
Lecture 8 – Group theory, Diffie-Hellman key exchange

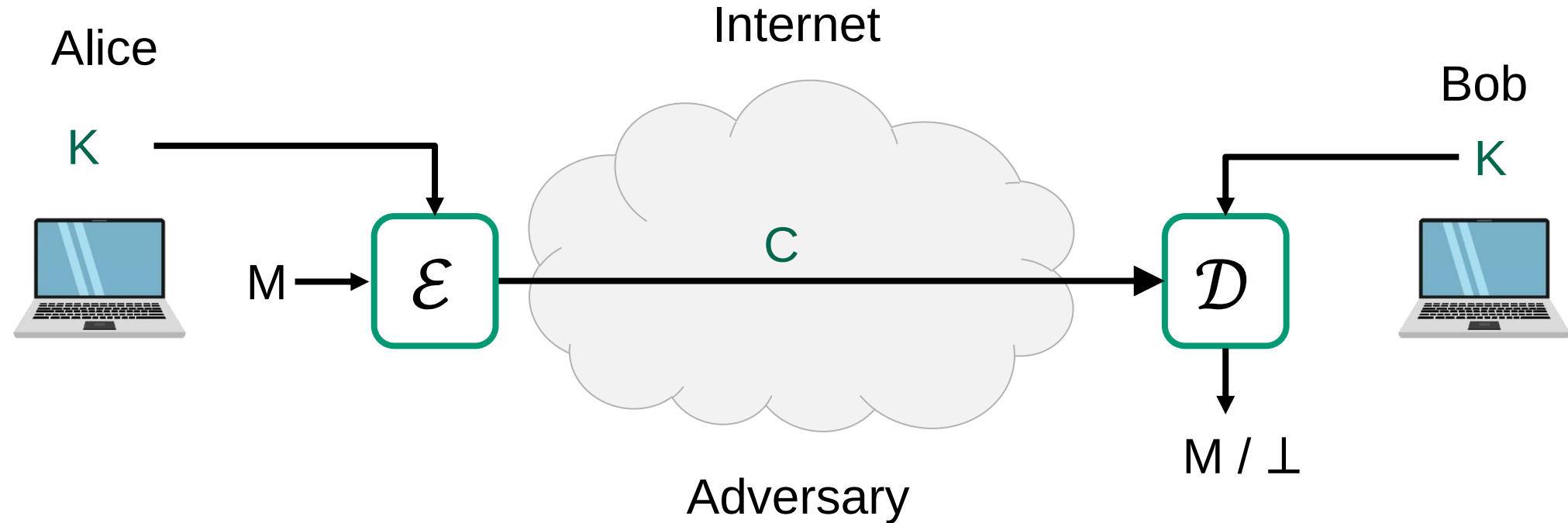
TEK4500

16.10.2024

Håkon Jacobsen

hakon.jacobsen@its.uio.no

Creating secure channels: encryption schemes

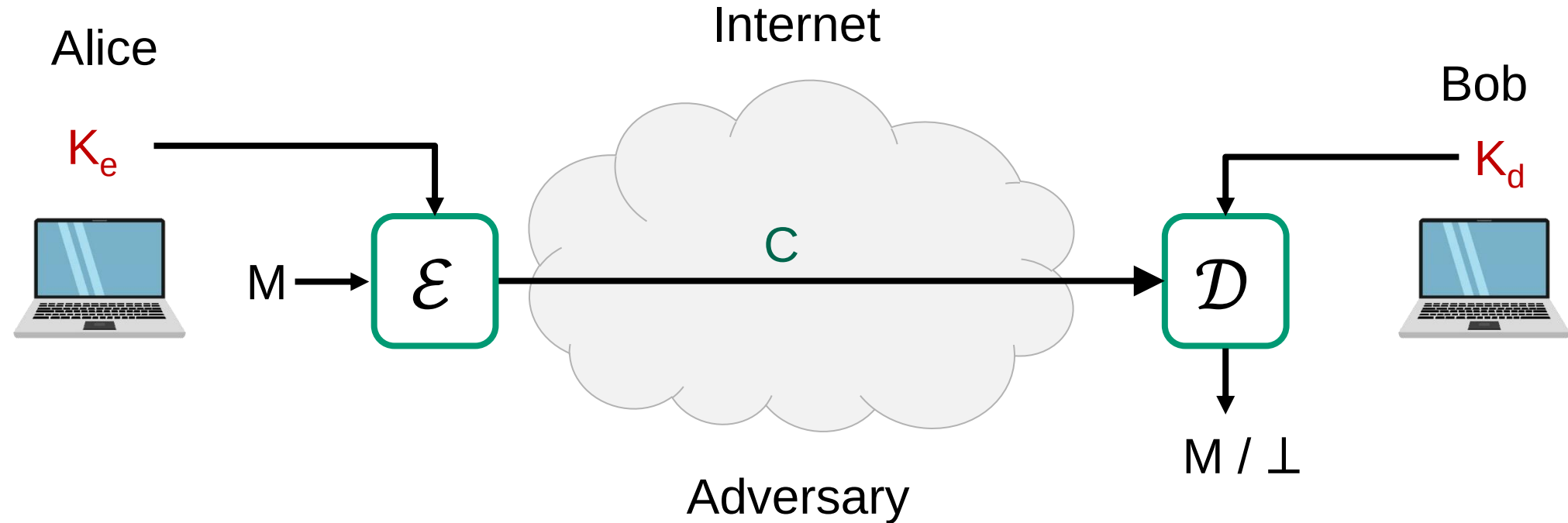


$\mathcal{E}$: encryption algorithm (public)

K : encryption / decryption key (secret)

$\mathcal{D}$: decryption algorithm (public)

Creating secure channels: encryption schemes



$\mathcal{E}$: encryption algorithm (public)

K_e : encryption key (public)

$\mathcal{D}$: decryption algorithm (public)

