Lecture 07: Graph Representation

Private-key encryption: correctness and privacy

Abstract and reading guide

Abstract

An encryption algorithm draws edges from messages to ciphertexts, with secret keys as labels. Correctness restricts which labels can meet at a ciphertext; privacy requires equal numbers of keys from every message. These two rules prove the lecture's ciphertext and key lower bounds, and explain the optimality of one-time pad.

Read the instructor's revised Lecture 07 first. It is the minimum course material. These companion notes expand its four properties and work through its closing questions.

Added examples are marked as additional reading. The note with Property 1 distinguishes the existence of a correct decoder from its implementation.

Primary source: [Maj26], slides 1–15.

The model: two simplifying assumptions

The message, key, and ciphertext spaces are finite, nonempty sets:

$$m \in \mathcal{M}, \qquad \mathsf{sk} \in \mathcal{K}, \qquad c \in \mathcal{C}.$$

Uniform key generation

Gen samples each $\mathsf{sk} \in \mathcal{K}$ with probability $1/|\mathcal{K}|$.
The key is independent of the message.

Deterministic encryption

For every pair (m, sk), encryption returns exactly one ciphertext:

$$\mathsf{Enc}_{\mathsf{sk}} : \mathcal{M} \longrightarrow \mathcal{C}.$$

$\mathsf{Dec}_{\mathsf{sk}}(c)$ is Bob's decryption rule. We study **one encryption** under the sampled key.

Relaxing these restrictions is outside this lecture. A declared ciphertext need not actually be produced.

[Maj26], slides 3–4.

Correctness, privacy, and security

Correctness

For every $m \in \mathcal{M}$ and $\mathsf{sk} \in \mathcal{K}$,

$$\mathsf{Dec}_{\mathsf{sk}}(\mathsf{Enc}_{\mathsf{sk}}(m)) = m.$$

This is an exact requirement, not an average over messages.

Privacy

Let M be a random message and let $C = \mathsf{Enc}_{\mathsf{sk}}(M)$.
For every distribution of M, the random variables M and C are independent.

Course convention: security = correctness + privacy.

Privacy here is perfect privacy. We will prove its edge-count characterization rather than assume it.

[Maj26], slides 2 and 9; further reading: [BS23], Theorem 2.4, pp. 11–12.

Introduction: four properties we will prove

1. Correctness: distinct outputs

A correct decoder exists exactly when each key maps distinct messages to distinct ciphertexts.

2. Ciphertext-space lower bound

Every correct scheme satisfies

$$|\mathcal{C}| \geq |\mathcal{M}|.$$

3. Privacy: equal key counts

For each ciphertext, the number of compatible keys is the same for every message.

4. Key-space lower bound

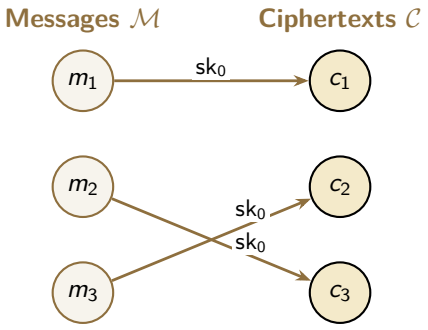
Every secure scheme satisfies

$$|\mathcal{K}| \geq |\mathcal{M}|.$$

One-time pad attains equality in both cardinality bounds.

[Maj26], slides 6–14. The decoder qualification is explained with Property 1.

Draw an edge for every encryption



Only the edges for one key are shown.

A **bipartite graph** has two vertex sets; each edge joins opposite sides.
Here, draw

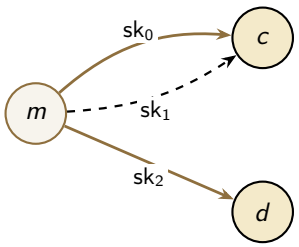
$$m \xrightarrow{\mathsf{sk}} c \quad \Longleftrightarrow \quad c = \mathsf{Enc}_{\mathsf{sk}}(m).$$

The key label witnesses an encryption.

For every pair (m, sk), exactly one edge leaves m with label sk.

Graph representation: [Maj26], slides 4–5. Illustration added.

Several keys may connect the same pair



$\mathcal{K} = \{\mathsf{sk}_0, \mathsf{sk}_1, \mathsf{sk}_2\}$; one message shown.

