USING STRUCTURED VARIANTS IN LATTICE-BASED CRYPTOGRAPHY

Adeline Roux-Langlois

Normandie Univ, UNICAEN, ENSICAEN, CNRS, GREYC, Caen, FRANCE

GREYC Electronics and Computer Science Laboratory

Normandie Université

UNICAEN

ENSI CAEN ÉCOLE PUBLIQUE D'INGÉNIEURS CENTRE DE RECHERCHE

cnrs

Cryptography

Let's start with a simple example: you want to send a message to someone.

Two possibilities:

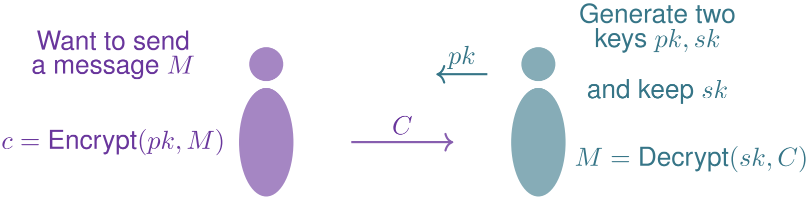
- Either you share a secret key (AES...),
- Either you don't $\Rightarrow$ public key cryptography.

Cryptography

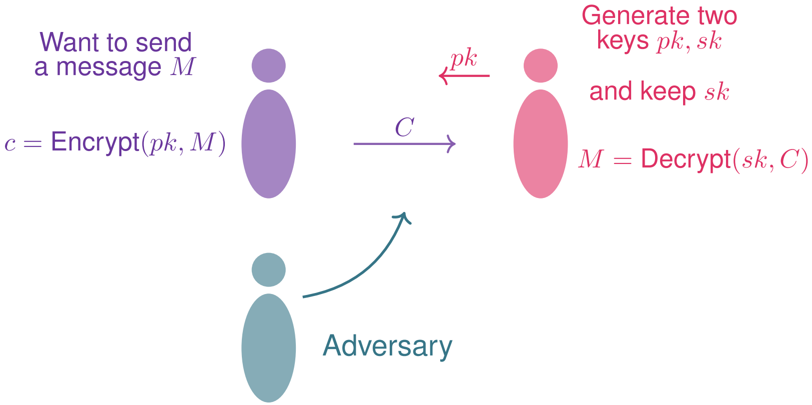
Let's start with a simple example: you want to send a message to someone.

Two possibilities:

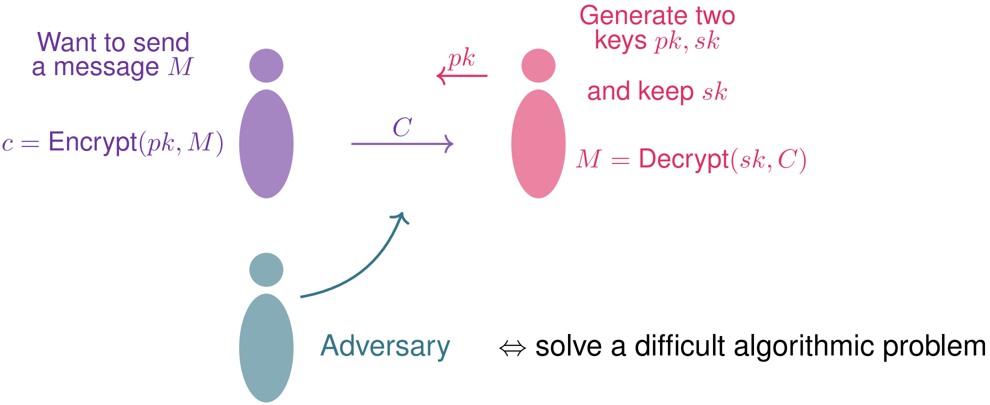
- Either you share a secret key (AES...),
- Either you don't $\Rightarrow$ public key cryptography.



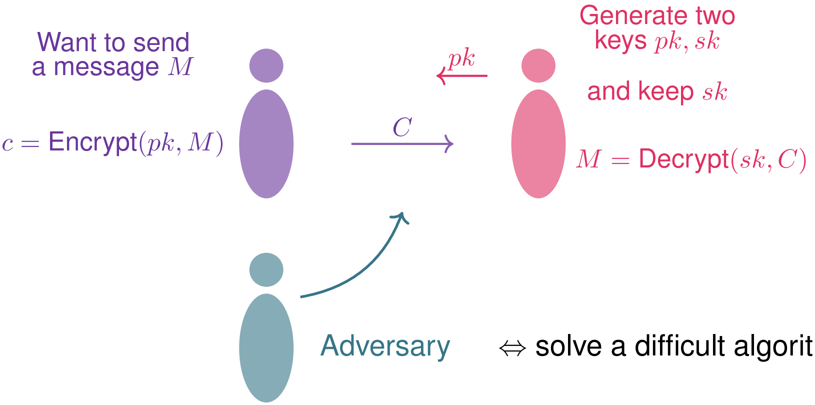
Public key cryptography



Public key cryptography



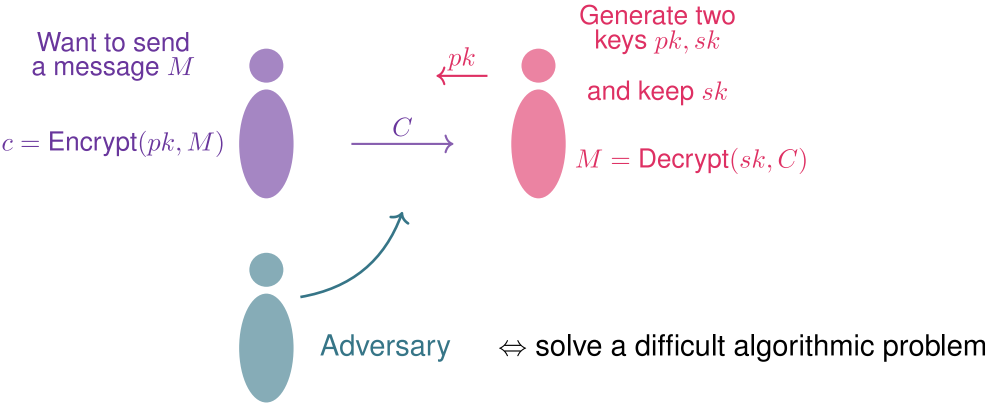
Public key cryptography



$\Leftrightarrow$ solve a difficult algorithmic problem

- Examples: factorisation (RSA), discrete log (El Gamal) ...
- Solving those problems needs an exponential complexity on a classical computer.

Public key cryptography



- Examples: factorisation (RSA), discrete log (El Gamal) ...
- Solving those problems needs an exponential complexity on a classical computer.
- Shor's algorithm (1997): **polynomial time on a quantum computer**.

Context

New goals in cryptography

- Resisting to quantum computers,
- Need of new functionalities,

$\rightarrow$ need alternatives

- Post-quantum secure,
- Efficient,
- New functionalities, different types of constructions.

