Quantum Computation

November 15, 2018

Contents

0 Introduction

asdasd

Exercise classes: Sat 3 Nov 11am MR4, Sat 24 Nov 11am MR4, early next term (tba).
Thursday 8 November lecture is moved to Saturday 10 November 11am (still MR4).

—Lecture 2—

1 1

Recall that we have an oracle U_f for $f : \mathbb{Z}_M \to \mathbb{Z}_{\mathbb{N}}$ periodic, with period r, $A = M/r$. We want to find r in $O(poly(m))$ time where $m = \log M$.

1.1 The quantum algorithm

Work on state space $\mathcal{H}_M \otimes \mathcal{N}$ with basis $\{|i\rangle|k\rangle\}_{i\in\mathbb{Z}_M, k\in\mathbb{Z}_N}$.
• Step 1. Make staet $\frac{1}{\sqrt{M}} \sum_{i=0}^{M-1} |i\rangle|0\rangle$.
• Step 2. Apply U_f to get $\frac{1}{\sqrt{M}} \sum_{i=0}^{M-1} |i\rangle|f(i)\rangle$.
• Step 3. Measure the 2nd register to get a result y. By Born rule, the first register collapses to all those i's (and only those) with $f(i)$ equal to the seen y, i.e. $i = x_0, x_0+r, ..., x_0+(A-1)r$, where $0 \le x_0 < r$ in 1st period has $f(m) = y$. Discard 2nd register to get $|per\rangle = \frac{1}{\sqrt{A}} \sum_{j=0}^{A-1} |x_0 + jr\rangle$.
Note: each of the r possible function values y occurs with same probability $1/r$, so $0 \le x_0 < r$ has been chosen uniformly at random.
If we now measure $|per\rangle$, we'd get a value $x_0 + jr$ for uniformly random j, i.e. random element (x_0^{th}) of a random period (j^{th}), i.e. random element of $\mathbb{Z}_m$, so we could get no information about r.
• Step 4. Apply quantum Fourier transform mod M (QFT) to $|per\rangle$. Recall the definition of QFT: $QFT : |x\rangle \to \sum_{y=0}^{M-1} \omega^{xy}|y\rangle$ for all $x \in \mathbb{Z}_M$ where $\omega = e^{2\pi i/M}$ is the Mth root of unity. The existing result is that QFT mod M can be implemented in $O(M^2)$ time.
Then we get

$$\begin{aligned} QFT|per\rangle &= \frac{1}{\sqrt{MA}} \sum_{j=0}^{A-1} \left(\sum_{y=0}^{M-1} \omega^{(x_0+jr)y}|y\rangle \right) \\ &= \frac{1}{\sqrt{MA}} \sum_{y=0}^{M-1} \omega^{x_0 y} \left[\sum_{j=0}^{A-1} \omega^{jry} \right] |y\rangle \ (*) \end{aligned}$$

where we group all the terms with the same $|y\rangle$ together. One good thing is that the sum inside the square bracket is a geometric series, with ratio $\alpha = \omega^{ry} = e^{2\pi iry/M} = (e^{2\pi i/A})^y$.
Hence term inside bracket $= A$ if $\alpha = 1$, i.e. $y = kA = k\frac{M}{r}$, $k = 0, 1, ..., (r-1)$, and equals 0 otherwise when $\alpha \neq 1$. Now

$$QFT|per\rangle = \sqrt{\frac{A}{M}} \sum_{k=0}^{r-1} \omega^{x_0 k \frac{M}{r}} |k\frac{M}{r}\rangle$$

The random shift x_0 now appears only in phase, so measurement probabilities are now independent of x_0!

Measuring $QFT|per\rangle$ gives a value c, where $c = k_0\frac{M}{r}$ with $0 \leq k_0 \leq r-1$ chosen uniformly at random. Thus $\frac{k_0}{r} = \frac{c}{M}$, note that c, M are known, r is unknown (what we want), and k_0 is unknown but uniformly random.

So note that if we are lucky and get a k_0 that is coprime to r then we could just simplify $\frac{c}{M}$ to get r. Obviously we cannot be always lucky every time, but by theorem in number theory, the number of integers $< r$ coprime to r grows as $O(r/\log\log r)$ for large r, so we know probability of k_0 coprime to r is $O(\frac{1}{\log\log r})$.

Then by some probability calculation we know that $O(1/p)$ trials are enough to achieve $1-\varepsilon$ probability of success.

So afer Step 4, cancel c/M to the lowest terms a/b, giving r as denominator b (if k_0 is coprime to r). Check b value by computing $f(0)$ and $f(b)$, since $b = r$ iff $f(0) = f(b)$.

Repeating $K = O(\log\log r)$ times gives r with any desired probability.

Further insights into utility of QFT here:
Write $R = \{0, r, 2r, ..., (A-1)r\} \subseteq \mathbb{Z}_M$. $|R\rangle = \frac{1}{\sqrt{A}}\sum_{k=0}^{A-1}|kr\rangle$, and $|per\rangle = |x_0 + R\rangle = \frac{1}{\sqrt{A}}\sum_{k=0}^{A-1}|x_0 + br\rangle$ where x_0 is the random shift that caused problem previously.
For each $x_0 \in \mathbb{Z}_M$, consider mapping $k \to k + x_0$ (shift by x_0) on $\mathbb{Z}_M$, which is a 1-1 invertible map.

So linear map $U(x_0)$ on $\mathcal{H}_M$ defined by $U(x_0) : |k\rangle \to |k + x_0\rangle$ is unitary, and $|x_0 + R\rangle = U(x_0)|R\rangle$.

Since $(\mathbb{Z}_M, +)$ is abelian, $U(x_0)U(x_1) = U(x_0 + x_1) = U(x_1)U(x_0)$ i.e. all $U(x_0)$'s commute as operators on $\mathcal{H}_M$.
So we have orthonormal basis of common eigenvectors $|\chi_k\rangle\}_{k\in\mathbb{Z}_M}$, called *shift invariant states*.

$U(x_0)|\chi_k\rangle = \omega(x_0, k)|\chi_k\rangle$ for all $x_0, k \in \mathbb{Z}_M$ with $|\omega(x_0, k)| = 1$. Now consider $|R\rangle$ written in $|\chi\rangle$ basis,
$|R\rangle = \sum_{k=0}^{M-1} a_k|\chi_k\rangle$ where a_k's depending on r (not x_0).
Then $|per\rangle = U(x_0)|R\rangle = \sum_{k=0}^{M-1} a_k\omega(x_0, k)|\chi_k\rangle$, and measurement in the χ-basis has $prob(k) = |a_k\omega(x_0, k)|^2 = |a_k|^2$ which is independent of x_0, i.e. giving information about r!

—Lecture 3—

Recall last time we had $\mathcal{H}_M$: shift operations $U(x_0)|y\rangle = |y + x_0\rangle$ for $x_0, y \in$

$\mathbb{Z}_M$, which all permute, so have a common eigenbasis (shift invariant states) $\{|\chi_k\rangle\}_{k\in\mathbb{Z}_M}$, $U(x_0)|x_k\rangle = \omega(x_0,k)|\chi_k\rangle$.
Measurement of $|x_0+R\rangle = \frac{1}{\sqrt{A}}\sum_{l=0}^{A-1}|x_0+l_r\rangle = U(x_0)|R\rangle$ in $|\chi\rangle$ basis has output distribution independent of x_0, therefore gives information about r.

Introduce QFT as the unitary mapping that rotates χ-basis to standard basis, i.e. define $QFT|\chi_k\rangle = |k\rangle$. So QFT followed by measurement implements χ-basis measurement.

Explicit form of $|\chi_k\rangle$ eigenspaces (!): consider

$$|\chi_k\rangle = \frac{1}{\sqrt{M}}\sum_{l=0}^{M-1} e^{-2\pi ikl/M}|l\rangle$$

Then

$$\begin{aligned} U(x_0)|\chi_k\rangle &= \frac{1}{\sqrt{M}}\sum_{l=0}^{M-1} e^{-2\pi ikl/M}|l+x_0\rangle \\ &= \frac{1}{\sqrt{M}}\sum_{\tilde{l}=0}^{M-1} e^{-2\pi ik(\tilde{l}-x_0)/M}|\tilde{l}\rangle \text{ where } \tilde{l} = l+x_0 \\ &= e^{2\pi ikx_0/M}\cdot|\chi_k\rangle \end{aligned}$$

i.e. these are the shift invariant staets, eigenvalues $\omega(x_0,k) = e^{2\pi ikx_0/M}$.

Matrix of QFT: So

$$[QFT^{-1}]_{lk} = \frac{1}{\sqrt{M}}e^{-2\pi ilk/M}$$

(componets of $|\chi_k\rangle = QFT^{-1}|k\rangle$ as k^{th} column). So

$$[QFT]_{kl} = \frac{1}{\sqrt{M}}e^{2\pi ilk/M}$$

as expected.

2 The hidden subgroup problem (HSP)

Let G be a finite group of size $|G|$. Given (oracle for) function $f : G \to X$ (X is some set), and promise that there is a subgroup $K < G$ such that f is constant on (left) cosets of K in G, and f is distinct on distinct cosets.
The problem: determine the *hidden subgroup* K (e.g. output a set of generators, or sample uniformly from K).
We want to solve in time $O(poly(\log |G|))$ (an efficient algorithm) with any constant probability $1 - \varepsilon$.