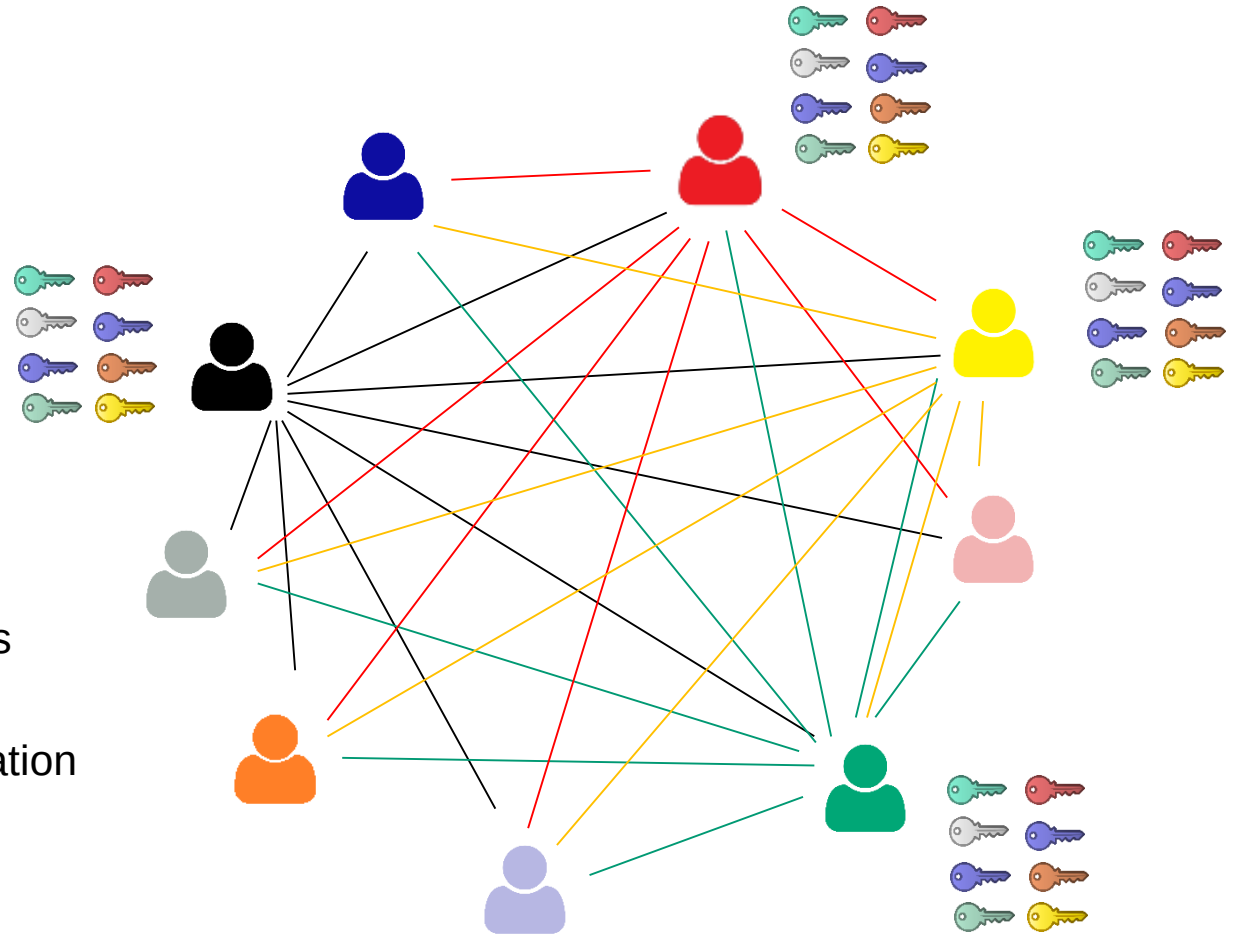
K_d : decryption key (secret)

Basic goals of cryptography

	Message privacy	Message integrity / authentication	
Symmetric keys	Symmetric encryption	Message authentication codes (MAC)	
Asymmetric keys	Asymmetric encryption (a.k.a. public-key encryption)	Digital signatures	Key exchange

Symmetric key distribution problem

- One user needs to store N symmetric keys when communicating with N other users
- $N(N - 1) = \mathcal{O}(N^2)$ keys stored in total
- Difficult to store and manage so many keys securely
- Partial solution: **key distribution centers**
 - One central authority hands out temporary keys
 - $\mathcal{O}(N)$ (long-term) keys needed (to the KDC)
 - Might be a feasible solution in a single organization
 - Single point of failure
 - **What about the internet?**