NIST competition

- Lattice-based cryptography
- Code-based cryptography
- Multivariate, Isogenies, Hash based ...

NIST competition

From 2017 to 2024, NIST competition to develop new standards
on post-quantum cryptography

Total: 69 accepted submissions (round 1)

- Signature (5 lattice-based),
- Public key encryption / Key Encapsulation Mechanism (21 lattice-based)

Other candidates: 17 code-based PKE, 7 multivariate signatures, 3 hash-based signatures, 7 from ”other” assumptions (isogenies, PQ RSA ...) and 4 attacked + 5 withdrawn.

$\Rightarrow$ lattice-based constructions are very serious candidates

5 over 7 finalists are lattice-based

2022 first results: **3 over 4 new standards** are lattice-based

Why lattice-based cryptography?

- Likely to resist attacks from quantum computers,
- Strong security guarantees,
  from well-understood hard problems on lattices.

- Novel and powerful cryptographic functionalities,
  - Public key encryption and signature scheme (practical),
  - Advanced signature (group signature ...),
    and encryption scheme (IBE, ABE, ...),
  - Fully homomorphic encryption.

- Efficiency

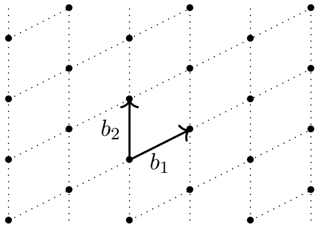
Today: an introduction to lattice-based cryptography

1. Lattices
   - Definition
   - Hard problem on lattices

2. The Learning With Errors problem
   - Definition
   - Difficulty
   - How to encrypt using LWE?

3. Practical scheme
   - Adding structure
   - Module-LWE
   - Kyber encryption scheme

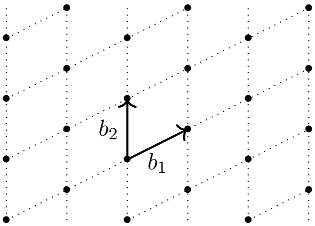
Lattices

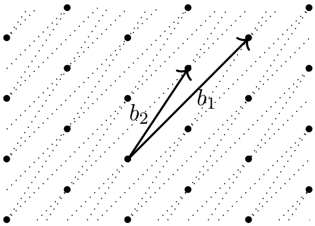


Lattice

$\mathcal{L}(\mathbf{B}) = \{\sum_{1=i}^{n} a_i \mathbf{b}_i, a_i \in \mathbb{Z}\}$, where the $(\mathbf{b}_i)_{1 \leq i \leq n}$'s, linearly independent vectors, are a basis of $\mathcal{L}(\mathbf{B})$.

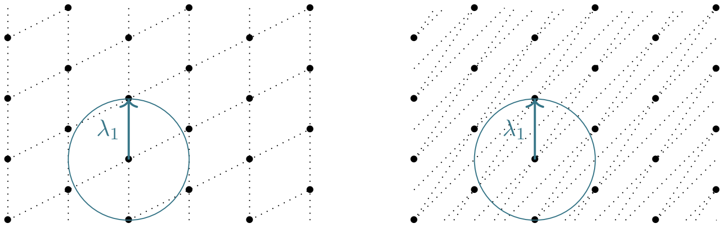
Lattices





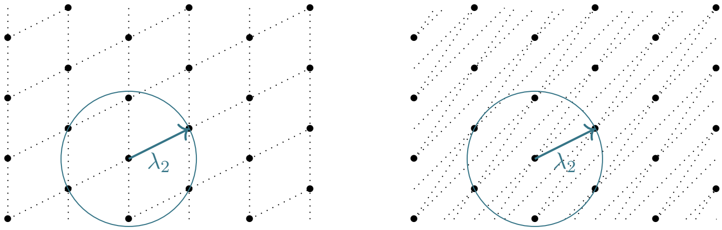
- Several basis define a lattice, some are better.

Lattices



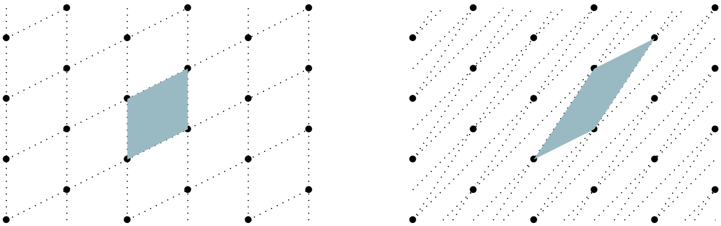
- Several basis define a lattice, some are better.
- The first minimum λ_1 is the norm of the smallest non-zero vector.

Lattices



- Several basis define a lattice, some are better.
- The first minimum λ_1 is the norm of the smallest non-zero vector.
- The n-th minima λ_n is the radius of a sphere which contains n linearly independent shortest vectors of the lattices.

Lattices



- Several basis define a lattice, some are better.
- The first minimum λ_1 is the norm of the smallest non-zero vector.
- The n-th minima λ_n is the radius of a sphere which contains n linearly independent shortest vectors of the lattices.
- The fundamental parallelepiped is defined by $\mathcal{P}(\mathbf{B}) = \{\sum_{i=1}^{n} c_i \mathbf{b}_i : c_i \in [0,1)\}$. Its volume defines the volume of the lattice: $\det(\Lambda) = |\det(\mathbf{B})|$.

Lattices

- The first minimum λ_1 is the norm of the smallest non-zero vector.
- The n-th minima λ_n is the radius of a sphere which contains n linearly independent shortest vectors of the lattices.
- The fundamental parallelepiped is defined by $\mathcal{P}(\mathbf{B}) = \{\sum_{i=1}^{n} c_i \mathbf{b}_i : c_i \in [0,1)\}$. Its volume defines the volume of the lattice: $\det(\Lambda) = |\det(\mathbf{B})|$.

- Minkowski Theorem:

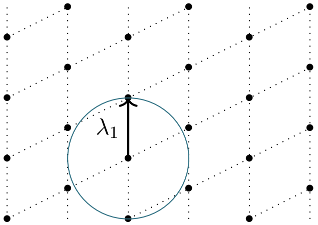
$$\lambda_1(\Lambda) \leq \sqrt{n} \cdot \det(\Lambda)^{1/n}$$

$$\left(\prod_{i=1}^{n} \lambda_i(\Lambda)\right)^{1/n} \leq \sqrt{n} \cdot \det(\Lambda)^{1/n}$$

Shortest Vector Problem (SVP)

Given a lattice $\mathcal{L}(\mathbf{B})$ of dimension n:

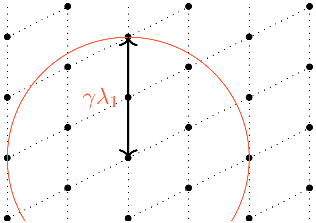
Output: find the shortest non-zero vector $\mathbf{x} \in \mathcal{L}(\mathbf{B})$.