Examples of problems that can be cast(?) as HSPs:
(i) periodicity: $f : \mathbb{Z}_M \to X$, periodic with period r. Let $G = (\mathbb{Z}_m, +)$, the hidden subgroup is $K = \{0, r, 2r, ...\} < G$, cosets $x_0 + K = \{x_0, x_0 + r, x_0 + 2r, ...\}$. The period r is generator of K.
(ii) discrete logarithm: for prime p, $\mathbb{Z}_p^* = \{1, 2, ..., p-1\}$ with multiplication mod p. $g \in \mathbb{Z}_p^*$ is a generator (or primitive root mod p). If powers generate all of $\mathbb{Z}_p^*$, $\mathbb{Z}_p^* = \{g^0 = 1, g^1, ..., g^{p-2}\}$, then also $g^{p-1} \equiv 1 \pmod p$ (easy number theory).
Fact: the generator always exists if p is prime. So any $x \in \mathbb{Z}_p^*$ can be written $x = g^y$ for some $y \in \mathbb{Z}_{p-1}$, write $y = \log_g x$ called the discrete log of x to base g.

Discrete log problem: given a generator g and $x \in \mathbb{Z}_p^*$, compute $y = \log_g x$ (classically hard).
To express as HSP, consider $f : \mathbb{Z}_{p-1} \times \mathbb{Z}_{p-1} \to \mathbb{Z}_p^*$: $f(a, b) = g^a x^{-b} \ mod\ p = g^{a-yb} \ mod\ p$.
Then check: $f(a_1, b_1) = f(a_2, b_2)$ iff $(a_2, b_2) = (a_1, b_1) + \lambda(y, 1)$ where $\lambda \in \mathbb{Z}_{p-1}$.

So if $G = \mathbb{Z}_{p-1} \times \mathbb{Z}_{p-1}$, $K = \{\lambda(y, 1) : \lambda \in \mathbb{Z}_{p-1}\} < G$. Then f is constant and distinct on the cosets of K in G, and generator $(y, 1)$ gives $y = \log_g x$.

(iii) graph problems (G non-abelian now): consider undirected graph $A = \{V, E\}$, $|V| = n$, with at most one edge between any two vertices. Label vertices by $[n] = \{1, 2, ..., n\}$.
Introduce the permutation group $\mathcal{P}_n$ of $[n]$. Define $Aut(A)$ to be the group of automorphisms of A, which is a subgroup of $\mathcal{P}_n$, containing exactly the permutations $\pi \in \mathcal{P}_n$ such that for all $i, j \in [n]$, $(i, j) \in E \iff (\pi(i), \pi(j)) \in E$, i.e. the labelled graph $\pi(A)$ obtained by permuting labels of A by π is the same *labelled* graph as A.

Associated HSP: Take $G = \mathcal{P}_n$. Let X be set of all labelled graphs on n vertices. Given A, consider $f_A : \mathcal{P}_n \to X$ by $f_A(\pi) = \pi(A)$, A with labels permuted by π. The associated hiiden subroup is $Aut(A) = K$.

Application: if we can sample uniformly from this K, then we can solve graph isomorphism problem (GI): two labelled graphs A, B are isomorphic if there is 1-1 map $\pi : [n] \to [n]$ such that for all $i, j \in [n]$, i, j is an edge in A iff $\pi(i), \pi(j)$ is an edge in B, i.e. A and B are the same graph but just labelled differently.

—Lecture 4—

Let's come back to the graph isomorphism problem.

Problem: given A, B, decide if $A \cong B$ or not. This can be expressed as anon-abelian HSP (on example sheet), no known classical polynomial time algorithm. However it is in NP, but it is not believed to be NP-complete.
Recent result (2017): a quasi-poly time classical algorithm (L.Babai).

Quantum algorithm for finite *abelian* HSP:
Write group $(G, +)$ additively.

Construction of shift invariant states and FT for G:
Let's introduce some representation theory for abelian group G. Consider mapping $\chi : G \to \mathbb{C}^* = (\mathbb{C} \setminus \{0\}, \cdot)$ satisfying $\chi(g_1 + g_2) = \chi(g_1)\chi(g_2)$, i.e. χ is a group homomorphism. Such χ's are called *irreducible* representations of G.
We have the following properties (without proof), which we'll call Theorem A later when we refer to it:
(i) any value $\chi(g)$ is a $|G|^{th}$ root of unity (so χ: $G \to S^1$ = unit circle in $\mathbb{C}$);
(ii) (Schur's lemma, orthogonality): If χ_i and χ_j are representations, then $\sum_{g\in G} \chi_i(g)\bar{\chi}_j(g) = \delta_{ij}|G|$;
(iii) there are always exactly $|G|$ different representations χ (well, this is a special case of general representation theory).

By (iii), we can label χ's as χ_g for $g \in G$. For example, $\chi(g) = 1$ for all $g \in G$ is always an irreducible representation (the trivial representation), labelled χ_0;
Then by orthogonality (ii) for any $\chi \neq \chi_0$ gives $\sum_{g\in G} \chi(g) = 0$.

Shift invariant states: in space $\mathcal{H}_{|G|}$ with basis $\{|g\rangle\}_{g\in G}$, introduce *shift operators* $U(k)$ for $k \in G$ defined by $U(k) : |g\rangle \to |g + k\rangle$. Clearly these all commute, so there is simultaneous eigenbasis:
For each χ_k, $k \in G$, consider state $|\chi_k\rangle = \frac{1}{\sqrt{|G|}} \sum_{g\in G} \bar{\chi}_k(g)|g\rangle$. Then theorem A(ii) implies these form orthonormal basis, and $U(g)|\chi_k\rangle = \chi_k(g)|\chi_k\rangle$.

Proof.

$$U(g)|\chi_k\rangle = \frac{1}{\sqrt{|G|}} \sum_{h\in G} \bar{\chi_k}(h)|h + g\rangle$$
$$\overset{h'=h+g}{=} \frac{1}{\sqrt{|G}} \sum_{h'\in G} \chi_k(\bar{h'} - g)|h'\rangle$$

This implies that

$$\chi_k * -g) = (\chi_k(g))^{-1} = \bar{\chi_k}(g),$$
$$\chi_k(\bar{h'} - g) = \bar{\chi_k}(h')\chi_k(\bar{-g}) = \chi_k(h')\chi_k(g)$$

So

$$U(g)|\chi_k\rangle = \frac{1}{\sqrt{|G|}} = \sum_{h'\in G} \chi_k(g)\bar{\chi}_k(h')|h'\rangle = \chi_k(g)|\chi_k\rangle$$

□

So $|\chi_k\rangle$'s are common eigenspaces, called *shift-invariant states.*
Introduce (define) Fourier transform QFT for group G as the unitary that

$QFT|\chi_g\rangle = |g\rangle$ for all $g \in G$.
In $|g\rangle-$basis matrices, k^{th} column of (QFT^{-1}) =components of $|\chi_k\rangle$, i.e. $\frac{1}{\sqrt{|G|}}\bar{\chi_k}(g) = [QFT^{-1}]_{gk}$.
So $[QFT]^{\dagger}_{kg} = \frac{1}{\sqrt{|G|}}\chi_k(g)$, and so $QFT|g\rangle = \frac{1}{\sqrt{|G|}}\sum_{k\in G}\chi_k(g)|k\rangle$.

Example. $G = \mathbb{Z}_M$. Check $\chi_a(b) = e^{2\pi iab/M}$, $a, b \in \mathbb{Z}_M$ is a representation. Similarly, for $G = \mathbb{Z}_{M_1} \times ... \times \mathbb{Z}_{M_r}$, $(a_1, ..., a_r) = g_1, (b_1, ..., b_r) = g_2$ where $g_1, g_2 \in G$,

$$\chi_{g_1}(g_2) \stackrel{def}{=} e^{2\pi i\left(\frac{a_1b_1}{M_1}+...+\frac{a_rb_r}{M_r}\right)}$$

is a representation of G. And we get

$$QFT_G = QFT_{M_1} \otimes ... \otimes QFT_{M_r}$$

on $\mathcal{H}_{|G|} = \mathcal{H}_{M_1} \otimes ... \otimes \mathcal{H}_{M_r}$.
This is exhaustive, since by classification theorem, every finite abelian group G is isomorphic to a direct product of the form $G \cong \mathbb{Z}_{M_1} \times ... \times \mathbb{Z}_{M_r}$. Furthermore, we can insist that M_i are prime powers $p_i^{s_i}$, where p_i are not necessarily distinct.

Quantum algorithm for finite abelian HSP:
Let $f : G \to X$, hidden subgroup $K < G$. We have cosets $K = 0 + K, g_2 + K, ..., g_m + K$, where $m = |G|/|K|$. State space as usual, with basis $\{|g\rangle, |x\rangle\}_{g\in G, x\in X}$.
• make the state $\frac{1}{\sqrt{|G|}}\sum_{g\in G}|g\rangle|0\rangle$;
• Apply oracle U_f, get $\frac{1}{\sqrt{|G|}}\sum_{g\in G}|g\rangle|f(g)\rangle$;
measure second register to see a value $f(g_0)$.
Then first register gives coset state (remember the function is constant on each coset). $|g_0 + K\rangle = \frac{1}{\sqrt{|K|}}\sum_{k\in K}|g_0 + K\rangle = U(g_0)|K\rangle$.
Apply QFT and measure to obtain result $g \in G$.

—Lecture 5—

Last time we disccused how to solve the abelian HSP problem. Now how does the output g related to K?
• the output distribution of g is independent of g_0, so same as that obtained from $QFT|K\rangle$ (i.e. $g_0 = 0$) since:
write $|K\rangle$ in shift invariant basis $|\chi_g\rangle$'s, $|K\rangle = \sum_g a_g|\chi_g\rangle$, then $|g_0 + K\rangle = U(g_0)|K\rangle = \sum a_g \underbrace{\chi_g(g_0)|\chi_g\rangle}_{=U(g_0)|\chi_g\rangle}$; but $QFT|\chi_g\rangle = |g\rangle$, so $Prob(g) = |a_g\chi_g(g_0)|^2 = |a_g|^2$ as $\chi_g(g_0)| = 1$.