A graph allowing parallel edges is a **multigraph**.

The **edge weight** counts compatible keys:

$$\mathrm{wt}(m, c) = \big|\{\mathsf{sk} \in \mathcal{K} : \mathsf{Enc}_{\mathsf{sk}}(m) = c\}\big|.$$

In the picture, $\mathrm{wt}(m, c) = 2$ and $\mathrm{wt}(m, d) = 1$.

Count all edges leaving a message

$$\sum_{c \in \mathcal{C}} \mathrm{wt}(m, c) = |\mathcal{K}| \qquad (m \in \mathcal{M}).$$

Every key contributes exactly once.

[Maj26], slide 4. Here wt means edge multiplicity, not Hamming weight.

A running picture: the one-bit pad

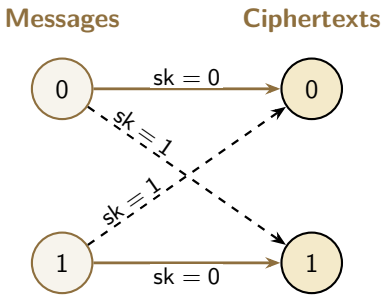
Worked illustration of the graph definition; one-time pad is revisited after the bounds.

Take $\mathcal{M} = \mathcal{K} = \mathcal{C} = \{0,1\}$ and $\mathsf{Enc}_{\mathsf{sk}}(m) = m \oplus \mathsf{sk}$, where $\oplus$ is XOR.

Message	$\mathsf{sk}=0$	$\mathsf{sk}=1$
0	0	1
1	1	0

Every message–ciphertext pair has exactly one compatible key.

$$\mathrm{wt}(m, c) = 1.$$



[Maj26], slides 4 and 14.

Property 1: correctness is fixed-key uniqueness

Theorem: when does a correct decoder exist?

For a deterministic encryption family, a correct decryption rule exists if and only if each $\mathsf{Enc}_{\mathsf{sk}}$ is **one-to-one**:

$$\mathsf{Enc}_{\mathsf{sk}}(m) = \mathsf{Enc}_{\mathsf{sk}}(m') \quad \Longrightarrow \quad m = m'.$$

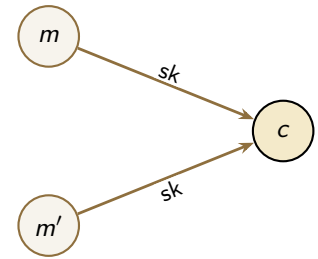
Proof idea. A repeated label at one ciphertext makes decoding ambiguous. Without a repeated label, reverse the unique edge.

Equivalent graph rule. Two edges with the same key label cannot arrive at one ciphertext from distinct messages.

Precision note on source slide 6: this characterizes the *existence* of a correct decoder. An incorrectly implemented Dec can fail even when all encryption maps are one-to-one.

[Maj26], slide 6, with the stated decoder-existence qualification.

A repeated key label makes decoding impossible



$m \neq m'$

Suppose distinct messages satisfy

$$\mathrm{Enc}_{\mathsf{sk}}(m) = c = \mathrm{Enc}_{\mathsf{sk}}(m').$$

Bob receives the same input (sk, c) in both cases.

Correctness would require both

$$\mathrm{Dec}_{\mathsf{sk}}(c) = m \quad \text{and} \quad \mathrm{Dec}_{\mathsf{sk}}(c) = m'.$$

That is impossible because $m \neq m'$. □

The obstruction is the **same key** on both edges. Different keys do not create ambiguity for Bob, who knows his key.

[Maj26], slide 6.

Conversely, reverse the unique edge

Assume that no ciphertext has two incoming edges with the same key label.

Construct the decoder

Given (sk, c), look for a message m such that

$$m \xrightarrow{\mathsf{sk}} c.$$

If such a message exists, it is unique: output it. If no such message exists, choose any fixed message as the output.

For an honestly generated ciphertext $c = \mathsf{Enc}_{\mathsf{sk}}(m)$, the edge exists. Its unique left endpoint is exactly m. Hence

$$\mathsf{Dec}_{\mathsf{sk}}(\mathsf{Enc}_{\mathsf{sk}}(m)) = m. \qquad \square$$

This is an existence argument, not an efficiency claim. Correctness places no requirement on a pair (sk, c) that encryption never produces.