Approx Shortest Vector Problem (Approx SVP$_\gamma$)

Given a lattice $\mathcal{L}(\mathbf{B})$ of dimension n:

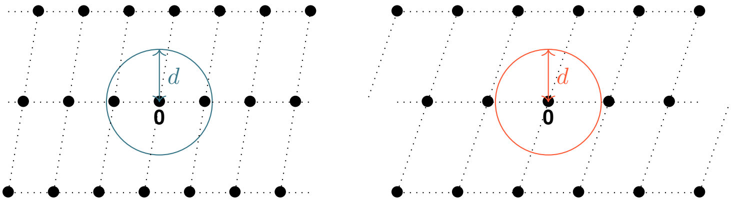
Output: find a non-zero vector $\mathbf{x} \in \mathcal{L}(\mathbf{B})$ such that $\|\mathbf{x}\| \leq \gamma\lambda_1(\mathcal{L}(\mathbf{B}))$



Gap Shortest Vector Problem (GapSVP)

Given a lattice $\mathcal{L}(\mathbf{B})$ of dimension n and $d > 0$:

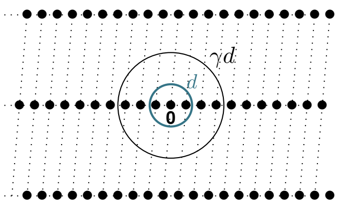
Output:
- YES: there is $\mathbf{z} \in \mathcal{L}(\mathbf{B})$ non-zero such that $\|\mathbf{z}\| < d$,
- NO: for all non-zero vectors $\mathbf{z} \in \mathcal{L}(\mathbf{B})$: $\|\mathbf{z}\| \geq d$.

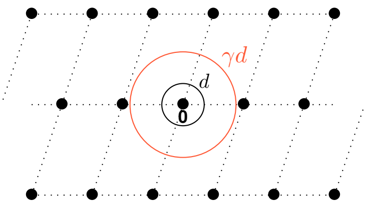


Gap Shortest Vector Problem (GapSVP$_\gamma$)

Given a lattice $\mathcal{L}(\mathbf{B})$ of dimension n and $d > 0$:

Output:
- YES: there is $\mathbf{z} \in \mathcal{L}(\mathbf{B})$ non-zero such that $\|\mathbf{z}\| < d$,
- NO: for all non-zero vectors $\mathbf{z} \in \mathcal{L}(\mathbf{B})$: $\|\mathbf{z}\| \geq \gamma d$.



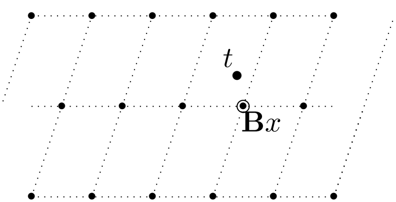


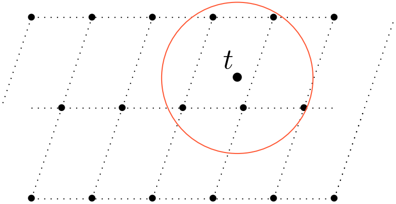
Closest Vector Problem

Given a lattice $\mathcal{L}(\mathbf{B})$ of dimension n and $\mathbf{t} \in \mathbb{Z}^m$:

Output: find $\mathbf{x} \in \mathbb{Z}^n$ minimizing $||\mathbf{Bx} - \mathbf{t}||$.

Approx variant: find $\mathbf{x} \in \mathbb{Z}^n$ such that $||\mathbf{Bx} - \mathbf{t}|| \leq \gamma \cdot \text{dist}(\mathbf{t}, \Lambda(\mathbf{B}))$.



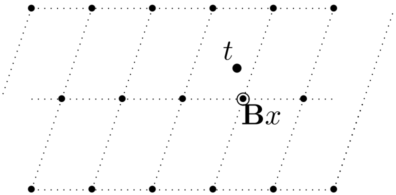


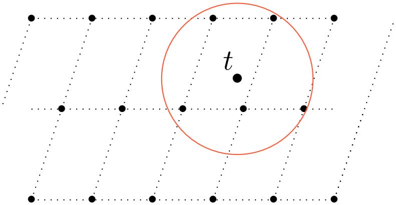
Closest Vector Problem

Given a lattice $\mathcal{L}(\mathbf{B})$ of dimension n and $\mathbf{t} \in \mathbb{Z}^m$:

Output: find $\mathbf{x} \in \mathbb{Z}^n$ minimizing $||\mathbf{Bx} - \mathbf{t}||$.

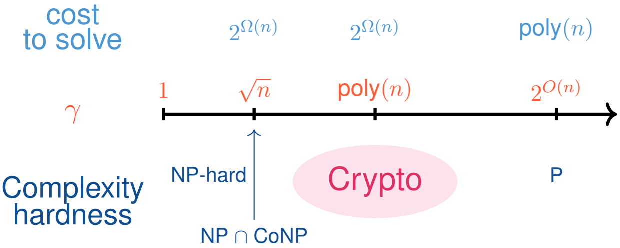
Approx variant: find $\mathbf{x} \in \mathbb{Z}^n$ such that $||\mathbf{Bx} - \mathbf{t}|| \leq \gamma \cdot \mathrm{dist}(\mathbf{t}, \Lambda(\mathbf{B}))$.





How hard is it to solve those problems?

Hardness of Approx SVP$_\gamma$



Conjecture

There is no polynomial time algorithm that approximates this lattice problem and its variants to within polynomial factors.

At the heart of lattice-based cryptography
the Learning With Errors problem

- Introduced by Regev in 2005

Problem: solve a linear system with m equations and n variables ($m \geq n$), with noise, and modulo an integer q.

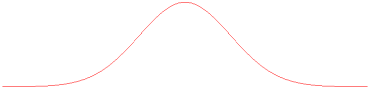
Find $(s_1, s_2, s_3, s_4, s_5)$ such that:

$$
\begin{array}{rcrl}
s_1 + 22s_2 + 17s_3 + 2s_4 + s_5 & \approx & 16 & \mod 23 \\
3s_1 + 2s_2 + 11s_3 + 7s_4 + 8s_5 & \approx & 17 & \mod 23 \\
15s_1 + 13s_2 + 10s_3 + 3s_4 + 5s_5 & \approx & 3 & \mod 23 \\
17s_1 + 11s_2 + 20s_3 + 9s_4 + 3s_5 & \approx & 8 & \mod 23 \\
2s_1 + 14s_2 + 13s_3 + 6s_4 + 7s_5 & \approx & 9 & \mod 23 \\
4s_1 + 21s_2 + 9s_3 + 5s_4 + s_5 & \approx & 18 & \mod 23 \\
11s_1 + 12s_2 + 5s_3 + s_4 + 9s_5 & \approx & 7 & \mod 23
\end{array}
$$

Gaussian distributions

Continuous Gaussian distribution of center c and parameter s:

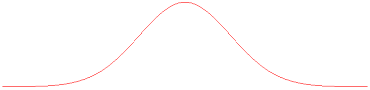
$D_{s,c}(x) \sim \frac{1}{s} \exp\left(-\pi \frac{||x-c||^2}{s^2} \right)$
$\forall x \in \mathbb{R}$



Gaussian distributions

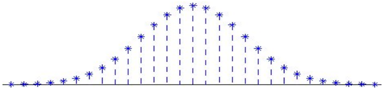
Continuous Gaussian distribution of center c and parameter s:

$$D_{s,c}(x) \sim \frac{1}{s} \exp\left(-\pi \frac{||x-c||^2}{s^2} \right)$$
$$\forall x \in \mathbb{R}$$



Gaussian distribution on $\mathbb{Z}$ of center c with parameter s:

$$D_{\mathbb{Z},s,c}(x) \sim \frac{1}{s} \exp\left(-\pi \frac{||x-c||^2}{s^2} \right)$$
$$\forall x \in \mathbb{Z}$$



- It is not the rounding of the continuous Gaussian.
- We now how to sample it efficiently.
- Almost all samples are in $[-t \cdot s, +t \cdot s]$ for a constant t, if s is not to small.

Discrete gaussian on lattices

Theorem (Gentry, Peikert, Vaikuntanathan 2008)

There exists a PPT algorithm which, given a basis $\mathbf{B}$ *of a lattice* $\Lambda(\mathbf{B})$ *of dimension* n*, a parameter* $s \geq \|\tilde{\mathbf{B}}\| \cdot \omega(\sqrt{\log n})$*, an a center* $c \in \mathbb{R}^n$*, outputs a sample from a distribution statistically close from* $D_{\Lambda,s,c}$*.*

Intuition: sampling on $\mathbb{Z}$ is quite easy, it is more complicated on a general lattice.

Important: Better is the basis (with short vectors), smaller is the parameter we can sample with, and then have short vectors.

Smoothing parameter

Definition

For all $\varepsilon > 0$, the *smoothing parameter* of a lattice Λ with parameter ε is the smallest s such that $\rho_{1/s}(\Lambda^* \setminus \{0\}) \leq \varepsilon$, we denote it by $\eta_\varepsilon(\Lambda)$.

When the Gaussian's parameter is bigger than smoothing parameter, the discrete gaussian distribution has the same properties than a continuous one. In particular:

- the discrete gaussian distribution $D_{\Lambda,s,c}$ is mainly concentrated in a sphere of radius $\sqrt{n}s$ around its center c.

  If $s > \eta_\varepsilon(\Lambda)$,

  $$\Pr_{x \hookleftarrow D_{\Lambda,s,c}} \left[\|x - c\| > \sqrt{n}s \right] \leq 2^{-n}.$$

- Addition: if $s, t > \eta_\varepsilon(\Lambda)$, we can also add two gaussian **on the same lattice** :

  $$D_{\Lambda,s} + D_{\Lambda,t} = D_{\Lambda,\sqrt{s^2+t^2}}.$$

Size of the smoothing parameter

The size of the smooting parameter can be compared to the size of the n-th minima.

- Micciancio, Regev 2004 and Regev 2005:
  For any lattice Λ and $\varepsilon > 0$

$$\sqrt{\frac{\ln(1/\varepsilon)}{\pi}} \cdot \frac{\lambda_n(\Lambda)}{n} \leq \eta_\varepsilon(\Lambda) \leq \sqrt{\frac{\ln(2n(1+1/\varepsilon))}{\pi}} \cdot \lambda_n(\Lambda).$$

The Learning With Errors problem [Regev 05]

Let $n > 1$, $q \geq 2$ and $\alpha \in]0, 1[$.
For any $\mathbf{s} \in \mathbb{Z}_q^n$, we define the distribution $\mathcal{D}_{n,q,\alpha}(\mathbf{s})$ by:

$$(\mathbf{a}, \langle \mathbf{a}, \mathbf{s} \rangle + e)\ , \text{ with } \mathbf{a} \leftarrow U(\mathbb{Z}_q^n) \text{ and } e \leftarrow D_{\mathbb{Z},\alpha q}.$$

- Search LWE
  For any $\mathbf{s}$: find $\mathbf{s}$ given an arbitrary number of samples from $\mathcal{D}_{n,q,\alpha}(\mathbf{s})$.

- Decision LWE
  With non-negligible probability on $\mathbf{s} \leftarrow U(\mathbb{Z}_q^n)$: distinguish between the distributions $\mathcal{D}_{n,q,\alpha}(\mathbf{s})$ and $U(\mathbb{Z}_q^{n+1})$.

Decision version

Let $n > 1$, $q \geq 2$ and $\alpha \in]0, 1[$.
For any $\mathbf{s} \in \mathbb{Z}_q^n$, we define the distribution $\mathcal{D}_{n,q,\alpha}(\mathbf{s})$ by:

$$(\mathbf{a}, \langle \mathbf{a}, \mathbf{s} \rangle + e)\,, \text{ with } \mathbf{a} \leftarrow U(\mathbb{Z}_q^n) \text{ and } e \leftarrow D_{\mathbb{Z},\alpha q}.$$

- Decision LWE
  With non-negligible probability on $\mathbf{s} \leftarrow U(\mathbb{Z}_q^n)$: distinguish between the distributions $\mathcal{D}_{n,q,\alpha}(\mathbf{s})$ and $U(\mathbb{Z}_q^{n+1})$.

We consider an oracle $\mathcal{O}$ which produces independant samples, all from the same distribution being:

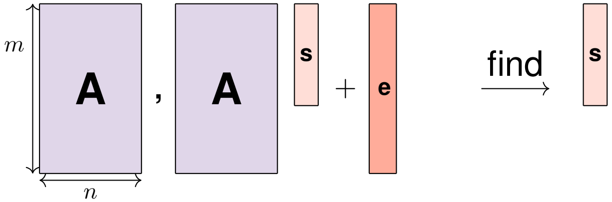
- either $\mathcal{D}_{n,q,\alpha}(\mathbf{s})$ for a fixed $\mathbf{s}$,
- either $U(\mathbb{Z}_q^{n+1})$.

The goal is to decide which one with a non-negligeable advantage.

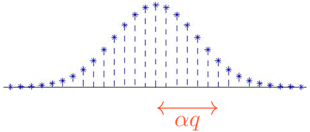
The Learning With Errors problem

$\mathsf{LWE}^n_{\alpha,q}$

Given



- $\mathbf{A} \leftarrow U(\mathbb{Z}_q^{m\times n})$,
- $\mathbf{s} \leftarrow U(\mathbb{Z}_q^{n})$,
- $\mathbf{e} \leftarrow D_{\mathbb{Z}^m,\alpha q}$, small compared to q.



Discrete Gaussian error $D_{\mathbb{Z},\alpha q}$

Search version: Given $(\mathbf{A}, \mathbf{b} = \mathbf{A}\mathbf{s} + \mathbf{e})$, find $\mathbf{s}$.
Decision version: Distinguish from $(\mathbf{A}, \mathbf{b})$ with $\mathbf{b}$ uniform.