Thus look at $QFT|K\rangle$. Recall $QFT|k\rangle = \frac{1}{\sqrt{|G|}}\sum_{l\in G}\chi_l(k)|l\rangle$, so $QFT|K\rangle = \frac{1}{\sqrt{|G|}}\frac{1}{\sqrt{|K|}}\sum_{l\in G}\left[\sum_{k\in K}\chi_l(k)\right]|l\rangle$.
The terms in [...] involves irreducible representation χ_l of G restricted to subgroup $K < G$, which is an irreducible representation of K. Hence

$$\sum_{k\in K}\chi_l(k) = \begin{cases} |K| & \chi_l \text{ restricts to trivial irreducible representation on } K \\ 0 & \text{otherwise} \end{cases}$$

and

$$QFT|K\rangle = \sqrt{\frac{|K|}{|G|}} \sum_{l\in G \text{ with } \chi_l \text{ reducing to trivial irreducible representation of } K} |l\rangle$$

So measurement gives a uniformly random choice of l such that $\chi_l(k) = 1$ for all $k \in K$.
e.g. If K has generators $k_1, k_2, ..., k_M$, $M = O(\log |K|) = O(\log |G|)$, then output has $\chi_l(k_i) = 1$ for all i.

It can be shown that if $O(\log |G|)$ such l's are chosen uniformly at random, then with probability $> 2/3$ they suffice to determine a generating set for K via equations $\chi_l(k) = 1$.
(see example sheet 1 for particular examples).

Example. If $G = \mathbb{Z}_{M_1} \times ... \times \mathbb{Z}_{M_q}$.
We had for $l = (l_1, ..., l_q)$, $g \in (b_1, ..., b_q) \in G$,

$$\chi_l(g) = e^{2\pi i(\frac{l_1 k_1}{M_1} + ... + \frac{l_q b_q}{M_q})}$$

So for $k = (k_1, ..., k_q)$, $\chi_l(k) = 1$ becomes

$$\frac{l_1 k_1}{M_1} + ... + \frac{l_q k_q}{M_q} \equiv 0 \pmod 1$$

(i.e. is an integer), a homogeneous linear equation on K, and $O(\log |K|)$ is independent such that equations determine K as null space.

Some remarks on HSP for non-abelian groups G (write multiplicatively):
As before, can easily generate coset states

$$|g_0 K\rangle = \frac{1}{\sqrt{|K|}} \sum_{k\in K} |g_0 K\rangle$$

where g_0's are randomly chosen. But problems arise with QFT construction, because now there's no basis of shift-invariant states exists! (this is since $U(g_0)$'s don't commute anymore, so no common full eigenbasis).

Construction of non-abelian Fourier Transform (some more representation theory):
• d-dimensional representation of G is a group homomorphism $\chi : G \to U(d)$ where $U(d)$ is the space of $d \times d$ unitary matrices acting on $\mathbb{C}^d$, by $\chi(g_1 g_2)\chi(g_1)\chi(g_2)$. (see part II representation theory for the general form)
• χ is irreducible representation if no subspace of $\mathbb{C}^d$ is left invariant under $\chi(g)$ for all $g \in G$ (i.e. cannot simultaneously block diagonalise all $\chi(g)$'s by a basis change).
• a complete set of irreducible representation: set $\chi_1, ..., \chi_m$ such that any irreducible representation is unitarily equivalent to one of them (equivalence $\chi \to \chi' = V\chi V^T$).

Theorem. (non-abelian version of theorem A – properties of representations)
If $d_1, ..., d_m$ are dimensions of a complete set of irreducible representations

$\chi_1, ..., \chi_m$, then:
(i) $d_1^2 + ... + d_m^2 = |G|$;
(ii) Write $\chi_i(g)_{jk}$ for the $(j,k)^{th}$ entry of matrix $\chi_i(g)$, where $j, k = 1, ..., d_i$.
Then (Schur orthogonality):

$$\sum_g \chi_i(g)_{jk} \bar{\chi}_{i'}(g)_{j'k'} = |G| \delta_{ii'} \delta j j' \delta k k'$$

Hence states

$$|\chi_{i,jk}\rangle \equiv \frac{1}{\sqrt{|G|}} \sum_{g \in G} \bar{\chi}_i(g)_{jk} |g\rangle$$

is an orthonomal basis.

• QFT on G *defined* to be the unitary that rotates $\{|\chi_{ijk}\rangle\}$ basis into standard basis $\{|g\rangle\}$. However, $|\chi_{ijk}\rangle$ are *not* shift invariant for all $U(g_0)$'s, and consequently measurement of coset state $|g_0 K\rangle$ in $|\chi\rangle$-basis gives an output distribution *not* independent of g_0.

However, *partial* shift invariance survives: Consider the incomplete measurement M_{rep} on $|g_0 K\rangle$ that distinguishes only the irreducible representations (i.e. i values) and not all (i, j, k)'s.
i.e. with measurement outcome i associated to d_i^2-dimensional orthogonal subspaces spanned by $\{|\chi_{(i),jk}\rangle\}_{j,k=1,...,d_i}$.
Then $\chi_i(g_1, g_2) = \chi_i(g_1)\chi_i(g_2)$ implies output distribution of i values is independent of g_0, giving direct, albeit imcomplete, information about K.
E.g. conjugate subgroups K and $= g_0 K g_0^{-1}$ for some $g_0 \in G$ give *same* output distribution.

—Lecture 6—
Non-abelian HSP/FT remarks:
For efficient HSP algorithm, we also need QFT to be efficiently implementable, i.e. $poly(\log |G|)$-time.
This is true for any abelian G and some non-ablien G's (such as $\mathcal{P}_n$), but even in latter case there's no known efficient HSP algorithm.

Some known result:
for normal subgroups, i.e. $gK = Kg$ for all $g \in G$:

Theorem. (Hallgrer, Russell, Tashma, SIAM J.Comp 32 p916-934 (2003))
Suppose G has efficient QFT. Then if hidden subgroup K is normal, then there is an efficient HSP quantum algorithm.
(Construct coset state $|g_0 K\rangle$, perform M_{rep} on it.)
Repeat $O(\log |G|)$ times. Then K normal implies outputs suffice to determine K.

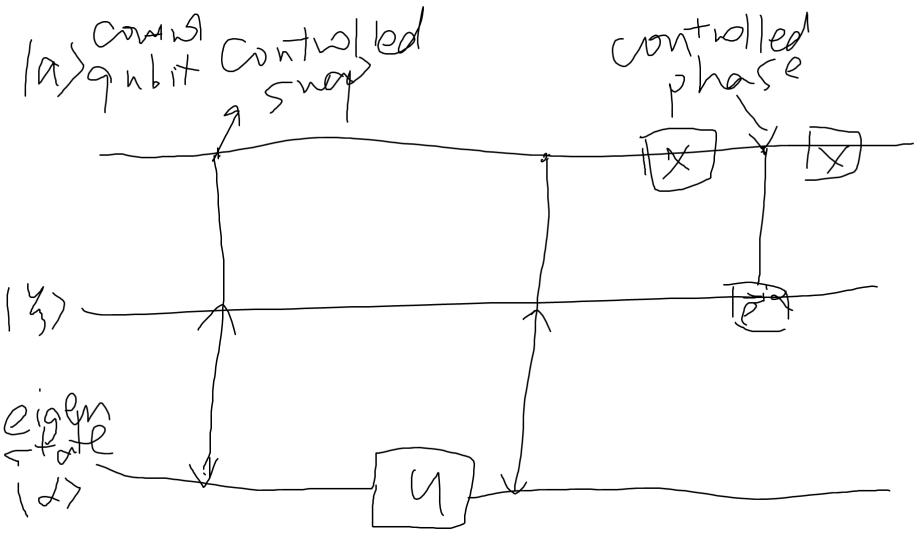
Theorem. (Ettinger, Hoyer, Knill)
For general non-abelian HSP, $M = O(poly(\log |G|))$ random coset states $|g_1 K\rangle$,...,$g_M K\rangle$ suffice to determine K from M coset states, but it's not efficient.
See example sheet for a proof – construct a measurement procedure on $|g_1 K\rangle \otimes ... \otimes g_M K\rangle$ to determine K, but it takes exponential time in $\log |G|$.

The phase estimation algorithm:
• a unifying principle for quantum algorithms, uses QFT_{2^n} again.
• many applications, e.g. an alternative efficient factoring algorithm (A.Kitaev).

Given unitary operator $\mathcal{U}$ and eigenstate $|v_\phi\rangle \cdot \mathcal{U}|v_\phi\rangle = e^{2\pi i\phi}|v_\phi\rangle$, we want to estimate phase ϕ, where $0 \leq \phi < 1$ (to some precision, say to n binary digits).

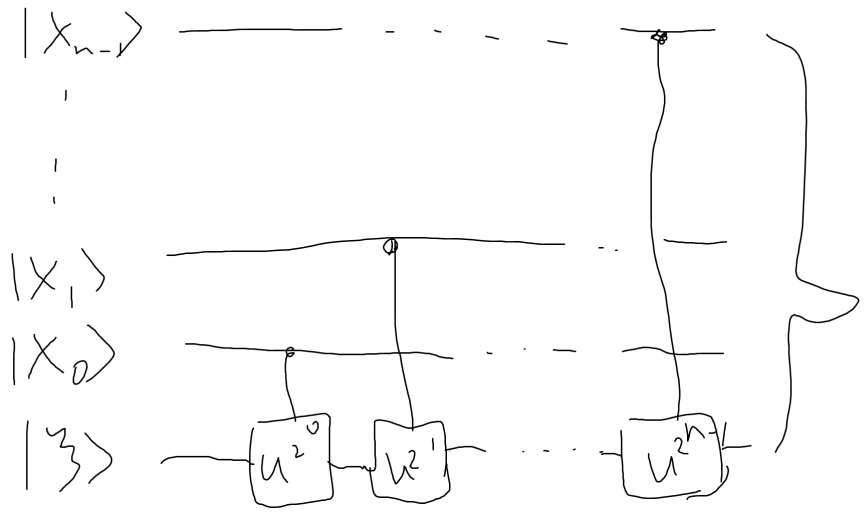
We'll need *controlled-U^k* for integers k, writte $C - U^k$, which satisfies $c - U^k|0\rangle|\xi\rangle = |0\rangle|\xi\rangle$, $C - U^k|1\rangle|\xi\rangle = |1\rangle U^k|\xi\rangle$, where $|\xi\rangle$ in general has dimension d.
Note $U^k|v_\phi\rangle = e^{2\pi ik\phi}|v_\phi\rangle$, $C - (U^k) = (C - U)^k$.