[Maj26], slides 5–6.

Property 2: correct encryption cannot compress

Theorem

Every correct private-key encryption scheme satisfies

$$\boxed{|\mathcal{C}| \geq |\mathcal{M}|.}$$

Proof idea. Fix one key and count its distinct outputs.

Fix $\mathsf{sk} \in \mathcal{K}$. Property 1 says that the list

$$\{\mathsf{Enc}_{\mathsf{sk}}(m) : m \in \mathcal{M}\}$$

contains $|\mathcal{M}|$ distinct elements of $\mathcal{C}$. Thus $|\mathcal{C}| \geq |\mathcal{M}|$. □

The source's pigeonhole proof, in one sentence. If fewer ciphertexts than messages were available, two messages would share an output under this fixed key, contradicting correctness.

[Maj26], slides 7–8. No privacy assumption is used.

Counting keys gives ciphertext probabilities

Fix a message m. Define its ciphertext distribution by

$$q_m(c) := \Pr_{\mathsf{sk} \leftarrow \mathsf{Gen}}\left[\mathsf{Enc}_{\mathsf{sk}}(m) = c\right].$$

The bridge from a graph to a probability

$$\boxed{q_m(c) = \frac{\mathrm{wt}(m, c)}{|\mathcal{K}|}.}$$

The numerator counts successful keys; each key has probability $1/|\mathcal{K}|$.

For a random message M, independence of the key and message gives

$$\Pr[C = c \mid M = m] = q_m(c) \qquad \text{when } \Pr[M = m] > 0.$$

The definition of q_m remains valid even when the chosen message distribution assigns probability zero to m.

[Maj26], slide 9.

Property 3: privacy is equality of edge weights

Theorem

The scheme is private if and only if, for every $c \in \mathcal{C}$,

$$\boxed{\mathrm{wt}(m, c) = \mathrm{wt}(m', c) \quad \text{for all } m, m' \in \mathcal{M}.}$$

The common weight may depend on c.

Proof idea. First show that privacy is equivalent to all q_m being identical. Then use $q_m(c) = \mathrm{wt}(m, c)/|\mathcal{K}|$.

For a fixed ciphertext, compare **different messages**. There is no requirement that different ciphertexts have the same probability.

We now solve the equivalence exercise on source slide 9. No correctness assumption is needed for Property 3.

[Maj26], slide 9; further reading: [BS23], Theorem 2.1, pp. 8–9.

Equal ciphertext laws imply privacy

Proof idea. Factor the joint distribution into its two marginals.

Suppose $q_m(c) = q(c)$ for every message m. Give M any distribution, and write $p(m) = \Pr[M = m]$.

Independence of the key and message gives, for every m, c,

$$\Pr[M = m, C = c] = p(m)q_m(c) = p(m)q(c).$$

Summing over the messages,

$$\Pr[C = c] = \sum_{m \in \mathcal{M}} p(m)q(c) = q(c).$$

Thus the joint probability is $\Pr[M = m] \Pr[C = c]$, proving independence. □

Equivalently, whenever $q(c) > 0$, $\Pr[M = m \mid C = c] = p(m)$: the observation leaves the prior unchanged.

Expanded solution of [Maj26], slide 9; compare [BS23], Theorem 2.4, pp. 11–12.

Proof idea. Use a message distribution in which every message is possible.

Assume privacy, and choose M uniformly from $\mathcal{M}$, independently of the key. Every message now has positive probability.

For all $m \in \mathcal{M}$ and $c \in \mathcal{C}$, independence of M and C gives

$$q_m(c) = \Pr[C = c \mid M = m] = \Pr[C = c].$$

The right-hand side does not depend on m. Hence all q_m are identical, and the edge weights are identical across messages. □

What was used?

The definition requires privacy for *every* message distribution. We used the uniform distribution only to prove the converse.

Expanded solution of [Maj26], slide 9.

A used ciphertext must be compatible with every message

The first consequence of privacy

Suppose the scheme is private and $m \xrightarrow{\mathsf{sk}} c$ for some m, sk. Then, for every $m' \in \mathcal{M}$, there is some $\mathsf{sk}' \in \mathcal{K}$ such that

$$m' \xrightarrow{\mathsf{sk}'} c.$$

Proof idea. A positive edge count stays positive when the message changes.

