# 1.1. Generalities on groups, order of an element, characters.

- A group is a set $G$ with a binary operat°: $G \times G \longrightarrow G$, $(a,b) \longmapsto a \cdot b$

such that . it is associative: $(a \cdot b) \cdot c = a \cdot (b \cdot c)$

. it has an identity/neutral elt: $\exists e \in G$ such that for all $a \in G$, $e \cdot a = a \cdot e = a$ (such an $e$ is unique)

. Every element has an inverse: $\forall a \in G$, $\exists b \in G$ such that $a \cdot b = b \cdot a = e$

- An abelian group is a group such that $\cdot$ is commutative ($ab = ba$)
- Def of the order of an element in a group
- Lagrange thm: the order of an element in a group divides the cardinality of the group.
- Def cyclic group.
- A group homomorphism from $(G, \cdot)$ to $(H, *)$ is a map $f: G \longrightarrow H$ such that $\forall x, y \in G$, $f(x \cdot y) = f(x) * f(y)$.
- An isomorphism of groups is a group homomorphism which is bijective.
- First isomorphism th: $G/\ker f \simeq \mathrm{Im} f$. $\Rightarrow$ cardinality.
- Dual group of a finite abelian group, or group of characters:

$$\hat{G} = \{ \chi : G \longrightarrow \mathbb{C}^\times \text{ group homomorphisms} \}$$

it is an abelian gp for $\chi_1 \chi_2 (g) := \chi_1(g) \chi_2(g)$.

- We'll see in the exercises that if $G$ is a cyclic group, then $G \simeq \hat{G}$. Now structure thm of finite abelian groups imply that if $G$ is a finite abelian gp then

$$G \simeq \prod \text{cyclic groups}$$

hence $G \simeq \hat{G}$ for any $G$ finite abelian.

- See ex sheet 1. "orthogonality relat°" $\forall g \in G$ ← finite abelian

$$\frac{1}{|G|} \sum_{\chi \in \hat{G}} \chi(g) = 1_{g=e}$$

## 1.2. The example of $\mathbb{Z}/n\mathbb{Z}$

← ring with the usual +, × and reduct° modulo n.

- There are two different group structures related to integers modulo n : $(\mathbb{Z}/n\mathbb{Z}, +)$ : cyclic generated by $\bar{1}$

$((\mathbb{Z}/n\mathbb{Z})^\times, \times)$ : not always cyclic

↳ $\bar{k}$ is invertible mod n $\Leftrightarrow \gcd(k,n) = 1$.

Pf: $\Rightarrow$ if $\bar{k}\bar{k'} \equiv 1 \pmod n$ then there exists $m \in \mathbb{Z}$ such that $kk' = 1 + mn$

$\Rightarrow$ $\underbrace{kk' - nm = 1}_{\text{Bézout relat°}} \Rightarrow \gcd(k,n) = 1$

$\Leftarrow$ Let $u, v \in \mathbb{Z}$ be such that $ku + nv = 1$

then $\bar{k}\bar{u} \equiv 1 \pmod n$ so $\bar{k} \in (\mathbb{Z}/n\mathbb{Z})^\times$.

In particular $|(\mathbb{Z}/n\mathbb{Z})^\times| = \#\{k,\ 1 \le k \le n,\ \gcd(k,n) = 1\}$

$= \varphi(n)$: Euler totient funct°

**Th** (Chinese Remainder theorem) If $n = \prod_{i=1}^{r} p_i^{\alpha_i}$ then

$$\mathbb{Z}/n\mathbb{Z} \underset{\text{ring}}{\simeq} \mathbb{Z}/p_1^{\alpha_1}\mathbb{Z} \times \cdots \times \mathbb{Z}/p_r^{\alpha_r}\mathbb{Z}$$

$$k \bmod n \longmapsto (k \bmod p_1^{\alpha_1}, \ldots, k \bmod p_r^{\alpha_r})$$

Rem: It also induces a group isomorph: $(\mathbb{Z}/n\mathbb{Z})^\times \simeq \prod_{i=1}^{r} (\mathbb{Z}/p_i^{\alpha_i}\mathbb{Z})^\times$

⇝ Reduces problem to arithmetic modulo prime powers.

Moreover:

Prop: If $p$ is an odd prime, and $\alpha \geq 1$, then $(\mathbb{Z}/p^\alpha\mathbb{Z})^\times$ is a cyclic group

In particular, for $\alpha = 1$, we denote by $\mathbb{F}_p := \mathbb{Z}/p\mathbb{Z}$ (it is a field), and $\mathbb{F}_p^\times$ is a cyclic group.

⇝ So $\mathbb{F}_p^\times = \langle g \rangle$, means all elements of $\mathbb{F}_p^\times$ can be written as $x = g^a$ for some $a \in \{0, \dots, p-2\}$

here $a = \log_g(x)$ is called the discrete logarithm of $x$ in base $g$.

## 1.3. RSA cryptosystem (Rivest-Shamir-Adleman) ≃ 1977

(previously developed by Clifford Cocks 1973 but it was classified).

Generalities on public-key cryptosystems: Let $\mathcal{M}$ denote the set of possibles messages to be sent. For instance $\mathcal{M} = \{0,1\}^n$ : words of $n$-bits.

- Alice has a <u>public</u> key $P_A$, and an associated transformat $P_A(-)$ which is a bijection from $\mathcal{M}$ to $\mathcal{M}$. Everyone has access to $P_A$, and can perform $P_A(-)$.

- Alice also has a <u>secret</u> key $S_A$, and an associated transformat $S_A(-)$. which must be the inverse of $P_A(-)$. ie $P_A \circ S_A = S_A \circ P_A = Id_{\mathcal{M}}$

Now, here's how Bob can safely send a message to Alice:

- Bob has a message $M \in \mathcal{M}$.
- He can cipher it using Alice's public key: $C := P_A(M)$ is the ciphered message.
- He sends $C$ to Alice
- Alice uses the secret key $S_A$ to recover the original message $M$ (since $S_A(C) = S_A(P_A(M)) = M$ because $S_A \circ P_A = Id$)

The safety of the protocol is only guaranteed if the knowledge of $P_A$ still makes it "impossible" to determine $S_A$. Then an attacker who intercepts $C$ could not decipher it.

Now let's describe more concretely $P_A$ and $S_A$ in the RSA cryptosystem:

- Alice chooses <u>secretly</u> two very large distinct prime numbers $p$ and $q$.
- Alice computes their product $N := pq$ and $\varphi(N) = (p-1)(q-1)$.
- She publishes the <u>public</u> key $(N, e)$ where $e$ is an integer coprime to $\varphi(N)$.

- Using the knowledge of $\varphi(N)$, she computes the inverse of $e \bmod \varphi(N)$ : $d$.
- Her <u>secret</u> key is $(N, d)$

Then, taking $\mathcal{M} = \mathbb{Z}/N\mathbb{Z}$ (if $N > 2^n$, this allows to send all $n$-bits words)

we have the following protocol for sending a message to Alice-

- Bob chooses his message $M \in \mathcal{M}$.
- Using the public key, he computes $C = P_A(M) := M^e \pmod N$
- Alice receives $C$ and uses her secret key to compute $C^d \pmod N$: this is Bob's original message.

<u>Proof of correctness</u> : If $M \equiv 0 \pmod p$ then $M^{ed} \equiv 0 \pmod p$.

Otherwise, since $ed \equiv 1 \pmod{\varphi(N)}$, there exists $k \in \mathbb{Z}$ such that $ed = 1 + k\varphi(N) = 1 + k(p-1)(q-1)$.

So $M^{ed} = M^1 (M^{p-1})^{k(q-1)} \pmod p$. But Fermat's little thm implies $M^{p-1} \equiv 1 \pmod p$, hence $M^{ed} \equiv M \pmod p$.

In any case, $M^{ed} \equiv M \pmod p$. The same proof shows that $M^{ed} \equiv M \pmod q$. Therefore $\begin{cases} M^{ed} \equiv M \pmod p \\ M^{ed} \equiv M \pmod q \end{cases}$

hence (CRT, or elementary arithmetic ($p|\cdots, q|\cdots \Rightarrow pq|\cdots$))

$$M^{ed} \equiv M \pmod N \qquad (N = pq)$$

□

<u>Security?</u> It is believed (but not proved) that breaking RSA is as hard as the factorizable problem (because if one can find $p$ and $q$ such that $N = pq$, then one can compute $\varphi(N) = (p-1)(q-1)$ and find Alice's secret key by finding the inverse of the public key $e \bmod \varphi(N)$. )

## 1.4. Diffie-Hellman-Merkle key exchange.

- Alice and Bob publicly choose a large prime number $p$ and a generator $g$ of $\mathbb{F}_p^\times$.
- Alice secretly chooses $a \in \{1, \dots, p-1\}$ and sends $\overbrace{(g^a \bmod p)}^{A}$ to Bob
- Bob secretly chooses $b \in \{1, \dots, p-1\}$ and sends $\underbrace{(g^b \bmod p)}_{B}$ to Alice.
- Then Alice computes $K := B^a \bmod p$
  Bob computes $K = A^b \bmod p$

  (they are the same since $(g^a)^b = (g^b)^a$)

At the end Alice and Bob now share the same secret key $K$.

Secure? If an attacker intercepts $g^a$ and $g^b$ ($\hat{=} A$ and $B$), he cannot find the key $K$ because he misses the knowledge of $a$ or $b$. And retrieving $a$ or $b$ from $g^a$ or $g^b$ amounts to solving the discrete logarithm problem, which believed to be hard.

Variants: Nowadays, this method is used in more complicated groups than $\mathbb{F}_p^\times$, such as the group associated with an elliptic curve.

# 2.1. Some notions of quantum physics

Quantum physics is a model that describes the behavior of systems at atomic and subatomic scales. A fundamental aspect of the theory is that the state of the system is not known before measurement, and that we rather describe the state by a complex-valued function, called the wave function, that encodes the probabilities to find the system in a certain state after measurement.

Ex: ① Position of a particle in the space:

$$\Psi : \mathbb{R}^4 \longrightarrow \mathbb{C}$$
$$(x,y,z,t) \longmapsto \underbrace{\Psi(x,y,z,t)}$$

This complex number has no physical interpretation. But its square magnitude does!

if $A \subset \mathbb{R}^3$ is a region of the space, the probability to find the particle in $A$ at time $t = \int_A |\Psi(x,y,z,t)|^2 \, dx\,dy\,dz$

$\Psi$ satisfies the normalization property $\int_{\mathbb{R}^3} |\Psi|^2 = 1$.

Rem: The Schrödinger equation is concerned with the evolution of $\Psi$ with respect to $t$.

② The spin of a particle: $\Psi : \{\uparrow, \downarrow\} \longrightarrow \mathbb{C}$

Denote by $|\downarrow\rangle$: $\uparrow \longmapsto 0$, $\downarrow \longmapsto 1$ and $|\uparrow\rangle$: $\uparrow \longmapsto 1$, $\downarrow \longmapsto 0$

Then $\Psi = \alpha|\uparrow\rangle + \beta|\downarrow\rangle$ where $\alpha = \Psi(\uparrow)$, $\beta = \Psi(\downarrow)$

must satisfy the normalizat property $|\alpha|^2 + |\beta|^2 = 1$

and $|\alpha|^2$ is interpreted as the probability that the system is in the state $\uparrow$, while $|\beta|^2$ is the proba that the system is in the state $\downarrow$.