Remark. Given U as a formula or (arant?) description, we can readily implement $C - U$, e.g. just control each gate of U's circuit.
However, if U is given as a *black box*, we need further info:
• it suffices to have an eigenstate $|\alpha\rangle$ with known eigenvalue $U|\alpha\rangle = e^{i\alpha}|\alpha\rangle$:
We can consider



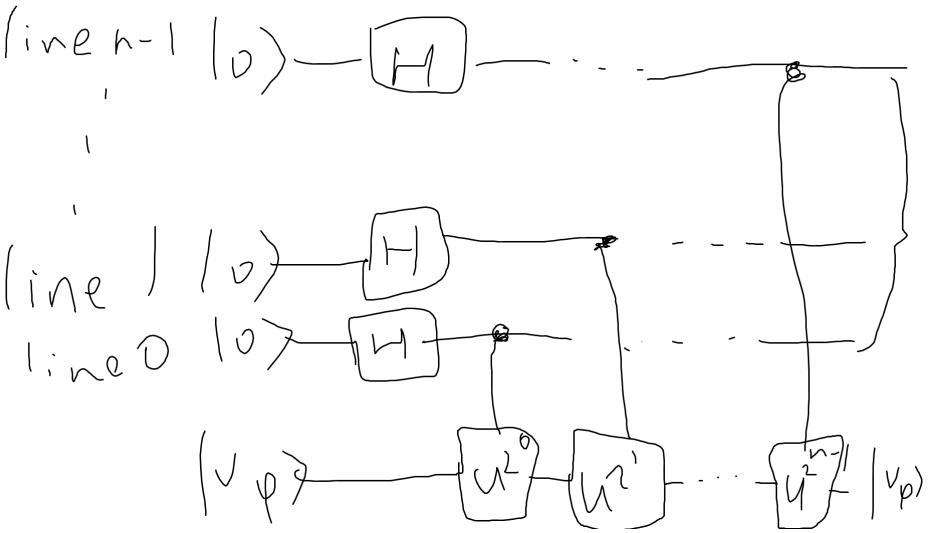
Where we get $CU|a\rangle|\xi\rangle$ at the first two row and the third row $|\alpha\rangle$ is always unchanged.
To see how it works, just check circuit action. (...)

We'll actually want *generalised controlled-U* with $|x\rangle|\xi\rangle \to |x\rangle U^x|\xi\rangle$, where $|x\rangle$ has n qubits, i.e. $x \in \mathbb{Z}_{2^n}$.
We can make this thing from $C - (U^k)$ as follows:



We get $|x\rangle U^x|\xi\rangle$, where $x = x_{n-1}...x_1x_0$ binary, $U^x = U^{2^{x_{n-1}}}...U^{2^{x_1}}U^{2^{x_0}}$.
Note: if input $|\xi\rangle = |v_\phi\rangle$, then get $e^{2\pi i\phi x}|v_\phi\rangle$.

Now suppose over all $x = 0, 1, ..., 2^{n-1}$ and use $|\xi\rangle = |v_\phi\rangle$,



Where the output is $\frac{1}{\sqrt{2^n}}\sum_x e^{2\pi i\phi x}|x\rangle$, we call this state $|A\rangle$.

Finally apply $QFT_{2^n}^{-1}$ to $|A\rangle$ and measure to see $y_0, ..., y_{n-1}$ on lines $0, 1, ..., n-1$.
Then output $0.y_0...y_{n-1} = \frac{y_0}{2} + ... + \frac{y_{n-1}}{2^{n-1}}$, as the estimate of ϕ.
That's the phase estimation algorithm (for given U and $V_\phi\rangle$).

Suppose ϕ actually had only n binary digits, i.e. ϕ exactly equals $0.z_0z_1...z_{n-1}$ for some $z_k = 0, 1$ for all k.

Then $\phi = \frac{z_0 \ldots z_{n-1}}{2^n} = \frac{z}{2^n}$ where z is n-bit integer in $\mathbb{Z}_{2^n}$, and

$$|A\rangle = \frac{1}{\sqrt{2^n}} \sum_x e^{2\pi i x z/2^n} |x\rangle$$

is QFT_{2^n} of $|z\rangle$.
So $QFT^{-1}|A\rangle = |z\rangle$ and get ϕ exactly, with certainty.
In this case the algorithm up to (not including) final measurements is a unitary operation, mapping $|0\rangle ... |0\rangle |v_\phi\rangle \to |z_0\rangle ... |z_{n-1}\rangle |v_\phi\rangle$.

—Lecture 7— Phase Estimation (continued):

U is a $d \times d$ unitary operation/matrix with eigenstate $U|v_\phi\rangle = 2^{2\pi i\phi}|v_\phi\rangle$, and we want to estimate ϕ.
U as a quantum physical operation is equivalent to $\tilde{U} = e^{i\alpha}U$ for any α and $\tilde{U}$ has $\phi \to \phi + \alpha/2\pi$.
So if U given as quantum physical operation alone, we cannot determine ϕ.
But controlled versions different: $C-U$ and $C-\tilde{U}$ are different as physical operations (set $\{e^{i\alpha}C-U\}_\alpha \neq \{e^{i\alpha}C-\tilde{U}\}_\alpha$), and $C-U/\tilde{U}$ *does* fix ϕ associatied to choice of phase α.
So quantum phase estimation algorithm use $C-U$ ($C-U^{2^k}$) physical operations (not just U's).

We had $\underbrace{|0\rangle ... |0\rangle}_{n} |v_\phi\rangle \overset{\text{unitary}}{\underset{C-U's}{\to}} |A\rangle = \frac{1}{\sqrt{2^n}} \sum_{x=0}^{2^n-1} e^{2\pi i\phi x}|x\rangle$ (n qubits).
Apply QFT^{-1} we get $QFT^{-1}|A\rangle$, measure to see $y_0, ..., y_{n-1}$; output $\phi = \frac{(y_0 y_1 \ldots y_{n-1})}{2^n}$, $0 \leq y < 2^{n-1}$, where the numerator is a n-bit integer.
If $\phi = \frac{z}{2^n}$ for integer $0 \leq z < 2^n$, i.e. ϕ has exactly n binary digits, then $|A\rangle = QFT|z\rangle$, so we get z with certainty in the measurement.

Now suppose ϕ has *more* than n bits, say $\phi = 0.z_0 z_1 z_2 ... z_{n-1}|z_n z_{n+1} ....$ Then we have:

Theorem. (PE) If measurement in above algorithm give $y_0, ...y_{n-1}$ (so output is $\theta = 0.y_0 ... y_{n-1}$), then
(a) $\mathbb{P}(\theta$ is closet n binary digit approximate to $\phi) \geq 4\pi^2$;
(b) $\mathbb{P}(|\theta - \phi| \geq \varepsilon)$ is at most $P(\frac{1}{2^n \varepsilon})$ (we'll show it's at most $\frac{1}{2^{n+1}\varepsilon}$).

Remark. In (a), we have probability $\frac{4}{\pi^2}$ that all n lines of n-line QPE process are *good*.
But, if we want ϕ accurate to m bits with probability $1-\eta$, then we use theorem (PE) (b) with $\varepsilon = 1/2^m$. Then we'll use $n > m$ lines with

$$\frac{1}{2^{n+1}}\varepsilon = \eta, \varepsilon = \frac{1}{2^m}$$

i.e. $n = m + \log(1/\eta) + 1$. In words, number of lines needed is only number of bits wanted with good probability $1-\eta$ plus a modest polynomial increase for exponetial reduction in η.

Proof. We have

$$QFT^{-1}|x\rangle = \frac{1}{\sqrt{2^n}} \sum_{y=0}^{2^n-1} e^{-2\pi i y x/2^n} |y\rangle$$

So

$$QFT^{-1}|A\rangle = \frac{1}{2^n} \sum_y \left[\sum_x e^{2\pi i(\phi - y/2^n)x} \right] |y\rangle$$

So for measurement,

$$\mathbb{P}(\text{see } n- \text{ bit integer } y = y_0 y_1 ... y_{n-1}) = \frac{1}{2^{2n}} \left| \sum_{x=0}^{2^n-1} e^{2\pi i \underbrace{(\phi - \frac{y}{2^n})}_{:=\delta(y)} x} \right|^2$$

Note that this is a geometric series $e^{2\pi i \delta(y)}$, so

$$\mathbb{P}(\text{see } y) = \frac{1}{2^{2n}} \left| \frac{1 - e^{2^n 2\pi i \delta(y)}}{1 - e^{2\pi i \delta(y)}} \right|^2$$

Let's call this equation (P) (maybe for *phase*).
We want to bound/estimate this expression.
For (a): Let $y = a = a_0 a_1 ... a_{n-1}$ give *closest* n-bit approximation to ϕ, i.e. $|\phi - \frac{a}{2^n}| \leq \frac{1}{2^{n+1}}$, i.e. $\delta(a) \leq \frac{1}{2^{n+1}}$.
Now we bounds:
(i) $|1 - e^{i\alpha}| = |2 \sin \frac{\alpha}{2}| \geq \frac{2}{\pi}|\alpha|$ if $|\alpha| < \pi$;
(ii) $|1 - e^{2\pi i \beta}| \leq 2\pi\beta$.