Equivalence between the two variants

LWE sample: $(\mathbf{A}, \mathbf{b} = \mathbf{A}\mathbf{s} + \mathbf{e} \mod q)$ with short $\mathbf{e}$.

- Easy reduction : from **decision** to **search**
  - find **s** $\Rightarrow$ distinguish **b** uniform or **b** LWE sample,

Equivalence between the two variants

LWE sample: $(\mathbf{A}, \mathbf{b} = \mathbf{A}\mathbf{s} + \mathbf{e} \mod q)$ with short $\mathbf{e}$.

- Easy reduction : from **decision** to **search**
  - find $\mathbf{s} \Rightarrow$ distinguish $\mathbf{b}$ uniform or $\mathbf{b}$ LWE sample,
  - Given $(\mathbf{A}, \mathbf{b})$, find the oracle to find $\mathbf{s}$, compute $\mathbf{b} - \mathbf{A}\mathbf{s}$:

Equivalence between the two variants

LWE sample: $(\mathbf{A}, \mathbf{b} = \mathbf{A}\mathbf{s} + \mathbf{e} \mod q)$ with short $\mathbf{e}$.

- Easy reduction : from **decision** to **search**
  - find **s** $\Rightarrow$ distinguish **b** uniform or **b** LWE sample,
  - Given $(\mathbf{A}, \mathbf{b})$, find the oracle to find **s**, compute $\mathbf{b} - \mathbf{A}\mathbf{s}$:
    - if it is small, then **b** is an LWE sample,
    - if it looks uniform, then **b** is uniform.

Equivalence between the two variants

LWE sample: $(\mathbf{A}, \mathbf{b} = \mathbf{A}\mathbf{s} + \mathbf{e} \bmod q)$ with short $\mathbf{e}$.

- Easy reduction : from **decision** to **search**
  - find **s** $\Rightarrow$ distinguish **b** uniform or **b** LWE sample,
  - Given $(\mathbf{A}, \mathbf{b})$, find the oracle to find **s**, compute $\mathbf{b} - \mathbf{A}\mathbf{s}$:
    - if it is small, then **b** is an LWE sample,
    - if it looks uniform, then **b** is uniform.

- 2nd reduction: from **search** to **decision**
  - Distinguish **b** uniform from **b** LWE sample $\Rightarrow$ find **s**,

Equivalence between the two variants

LWE sample: $(\mathbf{A}, \mathbf{b} = \mathbf{A}\mathbf{s} + \mathbf{e} \mod q)$ with short $\mathbf{e}$.

- Easy reduction : from **decision** to **search**
  - find **s** $\Rightarrow$ distinguish **b** uniform or **b** LWE sample,
  - Given $(\mathbf{A}, \mathbf{b})$, find the oracle to find **s**, compute $\mathbf{b} - \mathbf{A}\mathbf{s}$:
    - if it is small, then **b** is an LWE sample,
    - if it looks uniform, then **b** is uniform.

- 2nd reduction: from **search** to **decision**
  - Distinguish **b** uniform from **b** LWE sample $\Rightarrow$ find **s**,
  - Given $(\mathbf{A}, \mathbf{b})$ use the oracle to find each coordinate of **s**: for all s_1^*, choose u uniform in $\mathbb{Z}_q$ and modify $(\mathbf{A}, \mathbf{b})$ as follow:

$$(\mathbf{a}, b) + (u, 0, \ldots, 0, us_1^*) = (\mathbf{a}', \langle \mathbf{a}', \mathbf{s} \rangle + e + u(s_1^* - s_1)),.$$

    - if $s_1^* = s_1$ it stays a LWE sample,
    - else **b** will be uniform.

Short Integer Solution problem [Ajtai 1996]

For $\mathbf{A} \leftarrow U(\mathbb{Z}_q^{m\times n})$:

SIS$_\beta$

$$\mathbf{x}\,\mathbf{A} = \mathbf{0} \bmod q$$

Goal: Given $\mathbf{A} \leftarrow U(\mathbb{Z}_q^{m\times n})$,
find $\mathbf{x}$ s.t. $0 < \|\mathbf{x}\| \leq \beta$.

LWE$_\alpha$

$$\left(\underbrace{\mathbf{A}}_{n}, \; \mathbf{A}\,\mathbf{s} + \mathbf{e}\right) \quad (\mathbf{A} \text{ with } m \text{ rows})$$

$\mathbf{s} \leftarrow U(\mathbb{Z}_q^n)$, $\mathbf{e} \leftarrow D_{\mathbb{Z}^m,\alpha q}$.

Goal: Given $(\mathbf{A}, \mathbf{A}\,\mathbf{s} + \mathbf{e})$,
find $\mathbf{s}$.

Short Integer Solution problem [Ajtai 1996]

For $\mathbf{A} \leftarrow U(\mathbb{Z}_q^{m\times n})$:

SIS$_\beta$

$$\mathbf{x} \; \mathbf{A} = \mathbf{0} \bmod q$$

Solve SVP in

$$\Lambda_q^\perp(\mathbf{A}) = \{\mathbf{x} \in \mathbb{Z}^m | \mathbf{x}^T \mathbf{A} = 0 \bmod q\}$$

LWE$_\alpha$

$$\left(\underbrace{\mathbf{A}}_{n}, \; \mathbf{A}\,\mathbf{s} + \mathbf{e} \right), \quad \mathbf{A} \text{ of size } m \times n$$

$\mathbf{s} \leftarrow U(\mathbb{Z}_q^n)$, $\mathbf{e} \leftarrow D_{\mathbb{Z}^m,\alpha q}$.

Solve CVP in

$$\Lambda_q(\mathbf{A}) = \{\mathbf{y} \in \mathbb{Z}^m : \mathbf{y} = \mathbf{A}\,\mathbf{s} \bmod q$$
for some $\mathbf{s} \in \mathbb{Z}^n\}$

Hardness of LWE

- Exhaustive search
  - Try all the $\mathbf{s} \in \mathbb{Z}_q^n \to$ is $\mathbf{b} - \mathbf{A}\mathbf{s}$ small?
  - $\Rightarrow$ cost around q^n.

Hardness of LWE

- Exhaustive search
  - Try all the $\mathbf{s} \in \mathbb{Z}_q^n \to$ is $\mathbf{b} - \mathbf{A}\mathbf{s}$ small?
  - $\Rightarrow$ cost around q^n.
  - Other possibility: guess the n first errors, find $\mathbf{s} \to$ is $\mathbf{b} - \mathbf{A}\mathbf{s}$ small?
  - $\Rightarrow$ cost around $(\alpha q \sqrt{n})^n$.