③ Spin of two particles: $\Psi: \{\uparrow\uparrow, \uparrow\downarrow, \downarrow\uparrow, \downarrow\downarrow\} \longrightarrow \mathbb{C}$

$$\Psi = a|\uparrow\uparrow\rangle + b|\uparrow\downarrow\rangle + c|\downarrow\uparrow\rangle + d|\downarrow\downarrow\rangle$$

with $|a|^2 + |b|^2 + |c|^2 + |d|^2 = 1$

Ex: $\Psi = \frac{1}{\sqrt{2}}\left(|\uparrow\uparrow\rangle + |\downarrow\downarrow\rangle\right)$ : this state corresponds to a system of two <u>entangled</u> particles.

These entangled states play an important role in quantum computing. They were also important for the development of quantum mechanics, see for instance the Wikipédia page on Einstein-Podolsky-Rosen paradox.

Another fundamental feature of quantum mechanics is the fact that <u>measurements affect the system</u>.

Mathematically, an <u>observable</u> is a self-adjoint operator on the space where our wave functions live (ie $\langle \Psi_1, H(\Psi_2)\rangle = \langle H(\Psi_1), \Psi_2\rangle$)

Such an $H$ is diagonalizable with only real eigenvalues (we assume here that the hermitian space in which our wave functions live is finite dimensional, just for simplicity); thus in a certain orthonormal basis:

$$M_B(H) = \begin{pmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_r \end{pmatrix} \qquad \lambda_1, \dots, \lambda_r \in \mathbb{R}.$$

Denote by $E_i := \ker(H - \lambda_i \mathrm{Id})$. the eigenspace corresponding to $\lambda_i$.

**Born's rule:**

- The outcome of the measurement is $\lambda_i$ with probability $\| \underbrace{\text{orthogonal projection of } \Psi \text{ on } E_i}_{P_i(\Psi)} \|^2$
- After the measurement, the system is in a new state

$$\Psi_{new} = \frac{P_i(\Psi)}{\| P_i(\Psi) \|}$$

(this is called "wave function collapse")

## 2.2 Qubits as elements of an Hermitian space

To apply the formalism of quantum physics to bits, we need to allow superpositⁿ. A bit will no longer be 0 or 1, it will be represented by a wave function that takes the value 0 or 1 with certain probabilities after measurement.

Allowing superpositⁿ means that we allow linear combinations of the classic bits 0, 1, and measuring means taking orthogonal projections in quantum physics. So we want 0 and 1 to be seen as orthogonal vectors in some $\mathbb{C}$-vector space. This may be done via the following general constructⁿ.

Let $X$ be a finite set. Then $\mathbb{C}^X = \mathcal{F}(X, \mathbb{C})$ is a $\mathbb{C}$-vect. space, and the functions $|x\rangle : X \longrightarrow \mathbb{C}$, $y \longmapsto \begin{cases} 1 \text{ if } y = x \\ 0 \text{ otherwise} \end{cases}$ form a basis of $\mathbb{C}^X$.

(see ex 2.1). Moreover, if we endow $\mathbb{C}^X$ with the hermitian product

$$\langle \varphi, \psi \rangle := \sum_{x \in X} \overline{\varphi(x)}\, \psi(x)$$

then the family $(|x\rangle)_{x \in X}$ forms an orthonormal basis. Indeed, $\forall x, y \in X$,

$$\langle |x\rangle, |y\rangle \rangle = \sum_{z \in X} \overline{|x\rangle(z)}\, |y\rangle(z)$$

$$= \sum_{z \in X} \mathbb{1}_{x=z}\, \mathbb{1}_{y=z} = \mathbb{1}_{x=y}$$

We apply this construction to the set of "booleans" $B_n = \{0,1\}^n$ (that we can think of as $(\mathbb{Z}/2\mathbb{Z})^n$. this is often useful).

We denote by $Q_n := \mathbb{C}^{B_n}$ the vect. space with basis indexed by the elements of $B_n$. It is a hermitian space with respect to the hermitian product above, and the "elements of $B_n$" (their indicator funct$^{\text{s}}$ more precisely) form an orthonormal basis of $Q_n$.

Example : • For $n=1$, $B_1 = \{0,1\}$ $\quad Q_1 = \mathbb{C}|0\rangle \oplus \mathbb{C}|1\rangle$

and if $\varphi = a|0\rangle + b|1\rangle$ and $\psi = c|0\rangle + d|1\rangle$

we have $\langle \varphi, \psi \rangle = \bar{a}c + \bar{b}d$ and $\|\varphi\|^2 = |a|^2 + |b|^2$

$Q_1 \longmapsto \mathbb{C}^2$

$|0\rangle \longmapsto \begin{pmatrix} 1 \\ 0 \end{pmatrix}$

$|1\rangle \longmapsto \begin{pmatrix} 0 \\ 1 \end{pmatrix}$

is an isomorphism of $\mathbb{C}$ vector spaces that respects the hermitian product.

So one can think of $|0\rangle$ and $|1\rangle$ as the two canonical vectors of $\mathbb{C}^2$.

• For $n=2$, $B_2 = \{0,1\}^2$ so $Q_2 = \mathbb{C}|00\rangle \oplus \mathbb{C}|01\rangle \oplus \mathbb{C}|10\rangle \oplus \mathbb{C}|11\rangle$

elements of $Q_2$ are of the form $a|00\rangle + b|01\rangle + c|10\rangle + d|11\rangle$

their norm is $\sqrt{|a|^2 + |b|^2 + |c|^2 + |d|^2}$ and

$$|00\rangle \longmapsto \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \quad |01\rangle \longmapsto \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \end{pmatrix}, \quad |10\rangle \longmapsto \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \end{pmatrix}, \quad |11\rangle \longmapsto \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix}$$

is again an isomorphism $Q_2 \longrightarrow \mathbb{C}^4$ that preserves orthogonality.

Def : A $n$-qubit is a vector of norm 1 in $Q_n$; that is: a linear combinat $\sum_{x \in B_n} \lambda_x |x\rangle$ where $\sum_{x \in B_n} |\lambda_x|^2 = 1$.

Example : • $|0\rangle$ is a 1-qubit

• $\frac{1}{\sqrt{2}}(|0\rangle + |1\rangle)$ is also a 1-qubit.

• $\frac{1}{\sqrt{2}}(|00\rangle + |11\rangle)$ is a 2-qubit : it corresponds to the entanglement in the EPR experiment. The power of quantum computing relies on the possibility to create such entangled qubits.

when measured (we will give later a more precise sense to the notion of measure) this qubit can take the value $|00\rangle$ with probability $\frac{1}{2}$ and the value $|11\rangle$ with probability $\frac{1}{2}$.

**Def:** If $n, m \geq 1$, we define $Q_n \times Q_m \xrightarrow{\otimes} Q_{n+m}$ by $(\varphi, \psi) \longmapsto \varphi \otimes \psi$ the formula

$$|x_1 \cdots x_n\rangle \otimes |y_1 \cdots y_m\rangle = |x_1 \cdots x_n\, y_1 \cdots y_m\rangle$$

on $B_n \times B_m$, and then extend it by bilinearity

**Ex:**

- $|0\rangle \otimes |0\rangle = |00\rangle \in Q_2$
- $\frac{1}{\sqrt{2}}(|0\rangle + |1\rangle) \otimes \frac{1}{\sqrt{2}}(|0\rangle - |1\rangle) = \frac{1}{2}(|00\rangle - |01\rangle + |10\rangle - |11\rangle)$

**Rem:** $\otimes$ is not surjective (see ex 2.2)

## 2.3 Operations on qubits: only unitary operators make sense!

Let $V : Q_n \to Q_m$ be a $\mathbb{C}$-linear map. Since $n$-qubits are the elements of norm 1 in $Q_n$ and $m$-qubits are the elements of norm 1 in $Q_m$, if we want $V$ to turn qubits into qubits, it is natural to ask it to be norm preserving.

**Def:** If $E, E'$ are two hermitian spaces, a linear map $V : E \to E'$ is said to be unitary if $\forall x \in E$, $\|V(x)\|_{E'} = \|x\|_E$

(iff $\forall x, y \in E$, $\langle V(x), V(y)\rangle_{E'} = \langle x, y\rangle_E$)

(ADMISE)

**Prop:** If $M$ denotes the matrix of $V$ in orthonormal bases of $E, E'$ then $V$ is unitary $\Leftrightarrow M^*M = Id$, where $M^* := {}^t\overline{M}$ $\Leftrightarrow$ the columns of $M$ are orthonormal.

Since we want to do computer science, we want to extend any boolean funct $f: \{0,1\}^n \longrightarrow \{0,1\}^m$ to a linear map $Q_n \xrightarrow{V_f} Q_m$. (that needs to be unitary to act on qubits).

**Problem!** AND: $B_2 \longrightarrow B_1$

$$00 \longmapsto 0$$
$$01 \longmapsto 0$$
$$10 \longmapsto 0$$
$$11 \longmapsto 1$$

The matrix of AND in the qubits basis of $Q_2$ and $Q_1$ is

$$\begin{array}{c} \begin{matrix} |00\rangle & |01\rangle & |10\rangle & |11\rangle \end{matrix} \\ \begin{pmatrix} 1 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{matrix} |0\rangle \\ |1\rangle \end{matrix} \end{array}$$

So it is not unitary! (the first three columns are not orthogonal)

For instance, the image of $\frac{1}{\sqrt{2}}(|00\rangle + |10\rangle) \in Q_2$ via $V_f : Q_2 \to Q_1$ associated with $f = \text{AND}$ would be

$$V_f\left(\frac{1}{\sqrt{2}}(|00\rangle + |10\rangle)\right) = \frac{1}{\sqrt{2}}\left(f(|00\rangle) + f(|10\rangle)\right)$$

$$= \frac{1}{\sqrt{2}}(|0\rangle + |0\rangle) = \sqrt{2}\,|0\rangle$$ : it is not a 1-qubit since it has square-norm $= 2$.

