Idea: We will include encryptions of 0 in the public key and refresh ciphertexts by taking a subset sum of encryptions of 0:

Regev's public-key encryption scheme

- **Setup**: $A \xleftarrow{R} \mathbb{Z}_q^{n \times m}$, $s \xleftarrow{R} \mathbb{Z}_q^n$, $e \leftarrow \chi^m$; $b^T \leftarrow s^T A + e^T$; output $pk = (A, b^T)$, $sk = s$
  - ↑ can be viewed as m encryptions of 0 under the symmetric scheme with secret key s
- **Encrypt**(pk, μ): sample $r \xleftarrow{R} \{0,1\}^m$; output $(Ar, b^T r + \mu \cdot \lfloor \frac{q}{2} \rfloor)$
- **Decrypt**(sk, ct): output $\lfloor ct_2 - s^T ct_1 \rceil_2$

More compact form: $pk = A = \begin{bmatrix} \bar{A} \\ s^T \bar{A} + e^T \end{bmatrix}$, $ct = C = Ar + \begin{bmatrix} 0^n \\ \mu \cdot \lfloor q/2 \rfloor \end{bmatrix}$

Decryption: $[-s^T \mid 1] \cdot C = e^T r + \mu \cdot \lfloor \frac{q}{2} \rfloor$ → then round

Correctness: $ct_2 - s^T ct_1 = b^T r + \mu \cdot \lfloor \frac{q}{2} \rfloor - s^T A r = s^T A r + e^T r + \mu \cdot \lfloor \frac{q}{2} \rfloor - s^T A r = \mu \cdot \lfloor \frac{q}{2} \rfloor + e^T r$

if $|e^T r| < \frac{q}{4}$, then decryption succeeds (since e is small and r is binary, $e^T r$ is not large: $|e^T r| < m \|e\| \|r\| = m \|e\|$)

Security: Under LWE assumption public key

$(A, s^T A + e^T) \approx (A, u)$ where $A \xleftarrow{R} \mathbb{Z}_q^{n \times m}$, $u \xleftarrow{R} \mathbb{Z}_q^m$

By the "leftover hash lemma" if we sample $A \xleftarrow{R} \mathbb{Z}_q^{n \times m}$, $u \xleftarrow{R} \mathbb{Z}_q^n$, $r \xleftarrow{R} \{0,1\}^m$ where $m > 2n \log q$

$(Ar, u^T r) \approx (v, w)$ where $v \xleftarrow{R} \mathbb{Z}_q^n$ and $w \xleftarrow{R} \mathbb{Z}_q$

$\Rightarrow b^T r$ in ciphertext functions as a one-time pad

typically χ is B-bounded distribution ($\Pr[|e| \le B : e \leftarrow \chi] = 1$)

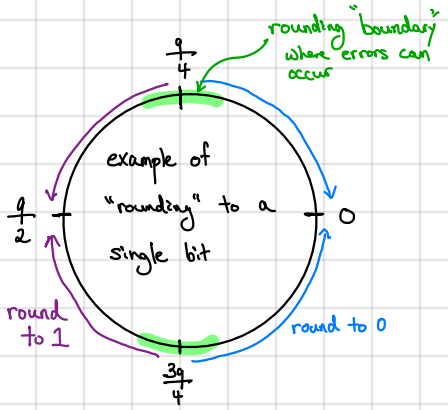
Formally: We will show encryption of 0 is computationally indistinguishable from encryption of 1

- **LWE** — Hyb_0: Challenger samples $A \xleftarrow{R} \mathbb{Z}_q^{n \times m}$, $s \xleftarrow{R} \mathbb{Z}_q^n$, $e \leftarrow \chi^m$ and computes $b^T = s^T A + e^T$. It sets $pk = (A, b)$. Then it samples $r \xleftarrow{R} \{0,1\}^m$ and computes $ct = (ct_0, ct_1) = (Ar, b^T r)$. It gives pk and ct to A.
- **LHL** — Hyb_1: Challenger samples $A \xleftarrow{R} \mathbb{Z}_q^{n \times m}$ and $b \xleftarrow{R} \mathbb{Z}_q^m$.
- **LHL** — Hyb_2: Challenger samples $c_0 \xleftarrow{R} \mathbb{Z}_q^n$, $c_1 \xleftarrow{R} \mathbb{Z}_q$.
- **LWE** — Hyb_3: Challenger sets $c_0 = Ar$ and $c_1 = b^T r + \lfloor \frac{q}{2} \rfloor$.
- Hyb_4: Challenger sets $b^T = s^T A + e^T$.