The given edge implies $\mathsf{wt}(m, c) \geq 1$. Property 3 therefore gives

$$\mathsf{wt}(m', c) = \mathsf{wt}(m, c) \geq 1 \qquad \text{for every } m' \in \mathcal{M}. \qquad \square$$

A *used* ciphertext is one produced by at least one message–key pair. An unused ciphertext has weight zero for every message. Privacy alone does not say the witnessing keys are distinct.

[Maj26], slide 10, first question.

Property 4: correct and private schemes need many keys

Theorem

Every secure private-key encryption scheme satisfies

$$\boxed{|\mathcal{K}| \geq |\mathcal{M}|.}$$

Proof idea. Fix one used ciphertext. Privacy supplies a key for every message; correctness forces those keys to be different.

Privacy supplies the edges

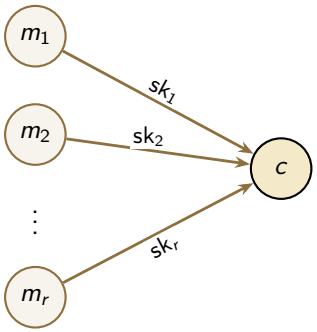
Every message has an edge into the chosen ciphertext.

Correctness separates the labels

Edges from distinct messages cannot carry the same key.

[Maj26], slides 11–13; further reading: [BS23], Theorem 2.5, pp. 12–13.

The key bound, proved at one ciphertext



Let $r = |\mathcal{M}|$, and choose any used ciphertext c.
By privacy, choose a key sk_i with $\mathsf{Enc}_{\mathsf{sk}_i}(m_i) = c$ for each message m_i.
If $\mathsf{sk}_i = \mathsf{sk}_j$, then the same key maps both m_i and m_j to c. Correctness forces $i = j$.
There are therefore r distinct keys in $\mathcal{K}$:

$$|\mathcal{K}| \geq r = |\mathcal{M}|. \qquad \square$$

[Maj26], slides 11–13. Historical comparison: [Sha49], Section 10, p. 681.

A sharper count: several compatible keys per message

The weighted bound from the source's second question

Suppose the scheme is **correct and private**. If $\mathrm{wt}(m, c) = w$ for one message m, then

$$\boxed{|\mathcal{K}| \geq w|\mathcal{M}|.}$$

Proof idea. Count disjoint sets of compatible keys, not just one witness per message.

For each $m' \in \mathcal{M}$, let $A_{m'} = \{\mathsf{sk} : \mathsf{Enc}_{\mathsf{sk}}(m') = c\}$. Privacy gives $|A_{m'}| = w$. Correctness says these sets are pairwise disjoint. Thus

$$|\mathcal{K}| \;\geq\; \left| \bigcup_{m' \in \mathcal{M}} A_{m'} \right| = \sum_{m' \in \mathcal{M}} |A_{m'}| = w|\mathcal{M}|. \qquad \square$$

For a used ciphertext, $w \geq 1$, so this independently recovers Property 4. For an unused ciphertext, $w = 0$ and the bound is immediate.

Expanded proof of [Maj26], slide 10, second question.

One-time pad attains both bounds

Let $\mathcal{M} = \mathcal{K} = \mathcal{C} = \{0,1\}^n$, with a uniform key, and set

$$\mathsf{Enc}_{\mathsf{sk}}(m) = m \oplus \mathsf{sk}, \qquad \mathsf{Dec}_{\mathsf{sk}}(c) = c \oplus \mathsf{sk}.$$

Correctness

XOR with the key twice cancels:

$$(m \oplus \mathsf{sk}) \oplus \mathsf{sk} = m.$$

Privacy

For every m, c, the unique key is

$$\mathsf{sk} = m \oplus c.$$

Thus $\mathsf{wt}(m, c) = 1$ for every pair.

$$|\mathcal{K}| = |\mathcal{C}| = |\mathcal{M}| = 2^n.$$

One-time pad is optimal for both **key-space size** and **ciphertext-space size**.

[Maj26], slide 14; further reading: [BS23], Theorem 2.2, p. 9.

Why both correctness and privacy are necessary

Assume $|\mathcal{M}| \geq 2$. Both examples use a single key: $|\mathcal{K}| = 1$.