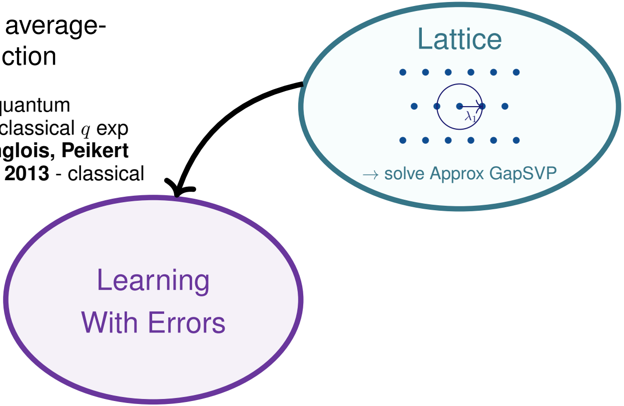
Hardness of LWE

- Exhaustive search
  - Try all the $\mathbf{s} \in \mathbb{Z}_q^n \to$ is $\mathbf{b} - \mathbf{A}\mathbf{s}$ small?
  - $\Rightarrow$ cost around q^n.
  - Other possibility: guess the n first errors, find $\mathbf{s} \to$ is $\mathbf{b} - \mathbf{A}\mathbf{s}$ small?
  - $\Rightarrow$ cost around $(\alpha q \sqrt{n})^n$.
- How to do better?
  - LWE is a lattice problem: consider
    $$\Lambda_q(\mathbf{A}) = \{\mathbf{y} \in \mathbb{Z}^m : \mathbf{y} = \mathbf{A}\mathbf{s} \bmod q \text{ for } \mathbf{s} \in \mathbb{Z}^n\}.$$
    Solving LWE $\Leftrightarrow$ solving CVP in this lattice.
  - Cost: $\left(\frac{n \log q}{\log^2 \alpha}\right)^{\frac{n \log q}{\log^2 \alpha}}$.

Hardness of the Learning With Errors problem

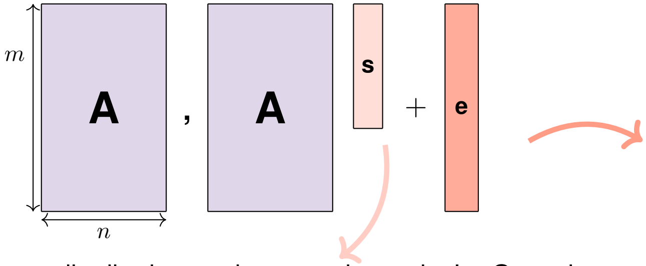
Worst-case to average-case reduction

- **Regev 2005** - quantum
- **Peikert 2009** - classical q exp
- **Brakerski, Langlois, Peikert Regev, Stehlé 2013** - classical



LWE variants

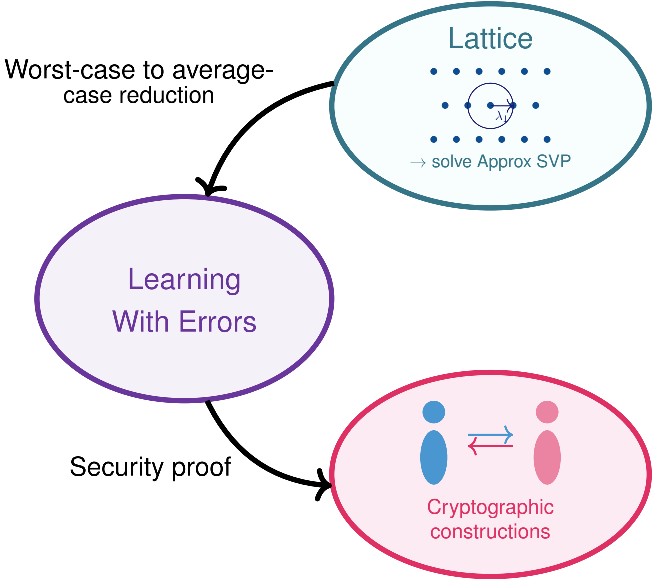
Choose another distribution for the secret or the error.
Regev 2009: uniform secret and gaussian error.



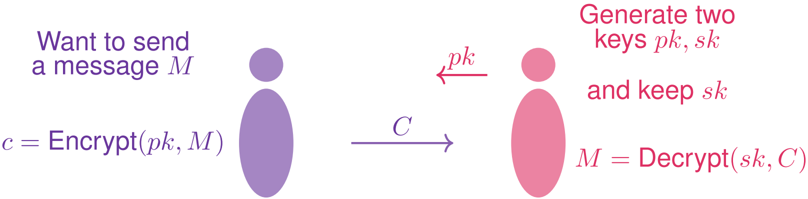
- Gaussian (continue, discretize, discrete ...),
- Uniform in small interval,
- Binary under conditions.

- Same distribution as the error: in particular Gaussian,
- Binary (Unif in $\{0,1\}^n$),
- Entropic.

Using LWE to build provable constructions - theory



Public key encryption - definition



An encryption scheme is defined by three algorithms **(KeyGen, Enc, Dec)**:

- The key generation algorithm **KeyGen** takes as input a security parameter λ and outputs the public and the secret keys (pk, sk).
- The encryption algorithm **Enc** takes as input the public key pk and a message m and outputs $c = Enc(pk, m)$,
- The decryption algorithm **Dec** takes as input the secret key sk and a ciphertext c and outputs $m = Dec(sk, c)$,

such that $Dec(sk, (Enc(pk, m)) = m$.

Regev's encryption scheme

- **Parameters**: $n, m, q \in \mathbb{Z}$, $\alpha \in \mathbb{R}$,
- **Keys**: sk = $\mathbf{s}$ and pk = $(\mathbf{A}, \mathbf{b})$, with $\mathbf{b} = \mathbf{A}\mathbf{s} + \mathbf{e} \bmod q$
  where $\mathbf{s} \hookleftarrow U(\mathbb{Z}_q^n)$, $\mathbf{A} \hookleftarrow U(\mathbb{Z}_q^{m\times n})$, $\mathbf{e} \hookleftarrow D_{\mathbb{Z}^m,\alpha q}$.