Many generalizations:

- Encrypting elements mod p: $(Ar, b^T r + \lfloor \frac{q}{p} \rfloor \cdot \mu)$
  Correctness requires that $|e^T r| < \frac{q}{2p}$
- Encrypting vectors $\vec{\mu} \in \{0,1\}^k$.
  Sample secret key $S \xleftarrow{R} \mathbb{Z}_q^{n \times k}$ and $E \leftarrow \chi^{m \times k}$. Let $B^T = S^T A + E^T \in \mathbb{Z}_q^{k \times m}$
  Public key is $pk = (A, B)$
  Ciphertext is $ct = (Ar, B^T r + \lfloor \frac{q}{2} \rfloor \cdot \vec{\mu})$
  ($Ar \in \mathbb{Z}_q^n$, $B^T r + \lfloor \frac{q}{2} \rfloor \cdot \vec{\mu} \in \mathbb{Z}_q^k$)
  Observe: $|ct| = (n+k) \log q$ bits to encrypt k bits
  - ↳ Naively encrypting bit-by-bit would require $kn \log q$ bits.
  - ↳ Approach essentially is reusing the encryption randomness (e.g., same r used to encrypt each bit).

So far... we have developed public-key encryption; what about key agreement?



Alice

$A \xleftarrow{R} \mathbb{Z}_q^{n \times n}$

$S, E \leftarrow \chi^{n \times n}$

$A, \overbrace{AS+E}^{B}$ →

← $S'A + E'$

↓ compute $(S'A+E')S$ and "round"

Bob

$S', E' \leftarrow \chi^{n \times n}$

$E'' \leftarrow \chi^{n \times n}$

↳ compute $S'B + E''$ and "round"

↳ drop least significant bits (similar to decryption in Regev encryption)

correctness requires that entries of $E'S$ be small (so S must be small)

Under the LWE assumption:

$(A, AS+E) \approx U$ where $U \xleftarrow{R} \mathbb{Z}_q^{n \times m}$ [note: requires that LWE holds even if S is sampled from error distribution]

→ shared key then derived by $S'B + E''$ → by LWE, $(B, S'B+E'') \approx (B, U')$

→ shared key is derived from random matrix (similar to Diffie-Hellman, the key material is hashed to derive a symmetric key)

Practical considerations:

- Key reconciliation: presence of noise means Alice and Bob may end up with inconsistent keys. Bob sends a "hint" with his message to reconcile any errors and ensure exact key agreement
- Message size: large matrix A is uniform — can be derived from a short seed (using PRG)
  ↳ justifiable using the random oracle model

Above construction relies on security of LWE where the secret key is sampled from error distribution

↳ This is LWE in "Hermite normal form" and is just as hard as standard LWE

↑ basis of Kyber protocol (NIST post-quantum standard)

LWE (Hermite normal form): For $A \xleftarrow{R} \mathbb{Z}_q^{n \times m}$, $s \leftarrow \chi^n$, $e \leftarrow \chi^m$, $u \xleftarrow{R} \mathbb{Z}_q^m$

$(A, s^T A + e^T)$ is computationally indistinguishable from (A, u^T)

[In normal LWE, we sample $s \xleftarrow{R} \mathbb{Z}_q^n$.]

Note: Kyber uses module-LWE for better efficiency

LWE in Hermite normal form is as hard as LWE (up to small loss in parameters).

To see this, suppose (A, b^T) is an LWE challenge where $A \xleftarrow{R} \mathbb{Z}_q^{n \times m}$ and $m > n$.

1. Parse $A = [A_1 \mid A_2]$ where $A_1 \in \mathbb{Z}_q^{n \times n}$. Suppose A_1 is invertible (it suffices that A is full rank, which holds w.h.p. when m sufficiently large)
2. Then $(\bar{A}, \bar{b})$ is an LWE-HNF instance where $\bar{A} = -A_1^{-1} A_2$ and $\bar{b} = b_1^T \bar{A} + b_2^T$.

First, distribution of $\bar{A}$ is uniform (since A_2 is uniform). Next:

$$\begin{aligned} \bar{b} = b_1^T \bar{A} + b_2^T &= (s^T A_1 + e_1^T)(-A_1^{-1} A_2) + (s^T A_2 + e_2^T) \\ &= -s^T A_2 - e_1^T A_1^{-1} A_2 + s^T A_2 + e_2^T \\ &= e_1^T \bar{A} + e_2^T \end{aligned}$$

where $e_1 \leftarrow \chi^n$ and $e_2 \leftarrow \chi^{m-n}$. This is an LWE instance in Hermite normal form with secret key e_1 and error e_2.

Note that if b is uniform, then so is $\bar{b}$ (since $b_2 \xleftarrow{R} \mathbb{Z}_q^{m-n}$).