Therefore, we need to use a trick to be able to extend boolean functions to unitary transformations and this trick consists in using two "registers", the input register and the output register.
In other words, instead of associating $f$ with $V_f$ in the "naive way"

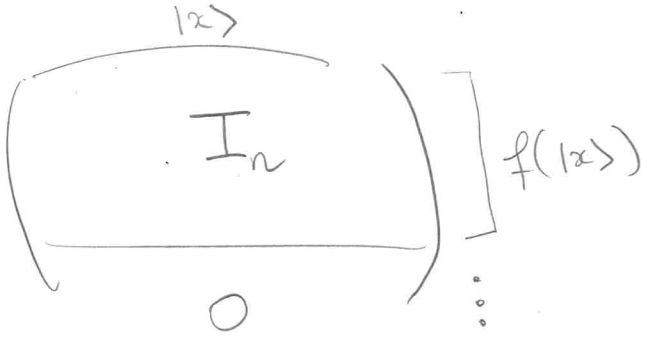
we are going to associate $f$ with a unitary gerator $U_f$ which "keeps memory of the input". Before doing that let us do a preparatory lemma

**Lem:** Let $f: B_n \to B_m$. Then the "naive" extension $V_f$ is unitary $\iff$ $f$ is injective.

**Proof:**

- If $V_f$ is unitary, then its matrix in the bases $(|x\rangle)_{x \in B_n}$ and $(|y\rangle)_{y \in B_m}$ has its columns of the form $\begin{pmatrix} 0 \\ \vdots \\ 1 \\ \vdots \\ 0 \end{pmatrix}$ and they must be orthogonal, meaning that no two 1's are on the same line, so $f$ is injective.

- Conversely, if $f$ is injective, then $(f(|x\rangle))_{x \in B_n}$ is a subset of the basis $(|y\rangle)_{y \in B_m}$ made of distinct vectors, so if we arrange $(|y\rangle)_{y \in B_m}$ in the order $(f(|x\rangle))_{x \in B_n}, \dots)$ then the matrix of $V_f$ is




So it is unitary. □

The following lemma gives a way to turn any boolean function into an injective one (the price to pay is to store more bits).

**Lem:** Let $f: B_n \longrightarrow B_m$. Then $\widehat{f}: B_{n+m} \longrightarrow B_{n+m}$

$$|x\ y\rangle \longmapsto |x\ \ y + f(x)\rangle$$

(sum modulo 2)

n bit "input register" — m bit "output register"

is injective (in fact even bijective)

Proof: $\tilde{f} \circ \tilde{f} = id_{B_{n+m}}$ !

Indeed $\tilde{f}(\tilde{f}(|xy\rangle)) = \tilde{f}(|x \quad y+f(x)\rangle)$

$= |x \quad y+\underbrace{f(x)+f(x)}_{0 \bmod 2}\rangle = |xy\rangle$

Thus, if $\tilde{f}(|xy\rangle) = \tilde{f}(|x'y'\rangle)$ then by applying $\tilde{f}$ again we deduce that $|xy\rangle = |x'y'\rangle$. This proves that $\tilde{f}$ is injective. □
(it also proves it is bijective)

Def: Given a boolean map $f: B_n \longrightarrow B_m$, we denote by $U_f$ the unitary operator $Q_{n+m} \longrightarrow Q_{n+m}$ naturally associated with $\tilde{f}: B_{n+m} \longrightarrow B_{n+m}$

Ex: • $f = AND: B_2 \longrightarrow B_1$ was not injective. The corresponding $\tilde{f}$ is

$\tilde{f}: B_3 \longrightarrow B_3$

$|x,y,z\rangle \longmapsto |x, y, z+f(|xy\rangle)\rangle$

For instance $\tilde{f}(|000\rangle) = |000\rangle$ $\quad \tilde{f}(|111\rangle) = |110\rangle$

$\tilde{f}(|001\rangle) = |001\rangle$

$\tilde{f}(|110\rangle) = |111\rangle$

So the matrix of $U_f$ in the basis $|000\rangle, |001\rangle, |110\rangle, |111\rangle, \ldots$ of $Q_3$ starts with

$$\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{pmatrix}$$

- The Hadamard gate : $H: Q_1 \longrightarrow Q_1$

$$|0\rangle \longmapsto \frac{1}{\sqrt{2}}(|0\rangle + |1\rangle)$$
$$|1\rangle \longmapsto \frac{1}{\sqrt{2}}(|0\rangle - |1\rangle)$$

- The S-gate $S: Q_1 \longrightarrow Q_1$

$$|0\rangle \longmapsto |0\rangle$$
$$|1\rangle \longmapsto i|1\rangle$$

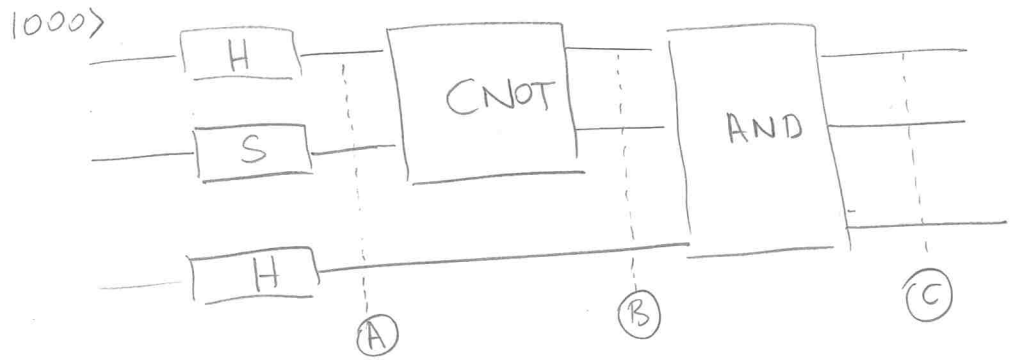
- The CNOT-gate $CNOT: Q_2 \longrightarrow Q_2$

$$|xy\rangle \longmapsto |x \quad x+y\rangle$$

(it is the oracle associated with $f = id: B_1 \longrightarrow B_1$)

Quantum algorithms can be represented by circuits:

Ex:




Notation To apply H to the 1st bit, S to the second, and H to the third, we often denote $(H \otimes S \otimes H)(|xyz\rangle)$

which is to be understood (by definition of this notation) as

$$H(x) \otimes S(y) \otimes H(z)$$

(Then $H \otimes S \otimes H$ is extended by linearity)

at Ⓐ : $H(|0\rangle) \otimes S(|0\rangle) \otimes H(|0\rangle)$

$$= \frac{1}{\sqrt{2}}(|0\rangle + |1\rangle) \otimes |0\rangle \otimes \frac{1}{\sqrt{2}}(|0\rangle + |1\rangle)$$

$$= \frac{1}{\sqrt{2}}(|00\rangle + |10\rangle) \otimes \frac{1}{\sqrt{2}}(|0\rangle + |1\rangle)$$

$$= \frac{1}{2}(|000\rangle + |001\rangle + |100\rangle + |101\rangle)$$

Then we apply the CNOT-gate at the first two bits, and do nothing on the third.

at Ⓑ : $\frac{1}{2}(|000\rangle + |001\rangle + |110\rangle + |111\rangle)$ (ici CNOT a agi.)

Finally we apply the unitary transformation associated with AND : $|x\,y\,z\rangle \longmapsto |x\,y\,z + f(x,y)\rangle$

at Ⓒ : $\frac{1}{2}(|000\rangle + |001\rangle + |111\rangle + |110\rangle)$

So the image of $|000\rangle$ under the unitary transformation $U : Q_3 \longrightarrow Q_3$ represented by the circuit and explicitly given by :

$$AND \circ (CNOT \otimes Id) \circ (H \otimes S \otimes H)$$

is $\frac{1}{2}(|000\rangle + |001\rangle + |111\rangle + |110\rangle)$.

Vague statement : Any unitary transformation $Q_n \longrightarrow Q_n$ can be approximated by applying successively gates belonging to the family

$$\{T, \underbrace{CNOT, H, S}\}$$

This is called the Clifford set

$Q_2 \longrightarrow Q_2$

$$\begin{pmatrix} 1 & 0 \\ 0 & e^{i\pi/4} \end{pmatrix}$$

Keywords : • Universal quantum gates.
• Solovay-Kitaev theorem (1995, 1997)

This has a practical interest because it is easier to build superpositions / gates of 1 or 2 qubits.

## 2.4 Measuring qubits.

If we have a 1-qubit $\varphi = \alpha|0\rangle + \beta|1\rangle$ (amplitudes), there is no way to extract information on $\alpha$ and $\beta$. Any measurement will return 0 or 1.

Born rule: "The probability of getting a particular result is given by the squared magnitude of the amplitude of the particular state in the expansion of $\varphi$ in the $\mathcal{B}_1$-basis"

here: $p(0) = |\alpha|^2$ and $p(1) = |\beta|^2$

For $n$-qubits with $n > 1$, we can of course measure the $n$ bits and similarly if $\varphi = \sum_{x \in \mathcal{B}_n} \alpha_x |x\rangle$ then the outcome of the measure is $|x\rangle$ with probability $p(x) = |\alpha_x|^2$. However, we can also do a measure only on the first bits.

Q°: What is the outcome of such a measure (with what probability)? What is the new state of the qubit after the measure?

Let's start with an example: $|\psi\rangle = \frac{1}{\sqrt{2}}|00\rangle + \frac{1}{2}|01\rangle + \frac{1}{2}|11\rangle \in \mathcal{Q}_2$

$$= |0\rangle \otimes \left(\frac{1}{\sqrt{2}}|0\rangle + \frac{1}{2}|1\rangle\right) + |1\rangle \otimes \frac{1}{2}|1\rangle$$

- If the outcome of the measure on the first bit is 1 (happens with proba $\left(\frac{1}{2}\right)^2$) necessarily the second bit is 1 because $|11\rangle$ is the only 2-qubit in the superposition defining $|\psi\rangle$ which has a 1 as a first bit $\Rightarrow |\psi_{new}\rangle = |11\rangle$
- On the other hand, if the outcome is 0 (happens with proba $\left(\frac{1}{\sqrt{2}}\right)^2 + \left(\frac{1}{2}\right)^2$) then $|\psi_{new}\rangle$ is still a superposition of $|00\rangle$ and $|01\rangle$. We want to say it is the first part of the expansion, that is. $|0\rangle \otimes \left(\frac{1}{\sqrt{2}}|0\rangle + \frac{1}{2}|1\rangle\right) = \frac{1}{\sqrt{2}}|00\rangle + \frac{1}{2}|01\rangle$

but ⚠ it is not a 2-qubit since it does not have norm 1. So we need to renormalize : $|\psi_{new}\rangle = \frac{\frac{1}{\sqrt{2}}|00\rangle + \frac{1}{2}|01\rangle}{\sqrt{\left(\frac{1}{\sqrt{2}}\right)^2 + \left(\frac{1}{2}\right)^2}}$

$$|\psi_{new}\rangle = \frac{\sqrt{2}}{\sqrt{3}}|00\rangle + \frac{1}{\sqrt{3}}|01\rangle$$

Generalized Born Rule: Let $n \geq 1$, $k \leq n$. Let $|\psi\rangle \in Q_n$.
We want to measure the first $k$ bits of $|\psi\rangle$.

- We write $|\psi\rangle = \sum_{x \in B_k} |x\rangle \otimes |\psi_x\rangle$ ($|\psi_x\rangle$: may be a superposition as we saw in the example)

- 


notation for a measurement gate

the outcome of the measure is $|x\rangle$ with probability $\| |\psi_x\rangle \|^2$, and if it is $x$, then

$$|\psi_{new}\rangle = \frac{|x\rangle \otimes |\psi_x\rangle}{\| |\psi_x\rangle \|^2}$$

## 2.5. Deutsch-Jozsa algorithm.

We present a simple case, for the general case see ex 2.7
We are given a map $f : B_1 \to B_1$ and the question is "Is $f(0)$ equal to $f(1)$ ?"
In the classical setting, we need to compute $f(0)$ and $f(1)$, so we need two calls to the map $f$.
The Deutsch-Jozsa algorithm shows that in the quantum setting, one needs only one call to the oracle $U_f$ associated with $f$ to answer the question!