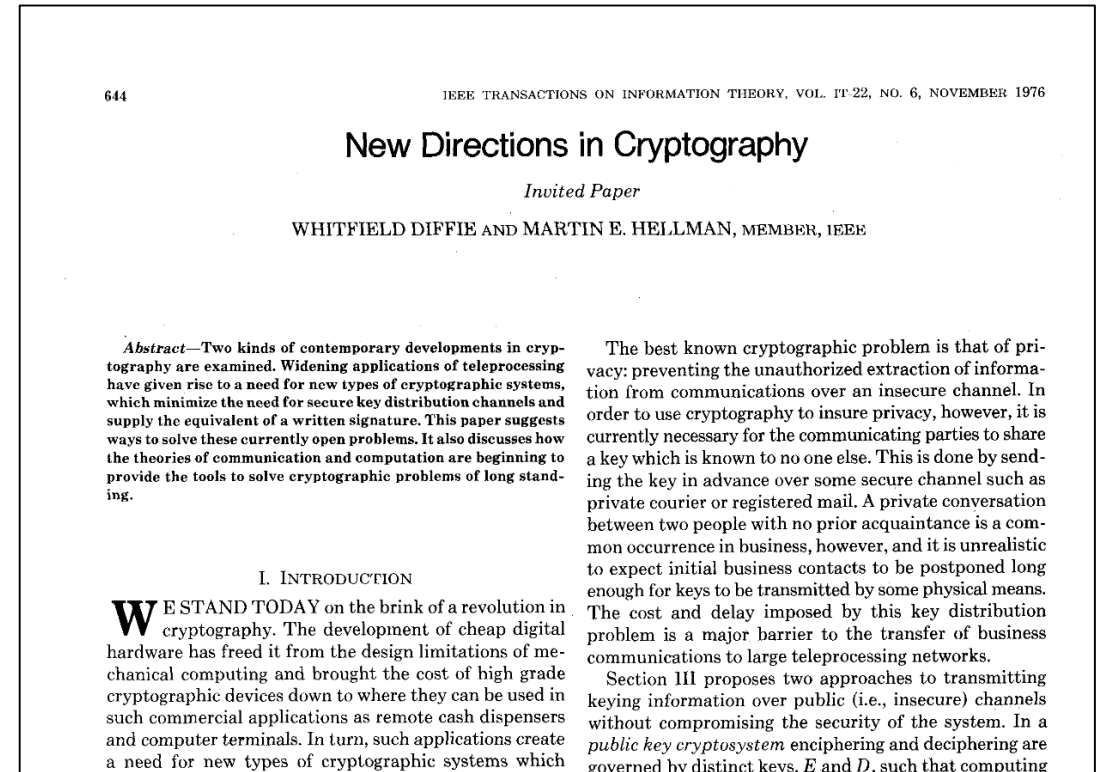
The public-key revolution

Diffie-Hellman key exchange – idea



Diffie-Hellman key exchange

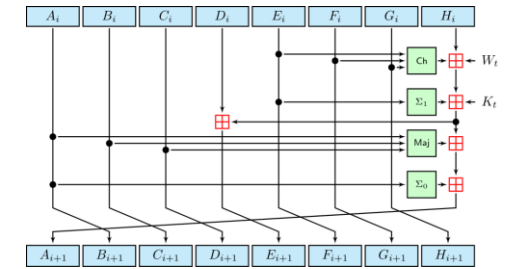
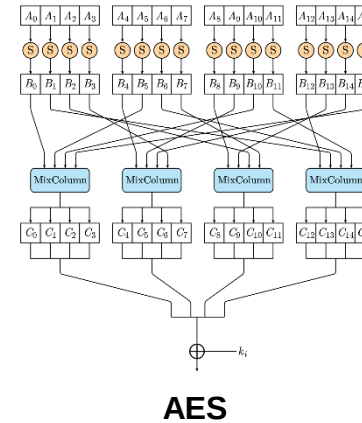
- Discovered in the 1970's
- Allows two parties to establish a shared secret without ever having met
- Diffie & Hellman paper also introduced:
 - Public-key encryption
 - Digital signatures



Ralph Merkle Whitfield Diffie
Martin Hellman

A different kind of primitives

- Symmetric crypto boils down to a few *primitives*
 - **Block ciphers** and **hash functions**
 - Why are these considered secure?
 - Lots and lots of cryptanalysis (*well-studied!*)
 - Artificial and man-made



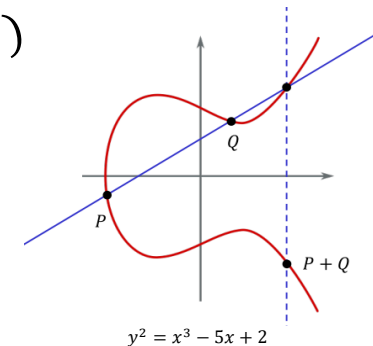
SHA2-256

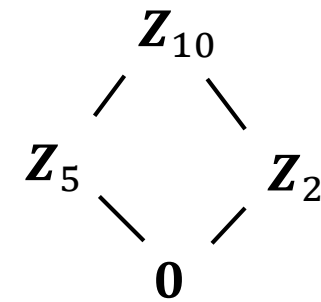
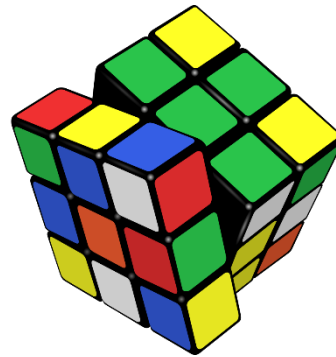
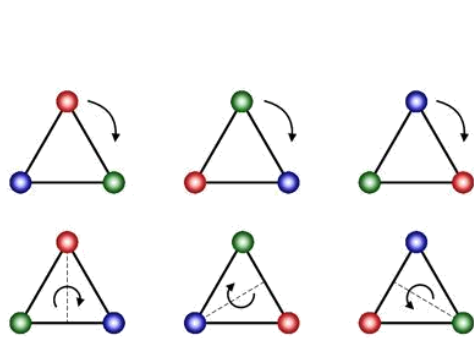
- Want asymmetric crypto to be based on a few well-studied primitives too
- Candidates come from a different place:
 - Hard *mathematical* problems
 - Good candidates: **discrete logarithms**, **factoring**
 - Much more algebraic structure

$$\mathbf{Z}_n^* \simeq \mathbf{Z}_{p_1}^* \times \mathbf{Z}_{p_2}^* \times \cdots \times \mathbf{Z}_{p_t}^*$$

$$C \leftarrow M^e \pmod{N}$$

$$e : G_1 \times G_2 \rightarrow G_T$$





Group theory + number theory

Preliminaries

(integers) $\mathbf{Z} = \{\dots, -2, -1, 0, 1, 2, 3, \dots\}$

An integer $p > 1$ is **prime** if it's only divisible by 1 and p

(reals) $\mathbf{R}$ = the real numbers

$$\mathbf{R}^* = \mathbf{R} \setminus \{0\}$$

(integers “mod n ”) $\mathbf{Z}_n = \{0, 1, 2, \dots, n - 1\}$

(integers “mod p ”) $\mathbf{Z}_p = \{0, 1, 2, \dots, p - 1\}$ ^{p prime}

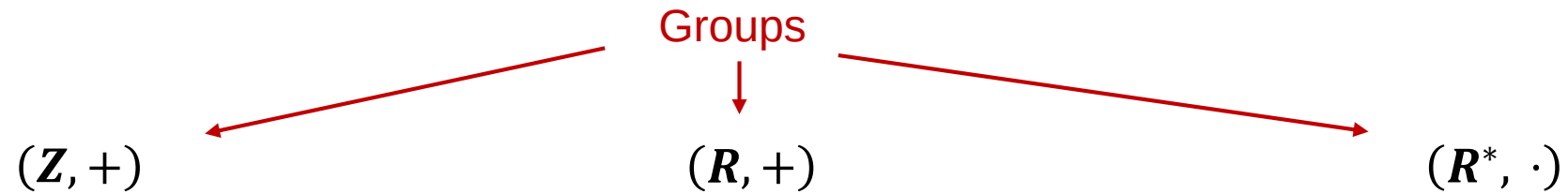
$$\mathbf{Z}_p^* = \mathbf{Z}_p \setminus \{0\}$$

Examples:

$$\mathbf{Z}_{11} = \{0, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10\}$$

$$\mathbf{Z}_{11}^* = \{1, 2, 3, 4, 5, 6, 7, 8, 9, 10\}$$

Groups – motivation



Associativity

$$(1 + 2) + 3 = 1 + (2 + 3)$$

$$(\sqrt{2} + e) + \pi = \sqrt{2} + (e + \pi)$$

$$(\sqrt{2} \cdot e) \cdot \pi = \sqrt{2} \cdot (e \cdot \pi)$$

Identity

$$6 + 0 = 6$$

$$\sqrt{2} + 0 = \sqrt{2}$$

$$\sqrt{2} \cdot 1 = \sqrt{2}$$

Inverses

$$6 + (-6) = 0$$

$$\sqrt{2} + (-\sqrt{2}) = 0$$

$$\sqrt{2} \cdot \frac{1}{\sqrt{2}} = 1$$

Groups – definition

Definition: A **group** $(G, \circ)$ is a set G together with a binary operation $\circ$ satisfying the following axioms

$\mathcal{G}1$: $(a \circ b) \circ c = a \circ (b \circ c)$ for all $a, b, c \in G$ (associativity)

$\mathcal{G}2$: $\exists e \in G$ such that $e \circ a = a \circ e = a$ for all $a \in G$ (identity)

$\mathcal{G}3$: $\forall a \in G$ there exists $a^{-1} \in G$ such that $a \circ a^{-1} = a^{-1} \circ a = e$ (inverse)

A group is **abelian/commutative** if $a \circ b = b \circ a$ for all $a, b \in G$

The **order** of a group is the number of elements in G , denoted $|G|$

Examples

Definition: A group $(G, \circ)$ is ...