In equation (P), use (i) with $\alpha = 2^n \cdot 2\pi\delta(a) \leq 2^n 2\pi \frac{1}{2^{n+1}} \leq \pi$ to lower bound top line, and (ii) with $\beta = \delta(a)$ to upper bound bottom line, get

$$\mathbb{P}(\text{see } a) \geq \frac{1}{2^{2n}} \left(\frac{2^{n+1}\delta(a)}{2\pi\delta(a)} \right)^2 = \frac{4}{\pi^2}$$

For (b), we want to upper bound equaiton (P): for top line, $|1 - e^{i\alpha}| \leq 2$ for any α; for bottom, use (i) get $|1 - e^{2\pi i \delta(y)}| \geq 4\delta(y)$. So

$$\mathbb{P}(y) \leq \frac{1}{2^{2n}} \left(\frac{2}{4\delta(y)} \right)^2 = \frac{1}{2^{2n+2}} \delta(y)^2$$

Now sum this for all $|\delta(y)| > \varepsilon$, $\delta(y)$ values spaced by $1/2^n$'s. Let δ_+ be first $\delta(y)$ (jumps?) with $\delta(y) \geq \varepsilon$, δ_- be that with $\delta(y) \leq -\varepsilon$. So $|\delta_+|, |\delta_-| \geq \varepsilon$.
Then if $|\delta(y)| \geq \varepsilon$, we have $\delta(y) = \delta_+ + \frac{k}{2^n}$, $k = 0, 1, ...$, or $= \delta_- - \frac{k}{2^n}$, $k = 0, 1, ...$
So $|\delta(y)| \geq \varepsilon + \frac{k}{2^n}$ with $k = 0, 1, 2, ...$ in each case.

So

$$\begin{aligned}\mathbb{P}(|\delta(y)| > \varepsilon) &\leq 2\sum_{k=0}^{\infty} \frac{1}{2^{2n+2}} \frac{1}{(\varepsilon + \frac{k}{2^n})^2} \\ &\leq \frac{1}{2}\int_0^{\infty} \frac{1}{(2^n\varepsilon + k)^2} dk \\ &= \int_{2^n\varepsilon}^{\infty} \frac{dk}{k^2} \\ &= \frac{1}{2^{n+1}\varepsilon}\end{aligned}$$

□

Further remarks on QPE algorithm:
(1) If $C-U^{2^k}$ is implemented as $(C-U)^{2^k}$, the QPE algorithm needs exponential time in n as we have $1 + 2 + ... + 2^{n-1} = 2^n - 1$ $(C - U)$ gates.
However, for some special U's, $C - U^{2^k}$ can be implemented in $poly(k)$ time, so we get a poly time QPE algorithm.
It can be used to provide alternative facoring (order finding) algorithm (due to A. Kitaev) using PE.

—Lecture 8—

First exercise class: Saturday 3 Nov 11am MR4.

(2) If instead of $|v_\phi\rangle$, use general input state $|\xi\rangle$:

$$|\xi\rangle = \sum_j c_j |v_{\phi_j}\rangle$$

$$U|v_{\phi_j}\rangle = e^{2\pi i\phi_j}|v_{\phi_j}\rangle$$

Then we get in QPE (before final measurement) a unitary process U_{PE} with (lecturer had *that*) effect

$$|0...0\rangle|\xi\rangle \xrightarrow{U_{PE}} \sum_j c_j|\phi_j\rangle|v_{\phi_j}\rangle$$

and final measurement will give a choice of ϕ_j's (or approximation) chosen with probabilities $|c_j|^2$.

Example. Implement $QFT_{\mathcal{Q}}$ for $\mathcal{Q}$ not a power of 2, with a quantum curcuit of $1-$ and $2-$ qubit gates of circuit size $O(poly(\log \mathcal{Q}))$ (Kitaev's method).

Remark. For $\mathcal{Q} = 2^m$, we have explicit known circuit of $O(m^2)$. H and C-phase gate to implement QFT_{2^m} exactly (cf part II QIC Notes).

For $QFT_{\mathcal{Q}}$: Introduce

$$|\eta_a\rangle = QFT_{\mathcal{Q}}|a\rangle = \frac{1}{\sqrt{\mathcal{Q}}}\sum_{b=0}^{\mathcal{Q}-1} \omega^{ab}|b\rangle, a \in \mathbb{Z}_{\mathcal{Q}}, \omega = e^{2\pi i/\mathcal{Q}}$$

It suffices to make circuit hat does $|a\rangle \to |\eta_a\rangle$ (*).
Let $2^{m-1} < \mathcal{Q} < 2^m$, and set $M = 2^m$, view $\mathcal{H}_\mathcal{Q}$ as subspace of m qubits (spanned by $|a\rangle : 0 \leq a < \mathcal{Q} - 1 < 2^m$).
To achieve (*), consider instead on $\mathcal{H}_\mathcal{Q} \otimes \mathcal{H}_\mathcal{Q}$

$$|a\rangle|0\rangle \xrightarrow{(1)} |a\rangle|\eta_a\rangle \xrightarrow{(2)} |0\rangle|\eta_a\rangle$$

(1): get $\eta_a\rangle$ from $|a\rangle$ while *remembering* $|a\rangle$;
(2): *erase/forget* $|a\rangle$.
For (1), first do $|0\rangle \to |\xi\rangle = \frac{1}{\sqrt{\mathcal{Q}}} \sum_{b=0}^{\mathcal{Q}-1} |b\rangle$ as follows:
on m qubits $\mathcal{H}^{\otimes m}$ gives $\frac{1}{\sqrt{M}} \sum_{x=0}^{2^m-1} |x\rangle$. Then consider the step function $f(x) = 0$ if $x < \mathcal{Q}$ and 1 if $x \geq \mathcal{Q}$. It's classically efficiently computable, so can efficiently implement U_f on $(m+1)$ qubits.
So applying U_ρ to $(H^{\otimes m}|0\rangle)|0\rangle$ and measure output $(m+1^{st})$ qubit to get $|\xi\rangle$ on first n qubits if measurement result is 0.
Note that $prob(0) > 1/2$ as $\mathcal{Q} > 2^{m-1} = 2^m/2$, so we can use multiple trials to give $|\xi\rangle$.
We can do offline: failures/re-tries do not affect state to which we want to apply $QFT_\mathcal{Q}$. So now we have $|\tilde{\xi} = |a\rangle \left(\frac{1}{\sqrt{\mathcal{Q}}} \sum_{b=0}^{\mathcal{Q}-1} |b\rangle\right)$.
Next consider $V|a\rangle|b\rangle = \omega^{ab}|a\rangle|b\rangle$.
Then $V|\tilde{\xi}\rangle = |a\rangle|\eta_a\rangle$ as we want for (1).

To implement V, consider

$$U : |b\rangle \to \omega^b|b\rangle$$

If $|b\rangle$ in m qubits given by $|b_{m-1}\rangle...|b_0\rangle$, i.e. $b = b_{m-1}...b_0$ in binary, then $\omega^b = \omega^{b_{m-1}2^{m-1}}...\omega^{b_0 2^0}$. So U is product of 1-qubit phase gates

$$P(\omega^{2^{m-1}}) \otimes ... \otimes \mathbb{P}(\omega^{2^0})$$

where $P(\xi) = Diag(1, \xi)$, $|\xi| = 1$ is a phase gate.

Similarly, for $C - U^{2^k}$ (starting with $U \to U^{2^k}$ i.e. $\omega^b \to \omega^{2^k b}$), and $V =$ *generalised* $C - U$:

$$|a\rangle|b\rangle \xrightarrow{V} |a\rangle U^a|b\rangle$$

which is constructed as before, from $C - U^{2^k}$'s.
So now we have $|a\rangle|0\rangle \xrightarrow{(1)} |a\rangle|\eta_a\rangle$.

For (2), i.e. $|a\rangle|\eta_a\rangle \xrightarrow{(2)} |0\rangle|\eta_a\rangle$, *if* we had U with eigenstates $|\eta_a\rangle$, eigenvalues $\omega^a = e^{2\pi i a/\mathcal{Q}}$, then U_{PE} would give

$$|0\rangle|\eta_a\rangle \xrightarrow{U_{PE}} |a\rangle|\eta_a\rangle$$

(we are a bit loose on how information is presented – writing eigenvalue output as a, and note we are assuming that PE works exactly)
Hence U_{PE}^{-1} (*inverse gates taken in reverse order*) would give desired (2)!

Consider $U : |x\rangle \to |x - 1 \ mod \ \mathcal{Q}\rangle$, and check that $U|\eta_a\rangle = \omega^a|\eta_a\rangle$ as wanted.
Now note $x \to x - k \ mod \ \mathcal{Q}$ for $k \in \mathbb{Z}_\mathcal{Q}$ is classically computable in $poly(\log \mathcal{Q})$-time, thus we also have $U^k : |x\rangle \to |x - k \ mod \ \mathcal{Q}\rangle$, and PE algotirhm with

$m = O(\log(Q))$ lines.
Then implementing (1) then (2) gives $poly(\log \mathcal{Q})$ sized circuit for $QFT_{\mathcal{Q}}$.