## Remark

We would like to underline the fact that the generalized Born rule stated here really fits in the general framework of quantum measurement. Indeed define $c : B_k \to \mathbb{R}$ any injective map (for instance it maps $x_1 \cdots x_k \in B_k$ to the integer whose binary expansion is $x_1 \cdots x_k$). Then define

$$H : Q_n \longrightarrow Q_n$$

by defining it on $B_n$ by $H(|x_1 \cdots x_n\rangle) = c(x_1 \cdots x_k)|x_1 \cdots x_n\rangle$

Then $H$ has a diagonal matrix in the orthonormal basis $(|x\rangle)_{x \in B_n}$ of $Q_n$, so it is indeed the matrix of an observable.

By the Born rule stated in sect° 2.1:

- the outcome of the measurement will be $c(x_1, \ldots, x_k)$ (so essentially recovers $x_1 \cdots x_k$ since $c$ is injective)
  with probability = the square magnitude of the orthogonal projection of $\psi$ on the eigenspace associated with the eigenvalue $c(x_1 \cdots x_k)$: that is the space generated by the elements $|x_1 \cdots x_k\rangle \otimes |y\rangle$ for $y \in B_{n-k}$
- After the measurement, the new $\psi$ that equals the (renormalized) projection on this eigenspace.

- Now, this may seem like cheating, because if you want to write down the matrix of $U_f$ in the basis $|00\rangle, |01\rangle, |10\rangle, |11\rangle$ of $\mathbb{Q}_2$, you need to evaluate $f$ at 0 and 1.

(since $\tilde{f}(|xy\rangle) = |x \quad y+f(x)\rangle$ )

However, as far as I understand, the point is that IF we can build a quantum circuit (made of gates such as H, CNOT, AND ...) that acts like $U_f$ without evaluating $f$ at 0 and 1, then the Deutsch-Jozsa algorithm is useful because it only requires one call to this $U_f$.

Remark on the general case that is treated in the exercise sheet:

One might say that the simple case above is not so impressive because we just pass from two calls to $f$ to one call to $U_f$. However, the general case is more convincing:

Assume that you are given $f: B_n = \{0,1\}^n \longrightarrow \{0,1\} = B_1$ and you know that either $f$ is constant

or $f$ is balanced ie $|f^{-1}(\{0\})| = |f^{-1}(\{1\})| = 2^{n-1}$.

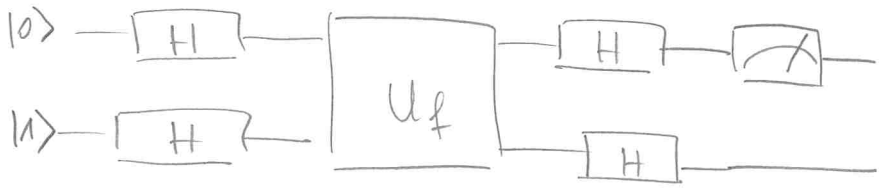
Q°: Is $f$ constant or balanced?

With a classical computer, answering this question would require $2^{n-1}+1$ evaluations of $f$. If $f$ takes the same value at these $2^{n-1}+1$ points then it must be constant

However, the Deutsch-Jozsa algorithm allows one to answer the quest° on a quantum computer with just one call to $U_f$.

So here we pass from $2^{n-1}+1$ to 1, so the gain is exponential!

Let us now describe the Deutsch-Jozsa algorithm in our simple case $f: B_1 \longrightarrow B_1$. We consider the following circuit




$$|\Psi\rangle \longrightarrow |\Psi_1\rangle \longrightarrow |\Psi_2\rangle \longrightarrow |\Psi_3\rangle \longrightarrow |\Psi_4\rangle$$

Let us compute the evolut° of the state of the 2-qubit $|\Psi\rangle = |01\rangle \in Q_2$ as it goes through this circuit.

$$|\Psi_1\rangle = (H\otimes H)(|01\rangle) = H(|0\rangle)\otimes H(|1\rangle) = \frac{1}{\sqrt{2}}(|0\rangle+|1\rangle)\otimes\frac{1}{\sqrt{2}}(|0\rangle-|1\rangle)$$

$$= \frac{1}{2}\left(|00\rangle - |01\rangle + |10\rangle - |11\rangle\right)$$

$$|\Psi_2\rangle = \frac{1}{2}\left(U_f(|00\rangle) - U_f(|01\rangle) + U_f(|10\rangle) - U_f(|11\rangle)\right)$$

$$= \frac{1}{2}\left(|0\ f(0)\rangle - |0\ 1+f(0)\rangle + |1\ f(1)\rangle - |1\ 1+f(1)\rangle\right)$$

because $U_f(|xy\rangle) = |x\ \ y+f(x)\rangle$.

- If $f(0)=f(1)$, then $|\Psi_2\rangle = \frac{1}{2}\left(|0\ f(0)\rangle - |0\ 1+f(0)\rangle + |1\ f(0)\rangle - |1\ 1+f(0)\rangle\right)$

$$= \frac{1}{2}\left(|0\rangle+|1\rangle\right)\otimes\left(|f(0)\rangle - |1+f(0)\rangle\right)$$

- If $f(1)=1+f(0)$ then $|\Psi_2\rangle = \frac{1}{2}\left(|0\rangle-|1\rangle\right)\otimes\left(|f(0)\rangle - |1+f(0)\rangle\right)$

Now we re-apply $H\otimes H$ to $|\Psi_2\rangle$.

Since we are going to measure the 1st bit, we just compute what happens on this bit:

- If $f(0) = f(1)$, then $H\left(\frac{1}{\sqrt{2}}(|0\rangle + |1\rangle)\right)$

$$= \frac{1}{\sqrt{2}}\left(\frac{1}{\sqrt{2}}(|0\rangle + |1\rangle) + \frac{1}{\sqrt{2}}(|0\rangle - |1\rangle)\right)$$

$$= |0\rangle$$

So $|\psi_2\rangle = |0\rangle \otimes (\dots)$ So the measure of the first bit returns $|0\rangle$ with probability 1.

- If $f(0) \neq f(1)$, then $H\left(\frac{1}{\sqrt{2}}(|0\rangle - |1\rangle)\right)$

$$= \frac{1}{\sqrt{2}}\left(\frac{1}{\sqrt{2}}(|0\rangle + |1\rangle) - \frac{1}{\sqrt{2}}(|0\rangle - |1\rangle)\right)$$

$$= |1\rangle$$ So the measure of the first bit returns $|1\rangle$ with probability 1.

Conclusion: The measure at the end of this circuit outputs

$|0\rangle$ if $f(0) = f(1)$

$|1\rangle$ if $f(0) \neq f(1)$

# (3) Shor's algorithm.

## 3.1 Overview of Shor's factoring algorithm.

- We are given an integer $N$ and we want an algorithm that outputs the factorizat° of $N$ as $p_1^{\alpha_1} \cdots p_r^{\alpha_r}$ (a product of prime numbers).

Best known algorithm for a classical computer: It is called the general number field sieve (GNFS) and its complexity is

$$O\left(\exp\left(C (\log N)^{\frac{1}{3}} (\log\log N)^{\frac{2}{3}}\right)\right)$$

Since $\log(N)$ is roughly the number of digits of $N$, it is the size of the input, so GNFS is super-polynomial in the size of the input (meaning that the complexity above is greater, as $N \to \infty$, than any polynomial in $\log N$)

- Shor's algorithm (1994) is a quantum algorithm whose number of computational steps is polynomial in $\log N$. It eventually relies on a quantum period finding algorithm. Let's explain what is the link between factorizat° and period finding.

In order to factor $N$, we are going to explain an algorithm that outputs two non-trivial factors (ie integers $d$ such that $1 < d < N$ and $d \mid N$) of $N$.

First, choose $a \in \{2, \dots, N-1\}$ at random.

- If $\gcd(a, N) > 1$, then $d := \gcd(a, N)$ is a non-trivial factor, so we output $d$ and $\frac{N}{d}$.

- If $\gcd(a, N) = 1$, then $a \in (\mathbb{Z}/N\mathbb{Z})^\times$, so it makes sense to speak of its multiplicative order $r := \mathrm{ord}(a)$.

Now, in order to go further, we need to make one assumption, and we will see in the exercises that it is in fact satisfied by a reasonable proportion of $a$ (so that if it is not satisfied, we can pick another $a'$ at random, and without too many repetitions we should end up picking a "good" element).

**Assumpt° A1** $r = \text{ord}(a)$ is even.

Thanks to this assumption, we can write $a^r \equiv 1 \pmod N$ as

$$\left(a^{\frac{r}{2}}\right)^2 - 1^2 \equiv 0 \pmod N, \text{ or equivalently } \left(a^{\frac{r}{2}} - 1\right)\left(a^{\frac{r}{2}} + 1\right) \equiv 0 \pmod N.$$

We now make a second assumpt° (again, we will see that it is "often" satisfied).

**Assumpt° A2** $a^{\frac{r}{2}} + 1 \not\equiv 0 \pmod N$.

Under these two assumptions, we have the following lemma which allows us to return the desired output:

**Lemma:** Under A1 and A2, $d := \gcd(a^{\frac{r}{2}} - 1, N)$ and $d' = \gcd(a^{\frac{r}{2}} + 1, N)$ are non-trivial divisors of $N$.

**Proof:**

- We have $d < N$ because if $d = N$ then $a^{\frac{r}{2}} - 1 \equiv 0 \pmod N$, contradicting the minimality of $r$.

  We have $d' < N$ thanks to A2

- We have $d > 1$ because if we had $d = 1$ then $N \mid (a^{\frac{r}{2}} - 1)(a^{\frac{r}{2}} + 1)$ (coprime)

  $\Rightarrow N \mid a^{\frac{r}{2}} + 1$ : contradicts A2.

  We have $d' > 1$ because if $d' = 1$ then $N \mid (a^{\frac{r}{2}} - 1)(a^{\frac{r}{2}} + 1)$ (coprime)

  $\Rightarrow N \mid a^{\frac{r}{2}} - 1$ : contradict°

□

However, we are only able to get good lower bounds for the probability that A1 and A2 are satisfied when $N$ has at least two distinct prime factors (note that this is the case in RSA) and $N$ is odd (however, this latter assumption is not really an issue since $\simeq \log_2(N)$ euclidean divisions are enough to determine the factor $2^\alpha$ which potentially appears in the factorization of $N$, so this can be done efficiently on a classical computer).

Summary of Shor's algorithm:

Input: An odd natural number $N$ that has at least two distinct prime factors.