$\mathcal{G}1$: $(a \circ b) \circ c = a \circ (b \circ c)$ (associativity)

$\mathcal{G}2$: $\exists e \in G: e \circ a = a \circ e = a$ (identity)

$\mathcal{G}3$: $\exists a^{-1} \in G: a \circ a^{-1} = a^{-1} \circ a = e$ (inverse)

Groups

$$(\mathbf{Z}, +) \quad e = 0 \quad "2^{-1}" = -2$$

$$(\mathbf{R}, +) \quad e = 0 \quad " \pi^{-1} " = -\pi \quad (\mathbf{R}^*, \cdot) \quad e = 1 \quad \pi^{-1} = 1/\pi$$

$$(\mathbf{Z}_6, +_6) \quad e = 0 \quad "2^{-1}" = 4: \quad 2 + 4 = 6 \equiv 0 \pmod{6}$$

$$(\mathbf{Z}_p^*, \cdot_p) \quad e = 1 \quad "3^{-1}" = x: \quad 3 \cdot x \equiv 1 \pmod{p}$$

$(G, \circ)$

$\circ$	e	a	b
e	e	a	b
a	a	b	e
b	b	e	a

$(\mathbf{Z}_3, +_3)$

$+_3$	0	1	2
0	0	1	2
1	1	2	0
2	2	0	1

$(G, \circ)$

$\circ$	e	a	b	c
e	e	a	b	c
a	a	b	c	e
b	b	c	e	a
c	c	e	a	b

Not groups

$$(\mathbf{Z}, \cdot) \quad 2^{-1} = ? \quad (\mathbf{Z}, -) \quad (1 - 2) - 3 \neq 1 - (2 - 3)$$

$$(\mathbf{R}, \cdot) \quad 0 \cdot x = 1?$$

$$(\mathbf{Z}_n, \cdot_n) \quad 2x = 1 \pmod{6}?$$

$$(\mathbf{Z}_p, \cdot_p)$$

$$(\mathbf{Z}_2 \times \mathbf{Z}_2, +_{2 \times 2})$$

$\cong$

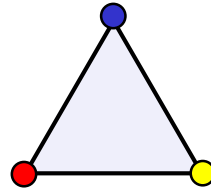
$(G, \star)$

$(\mathbf{Z}_4, +_4)$

$+_4$	0	1	2	3
0	0	1	2	3
1	1	2	3	0
2	2	3	0	1
3	3	0	1	2

$\star$	e	a	b	c
e	e	a	b	c
a	a	e	c	b
b	b	c	e	a
c	c	b	a	e

More groups



$$\sigma_5 \circ \sigma_2 = \sigma_4$$

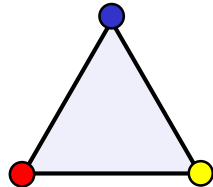
$\neq$

$$\sigma_2 \circ \sigma_5 = \sigma_6$$

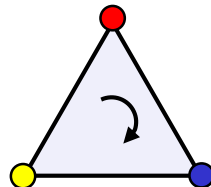
$$\sigma_4 \circ \sigma_4 = \sigma_1$$

$$\sigma_3 \circ \sigma_2 = \sigma_1$$

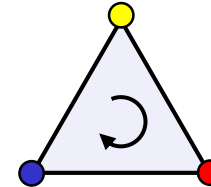
σ_1



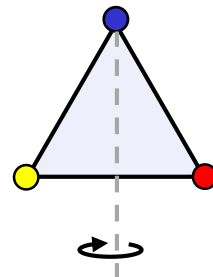
σ_2



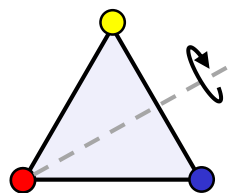
σ_3



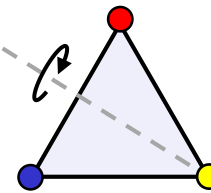
σ_4



σ_5



σ_6



$$D_3 = \{\sigma_1, \sigma_2, \dots, \sigma_6\}$$

$(D_3, \circ)$ is a group

Group exponentiation

$$g^0 \stackrel{\text{def}}{=} e$$

$$g^n \stackrel{\text{def}}{=} \overbrace{g \circ g \circ \cdots \circ g}^n$$

$$g^{-n} \stackrel{\text{def}}{=} (g^{-1})^n$$

Fact: $g^n \circ g^m = \underbrace{g \circ \cdots \circ g}_{n} \circ \underbrace{g \circ \cdots \circ g}_{m} = \underbrace{g \circ \cdots \circ g}_{n+m} = g^{n+m}$

Fact: $(g^n)^m = g^{nm} = (g^m)^n$

$$(\mathbb{Z}, +): "3^{15}" = \overbrace{3 + 3 + 3 + \cdots + 3}^{15} = 15 \cdot 3$$

Multiplicative notation

$$(G, \star): \begin{aligned} a \star b &= ab \\ a \star a \star \cdots a &= a^n \end{aligned}$$

just notation!