But PE is not exact. However, using more qubit lines ($O(\log 1/\varepsilon)$ lines), we can achieve (by theorem PE(b))

$$|0\rangle|\eta_a\rangle \xrightarrow{U_{PE}} (\sqrt{1-\varepsilon}|a\rangle + \sqrt{\varepsilon}|a^{\perp}\rangle)|\eta_a\rangle$$

(where $a^{\perp}$ is a state orthogonal to $|a\rangle$) for any (small) deserved ε. Then

$$|| \ |a\rangle - \sqrt{1-\varepsilon}|a\rangle + \sqrt{\varepsilon}|a^{\perp}\rangle \ || = O(\sqrt{\varepsilon})$$

So

$$|| \ U_{PE}^{-1}|a\rangle|\eta_a\rangle - |0\rangle|\eta_a\rangle \ || = O(\sqrt{\varepsilon})$$

(as unitaries preserve lengths). So we can approximate $QFT_{\mathcal{Q}}$ to any desired precision (omit details).

3 Amplitude Amplification

Note that this is a very good name – a fifth order literation (both starting with *Ampli*).
Apothesis of technique in Grover's algorithm.

Some background:
We'll make much use of *reflection operators*.

—Lecture 9—
A reminder that we don't have lecture next thursday.

Reflection operators:
• State $|\alpha\rangle$ in $\mathcal{H}_d \to$ 1-dimensional subspace L_α and $(d-1)$-dimensional orthogonal complement $L_\alpha^\perp$

$$I_{|\alpha\rangle} \stackrel{def}{=} I - 2|\alpha \times \alpha|$$

has $I_{|\alpha\rangle}|\alpha\rangle = -|\alpha\rangle$, $I_{|\alpha\rangle}|\beta\rangle = |\beta\rangle$ for any $|\beta\rangle \perp |\alpha\rangle$.
So $I_{|\alpha\rangle}$ is reflection in $(d-1)$-dimensional subspace $L_\alpha^\perp$.

Note that for any unitary U, $UI_{|\alpha\rangle}U^\dagger = I_{U|\alpha\rangle}$, since $U|\alpha \times \alpha|Y^\dagger = |\xi \times \xi|$ ofr $\xi = U|\alpha\rangle$ (basically a change of basis).

• Take k-dimensional subspace $A \subseteq \mathcal{H}_d$, and any orthonormal basis $|a_1\rangle, ..., |a_k\rangle$.
Then $P_A = \sum_{i=1}^k |a_i \times a_i|$ is projection operator into A.
Define $I_A = I - 2P_A$. Then we have $I_A|\xi\rangle = |\xi\rangle$ if $|\xi\rangle \in A^\perp$, and $I_A|\xi\rangle = -|\xi\rangle$ if $|\xi\rangle \in A$.
So I_A is reflection in $(d-k)$ dimensional mirror $A^\perp$.

Recap of Grover's algorithm (part II notes page 68-73):
• search for unique *good* item in unstructured database of $N = 2^n$ items formalised as: (write B_n to be the set of all n-bit strings, $N = 2^n$): Given oracle for $f : B_n \to B$, promised that there is unique $x_0 \in B_n$ with $f(x_0) = 1$, and we wish to find x_0.
This is closely related to class NP and Boolean satisfiability problem (see part II notes p 67-68).
Using one query to $(n+1)$-qubit $\mathcal{U}_f$, we can implement reflection operator $I_{|x_0\rangle} : |x\rangle \to |x\rangle$ if $x \neq x_0$, and to $-|x\rangle$ if $x = x_0$.
(viz. apply $\mathcal{U}_f$ to $|x\rangle(\frac{|0\rangle - |1\rangle}{\sqrt{2}})$ and discard the last qubit.)
Then consider *Grover iteration operator* on n qubits:

$$Q \stackrel{def}{=} -H_n I_{|0...0\rangle} H_n I_{|x_0\rangle} = -I_{|\psi_0\rangle} I_{|x_0\rangle}$$

here $H_n = H \otimes H \otimes ... \otimes H = H_n^\dagger$, and $|\psi_0\rangle = H^n|0...0\rangle = \frac{1}{\sqrt{2^n}} \sum_{x \in B_n} |x\rangle$.
So one application of Q uses 1 query to $\mathcal{U}_f$.

Theorem. (Grover, 1996)
In 2-dimensional span of $|\psi_0\rangle$ and (unknown) $|x_0\rangle$, the action of Q is rotation by angle 2α where $\sin\alpha = \frac{1}{\sqrt{N}}$.

Hence (Grover's algorithm) to find x_0 given U_f:
1. Make $|\psi_0\rangle$;
2. Apply Q m times where $m = \frac{\arccos(\frac{1}{\sqrt{N}})}{2\arctan(frac1\sqrt{N})}$ to rotate $|\psi_0\rangle$ very close to $|x_0\rangle$.
3. Measure to see x_0 with high probability $\sim 1 - \frac{1}{N}$.

For large N, $\arccos(\frac{1}{\sqrt{N}}) \approx \pi/2$, $\arcsin(\frac{1}{\sqrt{N}}) \approx \frac{1}{\sqrt{N}}$ so $m = \frac{\pi}{4}\sqrt{N}$ iterations/queries to U_f suffice.
Classically we need $O(N)$ queries to see x_0 with any *constant* probability (independent of N), so get *square-root* speed up quantumly.

Amplitude Amplification:
Let G be any subspace (*good subspace*) of state space $\mathcal{H}$, and $G^\perp$ is orthogonal complement (*bad subspace*) $\mathcal{J} = G \oplus G^\perp$.
Given any $|\psi\rangle \in \mathcal{H}$, we have unique decompoisiton with *real positive* coefficients

$$|\psi\rangle = \sin\theta|g\rangle + \cos\theta|b\rangle$$

where $|g\rangle \in G$, $|b\rangle \in G^\perp$ normalised. Introduce reflections: flip $|\psi\rangle$ and good vectors: $I_{|\psi\rangle} = I - 2|\psi\times\psi|$, $I_G = I - 2P_G$ (projection into G), so $\sin\theta = ||P_G|\psi\rangle||$ is the length of good projection.
Introduce $Q \stackrel{def}{=} -I_{|\psi\rangle}I_G$.

Theorem. (Amplitude Amplification)
In the 2-dimensional subspace spanned by $|g\rangle$ and $|\psi\rangle$ (or equivalently by orthonormal vectors $|g\rangle$ and $|b\rangle$), Q is rotation by 2θ where $\sin\theta$ is the length of good projection of $|\psi\rangle$.

Proof. We have $I_G|g\rangle = -|g\rangle$, $I_G|b\rangle = |b\rangle$. So $Q|g\rangle = +I_{|\psi\rangle}|g\rangle$, $Q|b\rangle = -I_{|\psi\rangle}|b\rangle$. Now

$$\begin{aligned} I_{|\psi\rangle} &= I - 2(\sin\theta|g\rangle + \cos\theta|b\rangle)(\sin\theta\langle g| + \cos\theta\langle b|) \\ &= I - 2[\sin^2\theta|g\times g| + \sin\theta\cos\theta|g\times b| + \sin\theta\cos\theta|b\times g| + \cos^2\theta|b\times b|]|b\rangle \end{aligned}$$

And direct calculation (using $\langle g|b\rangle = 0$, $\langle g|g\rangle = \langle b|b\rangle = 1$) gives

$$\begin{aligned} Q|b\rangle &= I_{|\psi\rangle}|b\rangle \\ &= 2\sin\theta\cos\theta|g\rangle - (1 - 2\cos^2\theta)|b\rangle \\ &= \cos 2\theta|b\rangle + \sin 2\theta|g\rangle \end{aligned}$$

and $Q|g\rangle = +I_{|\psi\rangle}|g\rangle = -\sin 2\theta|b\rangle + \cos 2\theta|g\rangle$.
So in $\{|b\rangle, |g\rangle\}$ basis, matrix of Q is exactly the matrix of rotation by 2θ. □

—Lecture 10—

Let's continue on Amplitude Amplification.

Last time we showed that $Q = -I_{|\psi\rangle}I_g$ is the rotation through 2θ in the plane of $|\psi\rangle$ and $|g\rangle$, i.e. in $|b\rangle$ and $|g\rangle$ (orthonormal).

So $Q^n|\psi\rangle = \sin(2n+1)\theta|g\rangle + \cos(2n+1)\theta|b\rangle$, and if we measure $Q^n|\psi\rangle$ for good vs bad, we get $prob(good) = \sin^2(2n+1)\theta$.
We want to maximize this: it is maximised when $(2n+1)\theta = \pi/2$, i.e. $n = \frac{\pi^2}{4\theta} - \frac{1}{2}$.

Example. *If* we had $\theta = 4/6$, then $n = \frac{\pi}{4\theta} - \frac{1}{2} = 1$ is an exact integer. So Q^1 rotates $|\psi\rangle$ *exactly* onto $|g\rangle$, so we see good result with certainty!

Generally, for given θ, n is not an integer. So we use n to be the nearest integer to $(\frac{\pi}{4\theta} - \frac{1}{2}) \approx \frac{\pi}{4\theta}$ (for small θ), which equals $O(\frac{1}{\theta} = O(\frac{1}{\sin\theta}) = O(\frac{1}{||good\ proj\ of\ |\psi||}))$, and $Q^n|\psi\rangle$ will be within angle $\pm\theta$ of $|g\rangle$, so probability of good result is at least $\cos^2\theta \approx 1 - O(\theta^2)$.

All this can be implemented oif $I_{|\psi\rangle}$ and I_G can be implemented. See example sheet – for I_G, suffices for G to be spanned by computational basis states $|x\rangle$'s, and indicator function $f(x) = 1$ for x good and 0 for x bad efficiently computable. For $I_{|\psi\rangle}$, usally have $|\psi\rangle = H_n|00...0\rangle$ (H is the Hadamard gate). Then $I_{|\psi\rangle}$ can be implemented in linear $O(n)$ time.