Regev's encryption scheme

- **Parameters**: $n, m, q \in \mathbb{Z}$, $\alpha \in \mathbb{R}$,
- **Keys**: sk = $\mathbf{s}$ and pk = $(\mathbf{A}, \mathbf{b})$, with $\mathbf{b} = \mathbf{A}\mathbf{s} + \mathbf{e} \mod q$
  where $\mathbf{s} \hookleftarrow U(\mathbb{Z}_q^n)$, $\mathbf{A} \hookleftarrow U(\mathbb{Z}_q^{m \times n})$, $\mathbf{e} \hookleftarrow D_{\mathbb{Z}^m, \alpha q}$.
- **Encryption** $(M \in \{0,1\})$**:** Let $\mathbf{r} \hookleftarrow U(\{0,1\}^m)$,

$$\mathbf{u}^T = \mathbf{r}\,\mathbf{A}, \quad v = \mathbf{r}\,\mathbf{b} + \lfloor q/2 \rceil \cdot M$$

Regev's encryption scheme

- **Parameters**: $n, m, q \in \mathbb{Z}$, $\alpha \in \mathbb{R}$,
- **Keys**: sk = $\mathbf{s}$ and pk = $(\mathbf{A}, \mathbf{b})$, with $\mathbf{b} = \mathbf{A}\mathbf{s} + \mathbf{e} \mod q$
  where $\mathbf{s} \hookleftarrow U(\mathbb{Z}_q^n)$, $\mathbf{A} \hookleftarrow U(\mathbb{Z}_q^{m \times n})$, $\mathbf{e} \hookleftarrow D_{\mathbb{Z}^m, \alpha q}$.
- **Encryption** $(M \in \{0,1\})$: Let $\mathbf{r} \hookleftarrow U(\{0,1\}^m)$,

$$\mathbf{u}^T = \mathbf{r}\,\mathbf{A}, \quad v = \mathbf{r}\,\mathbf{b} + \lfloor q/2 \rceil \cdot M$$

- **Decryption** of $(\mathbf{u}, v)$: compute $v - \mathbf{u}^T\mathbf{s}$,

$$\underbrace{\mathbf{r}\left[\mathbf{A}\,\mathbf{s} + \mathbf{e}\right] + \lfloor q/2 \rceil \cdot M}_{v} - \underbrace{\mathbf{r}\,\mathbf{A}\,\mathbf{s}}_{\mathbf{u}^T\mathbf{s}} = \text{small} + \lfloor q/2 \rceil \cdot M$$

If **close from** 0: return 0, if **close from** $\lfloor q/2 \rceil$: return 1.

Regev's encryption scheme

- **Parameters**: $n, m, q \in \mathbb{Z}$, $\alpha \in \mathbb{R}$,
- **Keys**: sk = $\mathbf{s}$ and pk = $(\mathbf{A}, \mathbf{b})$, with $\mathbf{b} = \mathbf{A}\,\mathbf{s} + \mathbf{e} \bmod q$
  where $\mathbf{s} \hookleftarrow U(\mathbb{Z}_q^n)$, $\mathbf{A} \hookleftarrow U(\mathbb{Z}_q^{m\times n})$, $\mathbf{e} \hookleftarrow D_{\mathbb{Z}^m,\alpha q}$.
- **Encryption** $(M \in \{0,1\})$**: Let** $\mathbf{r} \hookleftarrow U(\{0,1\}^m)$,

$$\mathbf{u}^T = \mathbf{r}\,\mathbf{A}, \quad v = \mathbf{r}\,\mathbf{b} + \lfloor q/2 \rceil \cdot M$$

- **Decryption** of $(\mathbf{u}, v)$: compute $v - \mathbf{u}^T\mathbf{s}$,

$$\mathbf{r}\left[\mathbf{A}\,\mathbf{s} + \mathbf{e}\right] + \lfloor q/2 \rceil \cdot M - \mathbf{r}\,\mathbf{A}\,\mathbf{s} = \text{small} + \lfloor q/2 \rceil \cdot M$$

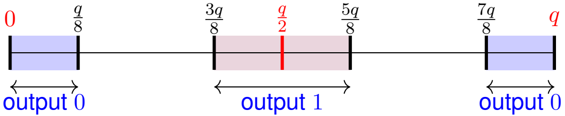
LWE hard $\Rightarrow$ Regev's scheme is IND-CPA secure.

Correction

The randomness **r** is uniformly chosen in $\{0,1\}^m$,
and **e** is sampled from a discrete gaussian of parameter $\alpha q \leq q/(8m)$,
then, with overwhealming probability,

$$\Big|\sum_{i\leq m} r_i e_i\Big| \leq \|\mathbf{r}\| \cdot \|\mathbf{e}\| \leq \sqrt{m} \cdot \frac{q}{8\sqrt{m}} = \frac{q}{8}.$$

$v - \mathbf{u}^T\mathbf{s}$ is either close from 0, either close from $\lfloor q/2 \rfloor$, which allows to find M.



IND-CPA security

To define the security, we use a game between a challenger and an adversary. We define two experiments Exp_b for $b \in \{0, 1\}$:

$\mathcal{C}$hallenger		$\mathcal{A}$dversary
Generate (pk, sk)	$\xrightarrow{pk}$	
	$\xleftarrow{M_0, M_1}$	Choose M_0, M_1
	$\xrightarrow{c \leftarrow \mathsf{Enc}(pk, M_b)}$	
		output b'

$$Adv^{CPA}(\mathcal{A}) = |\Pr[\mathcal{A} \rightarrow^{Exp_0} 1] - \Pr[\mathcal{A} \rightarrow^{Exp_1} 1]|.$$

IND-CPA security

Goal of the proof: show that if an adversary succeed in attacking the encryption scheme with a non-negligible advantage, then the challenger can use it to solve a difficult problem (here LWE).

Decision LWE can be seen as a game:

$\mathcal{C}$		$\mathcal{B}$
$\mathbf{A} \longleftarrow U(\mathbb{Z}_q^{m\times n})$		
RAND $(b=0)$: $\mathbf{b} \leftarrow U(\mathbb{Z}_q^{m})$		
LWE $(b=1)$: $\mathbf{b} = \mathbf{A}\mathbf{s} + \mathbf{e}$	$\xrightarrow{(\mathbf{A},\mathbf{b})}$	
		output b'

$$Adv(\mathcal{B}) = \left|\Pr[\mathcal{B} \xrightarrow{RAND} 1] - \Pr[\mathcal{B} \xrightarrow{LWE} 1]\right|.$$

Leftover Hash Lemma

Let $m, n, q \geq 1$ be integers such that $m \geq 4n \log q$ and q prime, and let $\mathbf{A} \hookleftarrow U(\mathbb{Z}_q^{m \times n})$ and $\mathbf{r} \hookleftarrow U(\{0,1\}^m)$. Then $(\mathbf{A}, \mathbf{r}^T\mathbf{A})$ has statistical distance $\leq 2^{-n}$ from the uniform distribution on $\mathbb{Z}_q^{m \times n} \times \mathbb{Z}_q^n$.

Leftover Hash Lemma

Let $m, n, q \geq 1$ be integers such that $m \geq 4n \log q$ and q prime, and let $\mathbf{A} \hookleftarrow U(\mathbb{Z}_q^{m \times n})$ and $\mathbf{r} \hookleftarrow U(\{0,1\}^m)$. Then $(\mathbf{A}, \mathbf{r}^T\mathbf{A})$ has statistical distance $\leq 2^{-n}$ from the uniform distribution on $\mathbb{Z}_q^{m \times n} \times \mathbb{Z}_q^n$.

- Statistical distance : $\Delta(D_1, D_2) = \frac{1}{2} \sum_x |D_1(x) - D_2(x)|$.