Additive notation

$$(G, \star): \begin{aligned} a \star b &= a + b \\ a \star a \star \cdots a &= n \cdot a \end{aligned}$$

Cyclic groups

Definition: A group is **cyclic** if there is a $g \in G$ such that

$$G = \{\dots, g^{-2}, g^{-1}, g^0, g^1, g^2, g^3, \dots\}$$

g is called a **generator** for G and we write $(G, \circ) = \langle g \rangle$

Examples:

$$(\mathbb{Z}, +) = \langle 1 \rangle = \langle -1 \rangle$$

$$(\mathbb{Z}_n, +_n) = \langle 1 \rangle$$

$$\begin{aligned} (\mathbb{Z}_7^*, \cdot) &= \langle 3 \rangle = \{3^0, 3^1, 3^2, 3^3, 3^4, 3^5\} = \{1, 3, 2, 6, 4, 5\} \\ &= \langle 5 \rangle = \{5^0, 5^1, 5^2, 5^3, 5^4, 5^5\} = \{1, 5, 4, 6, 2, 3\} \\ &\neq \langle 2 \rangle = \{2^0, 2^1, 2^2, 2^3, 2^4, 2^5\} = \{1, 2, 4, 1, 2, 4\} = \{1, 2, 4\} \end{aligned}$$

Fact: $(\mathbb{Z}_p^*, \cdot)$ is cyclic for all primes p

Not cyclic groups:

$$(\mathbb{R}, +) \quad (\mathbb{R}^*, \cdot)$$

$$(\mathbb{G}, \star)$$

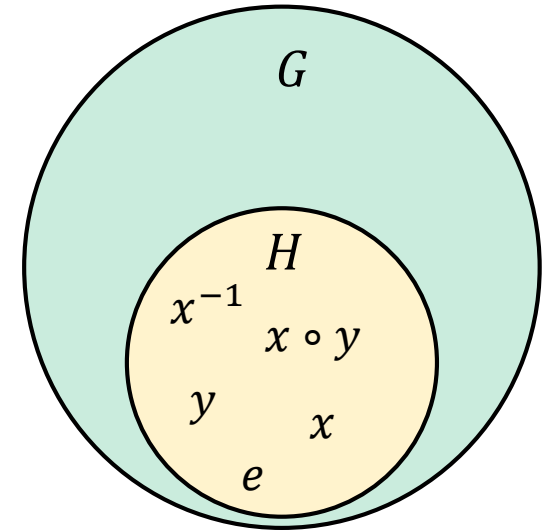
$\circ$	e	a	b	c
e	e	a	b	c
a	a	b	c	e
b	b	c	e	a
c	c	e	a	b

Subgroups

Definition: A subset $H \subseteq G$ is a **subgroup**, written

$$H < G$$

if H is a group under the binary operation from G



Examples:

$$\{e\} < G \text{ (for all groups)}$$

$$G < G \text{ (for all groups)}$$

$$2\mathbf{Z} = \{\dots, -2, 0, 2, 4, 6, \dots\} < (\mathbf{Z}, +)$$

$$3\mathbf{Z} = \{\dots, -3, 0, 3, 6, 9, \dots\} < (\mathbf{Z}, +)$$

positive real numbers

$$(\mathbf{R}_{>0}, \cdot) < (\mathbf{R}^*, \cdot)$$

$$(\{1, -1\}, \cdot) < (\mathbf{R}^*, \cdot)$$

$$\langle 20 \rangle < \langle 10 \rangle < \langle 5 \rangle < (\mathbf{Z}_{40}, +)$$

$$\langle 5 \rangle = \{0, 5, 10, \dots, 35\}$$

$$\langle 10 \rangle = \{0, 10, 20, 30\}$$

$$\langle 20 \rangle = \{0, 20\}$$

Cyclic groups

Theorem: if $(G, \circ)$ is a finite group, then for all $g \in G$:

$$g^{|G|} = e$$

Proof (finite cyclic groups):

$|G| = |\langle g \rangle| = n$

$$\begin{array}{ccccccccccc} & & & G & & & & & & & \\ & & & \underbrace{\hspace{10em}} & & & & & & & \\ e & g^1 & g^2 & g^3 & \dots & g^{n-1} & g^n & g^{n+1} & g^{n+2} & \dots \end{array}$$

$g^3 g^n = g^3 g^4 = g^7 \dots g^{n+3} = g^3 \dots g^6 = g^3 = e, j < n.. \quad \text{contradiction!}$

Theorem: $g^i = g^{i \pmod n} = g^{i \pmod{|G|}}$

Theorem (Lagrange's theorem): if $H < G$ then $|H|$ divides $|G|$

Groups of prime order

Theorem (Lagrange's theorem): if $H < G$ then $|H|$ divides $|G|$

Fact: any prime-order group is cyclic and any non-identity element ($\neq e$) is generator

Warning: $(\mathbf{Z}_p^*, \cdot)$ is *not* a prime-order group! $|\mathbf{Z}_p^*| = p - 1$

Suppose $p = 2q + 1$, with q being prime; what are the possible sub-groups of $(\mathbf{Z}_p^*, \cdot)$?

$$|\mathbf{Z}_p^*| = p - 1 = 2q$$

Example: $\mathbf{Z}_{11}^* = \{1, 2, 3, 4, 5, 6, 7, 8, 9, 10\}$ $\{1\} < \mathbf{Z}_{11}^*$

$$11 = 2 \cdot 5 + 1$$

$$\{1, -1\} = \{1, 10\} < \mathbf{Z}_{11}^*$$

$$\mathbf{Z}_p^* = \begin{cases} \{1\}, \\ \{1, -1\}, \\ H, \\ \mathbf{Z}_p^* \end{cases} \quad |H| = q$$

$$H = \langle 3 \rangle = \langle 4 \rangle = \langle 5 \rangle = \langle 9 \rangle = \{1, 3, 4, 5, 9\} < \mathbf{Z}_{11}^*$$

$$\mathbf{Z}_{11}^* < \mathbf{Z}_{11}^*$$

Why is $(\mathbb{Z}_p^*, \cdot)$ a group?

- $\mathbb{Z}_p^* = \{1, 2, \dots, p-1\}$
- Associativity ✓ $(a \cdot b) \cdot c = a \cdot (b \cdot c) \pmod p$
- Identity ✓ $1 \cdot a = a \cdot 1 = a \pmod p$
- Inverses ?

Given $a \in \mathbb{Z}_p^*$ can we always find $a^{-1} \in \mathbb{Z}_p^*$?

Need to solve: $x \cdot a = 1 \pmod p$

Claim: $\exists m, n \in \mathbb{Z}$ such that $ma + np = 1 \Rightarrow ma = 1 - np = 1 \pmod p \Rightarrow m = a^{-1} \in \mathbb{Z}_p^*$

Proof: $I = \{sa + tp \mid s, t \in \mathbb{Z}\}$

- 1) I contains a positive integer
- 2) I contains a smallest positive integer $d = ma + np$
- 3) $p = qd + r$ $0 \leq r < d$
- 4) $r = p - qd = p - q(ma + np) = -qma + (1 - qn)p \in I$
- 5) $r = 0$
- 6) $p = qd \Rightarrow d = 1$
- 7) $1 = ma + np$

Definition: A group $(G, \circ)$...