Notes:
(1) In AA process, relative amplitudes of good labels in $|g\rangle$ stay *same* as they were in $|\psi\rangle = \sin\theta|g\rangle + \cos\theta|b\rangle$.
(2) Final state is generally not exactly $|g\rangle$, but *if* $\sin\theta$ is *known*, then we can modify AA process to make it exact, i.e. giving $|g\rangle$ state exactly (see example sheet).

Applications of AA:
(1) *Grover Search* with one or *more* (k) good items in N:

$$\begin{aligned}|\psi\rangle = |\psi_0\rangle &= \frac{1}{\sqrt{2^n}} \sum_{x\in B_n} |x\rangle \\ &= \sqrt{\frac{k}{N}} \left(\frac{1}{\sqrt{k}} \sum_{good\ |g\rangle} |x\rangle \right) + \sqrt{\frac{N-k}{N}} \left(\frac{1}{\sqrt{N-k}} \sum_{x\ bad} |x\rangle \right)\end{aligned}$$

G spanned by good x's, $\sin\theta = \frac{k}{n}$ so Q is rotation through 2θ, $\theta = \arcsin\sqrt{k/N} \approx \sqrt{k/N}$, where $k \ll N$; and we only need $O(\sqrt{N/k})$ queries.

Note: for 2-bit case, $N = 4$ with $k = 1$ good item; we have $\theta = \arcsin(1/2) = \pi/6$, so one application of Q rotates $|\psi_0\rangle$ exactly onto $|x_{good}\rangle$, i.e. a *single* query suffices to find a unique good item in four, *with certainty*!

(2) Square-root sppedup of general quantum algorithms:
Let A be a quantum algorithm/circuit (sequence of unitary gates). on input, say $|0...0\rangle$. So final state is $A|0...0\rangle$.
Good labels = desired computational outcomes

$$A|0.000\rangle = \alpha|a\rangle + \beta|b\rangle, \alpha = \sin\theta$$

where $|a\rangle$ is normalised, genrally unequal superposition $\sum_{good\ x} c_x|x\rangle$.
So $Prob$(success in 1 run) $= |\alpha|^2$, so $O(\frac{1}{|\alpha|^2})$ repetitions of A needed to succeed with any *constant* high probability $1 - \varepsilon$.

Instead use AA: assumed we can check if answer is good or bad (e.g. factoring). So we can then implement $I_G : |x\rangle \to -|x\rangle$ if x is good, and $\to |x\rangle$ if x is bad. Consider $|\psi\rangle = A|0...0\rangle$ and $Q = -I_{A|0...0\rangle} I_G = -(AI_{|0...0\rangle}A^\dagger)I_G$. All parts are implementable (A is the algorithm, $A^\dagger$ is inverse gate in reverse order, and $I_{|0...0\rangle}$ see example sheet).
By AA theorem, Q is rotation through 2θ, where $\sin\theta = |\alpha|$. So after $n \approx \frac{\pi}{4\theta} = O(\frac{1}{\theta}) = O(\frac{1}{\sin\theta}) = O(\frac{1}{|\alpha|})$ (for small $|\alpha|$).
Repetitions $A|0...0\rangle$ will be rotate very near to $|g\rangle$, and final measurement will succeed with high probability.
Each application of $\mathcal{Q}$ needs one A and one $A^\dagger$; $A^\dagger$ is the *inverse gate in reverse order*, i.e. the time complexity is the same, i.e. $O(\frac{1}{|\alpha|})$ repetition of Q gives square root time speed up over direct method.

Also, if success probability of A (i.e. $|\alpha|^2$) is known, then *improved* modification of the AA process that is *exact* can be applied; we convert probabilistic algorithm A into *deterministic* one, giving a good outcome with certainty.

4 Quantum Counting

Given $f : B_n \to B$ a boolean function with an unknown number k good x's, we want ot *estimate* k (rather than just find some good x).

Recall that Grover operator Q_G for f is rotaion through 2θ in 2-dimensional space of $|\psi_0 = \frac{1}{\sqrt{2^n}} \sum_{x \in B_n} |x\rangle$ and its good projection $|g\rangle = \frac{1}{\sqrt{k}} \sum_{good\ x} |x\rangle$, with $\sin\theta = \sqrt{k/N} \approx \theta$ for $k \ll N$.

—Lecture 11—

(Did I miss a lecture?)

4.1 Hamiltonian Simulation

k-local Halmiltonians:

H on n qubits is a $2^n \times 2^n$ Hermitian matrix. We'll want to simulate $U = e^{-iHt}$ with a circuit of $poly(n, t)$ basic unitary gates, i.e. efficient simulation.
Not all H's can be efficently simulated, but some physically important classes can –

Definition. H is k-local (k is a fixed constant) on n qubits if $H = \sum_{j=1}^{m} H_j$ where each H_j is a Hermitian matrix acting on at most k qubits (not necessarily contiguous).
i.e. each $H_j = \tilde{H}_j \otimes I$ (on some k qubits, and identity on rest of qubits).

So $m \leq \binom{n}{i} = O(n^k) = poly(n)$ terms in H.

Example. (1) $H = X \otimes I \otimes I - 5Z \otimes I \otimes Y$ is 2-local on 3 qubits.
(2) Write $M_{(j)}$ to denote operator M acting on j^{th} qubit (and I on all others).
Physically important spin-spin interactions:
Ising model on $n \times n$ square lattice of qubits, $H = J \sum_{i,j=1}^{n-1} Z_{(i,j)} Z_{(i,j+1)} + Z_{(i,j)} Z_{(i+1,j)}$, i.e. all nearest neighbours on square.
Heisenberg model on a line: $H = \sum_{i=1}^{n-1} J_x X_{(i)} X_{(i+1)} + J_y Y_{(i)} Y_{(i+1)} + J_z Z_{(i)} Z_{(i+1)}$ (where J, J_x, J_y, J_z are all real constants). This is very relevatn on chemistry (in studying covalence bonds?)

Note: in general, $e^{-i \sum_j H_j t} \neq \prod_j (e^{-iH_j t})$ – if the H_j's don't commute.

But $e^{-iH_j t}$'s *are* local unitary gates (acting on k qubits each), and we'll simulate $U(t_0) = e^{-i \sum_j H_i t_0}$ in terms of these (for suitable t's), and we'll have a $poly(n, t_0)$-sized circuit too.

If we want to use some standard universal gate set (to further express the above gates), then use: *Solovay-Kitaev theorem*: Let U be a unitary operator on k (const) qubits, and S any universal set of quantum gates (S is the set that have property that if we look at all circuit and all finite sets, then it's dense in all

circuits?? Lecturer didn't write down).
Then U can be approximated to within ε using $O(\log^c(1/\varepsilon))$ gates from S with $c < 4$ (it's actually exponential in k, but here we consider k constant).
So all S-products of length $O(\log^c(1/\varepsilon))$ get within ε of any element of $U(k)$.

We'll also need a lemma about accumulation of errors (c.f. Example sheet).
We call this lemmma A: let $\{u_i\}, \{v_i\}$ be sets of m unitary operators, with $||u_i - v_i|| < \varepsilon$ for all $i = 1, 2, ..., m$, then

$$||u_m...u_1 - v_m...v_1|| \leq m\varepsilon$$

i.e. errors accumulate linearly.
Proof is an easy exercise – induction on m.

Warm up: the (easy) commuting case.

Proposition. (k-local Hamiltonioans with commuting terms)
$H = \sum_{j=1}^m H_j$ with H_j commuting – any local Hamiltonian with commuting terms.
Then for any t, e^{-iHt} can be approximated to within ε by a circuit of $O(m\ poly(\log \frac{m}{\varepsilon}))$ gates from any given universal set.
Note, as $m = O(n^k)$ this is $poly(n, \log \frac{1}{\varepsilon})$ too. Also, $\log(\frac{1}{\varepsilon})$ is the number of digits of precision in the approximation.

Proof. H_j's commute implies that $e^{-i\sum_j H_j t} = \prod_{j=1}^m (e^{-iH_j t})$. Then SK theorem implies, for each $e^{-iH_j t}$ can be approximated to within ε/m within $O(poly(\log \frac{m}{\varepsilon}))$ gates, so lemma A then implies that the full product $\prod_{j=1}^m$ is then approximated to within $m(\varepsilon/m) = \varepsilon$, with a total of $O(m\ poly(\log \frac{m}{\varepsilon}))$ gates (from the universal set). □

Now let's look at the full non-commuting case: for any matrix X, write $X + O(\varepsilon)$ for $X + E$ where $||E|| = O(\varepsilon)$.

Lemma. (B, Lie-Trotter product formula)
Let A, B be matrices with $||A|| \leq K$, $||B|| \leq k$ and $k < 1$ (small).
Then $e^{-iA}e^{-iB} = e^{-i(A+B)} + O(k^2)$.

Proof.

$$\begin{aligned} e^{-iA} &= I - iA + \sum_{k=2}^{\infty} \frac{(-iA)^k}{k!} \\ &= I - iA + (iA)^2 \sum_{k=0}^{\infty} \frac{(-iA)^k}{(k+2)!} \\ &= I - iA + O(k^2) \end{aligned}$$

since $||(iA)^2|| < k^2$, and the remainder term is at most 1. So

$$\begin{aligned} e^{-iA}e^{-iB} &= (I - iA + O(k^2))(I - iB + O(k^2)) \\ &= I - i(A + B) + O(k^2) \\ &= e^{-i(A+B)} + O(k^2) \end{aligned}$$

by applying the inverse of above □

Now apply this repeatedly to accumulate sums of $H_1, ..., H_m$ in exponent. Note thta if each $||H_i|| < k$, then $||H_1 + ... + H_l|| < lk$, we want this to be < 1 for all $l \leq m$.
So *for now*, we'll assume $||H_i|| < \frac{1}{m}$ to have Lie-Trotter for all stages. Also take $t = 1$ *for now.*
Then consider

$$\begin{aligned} e^{-iH_1}e^{-iH_2}...e^{-iH_m} &= \left[e^{-i(H_1+H_2)} + O(k^2)\right] e^{-iH^3}...e^{-iH_m} \\ &= e^{-i(H_1+H_2)}e^{-iH_3}...e^{-iH_m} + O(k^2) \\ &= ... \\ &= e^{-i(H_1+H_2+...+H_m)} + O(k^2) + O((2k)^2) + ... + O(((m-1)k)^2) \\ &= e^{-i(H_1+H_2+...+H_m)} + O(m^3k^2) \ (1) \end{aligned}$$

where in the second equality we used that $||AU|| = ||A||$ for any unitary U, and note that the sum of squares up to m is of order m^3.