Correct, but not private

Set $\mathcal{C} = \mathcal{M}$ and

$$\mathsf{Enc}_{\mathsf{sk}}(m) = m, \qquad \mathsf{Dec}_{\mathsf{sk}}(c) = c.$$

Bob recovers every message.
The ciphertext reveals the message itself, so privacy fails.

Private, but not correct

Set $\mathcal{C} = \{c_0\}$ and

$$\mathsf{Enc}_{\mathsf{sk}}(m) = c_0 \quad \text{for every } m.$$

The ciphertext is independent of the message.
One output cannot decode to two different messages.

Neither scheme is secure: the first lacks privacy, and the second lacks correctness. Security requires both properties.

Solutions to [Maj26], slide 15. The nontrivial-message assumption is explicit.

The core lecture, on one page

Requirement	What to inspect	Consequence
Correctness	Fix one key: its outputs on distinct messages must differ.	$\lvert\mathcal{C}\rvert \geq \lvert\mathcal{M}\rvert$.
Privacy	Fix one ciphertext: all messages have the same edge weight.	All used ciphertexts remain possible for every message.
Correctness + privacy	Fix one used ciphertext: incoming keys for different messages are disjoint.	$\lvert\mathcal{K}\rvert \geq \lvert\mathcal{M}\rvert$.

The key-bound proof is local: one ciphertext, all messages, distinct keys.

[Maj26], slides 6–14.

Additional reading: what the graph does not say

The four properties above complete the source lecture's main argument. The next examples sharpen three distinctions.

Possibility is weaker than privacy

Every message may be compatible with a ciphertext, yet their compatible-key counts may be different.

Equal across messages does not mean uniform over ciphertexts

Privacy permits different ciphertexts to have different probabilities.

A lower bound is not a divisibility condition. The key-space size need not be a multiple of the message-space size.

Added worked illustrations of Property 3 and the questions in [Maj26], slide 10.

Possibility alone does not imply privacy

Additional reading: a worked illustration added to the source lecture.

Take $\mathcal{M} = \mathcal{C} = \{0, 1\}$ and three equally likely keys.

	sk_0	sk_1	sk_2
$m = 0$	0	0	1
$m = 1$	1	1	0

Each key's column has distinct outputs, so a correct decoder exists.

Every message can produce either ciphertext. But

$q_m(c)$	$c = 0$	$c = 1$
$m = 0$	$2/3$	$1/3$
$m = 1$	$1/3$	$2/3$

The two distributions differ, so privacy fails.

With a uniform message, observing $C = 0$ changes the probability of $M = 0$ from $1/2$ to $2/3$.

Illustrates why the implication on [Maj26], slide 10, is not an equivalence.

Privacy does not require uniform ciphertexts

Additional reading: a worked illustration added to the source lecture.

Let $\mathcal{M} = \{0,1\}$, $\mathcal{C} = \{0,1,2,3\}$, and use six equally likely keys.

	sk_0	sk_1	sk_2	sk_3	sk_4	sk_5
$m=0$	0	0	1	1	2	3
$m=1$	1	1	0	0	3	2

Every fixed key maps the two messages to different ciphertexts.
Therefore a correct decoder exists.

For each message, the weight row is

c	0	1	2	3
$\mathrm{wt}(m,c)$	2	2	1	1.

Hence, for *both* messages,

$$q_m = (1/3,\ 1/3,\ 1/6,\ 1/6).$$

These laws are identical but not uniform.

Two different tables. In the encryption table, columns are keys. In the weight table, columns are ciphertexts.

Worked illustration of [Maj26], slide 9.

The key count need not be divisible by the message count

Worked solution of the divisibility question on source slide 10.

Take $\mathcal{M} = \{0,1\}$ and $\mathcal{K} = \mathcal{C} = \mathbb{Z}/3\mathbb{Z}$. Define

$$\mathrm{Enc}_{\mathsf{sk}}(m) = m + \mathsf{sk} \pmod 3.$$

	$\mathsf{sk}=0$	$\mathsf{sk}=1$	$\mathsf{sk}=2$
$m=0$	0	1	2
$m=1$	1	2	0

Correctness. Compute $c - \mathsf{sk} \pmod 3$. For an honestly generated ciphertext, this equals $m \in \{0,1\}$.