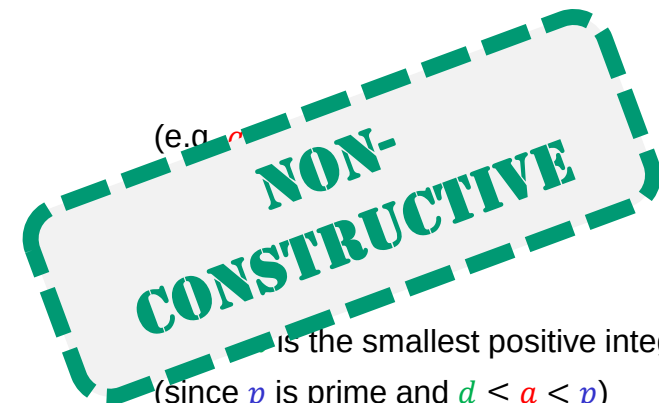
$\mathcal{G}1$: $(a \circ b) \circ c = a \circ (b \circ c)$ (associativity)

$\mathcal{G}2$: $\exists e \in G: e \circ a = a \circ e = a$ (identity)

$\mathcal{G}3$: $\exists a^{-1} \in G: a \circ a^{-1} = a^{-1} \circ a = e$ (inverse)

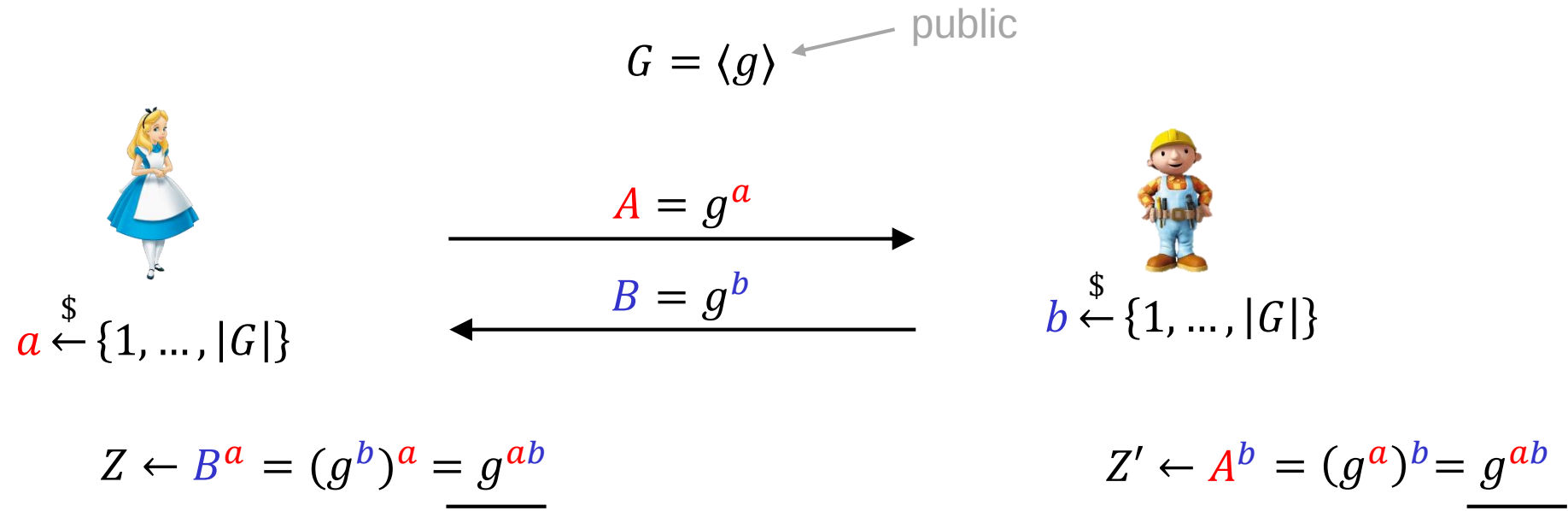
How do we actually find a^{-1} ?