—Lecture 12—

For general finite $||H_j||$'s and t values, $||H_jt|| < kt$ can be large, so interoduce N (large-ish, fix later), and note

$$\frac{H_jt}{N} \text{ has } \tilde{K} = ||\frac{H_jt}{N}|| < \frac{kt}{N} \ (*)$$

can be suitably small, i.e. divide t into (small) $\frac{1}{N}$ intervals,

$$U = e^{i(H_1+...+H_m)t} = [e^{i(\frac{H_1t}{N}+...+\frac{H_mt}{N})}]^N$$

We want final error for U to be $< \varepsilon$, so by lemma A, we want error for [...] to be $< \frac{\varepsilon}{N}$.
So by (1) and (*), $Cm^3\tilde{k}^2 < \frac{\varepsilon}{N}$, i.e. $Cm^3\frac{K^2t^2}{N^2} < \frac{\varepsilon}{N}$, i.e. $N > \frac{Cm^3K^2t^2}{\varepsilon}$ (2).
Then

$$||e^{-iH_1t/N}e^{-iH_2t/N}...e^{-iH_mt/N} - e^{-i\frac{(H_1+...+H_m)t}{N}}|| < \frac{\varepsilon}{N}$$

so by lemma A again,

$$||(e^{-iH_1t/N}...e^{-iH_mt/N})^N - e^{-i(H_1+...+H_m)t}|| < \varepsilon$$

(this is Nm gates if form $e^{-iH_jt/N}$, and N given by (2)). So circuit size is $O(\frac{m^4(Kt)^2}{\varepsilon})$.
Recall for n qubits and k-local Hamiltonians, $m = O(n^k)$, so circuit size is $O(n^{4k}\frac{(Kt)^2}{\varepsilon}) = O(n^{4k}, t^2, \frac{1}{\varepsilon})$.

We have circuit $\mathcal{C}$ of size $|\mathcal{C}| = O(m^4\frac{(Kt)^2}{\varepsilon})$ gates of form $e^{-i\tilde{H}_j/N}$ approximate to $O(t)$.
If we want to use standard universal set, lemma A gives each of these gates needs to be approximate to $O(\varepsilon/|\mathcal{C}|)$ to maintain an overall $O(\varepsilon)$ level approximation

to U.
So by SK theorem, need $O(\log^c(|\mathcal{C}|/\varepsilon))$ gates from the universal set for each, i.e. get (modest) multiplicative factor of $O(\log^c \frac{m^4(Kt)^2}{\varepsilon^2})$ in $|\mathcal{C}|$.

For fixed n, ε, and variable t, the quantum process e^{iHt} runs for time t; but our circuit simulation runs for time $O(t^2)$. By refining/improving Lie-Trotter formula, it can be shown that this can be improved to $O(t^{1+\delta})$ for any $\delta > 0$ (See example sheet 2 for an example).

Harrow-Hassidim-Lloyd (HHL) quantum algorithm for linear systems of equations:
We'll want to solve a linear system of equations $A\mathbf{x} = \mathbf{b}$, where $\mathbf{b}, \mathbf{x} \in \mathbb{C}^N$, and dimension N is potentially *very* large (we could set $N = 2^n$, where 2^n is the least power of 2 that is greater than N).
Rather than outputting the full solution $\mathbf{x}$ itself, which would take $O(N)$ time, we instead want to compute (suitable approximates to) the value of properties of the solution, such as quadratic expressions $\mathbf{x}^T M\mathbf{x}$ e.g. total weight of some subset of components.

Very large lineary systems are important in applications: data mining/machine learning on data sets of increasingly large size to discover pattern properties in data; in science/engineering , we have numerical solutions of PDEs, where discretisation techniques (finite element methods) lead to linear systems far larger than original problem description.

The best known classical techniques take $poly(N)$ time to solve such problems.

Important parameters (for both classical and quantum algorithms):
• the system size N;
• the desired approximate tolerance ε;
• the *condition number* κ of matrix A, defined as ratio of largest to smallest eigenvalue size $\kappa = |\frac{\lambda_{max}}{\lambda_{min}}|$;
It provides intrinsic scale of the linear transformation A, and is a measure of how close A is to being non-invertible.
If renormalize A to have $\lambda_{max} = 1$, then $|\lambda_{min}| = 1/\kappa$, and numerical computation of A^{-1} becomes less stable with increasing κ, needing more significant digits and correspondingly longer runtime.

Preliminary requirements for the HHL algorithm:
Aim: to compute ε-approximate to properties of the solution of an N-dimensional system $A\mathbf{x} = \mathbf{b}$ in time $poly(\log N) = poly(n)$. We'll take the property to be a quadratic expression $\mu = \mathbf{x}^T M\mathbf{x}$.

Immediate issue is how the defining ingredients viz A and $\mathbf{b}$ are actually given (as reading them componently would already have taken $O(poly(N))$ time). Thus we'll need a different presentation of the problem (cf below) that is still available in applications.

The $poly(\log N)$ run time will be achieved by using only $O(\log N)$ qubits and it will never be required to write down all of A or $\mathbf{b}$ or $\mathbf{x}$ in the course of the

algorithm (as that would be impossible given the complexity constraints).

5 Example Class 1

5.1 Question 1

Basic representation theory exercises.
For details see Part II Representation Theory.

5.2 Question 2

Let $G = (\mathbb{Z}_2^n, \oplus)$ where $\oplus$ is componentwise-addition. Subgroup K generated by $a_1...a_k$, $K = \{b_1a_1 \oplus ... \oplus b_ka_k : b_1, ..., b_k \in \mathbb{Z}_2\}$. Note that K has size 2^k if a_i's are LI, and so does any coset of K.
Then $f(x) = f(x \oplus a_i)$ for all a_i's $\equiv f$ constant on cosets of K. f is 2^k-to-1: a_k's all linearly independent and f different on different cosets.

Shift invariant states: For $\mathbb{Z}_2$ irreps are $\chi_a(x) = (-1)^{ax)}$, $a, x \in \mathbb{Z}_2$ (-1 is the 2nd root of unity). So irreps on $(\mathbb{Z}_2)^n$ are $\chi_a(x) = (-1)^{a_1x_1}...(-1)^{a_nx_n}$ where $a = a_1...a_n$, $x = x_1...x_n$ are in $\mathbb{Z}_2^n$.
We also introduce a dot product $a \cdot x = a_1x_1 \oplus ... \oplus a_nx_n \in \mathbb{Z}_2$ where $\oplus$ here is $+$ modulo 2.
So shift invariant states

$$|\chi\rangle = \frac{1}{\sqrt{|G|}} \sum_g \overline{\chi(g)}|g\rangle$$

So

$$|\chi_a\rangle = \frac{1}{\sqrt{2^n}} \sum_{b \in \mathbb{Z}_2^n} (-1)^{a \cdot b}|b\rangle$$

in n qubits. So

$$(QFT)_{alr(?)} = \frac{1}{\sqrt{2^n}}(-1)^{a \cdot b} = \frac{1}{\sqrt{2^n}}(-1)^{a_1b_1}...(-1)^{a_nb_n}$$

so $QFT = H \otimes H \otimes ... \otimes H$ where H is the Hadamard gate.
For second part, it's just calculation:
Probability that first string is LI is $1 - 2^{-m}$ (just exclude 0...0);
Probability that first 2 strings are LI given the first is is $1 - 2/2^m$ (i.e. as 1st string x_1 spans 2 strings, namely 0...0 and x_1).
...Probability that first j strings are LI given the first $j-1$ are is $1 - 2^{j-1}/2^m$.
So by Bayes rule, probability that all of them are LI is $(1 - 1/2^m)...(1 - 2^{m-2}/2^m)(1 - 2^{m-1}/2^m)$. Use the hint given we get that is at least $\frac{1}{2}(1 - 1/2) = 1/4$.

Stanrdard HSP algorithm:
1. query to f, get random coset state $|y \oplus K\rangle = \frac{1}{\sqrt{2^n}} \sum_{x \in K} |y \oplus x\rangle, y \in \mathbb{Z}_2^n$.
Apply $QFT = H^{\otimes n}$ and measure; our theory assures that output is then uniformly random $c \in (\mathbb{Z}_2)^n$ s.t. irrep χ_c of G restricted to K is *trivial irrep* of K, i.e $.\chi_c(a) = 1$ for all $a \in K$.

So $(-1)^{a \cdot c} = 1$ for all $a \in K$, i.e. $c \cdot a = 0 \pmod 2$ for all $a = K$, i.e. $c_1 a_1 \oplus ... \oplus c_n a_n = 0$ for $a = a_1 ... a_n$.
We know K viewed as subspace of $(\mathbb{Z}_2)^n$ (n dimensional vector space over field $\mathbb{Z}_2$) has dimension k. So $(n-k)$ LI c_i's with $c_1 \cdot a = 0$ suffice to determine K as null spaceof linear system of $(n-k)$ equations.
So run HSP algorithm $(n-k)$ times: by (b) we'll get $(n-k)$ LI $c's$ with probability at least $1/4$, and we can solve for elements of K.