Step 1: Take $a \in \{2, \dots, N-1\}$ at random

- if $\gcd(a,N) > 1$, output $\gcd(a,N)$ and $\frac{N}{\gcd(a,N)}$
- if $\gcd(a,N) = 1$ go to Step 2.

Step 2: Determine the order $r$ of $a$ in $(\mathbb{Z}/N\mathbb{Z})^\times$.

- If $r$ is odd, start again at Step 1
- If $r$ is even, go to Step 3.

Step 3: Determine $\gcd(a^{r/2}+1, N)$

- If $\gcd(a^{r/2}+1, N) = N$, start again at Step 1
- otherwise, output $\gcd(a^{r/2}+1, N)$ and $\gcd(a^{r/2}-1, N)$

Output: Two non-trivial divisors of $N$.

In order for this algo to be efficient, we need the "start again" steps to be rare. This can be done rigorously by estimating the probability that A1 and A2 fail for a random element $a \in \mathbb{Z}/N\mathbb{Z}$ (see the exercises)

The link with period finding lies in Step 2. Indeed, the order of $a \in (\mathbb{Z}/N\mathbb{Z})^\times$ is nothing else than the shortest period of the map

$$f : \mathbb{N} \longrightarrow \{0, \dots, N-1\}$$
$$k \longmapsto a^k \pmod N$$

## 3.2 Period-finding algorithm via Quantum Fourier Transform

We first define the Fourier transform on any finite abelian group before specializing to the case of interest to us.

**Def:** Let $G$ be a finite abelian group. If $f : G \to \mathbb{C}$, its Fourier transform is $\hat{f} : \hat{G} \to \mathbb{C}$

$$\chi \longmapsto \frac{1}{\sqrt{|G|}} \sum_{x \in G} f(x)\chi(x)$$

**Rem:** In the particular case $G = \mathbb{Z}/n\mathbb{Z}$, we've seen in the first exercise sheet that

$$\hat{G} \simeq G$$
$$\psi_a \longleftrightarrow a \qquad \text{where } \psi_a(x) = \exp\left(\frac{2i\pi a x}{n}\right)$$

Thus we rather view $\hat{f}$ as a map $G \to \mathbb{C}$:

$$\hat{f}(a) = \frac{1}{\sqrt{n}} \sum_{x \in \mathbb{Z}/n\mathbb{Z}} f(x) \exp\left(\frac{2i\pi a x}{n}\right) \qquad (*)$$

Let us go back to our question of finding the shortest period of $f : \mathbb{N} \to \mathbb{N}$, $k \mapsto a^k \pmod N$.

We take an integer $n$ such that $N^2 < 2^n$ (this is a technical condit°, the reason for this choice will become clearer later. For now, $N < 2^n$ would be sufficient)

We identify the integers $x \in \{0, \dots, 2^n - 1\}$ to elements of $\mathcal{B}_n$ via their binary expansion. In other words, $x = x_{n-1} 2^{n-1} + \dots + x_1 2^1 + x_0 2^0$ is identified to the $n$-qubit $|x_{n-1} \dots x_0\rangle$.

We denote by $QFT_{2^n}$ the linear map $\mathcal{Q}_n \longrightarrow \mathcal{Q}_n$ defined on $\mathcal{B}_n$ by

$$QFT_{2^n}(|x\rangle) = \frac{1}{2^{n/2}} \sum_{y=0}^{2^n-1} \exp\left(\frac{2i\pi xy}{2^n}\right) |y\rangle$$

[observe that this is just $(*)$ with $f = |x\rangle$ and $G = \mathbb{Z}/2^n\mathbb{Z}$.]

Lemma: $QFT_{2^n}$ is a unitary transformation $\mathcal{Q}_n \longrightarrow \mathcal{Q}_n$

Proof: $\left\langle QFT_{2^n}(|x\rangle), QFT_{2^n}(|x'\rangle) \right\rangle$

$$= \left\langle \frac{1}{2^{n/2}} \sum_{y=0}^{2^n-1} \exp\left(\frac{2i\pi xy}{2^n}\right) |y\rangle \;,\; \frac{1}{2^{n/2}} \sum_{y'=0}^{2^n-1} \exp\left(\frac{2i\pi x'y'}{2^n}\right) |y'\rangle \right\rangle$$

$$= \frac{1}{2^n} \sum_{y,y'} \exp\left(\frac{2i\pi(x'y' - xy)}{2^n}\right) \underbrace{\langle |y\rangle, |y'\rangle \rangle}_{= \mathbb{1}_{y=y'}}$$

$$= \frac{1}{2^n} \sum_{y=0}^{2^n-1} \exp\left(\frac{2i\pi y(x'-x)}{2^n}\right) = \mathbb{1}_{2^n \mid x'-x} = \mathbb{1}_{x'=x}$$

(see the exercise on the first exercise sheet) □

Since it is unitary, it belongs to the family of transformations that "make sense" or "act on qubits", so IF we can build a quantum circuit (made of "elementary" gates) that acts as $QFT_{2^n}$, then we would construct a period-finding algorithm as follows:

We consider the following circuit:


unitary transformation associated with $f: k \mapsto a^k \pmod N$ via the usual $|xy\rangle \mapsto |x\ y+f(x)\rangle$.

Explicitly implementing the gate $U_f$ in terms of elementary gates is also a difficult question that we shall study later.

Input: $|\psi\rangle = |0\rangle \otimes |0\rangle \in Q_{2n}$

Then $|\psi_1\rangle = H^{\otimes n}(|0\rangle) \otimes |0\rangle = \left( \frac{1}{\sqrt{2}}(|0\rangle + |1\rangle) \otimes \cdots \otimes \frac{1}{\sqrt{2}}(|0\rangle + |1\rangle) \right) \otimes |0\rangle$

$$= \frac{1}{2^{n/2}} \sum_{x=0}^{2^n-1} |x\rangle \otimes |0\rangle$$

and $$|\psi_2\rangle = \frac{1}{2^n} \sum_{x=0}^{2^n-1} |x\rangle \otimes |f(x)\rangle$$

Finally, $$|\psi_3\rangle = \frac{1}{2^{n/2}} \sum_{x=0}^{2^n-1} QFT_{2^n}(|x\rangle) \otimes f(x)$$

$$= \frac{1}{2^{n/2}} \sum_{x=0}^{2^n-1} \frac{1}{2^{n/2}} \sum_{y=0}^{2^n-1} \exp\left(\frac{2i\pi xy}{2^n}\right) |y\rangle \otimes |f(x)\rangle$$

$$= \frac{1}{2^n} \sum_{y=0}^{2^n-1} |y\rangle \otimes \left( \sum_{x=0}^{2^n-1} \exp\left(\frac{2i\pi xy}{2^n}\right) |f(x)\rangle \right)$$

So the outcome of the measure of the first n-bits is $y$ with probability

$$p(y) = \left\| \frac{1}{2^n} \sum_{x=0}^{2^n-1} \exp\left(\frac{2i\pi xy}{2^n}\right) |f(x)\rangle \right\|^2$$

Now the problem is that the $|f(x)\rangle$ that appear in this sum are not necessarily distinct (they typically aren't, since $f$ is periodic!), so we cannot easily compute the norm squared using the orthogonality of the basis vectors $|f(x)\rangle$.

Let us denote by $r$ the smallest period of $f$.

Then we have

$$\sum_{x=0}^{2^n-1} \exp\left(\frac{2i\pi xy}{2^n}\right) |f(x)\rangle$$

$$= \sum_{x=0}^{r-1} \left( \sum_{\substack{x'=0 \\ x' \equiv x \,(\mathrm{mod}\, r)}}^{2^n-1} \exp\left(\frac{2i\pi x'y}{2^n}\right) \right) |f(x)\rangle$$

and now the $|f(x)\rangle$ in this sum are distinct vectors of the basis of $\mathbb{Q}_n$ made of $(|y\rangle)_{y \in B_n}$.

Therefore, $p(y) = \frac{1}{2^{2n}} \sum_{x=0}^{r-1} \left| \sum_{\substack{x'=0 \\ x' \equiv x \,(\mathrm{mod}\, r)}}^{2^n-1} \exp\left(\frac{2i\pi x'y}{2^n}\right) \right|^2$

Let's compute the inner sum:

$$\sum_{\substack{x'=0 \\ x' \equiv x \,(\mathrm{mod}\, r)}}^{2^n-1} \exp\left(\frac{2i\pi x'y}{2^n}\right) \overset{x' = x + jr}{=} \sum_{j=0}^{J_x} \exp\left(\frac{2i\pi (x+jr)y}{2^n}\right)$$

$$\left( \begin{array}{c} 0 \leq x + jr \leq 2^n - 1 \\ \Leftrightarrow -x \leq jr \leq 2^n - 1 - x \\ \Leftrightarrow \underbrace{\frac{-x}{r}}_{\text{strictly greater than } -1} \leq j \leq \frac{2^n - 1 - x}{r} \end{array} \right)$$

where $J_x := \left\lfloor \frac{2^n - 1 - x}{r} \right\rfloor$

$$= \underbrace{\exp\left(\frac{2i\pi xy}{2^n}\right)}_{\zeta_x} \sum_{j=0}^{J_x} \exp\left(\frac{2i\pi jry}{2^n}\right)$$

$$= \zeta_x \sum_{j=0}^{J_x} \exp\left(\frac{2i\pi ry}{2^n}\right)^j$$

$$= \begin{cases} \zeta_x (J_x+1) & \text{if } 2^n \mid ry \quad (\Leftrightarrow \frac{2^n}{r} \mid y) \\ \zeta_x \dfrac{1 - \exp\left(\frac{2i\pi ry}{2^n}\right)^{J+1}}{1 - \exp\left(\frac{2i\pi ry}{2^n}\right)} & \text{otherwise} \end{cases}$$

Thus, $$p(y) = \begin{cases} \frac{1}{2^{2n}} \sum_{x=0}^{r-1} (J_x+1)^2 & \text{if } \frac{2^n}{r} \mid y \\ \frac{1}{2^{2n}} \sum_{x=0}^{r-1} \underbrace{\left| \dfrac{1 - \exp\left(\frac{2i\pi ry(J_x+1)}{2^n}\right)}{1 - \exp\left(\frac{2i\pi ry}{2^n}\right)} \right|^2}_{} & \text{otherwise} \end{cases}$$

$$= \frac{\sin^2\left(\frac{\pi ry(J_x+1)}{2^n}\right)}{\sin^2\left(\frac{\pi ry}{2^n}\right)}$$

The function $p$ can be studied as a function of the variable $y$, and one can prove (but this is quite technical) that it takes large values when $y$ is close to a multiple of $\frac{2^n}{r}$

Therefore, one can show that with high probability (which makes it very unlikely to need to run many times this circuit), the measure outputs a $y \in \{0, \dots, 2^n - 1\}$ such that

$$\text{dist}\left(y, \frac{2^n}{r}\mathbb{Z}\right) \leq \frac{1}{2}$$

i.e. there exists $\ell \in \mathbb{Z}$ such that

$$\left| y - \frac{2^n}{r}\ell \right| \leq \frac{1}{2} \qquad (\text{in fact } \ell \in \{0, \dots, r-1\})$$

Q°: How do we recover $r$ from the output of the measure, which is $y$? We just know that with high probability $y$ is close to a rational number with denominator $r$, but how can we recover $r$ from $y$?