Extended Euclidian Algorithm



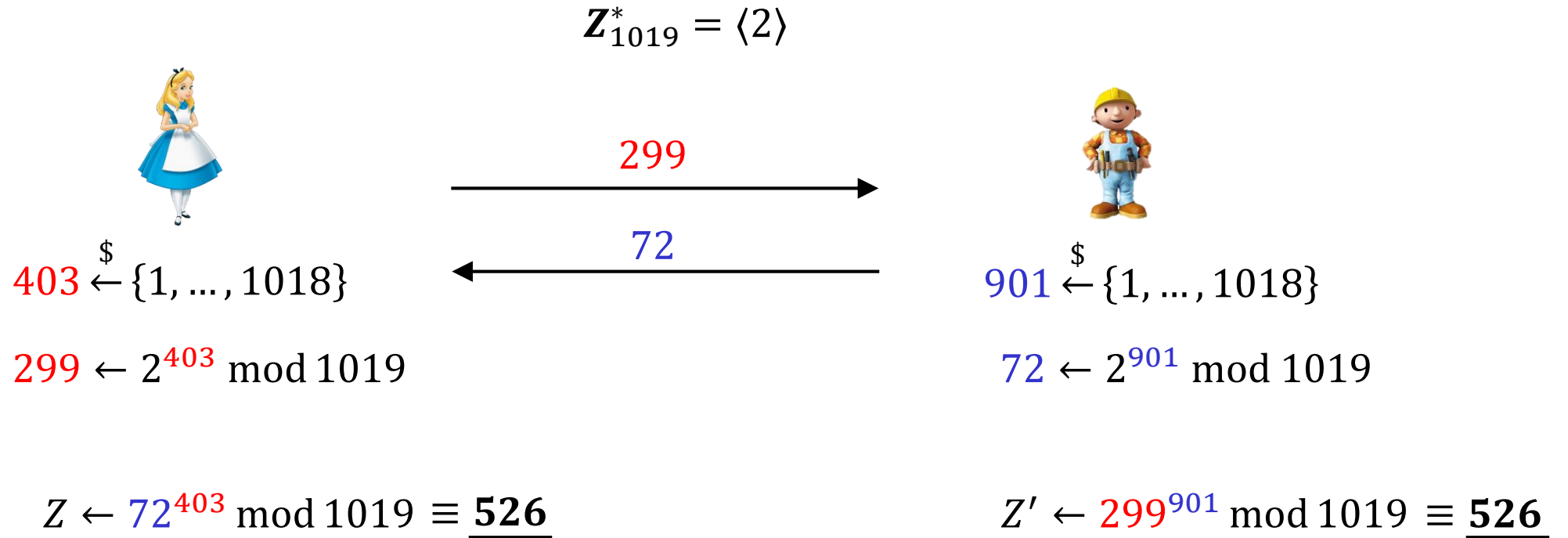
Q.E.D

Diffie-Hellman

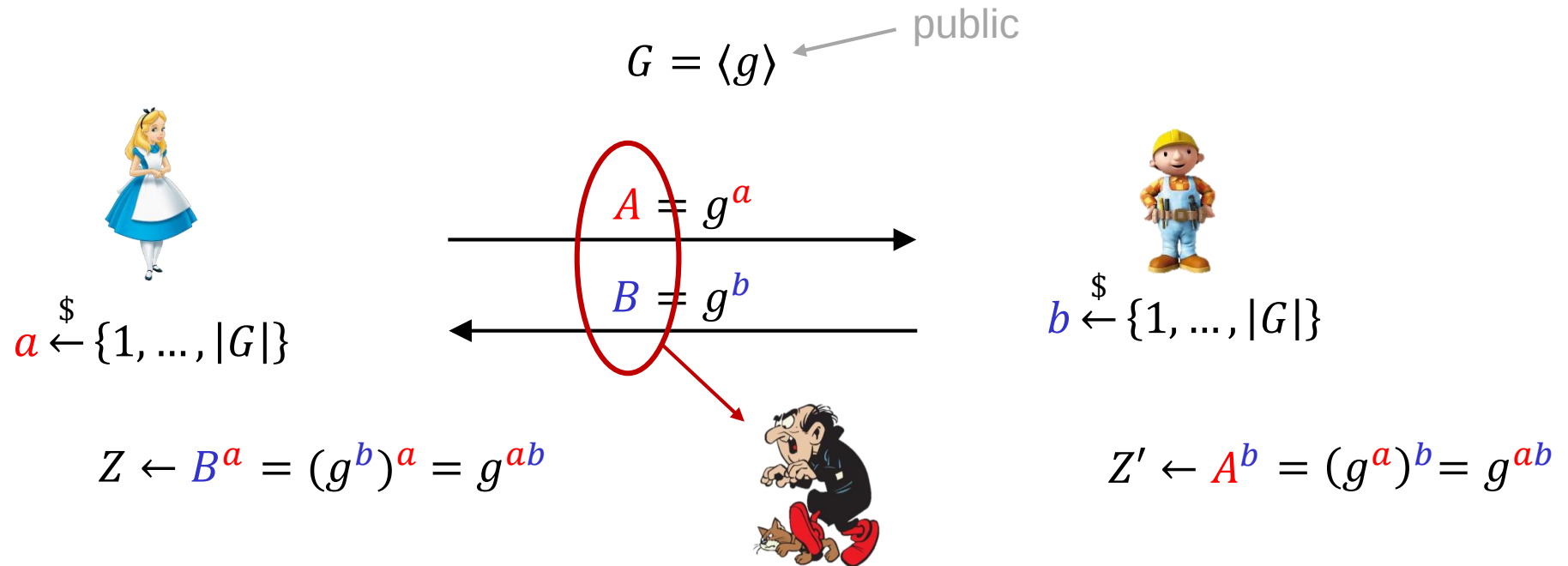


Claim: $Z = Z'$

Diffie-Hellman – example



Diffie-Hellman



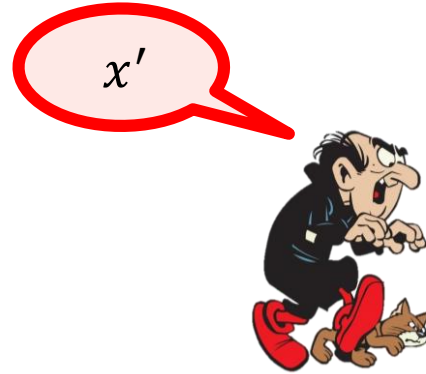
Security:

- Must be hard to compute $Z \leftarrow g^{ab}$ given g, A, B (DH assumption)
- Must be hard to find a (or b) given g, A, B (DLOG assumption)

Doesn't work: $A \circ B = g^a \circ g^b = g^{a+b} \neq g^{ab}$

Discrete logarithm (DLOG) problem

$\text{Exp}_{G,g}^{\text{dlog}}(A)$
1. $x \xleftarrow{\$} \{1, 2, \dots, G \}$
2. $X \leftarrow g^x$
3. $x' \leftarrow A(X)$
4. return $g^{x'} \stackrel{?}{=} X$



public

$$G = \langle g \rangle$$

Challenger

$$x \xleftarrow{\$} \{1, 2, \dots, |G|\}$$

$$X \leftarrow g^x$$

$$\xleftarrow{X}$$

Adversary wins if $g^{x'} = X$

In other words: $x' = \log_g X$

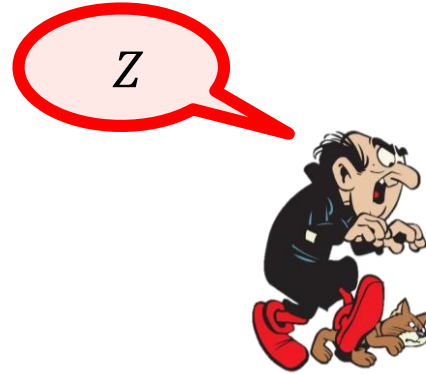
Definition: The **DLOG-advantage** of an adversary A is

$$\text{Adv}_{G,g}^{\text{dlog}}(A) = \Pr \left[\text{Exp}_{G,g}^{\text{dlog}}(A) \Rightarrow \text{true} \right]$$

Diffie-Hellman (DH) problem

$\text{Exp}_{G,g}^{\text{dh}}(A)$

1. $x, y \xleftarrow{\$} \{1, 2, \dots, |G|\}$
2. $X \leftarrow g^x$
3. $Y \leftarrow g^y$
4. $Z \leftarrow A(X, Y)$
5. **return** $Z \stackrel{?}{=} g^{xy}$



Adversary wins if $Z = g^{xy}$

public

$$G = \langle g \rangle$$

$\xleftarrow{X, Y}$

Challenger

$$x, y \xleftarrow{\$} \{1, 2, \dots, |G|\}$$

$$X \leftarrow g^x$$

$$Y \leftarrow g^y$$

Definition: The **DH-advantage** of an adversary A is

$$\text{Adv}_{G,g}^{\text{dh}}(A) = \Pr[\text{Exp}_{G,g}^{\text{dh}}(A) \Rightarrow \text{true}]$$

DLOG vs. DH

Exp_{G,g}^{dlog}(A)