Leftover Hash Lemma

Let $m, n, q \geq 1$ be integers such that $m \geq 4n \log q$ and q prime, and let $\mathbf{A} \hookleftarrow U(\mathbb{Z}_q^{m\times n})$ and $\mathbf{r} \hookleftarrow U(\{0,1\}^m)$. Then $(\mathbf{A}, \mathbf{r}^T\mathbf{A})$ has statistical distance $\leq 2^{-n}$ from the uniform distribution on $\mathbb{Z}_q^{m\times n} \times \mathbb{Z}_q^n$.

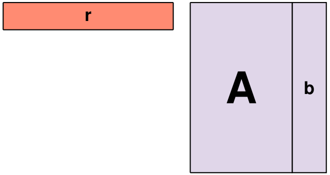
- Statistical distance : $\Delta(D_1, D_2) = \frac{1}{2}\sum_x |D_1(x) - D_2(x)|$.
- For any algorithm $\mathcal{A}$, we have $|\Pr[\mathcal{A}(D_1) = 1] - \Pr[\mathcal{A}(D_2) = 1]| \leq \Delta(D_1, D_2)$.
  $\Delta(D_1, D_2)$ small $\Rightarrow$ D_1 and D_2 are statistically indistinguishable.

Leftover Hash Lemma

Let $m, n, q \geq 1$ be integers such that $m \geq 4n \log q$ and q prime, and let $\mathbf{A} \hookleftarrow U(\mathbb{Z}_q^{m\times n})$ and $\mathbf{r} \hookleftarrow U(\{0,1\}^m)$. Then $(\mathbf{A}, \mathbf{r}^T\mathbf{A})$ has statistical distance $\leq 2^{-n}$ from the uniform distribution on $\mathbb{Z}_q^{m\times n} \times \mathbb{Z}_q^n$.

- Statistical distance : $\Delta(D_1, D_2) = \frac{1}{2}\sum_x |D_1(x) - D_2(x)|$.
- For any algorithm $\mathcal{A}$, we have $|\Pr[\mathcal{A}(D_1) = 1] - \Pr[\mathcal{A}(D_2) = 1]| \leq \Delta(D_1, D_2)$.
  $\Delta(D_1, D_2)$ small $\Rightarrow$ D_1 and D_2 are statistically indistinguishable.

The LHL implies that (**(A b)** , **r** **(A b)**) is indistinguishable from uniform.



IND-CPA security proof

Idea: we start from an LWE instance, and build an instance of the IND-CPA experiment, then we use the answer of the adversary to solve LWE.
We use the following IND-CPA game:

$\mathcal{B}$		$\mathcal{A}$
$(sk = \mathbf{s}, pk = (\mathbf{A}, \mathbf{b} = \mathbf{A}\mathbf{s} + \mathbf{e}) \leftarrow KeyGen(.)$	$\xrightarrow{pk=(\mathbf{A},\mathbf{b})}$	
chooses b	$\xleftarrow{m_0,m_1}$	Chooses m_0, m_1,
computes $(c_1, c_2) \leftarrow Enc(pk, m_b)$	$\xrightarrow{c_1,c_2}$	Computes a bit b'
		if $b = b'$ then output Win

We want to show that if LWE is hard, then there exists a negligible function $negl$ such that:

$$\Pr[\mathcal{A} \text{ Win}] \leq 1/2 + negl(n).$$

IND-CPA security proof

$\mathcal{B}$ wants to solve decisional LWE using $\mathcal{A}$.

$\mathcal{C}$		$\mathcal{B}$		$\mathcal{A}$
RAND: $\mathbf{b}$ unif				
LWE: $\mathbf{b} = \mathbf{A}\mathbf{s} + \mathbf{e}$	$\xrightarrow{(\mathbf{A},\mathbf{b})}$		$\xrightarrow{(\mathbf{A},\mathbf{b})}$	m_0, m_1,
		choose b	$\xleftarrow{m_0,m_1}$	
		$(\mathbf{r}^T\mathbf{A}, \mathbf{r}^T\mathbf{b} + q/2 \cdot m_b)$	$\xrightarrow{c_1,c_2}$	Computes b'
		if $b = b'$ output 1	$\xleftarrow{b'}$	
		else output 0		

For $\mathcal{B}$:

- RAND: $\mathbf{b}$ is uniform then c_2 is uniform. $\mathcal{A}$ cannot distinguish between the two cases, its advantage is equals to zero, the probability that $\mathcal{B}$ outputs 1 is $1/2$.

$$\Pr[\mathcal{B} \xrightarrow{RAND} 1] = 1/2,$$

IND-CPA security proof

$\mathcal{B}$ wants to solve decisional LWE using $\mathcal{A}$.

$\mathcal{C}$		$\mathcal{B}$		$\mathcal{A}$
RAND: **b** unif				
LWE: $\mathbf{b} = \mathbf{A}\mathbf{s} + \mathbf{e}$	$\xrightarrow{(\mathbf{A},\mathbf{b})}$		$\xrightarrow{(\mathbf{A},\mathbf{b})}$	m_0, m_1,
		choose b	$\xleftarrow{m_0,m_1}$	
		$(\mathbf{r}^T\mathbf{A}, \mathbf{r}^T\mathbf{b} + q/2 \cdot m_b)$	$\xrightarrow{c_1,c_2}$	Computes b'
		if $b = b'$ output 1	$\xleftarrow{b'}$	
		else output 0		

For $\mathcal{B}$:

- LWE: $\mathbf{b} = \mathbf{A}\mathbf{s} + \mathbf{e}$ and then the ciphertext is exactly a ciphertext from the Regev encryption scheme. The probability that $\mathcal{B}$ outputs 1 is exactly the success probability of $\mathcal{A}$ in the encryption scheme security game (as it has the same view of the experiment).

$$\Pr[\mathcal{B} \xrightarrow{LWE} 1] = \Pr[\mathcal{A}\ \text{win}],$$

IND-CPA security proof

To conclude, we have:

$$\Pr[\mathcal{B} \xrightarrow{RAND} 1] = 1/2,$$

$$\Pr[\mathcal{B} \xrightarrow{LWE} 1] = \Pr[\mathcal{A} \text{ win}],$$

then:

$$\begin{aligned} Adv(\mathcal{B}) &= |\Pr[\mathcal{B} \xrightarrow{RAND} 1] - \Pr[\mathcal{B} \xrightarrow{LWE} 1]| \\ &= |\Pr[\mathcal{A} \text{ win}] - 1/2| \end{aligned}$$

If $\mathcal{A}$ succeeds with a non-negligible probability, then there exists ε such that $\Pr[\mathcal{A} \text{ win}] \geq 1/2 + \varepsilon$, then $Adv(\mathcal{B}) \geq \varepsilon$ which implies that there exists a distinguisher able to solve the decisional LWE problem.

Using LWE

Hardness of LWE used as a foundation for many constructions.

Learning
With Errors

Problem: constructions based on LWE
enjoy a nice guarantee of security
but are too costly in practice.

Security proof

Cryptographic constructions

Signature, encryption
Advanced schemes
Fully Homomorphic Encryption

Solutions used today?