1 \cdot 53 + (-2) \cdot (79 - 1 \cdot 53) = (-2) \cdot 79 \\ &+ 3 \cdot 53 & &= (-2) \cdot 79 + 3 \cdot (211 - 2 \cdot 79) = 3 \cdot 211 + (-8) \cdot 79. \end{aligned}$$

Розширений алгоритм Евкліда.

Вправа 1. Довести рівність:

$$r_i = u_i a + v_i b, 0 \leq i \leq m;$$

$$u_0 = v_1 = 1;$$

$$u_1 = v_0 = 0;$$

$$u_{i+1} = u_{i-1} - q_i u_i;$$

$$v_{i+1} = v_{i-1} - q_i v_i;$$

де q_i отримується на відповідному кроці алгоритму Евкліда:

$$r_{i-1} = r_i q_i + r_{i+1}$$

$$r_m = \text{НСД}(a, b), u = u_m, v = v_m.$$

Вправа 2. Знайти цілі u та v такі, що $137u + 113v = 1$.



Розклад на прості співмножники

ТВЕРДЖЕННЯ 2.4. *Якщо просте число p ділить добуток ab двох натуральних чисел, то воно мусить ділити хоча б одне із чисел a і b .*

Доведення. Якщо a не ділиться на p , то a і p взаємно прості. Тоді за твердженням 2.2 для деяких цілих чисел u і v виконується рівність $ua + vp = 1$. Домноживши її на b , отримуємо $uab + vbp = b$. Оскільки $p \mid ab$, то ліва частина ділиться на p , а відтак $p \mid b$. ■

Зрозуміло, що кожне складене число можна записати як добуток простих. Такий добуток називається *розкладом числа на прості співмножники*. Вважаємо, що розклад простого числа складається з єдиного елемента — самого числа. Має місце

ТЕОРЕМА ПРО ОДНОЗНАЧНІСТЬ РОЗКЛАДУ НА ПРОСТІ СПІВМНОЖНИКИ. *Кожне більше від 1 ціле число однозначно розкладається на прості співмножники (якщо не враховувати їхнього порядку).*

Доведення. Припустимо, що $\prod_{i=1}^k p_i^{\alpha_i} = \prod_{i=1}^k p_i^{\beta_i}$ — два різні розклади одного й того ж числа на прості співмножники. Тоді мусить бути $\alpha_j \neq \beta_j$ для деякого j . Якщо $\alpha_j > \beta_j$, то маємо $p_j \mid \prod_{i \neq j} p_i^{\beta_i}$, що суперечить твердженню 2.4. Випадок $\alpha_j < \beta_j$ розглядається симетрично. ■

Розклад, в якому прості співмножники йдуть у неспадному порядку, називається *канонічним*.