1. $x \xleftarrow{\$} \{1, 2, \dots, |G|\}$
2. $X \leftarrow g^x$
3. $x' \leftarrow A(X)$
4. **return** $g^{x'} \stackrel{?}{=} x$

Exp_{G,g}^{dh}(A)

1. $x, y \xleftarrow{\$} \{1, 2, \dots, |G|\}$
2. $X \leftarrow g^x$
3. $Y \leftarrow g^y$
4. $Z \leftarrow A(X, Y)$
5. **return** $Z \stackrel{?}{=} g^{xy}$

DLOG security $\stackrel{?}{\Rightarrow}$ DH security

DLOG security $\Leftarrow$ DH security



DLOG *insecurity* $\Rightarrow$ DH *insecurity*

Algorithms for solving DLOG

- **Generic algorithms:** works for *all* (cyclic) groups
 - Brute-force
 1. Given g and $X \in G$
 2. for $i = 1, 2, \dots, |G|$ check if $g^i = X$
 - Are there better algorithms?
 - **Group-specific algorithms:** exploits algebraic features of given group
- run time: $\mathcal{O}(|G|) = (2^n)$, assuming $|G| \approx 2^n$

Solving DLOG: the baby-step giant-step algorithm

Given: $X \leftarrow g^x$

$$n \leftarrow \lceil \sqrt{|G|} \rceil$$

$Y \leftarrow g^n$

Find: x

Look for i, j such that:

$$Xg^i = Y^j$$

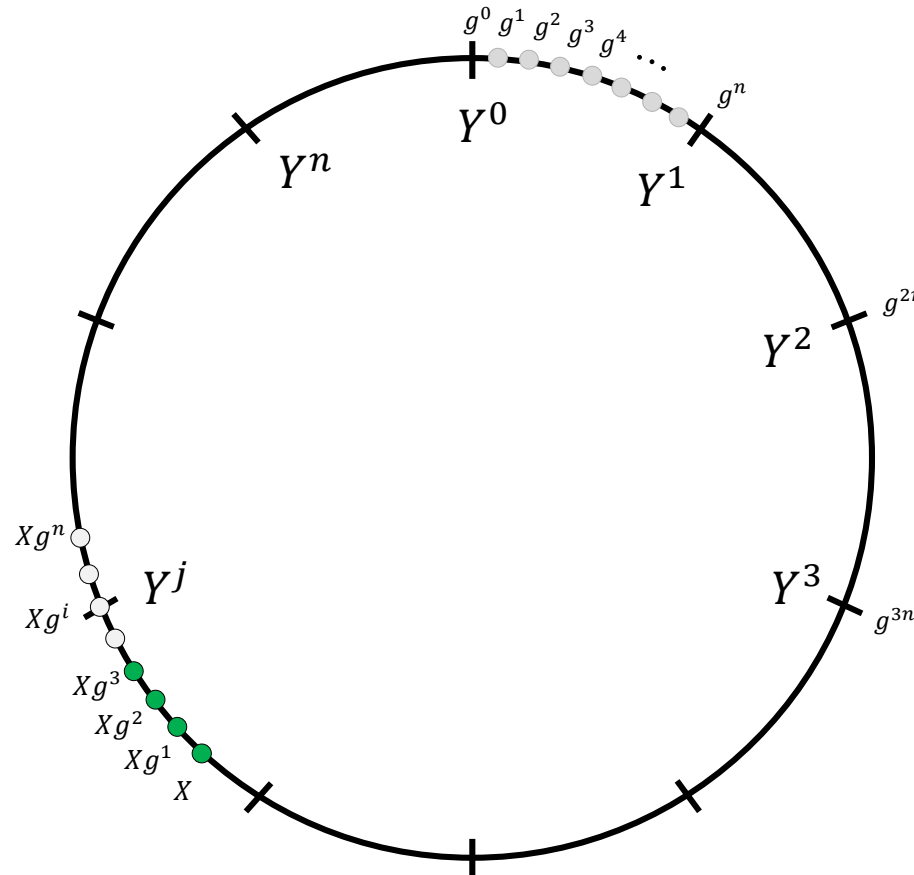
$$X = Y^j g^{-i}$$

$$X = g^{nj} g^{-i}$$

$$g^x = g^{nj} g^{-i}$$

$$g^x = g^{nj-i}$$

$$x = nj - i$$



$$\underbrace{O(\sqrt{|G|})}_{Y^0, Y^1, \dots, Y^n}$$

$$\underbrace{Xg^0, Xg^1, \dots, Xg^n}_{\tilde{O}(\sqrt{|G|})}$$

Time + memory: $\tilde{O}(\sqrt{|G|})$

Generic algorithms for solving DLOG

- Baby-step, giant-step: time $\mathcal{O}(\sqrt{|G|})$ memory $\mathcal{O}(\sqrt{|G|})$
- Pollard's rho: time $\mathcal{O}(\sqrt{|G|})$ memory $\mathcal{O}(1)$
- Pohlig-Hellman: time $\max_{p_i} \mathcal{O}(\sqrt{p_i})$ memory $\mathcal{O}(1)$ **One of the p_i must be large enough!**

$$p_1^{e_1} p_2^{e_2} \cdots p_t^{e_t}$$

- **Consequence:** for DLOG to be hard $\sqrt{|G|}$ must be large enough!
 - $|G| \approx 2^{128}$ only gives $\sqrt{2^{128}} = 2^{64}$ security
 - $|G| \approx 2^{256}$ only gives $\sqrt{2^{256}} = 2^{128}$ security
 - $|G| \approx 2^{512}$ only gives $\sqrt{2^{512}} = 2^{256}$ security
 - etc...
- **Nechaev'94 & Shoup'97:** Solving DLOG requires $\Omega(\sqrt{|G|})$ time in *generic* groups

Non-generic algorithms for DLOG

- Unfortunately, $(\mathbf{Z}_p^*, \cdot)$ is *not* a generic group!
- *Much* faster specific algorithms exist for solving DLOG in $\mathbf{Z}_p^*$
 - Index-calculus
 - Elliptic-curve method
 - Special number-field sieve (SNFS)
 - General number-field sieve (GNFS)

} *exceptionally complicated algorithms, requiring very advanced mathematics!*
- Current DLOG-solving record: $|\mathbf{Z}_p^*| \approx 2^{795}$ using GNFS (Heninger et al. '19)
 - Previous records: https://en.wikipedia.org/wiki/Discrete_logarithm_records
- $|\mathbf{Z}_p^*| \geq 2^{2048}$ typically required as a minimum today



Z_p^*

Group
where GNFS
doesn't work

Next week: better alternatives to Z_p^* ?