Dividing by $2^n$, we have $\left| \underbrace{\frac{y}{2^n}}_{\substack{\text{known by the measurement} \\ \text{and the knowledge of } n}} - \underbrace{\frac{\ell}{r}}_{\text{unknown}} \right| \leq \frac{1}{2^{n+1}}$

So we know a good approximate of $\frac{y}{2^n}$ by a rational with a smaller denominator, because $r \leq \varphi(N) \leq N \leq N^2 \leq 2^n$. Moreover, this approximation is so good that the theory of continued fractions will allow us to recover $r$!

Th: Let $x \in \mathbb{R}$. If $\frac{p}{q} \in \mathbb{Q}$ is such that $\left| x - \frac{p}{q} \right| < \frac{1}{2q^2}$, then $\frac{p}{q}$ is obtained as one of the fractions of the continued fraction expansion of $x$.

Here $\left|\frac{y}{2^n} - \frac{l}{r}\right| \leq \frac{1}{2 \times 2^n} < \frac{1}{2r^2}$ so the thm applies

and tells us that $\frac{l}{r}$ appears as one of the fractions of the continued fract° expansion of $\frac{y}{2^n}$ (known), so this allows us to recover $\frac{l}{r}$, and then $r$.

Example. Take $N = 21$, $n = 9$ $\left(2^9 = 512 > N^2 = 441\right)$

and $a = 2 \in (\mathbb{Z}/N\mathbb{Z})^\times$.

If the outcome of Shor's algorithm is 427, then we calculate the convergents of the continued fraction representation of $\frac{427}{512}$.

- $512 = 427 \times 1 + 85 \iff \frac{512}{427} = 1 + \frac{85}{427}$ so $\frac{427}{512} = \frac{1}{1 + \frac{85}{427}}$

- $427 = 85 \times 5 + 2 \iff \frac{427}{85} = 5 + \frac{2}{85}$ so $\frac{427}{512} = \frac{1}{1 + \frac{1}{5 + \frac{2}{85}}}$

- $85 = 2 \times 42 + 1 \iff \frac{85}{2} = 42 + \frac{1}{2}$ so $\frac{427}{512} = \frac{1}{1 + \frac{1}{5 + \frac{1}{42 + \frac{1}{2}}}}$

Cl: $\frac{427}{512} = 0 + \cfrac{1}{1 + \cfrac{1}{5 + \cfrac{1}{42 + \frac{1}{2}}}}$ : denoted by $[0, 1, 5, 42, 2]$

This gives the following list of approximations of $\frac{427}{512}$:

$[0,1]$ gives $\frac{427}{512} \simeq \frac{p_1}{q_1} = 0 + \frac{1}{1} = 1$

$[0,1,5]$ gives $\frac{427}{512} \simeq \frac{p_2}{q_2} = 0 + \frac{1}{1+\frac{1}{5}} = \frac{5}{6}$

$[0,1,5,42]$ gives $\frac{427}{512} \simeq \frac{p_3}{q_3} = 0 + \frac{1}{1+\frac{1}{5+\frac{1}{42}}} = \frac{211}{253}$

$[0,1,5,42,2]$ gives $\frac{427}{512}$

The only ones with denominator $\leq N$ are 1 and $\frac{5}{6}$. Since the order of a is not 1, it is 6. If we had more convergents of denominator $\leq N$, we would have to test them by calculating $a^q$ and see which one is the order of a (but this method still reduces drastically the number of q to test).

In the exercises, we will plot the function $p(y)$ in this particular case, and see that it takes large values only for those y that are close to $\frac{2^n \ell}{r}$ for $\ell \in \{0, \dots, r-1\}$.

## 3.3. Implementing the gates $U_f$ and $QFT_{2^n}$

In the previous section, we defined the unitary transformation $QFT_{2^n}: Q_n \longrightarrow Q_n$ by its values on $B_n$:

$$QFT_{2^n}(|x\rangle) = \frac{1}{2^{n/2}} \sum_{y=0}^{2^n-1} \exp\left(\frac{2i\pi xy}{2^n}\right) |y\rangle$$

$x = x_{n-1} \cdots x_0$ corresponds to the integer $x_{n-1}2^{n-1} + \cdots + a_0 2^0$

Let us denote by $w_j := \exp\left(\frac{2i\pi}{2^j}\right)$ for all $j \geq 1$.

Then $QFT_{2^n}(|x\rangle) = \frac{1}{2^{n/2}} \sum_{y=0}^{2^n-1} \exp\left(\frac{2i\pi xy}{2^n}\right)|y\rangle$

$$= \frac{1}{2^{n/2}} \sum_{(y_{n-1},\dots,y_0)\in B_n} \omega_1^{xy_{n-1}} \omega_2^{xy_{n-2}} \cdots \omega_n^{xy_0} |y_{n-1}\cdots y_0\rangle$$

(because if $y = y_{n-1}2^{n-1} + \cdots + y_0 2^0$

then $\exp\left(\frac{2i\pi xy}{2^n}\right) = \exp\left(\frac{2i\pi}{2^n} x(y_{n-1}2^{n-1} + \cdots + y_0 2^0)\right)$

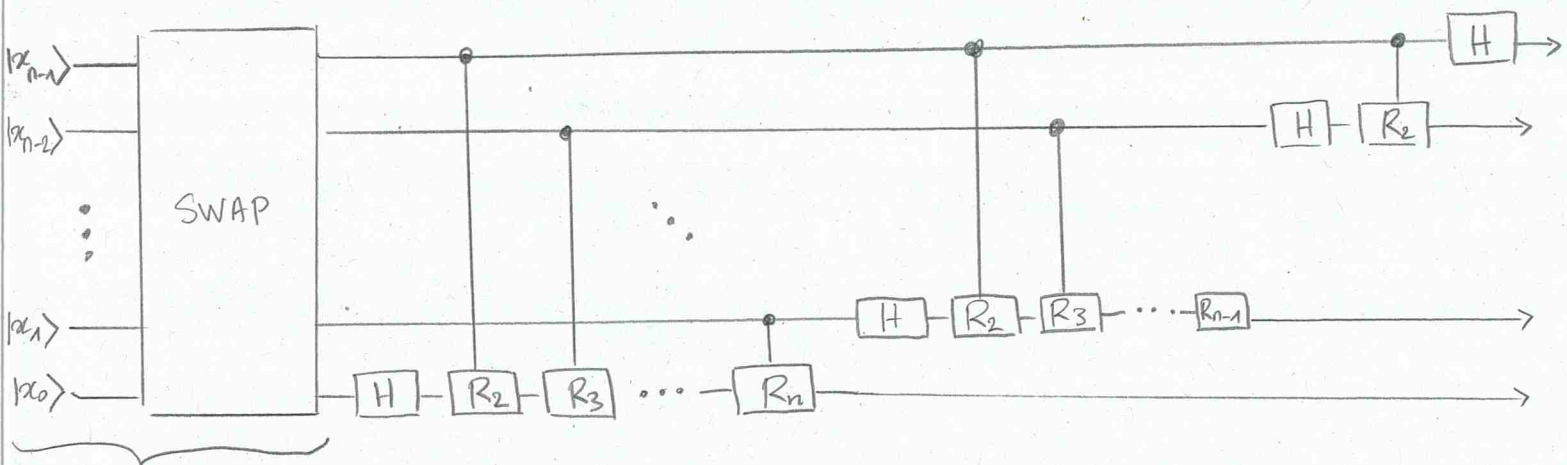
$= \omega_1^{xy_{n-1}} \omega_2^{xy_{n-2}} \cdots \omega_n^{xy_0}$)

So $QFT_{2^n}(|x\rangle) = \frac{1}{\sqrt{2}}(|0\rangle + \omega_1^x|1\rangle) \otimes \frac{1}{\sqrt{2}}(|0\rangle + \omega_2^x|1\rangle) \otimes \cdots \otimes \frac{1}{\sqrt{2}}(|0\rangle + \omega_n^x|1\rangle)$

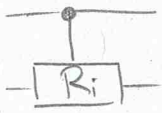
**Def:** For all $j \geq 1$, define $R_j : Q_1 \longrightarrow Q_1$ to be the unitary transformation whose matrix in the basis $(|0\rangle, |1\rangle)$ is

$$\begin{pmatrix} 1 & 0 \\ 0 & \omega_j \end{pmatrix} \quad \text{ie} \quad \begin{cases} R_j(|0\rangle) = |0\rangle \\ R_j(|1\rangle) = \omega_j|1\rangle \end{cases}$$

The formula for $QFT_{2^n}(|x\rangle)$ as a tensor product of terms of the form $\frac{1}{\sqrt{2}}(|0\rangle + \omega_j^x|1\rangle)$ guides us to the following circuit:

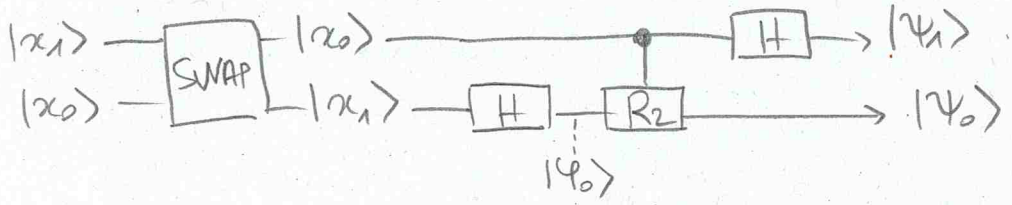



we will not focus on this step but it just turns $|x_{n-1}\cdots x_0\rangle$ into $|x_0\cdots x_{n-1}\rangle$

The notation  means that the second qubit is <u>controlled</u> by the first (it is called a controlled gate). If the first qubit is 1, then $R_i$ acts normally on the second one, but if the first qubit is 0, then $R_i$ does not act at all.