Privacy. For every m, c, exactly one key works: $\mathsf{sk} = c - m \pmod 3$. Thus $\mathrm{wt}(m,c) = 1$.

The promised example

$$|\mathcal{K}| = 3 \geq 2 = |\mathcal{M}|, \qquad 2 \nmid 3.$$

On an unused pair giving residue 2, output any fixed message.

[Maj26], slide 10.

Check your understanding

1. Which keys must be different?

Two distinct messages produce the same ciphertext under two *different* keys. Does this violate correctness? Explain who knows the key.

2. Which equality is the privacy test?

Should one compare $\mathrm{wt}(m, c)$ with $\mathrm{wt}(m, c')$, or with $\mathrm{wt}(m', c)$? State the quantifiers.

3. Count at one ciphertext. A correct and private scheme has five messages. One message has three keys that produce a particular ciphertext. How many keys must the scheme have?

Added checks on the arguments in [Maj26], slides 6, 9, and 10.

Answers: keep the two viewpoints separate

1. No. Bob is given (sk, c), not just c. A collision obstructs correctness only when the key is also the same.

2. Fix the ciphertext and vary the message. The condition is

$$\forall c \in \mathcal{C} \quad \forall m, m' \in \mathcal{M} : \qquad \mathsf{wt}(m, c) = \mathsf{wt}(m', c).$$

3. At least 15 keys. Privacy gives three compatible keys for each of five messages. Correctness makes the five key sets disjoint, so

$$|\mathcal{K}| \geq 3 \cdot 5 = 15.$$

Correctness tracks labels. **Privacy counts labels.**

Answers derived from Properties 1 and 3 and the weighted bound.

Conventions and hypotheses to keep in view

Property 1: existence of a decoder versus a given decoder

The graph characterizes whether correct decryption is possible. To conclude correctness of a supplied Dec, it must actually return the unique preimage on every generated ciphertext. This distinction still applies to source slide 6.

Weighted bound: correctness and privacy

Privacy gives equally sized compatible-key sets. Correctness makes them disjoint. These are the two hypotheses in source slide 10 for the bound $|\mathcal{K}| \geq w|\mathcal{M}|$.

Separation examples: correct or private is not enough

For $|\mathcal{M}| \geq 2$, cleartext encryption is correct but not private; constant-output encryption is private but not correct. Neither scheme is secure. These are the two examples requested on source slide 15.

Reading notes for [Maj26], slides 6, 10, and 15.

Sources and further reading

The revised lecture is the primary source

[Maj26]: graph definition, slides 4–5; correctness and ciphertext bound, slides 6–8; privacy, slides 9–10; key bound, slides 11–13; one-time pad and closing questions, slides 14–15.

Further reading, not additional required course material.

Boneh–Shoup, version 0.6. Theorem 2.1, pp. 8–9 (equal key counts); Theorem 2.2, p. 9 (one-time pad); Theorem 2.4, pp. 11–12 (independence); Theorem 2.5, pp. 12–13 (key bound). [BS23]

Shannon. Section 10, Theorem 6, p. 680 (perfect secrecy); p. 681 (the counting argument and a key-labelled graph in Figure 5). [Sha49]

All four properties are proved here. These references are parallels and further reading, not assumed technical results. Page numbers above are printed page numbers, not PDF-viewer indices.

References

[BS23] Dan Boneh and Victor Shoup.
A Graduate Course in Applied Cryptography.
Author-maintained textbook, 2023.
Version 0.6, January 2023. Further reading: Theorems 2.1–2.2, pp. 8–9; Theorem 2.4, pp. 11–12; Theorem 2.5, pp. 12–13.
URL: https://toc.cryptobook.us/book.pdf.

[Maj26] **Hemanta K. Maji.**
Lecture 07: Graph representation.
CS 35500, Purdue University. Instructor's lecture slides, 2026.
Supplied revised file: 07(1).pdf, 15 slides. Primary source for these companion notes.

[Sha49] **Claude E. Shannon.**
Communication theory of secrecy systems.
Bell System Technical Journal, 28(4):656–715, 1949.
Further reading: Section 10, Theorem 6, p. 680; cardinality argument and Figure 5, p. 681.
URL: https://pages.cs.wisc.edu/~rist/642-spring-2014/shannon-secrecy.pdf.