Let's check that this circuit works for $n=2$: (the general case can be proved by induction, but it is quite tedious).




we have $|\psi_1\rangle = H(|x_0\rangle) = \frac{1}{\sqrt{2}}\left(|0\rangle + (-1)^{x_0}|1\rangle\right)$

but $-1 = \exp\left(\frac{2i\pi}{2}\right) = \omega_1$ so $\boxed{|\psi_1\rangle = \frac{1}{\sqrt{2}}\left(|0\rangle + \omega_1^{x_0}|1\rangle\right)}$

On the other hand,

$|\varphi_0\rangle = \frac{1}{\sqrt{2}}\left(|0\rangle + (-1)^{x_1}|1\rangle\right)$ and then we need to split into two cases, depending on whether $x_0 = 1$ or not (because then $R_2$ acts or does not act)

$x_0 = 0$

Then $R_2$ is inhibited, so

$|\psi_0\rangle = |\varphi_0\rangle = \frac{1}{\sqrt{2}}(|0\rangle + (-1)^{x_1}|1\rangle)$

$x_0 = 1$

Then $R_2$ acts so

$|\psi_0\rangle = \frac{1}{\sqrt{2}}(|0\rangle + (-1)^{x_1} R_2(|1\rangle))$

$= \frac{1}{\sqrt{2}}(|0\rangle + (-1)^{x_1} \omega_2 |1\rangle)$

In both cases, $\boxed{|\psi_0\rangle = \frac{1}{\sqrt{2}}(|0\rangle + \omega_1^{x_1} \omega_2^{x_0} |1\rangle)}$

But now, we can observe that $|\psi_1\rangle = \frac{1}{\sqrt{2}}(|0\rangle + \omega_1^{x_0}|1\rangle)$

$= \frac{1}{\sqrt{2}}(|0\rangle + \omega_1^{2x_1+x_0}|1\rangle)$

because $\omega_1^2 = 1$

$\Rightarrow \boxed{|\psi_1\rangle = \frac{1}{\sqrt{2}}(|0\rangle + \omega_1^{x}|1\rangle)}$

and $|\psi_0\rangle = \frac{1}{\sqrt{2}}(|0\rangle + \omega_1^{x_1}\omega_2^{x_0}|1\rangle)$

$= \frac{1}{\sqrt{2}}(|0\rangle + \omega_2^{2x_1+x_0}|1\rangle)$ because $\omega_1 = \omega_2^2$

$\boxed{|\psi_0\rangle = \frac{1}{\sqrt{2}}(|0\rangle + \omega_2^{x}|1\rangle)}$

So thanks to the expansion of $QFT_{2^n}$ as a tensor product, we see that $|\psi_1\rangle \otimes |\psi_0\rangle = QFT_{2^n}(|x\rangle)$ : the output of the circuit is indeed $QFT_{2^n}(|x\rangle)$.

This circuit has $n + (n-1) + \dots + 1 = \frac{n(n+1)}{2} = O(n^2)$ gates : so it is polynomial in the size $(n)$ of the entry.

Dans l'algo de Shor, bien dire que c'est l'exponentiat qui coûte le plus en temps et en espace (cf l'article de Shor).

Mais elle peut être réalisée en $O(n^2 \log(n) \log(\log n))$

$= \tilde{O}(n^2)$

For large values of $n$, this is the best.

But for small values one rather uses "textbook" multiplicat in $O(n^3)$ (time).

# 4.1. Regev's variant as a multidimensional period-finding algorithm

We give an overview of the result proved by Regev in the article "An efficient quantum factoring algorithm".

Let $N < 2^n$ be the integer we want to factor. Let $d \simeq \sqrt{n}$ and $b_1, \dots, b_d$ be very small integers compared to $N$: ie

$b_1, \dots, b_d$ satisfy $|b_i| \leq d^{O(1)}$

(In fact, Pilatte explains in his paper that this condition can be relaxed to $|b_i| \leq \exp(\tilde{O}(d))$ )

We define two lattices:

$$\mathcal{L} := \{ (e_1, \dots, e_d) \in \mathbb{Z}^d \mid \left( \prod_{i=1}^{d} b_i^{e_i} \right)^2 \equiv 1 \pmod N \}$$

ie $\prod_{i=1}^{d} b_i^{e_i}$ is a square root of 1 modulo $N$

$$\mathcal{L}_0 := \{ (e_1, \dots, e_d) \in \mathbb{Z}^d \mid \left( \prod_{i=1}^{d} b_i^{e_i} \right) \in \{ \pm 1 \pmod N \} \}$$

ie $\prod_{i=1}^{d} b_i^{e_i}$ is a "trivial" square root of 1 mod $N$.

We saw that in Shor's algorithm, a key-step was to find $a \in (\mathbb{Z}/N\mathbb{Z})^\times$ of order $r$ such that $a^{r/2} \notin \{ \pm 1 \pmod N \}$.

In other words, $a^{r/2}$ is a non-trivial square root of 1 modulo $N$.

So here, we are looking for $(e_1, \dots, e_d) \in \mathcal{L} \setminus \mathcal{L}_0$ : indeed, this will give $b := \prod_{i=1}^{d} b_i^{e_i}$ a non-trivial square root of 1

and so as in Shor's algorithm, $\gcd(b-1, N)$ and $\gcd(b+1, N)$ will give two non-trivial divisors of $N$.

Now, Regev explains that at least heuristically, we can find a vector in $\mathcal{L} \setminus \mathcal{L}_0$ that is short in a precise sense.

Indeed: • first, $\mathcal{L}$ contains short vectors: consider the set

$$A_d = \{-2^{\frac{n}{d}-1}, \dots, 2^{\frac{n}{d}-1}\}^d$$ : its cardinality is $\left(2^{\frac{n}{d}}+1\right)^d$

which is greater that $\left(2^{\frac{n}{d}}\right)^d = 2^n > N$. Therefore, there exist at least two distinct vectors $e = (e_1, \dots, e_d)$ and $e' = (e'_1, \dots, e'_d)$ in $A_d$ such that

$$\left(\prod_{i=1}^{d} b_i^{e_i}\right)^2 \equiv \left(\prod_{i=1}^{d} b_i^{e'_i}\right)^2 \pmod N$$

(by the pigeon-hole principle).

Then $e - e' \in \mathcal{L}$, but $\|e-e'\|_2 \leq \|e\|_2 + \|e'\|_2$

$$\leq \sqrt{d}\left(\|e\|_\infty + \|e'\|_\infty\right)$$
$$\leq \sqrt{d}\left(2^{\frac{n}{d}-1} + 2^{\frac{n}{d}-1}\right)$$
$$= \sqrt{d}\, 2^{\frac{n}{d}}$$

• Second, we expect at least half of the vectors of $\mathcal{L}$ not to be in $\mathcal{L}_0$.

Indeed, think of the case $N = pq$ a product of two distinct odd primes.

Then by the CRT, $(\mathbb{Z}/N\mathbb{Z})^\times \simeq (\mathbb{Z}/p\mathbb{Z})^\times \times (\mathbb{Z}/q\mathbb{Z})^\times$

So there are four square roots of 1 in $(\mathbb{Z}/N\mathbb{Z})^\times$; corresponding to

$\underbrace{(1,1), (-1,-1)}_{\text{trivial roots}}$, $\underbrace{(1,-1), (-1,1)}_{\text{non-trivial roots}}$.

Assumption: being a "short" vector is independent of the fact that it belongs to $\mathcal{L}_0$ or not:

Then we expect that since at least half of the vectors of $\mathcal{L}$ do not belong to $\mathcal{L}_0$, it is the same for short vectors. So there must exist a vector in $\mathcal{L} \setminus \mathcal{L}_0$ of norm $\mathcal{O}(\sqrt{d}\, 2^{n/d})$.

This is the assumpt° that Regev makes:

Th 1.1 of his paper. Let $N$ be an $n$-bit number, and assume that for $d \asymp \sqrt{n}$ and $\mathcal{O}(\log n)$-bit numbers $b_1, \dots, b_d$, there exists $e \in \mathcal{L} \setminus \mathcal{L}_0$ such that $\|e\|_2 \leq \exp(\mathcal{O}(\sqrt{n}))$. Then there exists a classical polynomial time algorithm that outputs a non-trivial factor of $N$ using $\sqrt{n} + 4$ calls to a quantum circuit of size $\mathcal{O}(n^{3/2} \log n)$.

We will not prove this difficult theorem, but let us make a few remarks

- The improvement is due to the modular exponentiation step, which is the main contribution to the size of the quantum circuit in Shor's algorithm (the $\tilde{\mathcal{O}}(n^2)$ comes from this step).

  This is where the assumption that the $b_i$'s are small is used, to reduce the number of gates to $\tilde{\mathcal{O}}(n^{3/2})$

- The drawback is that Regev's variant uses more space $\mathcal{O}(n^{3/2})$ qubits instead of $\mathcal{O}(n)$ qubits in optimized versions of Shor's algorithm.

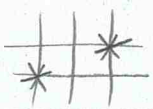
- Just like in Shor's algorithm, Th 1.1 relies on quantum Fourier transform, and the result of the last measurement does

not directly give a vector $(e_1, \dots, e_d)$ in $\mathcal{L} \setminus \mathcal{L}_0$. It gives an approximation of this vector.

↳ (here we hide a subtlety, in fact it returns an approximate of the vector in the dual lattice $\mathcal{L}^*$)

This is analogous to Shor's period-finding algorithm where the output is $\frac{y}{2^n}$ and we recover the period $r$ by saying that it appears as the denominator of one of the convergents of $y/2^n$.

Summary

| Shor's algorithm | Regev's variant |
|---|---|
| $a \in (\mathbb{Z}/N\mathbb{Z})^\times$ at random | $b_1, \dots, b_d$ small compared to $N$ |
| Find the smallest period of $\mathbb{N} \longrightarrow \mathbb{Z}/N\mathbb{Z}$, $k \longmapsto a^k \bmod N$  | Find a short vector $e_1, \dots, e_d$ such that $\left(b_1^{e_1} \cdots b_d^{e_d}\right)^2 \equiv 1 \pmod N$  $b_1^{e_1} \cdots b_d^{e_d} \not\equiv \pm 1 \pmod N$. |
| $a^{r/2}$ is a non-trivial square root of $1 \bmod N$ with high probability ⇝ deduce a non-trivial divisor of $N$ | $b_1^{e_1} \cdots b_d^{e_d}$ is a non-trivial square-root of 1 ⇝ // |

| Shor's algorithm | Regev's variant |
|---|---|
| Time: $\widetilde{O}(n^2)$ gates ($O(n^2 \log n \log\log n)$ in fact) | Time: $\widetilde{O}(n^{3/2})$ gates ($O(n^{3/2}\log n)$) |
| Space: $O(n)$ qubits | Space: $O(n^{3/2})$ qubits [Ragavan & Vaikuntanathan improved to $O(n\log n)$] |
| #calls: $O(1)$ | #calls: $O(\sqrt{n})$ |
| UNCONDITIONAL | CONDITIONAL ON THE ASSUMPTION THAT SHORT VECTORS IN $\mathcal{L}_{\|\cdot\|_\infty}$ EXIST. |

## 4.2. Analytic number theory techniques in Pilatte's proof of correctness.

Recall that the number theoretic assumption in Regev's algorithm is that for (sufficiently) small (compared to $N$) integers $b_1, \dots, b_d$, there exists a short vector $(e_1, \dots, e_d) \in \mathbb{Z}^d$ such that

$$\prod_{i=1}^{d} b_i^{e_i} \quad \text{is a non-trivial square root of } 1 \pmod N$$

A first observation made by Pilatte is that if the integers $b_i$ are fixed small integers, the result may be false due to what he refers to as the subgroup obstruction. Let us explain what this obstruction is. Assume that $N = p_1 p_2$ is a product of two distinct odd primes, both congruent to $3 \bmod 4$. For prime numbers $p$, denote by $n(p)$ the least quadratic non-residue modulo $p$:

that is: $n(p) = \min\left\{ k \in \{0, \dots, p-1\},\ (k \bmod p) \notin (\mathbb{F}_p^\times)^2 \right\}$

The best known unconditional bound on $n(p)$ is the Burgess bound:

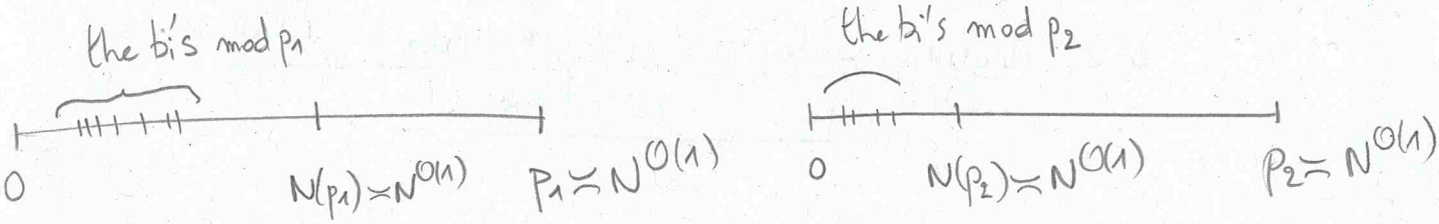
$$\boxed{n(p) \ll_\varepsilon p^{\frac{1}{4\sqrt{e}} + \varepsilon}}$$

The problem is that this bound does not prevent the $b_i$'s from being all squares modulo $N$! Indeed, if $p_1$ and $p_2$ have roughly the same size, then $p_1, p_2 \asymp \sqrt{N}$, so

$$n(p_1),\ n(p_2) \ll_\varepsilon \sqrt{N}^{\frac{1}{4\sqrt{e}}+\varepsilon}$$

So Burgess's bound does not exclude the possibility to have

$$n(p_1) \asymp n(p_2) \asymp N^{O(1)}$$




Now since Regev's algorithm relies on small integers $b_i$: namely $b_i$'s of $O(\log n)$ bits, that is $b_i \ll \exp(O(\log n))$

$$= n^{O(1)} = (\log N)^{O(1)}$$

we cannot rule out the possibility that all $b_i$'s are squares modulo $p_1$ and $p_2$. But then the subgroup of $(\mathbb{Z}/N\mathbb{Z})^\times$ generated by $b_1, \dots, b_d$ would be contained in the subgrap of squares:

$$\langle b_1, \dots, b_d \rangle \subset \{x^2,\ x \in (\mathbb{Z}/N\mathbb{Z})^\times\}$$

Now, the square roots of 1 in $\mathbb{Z}/N\mathbb{Z} \simeq \mathbb{Z}/p_1\mathbb{Z} \times \mathbb{Z}/p_2\mathbb{Z}$ are $(1,1)$, $(-1,1)$, $(1,-1)$ and $(-1,-1)$, but since $p_1, p_2 \equiv 3 \pmod 4$, $-1$ is not a square in $\mathbb{Z}/p_1\mathbb{Z}$ or $\mathbb{Z}/p_2\mathbb{Z}$ ( see ex sheet 1, it comes from the fact that $(-1)^{\frac{p_1-1}{2}} = -1$ if $p_1 \equiv 3 \pmod 4$ )

so the only one which belongs to $\langle b_1, \dots, b_d \rangle$ is $(1,1) = 1_{\mathbb{Z}/N\mathbb{Z}}$.

In other words, $\langle b_1, \dots, b_d \rangle$ only contains trivial square roots of one. So even if we can find a short vector $(e_1, \dots, e_d)$ in $\mathcal{L}$, it will always be in $\mathcal{L}_0$.

This is why Pilatte modifies Regev's assumption and the corresponding circuit, but still obtains (up to log factors) an algorithm of the same quality, working unconditionnally this time!

Rem: As pointed out by Pilatte, Burgess's bound can be improved under the Generalized Riemann Hypothesis (GRH). Under this hypothesis, Ankeny proved the much stronger bound $n(p) \ll (\log p)^2$. This bound suggests that the $b_i$'s can be chosen as small as needed for Regev's algorithm to work, while avoiding the subgroup obstruction.

To prove an unconditional result, Pilatte overcomes the subgroup obstruction by chosing an $x$ uniformly in $(\mathbb{Z}/N\mathbb{Z})^\times$ : this is not a small $b_i$ as in Regev's assumption, but it turns out that having a bounded number of large $b_i$'s still allows to conclude. Then $\langle b_1, \dots, b_d, x \rangle$ contains a non-trivial square root of 1 with probability $\gg 1$.

This resolves the difficulty of finding a non-trivial square-root of 1, but the difficulty that remains is to find a short vector

$$(e_1, \dots, e_d, f) \in \mathbb{Z}^{d+1} \quad \text{such that}$$

$$\left( \prod_{i=1}^{d} b_i^{e_i} \right) x^f \equiv 1 \pmod{N}$$

In fact, Pilatte even proves a stronger result : under the suitable

assumptions, the lattice

$$\{(e_1, \dots, e_d, f) \in \mathbb{Z}^{d+1} \mid \left(\prod_{i=1}^{d} b_i^{e_i}\right) x^f \equiv 1 \pmod N\}$$

admits a basis of vectors of "small" (in a precise sense) vectors.

What are the techniques involved in the proof ?

* In a first step, Pilatte relates the problem of counting points in a lattice to estimating character sums.

* In a second step, he uses classic theorems of analytic number theory that relate character sums to zeros of their associated Dirichlet L-functions $L(s, \chi)$

* Finally, once the estimates on the number of lattice points are obtained, he uses geometry of numbers to relate the point counting in boxes $[-H, H]^d$ to the existence of a short basis.

The precise statement of his theorem 2.18 is

Th (Pilatte) Let $N > 2$ be an integer, let $d := \lceil \log N \rceil$ and $X = d^{10^3 d}$. Let $b_1, \dots, b_d$ be iid random variables, each uniformly distributed in the set of primes $\leq X$ not dividing $N$. Let $r \geq 0$ and let $x_1, \dots, x_r$ be arbitrary random variables taking values in $(\mathbb{Z}/N\mathbb{Z})^\times$.
Then with proba $1 - O(d^{-d})$, the lattice

$$\mathcal{L} = \{(e_1, \dots, e_d, f_1, \dots, f_r) \in \mathbb{Z}^{d+r} : \prod_{i=1}^{d} b_i^{e_i} \prod_{i=1}^{r} x_i^{f_i} \equiv 1 \pmod N\}$$

has a basis consisting of vectors of Euclidean norm $\ll e^{42(d+r)}$

Some elements of the proof:

* Step 1: point counting $\longrightarrow$ character sums:

Preliminary on characters: For an abelian group $G$, denote by $G^M$ the subset of $M$-th powers in $G$: $G^M := \{x^M,\ x \in G\}$

**Lem:**
If $G$ is a finite abelian group, there is an isomorphism

$$i: \widehat{G^M} \xrightarrow{\sim} \widehat{G}^M$$

Proof: There is a surjective homomorphism $j: G \longrightarrow G^M$, $g \longmapsto g^M$

It induces an homomorphism $i: \widehat{G^M} \longrightarrow \widehat{G}$, $\chi \longmapsto \chi \circ j$

It easily follows from the surjectivity of $j$ that $i$ is injective.
Moreover, it is well-known (Pontryagin duality) that the restriction map $\widehat{G} \longrightarrow \widehat{G^M}$ is surjective, so we can find for each $\chi \in \widehat{G^M}$ a character $\widetilde{\chi}$ of $G$ such that $\widetilde{\chi}_{|G^M} = \chi$.

Then for all $g \in G$, $\chi(g^M) = \widetilde{\chi}(g^M) = \widetilde{\chi}^M(g)$

$\chi(g^M) = i(\chi)(g)$

So $i(\chi) = \widetilde{\chi}^M$, which shows that the image of $i$ is contained in $\widehat{G}^M$. We conclude by equality of the cardinality

since $|\widehat{G}^M| = |G^M| = |\widehat{G^M}|$.

$\widehat{G} \simeq G$ $\qquad$ $G^M \simeq \widehat{G^M}$

□

Using the previous lemma, we can relate lattice point counting to character sums:

**Prop** (Pillate, Lemma 2.7)

Let $N \geq 1$. Denote by $G := (\mathbb{Z}/N\mathbb{Z})^\times$. For all $H \geq 1$, $\chi \in \widehat{G}$, $d \geq 1$, $b_1, \dots, b_d \in G$, define

$$F_{\chi, H}(b_1, \dots, b_d) := \prod_{i=1}^{d} \sum_{|h| \leq H} \chi^h(b_i)$$

Then for all $M \geq 1$,

$$\#\left\{(e_1, \dots, e_d) \in \mathbb{Z}^d \cap [-H, H]^d \,,\ \prod_{i=1}^{d} b_i^{Me_i} \equiv 1 \pmod{N}\right\}$$

$$= \frac{1}{|\widehat{G}^M|} \sum_{\chi \in \widehat{G}^M} F_{\chi, H}(b_1, \dots, b_d)$$

**Proof**: By orthogonality of characters in $G^M$,

$$\mathbb{1}_{\prod_{i=1}^{d} (b_i^M)^{e_i} \equiv 1 \pmod{N}} = \frac{1}{|\widehat{G^M}|} \sum_{\chi \in \widehat{G^M}} \underbrace{\prod_{i=1}^{d} \chi\left((b_i^M)^{e_i}\right)}$$

$$= \chi\left(\left(b_i^{e_i}\right)^M\right)$$

$$= i(\chi)\left(b_i^{e_i}\right)$$

so by the isomorphism $i$, we have

$$\mathbb{1}_{\prod_{i=1}^{d} b_i^{Me_i} \equiv 1 \pmod{N}} = \frac{1}{|\widehat{G}^M|} \sum_{\chi \in \widehat{G}^M} \prod_{i=1}^{d} \chi(b_i^{e_i})$$

Therefore,

$$\#\left\{(e_1,\dots,e_d)\in\mathbb{Z}^d\cap[-H,H]^d,\ \prod_{i=1}^{d} b_i^{e_i M}\equiv 1 \pmod N\right\}$$

$$=\sum_{(e_1,\dots,e_d)\in\mathbb{Z}^d\cap[-H,H]^d} \mathbb{1}_{\prod_{i=1}^{d} b_i^{Me_i}\equiv 1 \pmod N}$$

$$=\sum_{(e_1,\dots,e_d)\in\mathbb{Z}^d\cap[-H,H]^d} \frac{1}{|\hat{G}^M|}\sum_{\chi\in\hat{G}^M}\prod_{i=1}^{d}\chi(b_i^{e_i})$$

$$=\frac{1}{|\hat{G}^M|}\sum_{\chi\in\hat{G}^M}\underbrace{\sum_{(e_1,\dots,e_d)\in\mathbb{Z}^d\cap[-H,H]^d}\prod_{i=1}^{d}\chi^{e_i}(b_i)}$$

$$=\prod_{i=1}^{d}\sum_{|h|\le H}\chi^h(b_i)=F_{\chi,H}(b_1,\dots,b_d)$$

□

hence the conclusion

* <u>Step 2</u>: character sums → zeros of Dirichlet L-functions: