From optimal measurement to efficient quantum algorithms for the hidden subgroup problem and beyond

Andrew Childs
with Dave Bacon
& Wim van Dam

Overall outline:

1. Hidden subgroup problem, optimal measurement for the dihedral HSP
2. General approach to optimal measurements for HSPs in semidirect product groups
3. Examples: dihedral, metacyclic, Heisenberg
4. ~~Generalized abelian hidden shift problem~~

Part 1

quant-ph/0501044
0504083
0507190

Outline:

- The HSP
- Applications
- Known algorithms
- Standard approach
- Fourier sampling
- HSP as state distinguishability
- Dihedral group
- Dihedral coset states; the subset sum problem
- Pretty good measurement
- Optimality of PGM
- Success probability of PGM
- Implementing the PGM
- Properties of subset sum

The hidden subgroup problem

Problem: Fix a group G (known) and a subgroup $H \leq G$ (unknown).
Given a black box function $f: G \to S$ that is

(i) constant on left cosets of H in G
(ii) distinct on different left cosets of H in G

Find (a generating set for) H.

An efficient algorithm has run time poly$(\log |G|)$.

Example: Simon's problem $G = \mathbb{Z}_2^n$. Fix a hidden bitmask $s \in G$.
Given ~~an unspecified function~~ a 2-to-1 function satisfying $f(x \oplus s) = f(x)$.
So $H = \{0, s\} \cong \mathbb{Z}_2$.

Even this very simple case is hard for classical computers: can prove that finding s requires exponentially many queries to f.

But there is an efficient quantum algorithm!

(Need a quantum black box for f: $|x\rangle \mapsto |x\rangle|f(x)\rangle$ (isometry)
$|x\rangle|y\rangle \mapsto |x\rangle|y \oplus f(x)\rangle$ (unitary))

Applications

- G abelian
  This can be used to solve factoring, discrete log, Pell's equation, etc.
  Can always be solved efficiently

- G dihedral
  This can be used to solve the poly(d) unique shortest vector problem [Regev 02]
  [It can be reduced to a certain average case subset sum problem [Regev 02]]
  There is an efficient quantum algorithm that produces data (classical) which info-theoretically determines the answer [Ettinger - Høyer 00]
  There is a quantum algorithm with run time $2^{O(\sqrt{\log |G|})}$ [Kuperberg 03]

- G symmetric
  This can be used to solve graph isomorphism
  No nontrivial algorithms

Standard approach

$$\frac{1}{\sqrt{|G|}} \sum_{g \in G} |g\rangle \longmapsto \frac{1}{\sqrt{|G|}} \sum_{g \in G} |g, f(g)\rangle \; ;$$ discard 2^{nd} register
(for one copy, this is WLOG. for >1 copy, ?)

then we get a <u>coset state</u>,

$$|gH\rangle := \frac{1}{\sqrt{|H|}} \sum_{h \in H} |gh\rangle$$ with $g \in G$ uniformly random (unknown)

equivalently, we have a <u>hidden subgroup state</u>,

$$\rho_H := \frac{1}{|G|} \sum_{g \in G} |gH\rangle\langle gH| .$$

Note that this is not the only way to query f. But it is natural, and all known algorithms use this approach.

Fourier sampling

The symmetry of these states tells us a lot about how to deal with them.

$$\rho_H = \frac{1}{|G|\cdot|H|} \sum_{g \in G} \sum_{h,h' \in H} |gh\rangle\langle gh'|$$

$$= \frac{1}{|G|\cdot|H|} \sum_{g \in G} \sum_{h,h' \in H} R(h^{-1}) |g\rangle\langle g| R(h')$$ where R is the right regular representation of G, $R(g_1)|g_2\rangle = |g_2 g_1^{-1}\rangle$

$$= \frac{1}{|G|\cdot|H|} \sum_{h,h' \in H} R(h^{-1}h')$$

$$= \frac{1}{|G|} \sum_{h \in H} R(h) .$$

Now the regular representation is block diagonalized by the Fourier transform

$$F_G = \sum_{g \in G} \sum_{\sigma \in \hat{G}} \sum_{\ell,m=1}^{d_\sigma} \sqrt{\frac{d_\sigma}{|G|}} \, \sigma(g)_{\ell m} \, |\sigma, \ell, m\rangle\langle g|$$

$$\rho_H = \frac{1}{|G|}\sum_{h\in H} R(h)$$

$$FT_G\,\rho_H\,FT_G^{\dagger} = \sum_{g,g'}\sum_{\sigma,\sigma'}\sum_{\substack{\ell,m\\ \ell',m'}}\sum_{h}\frac{\sqrt{d_\sigma d_{\sigma'}}}{|G|^2}\,\sigma(g)_{\ell m}\,\sigma'(g')^*_{\ell' m'}\,|\sigma,\ell,m\rangle\underbrace{\langle g|R(h)|g'\rangle}_{\langle gh|g'\rangle=\delta_{g',gh}}\langle\sigma',\ell',m'|$$

$$=\sum_{g}\sum_{\sigma,\sigma'}\sum_{\substack{\ell,m,\\ \ell',m',n'}}\sum_{h}\frac{\sqrt{d_\sigma d_{\sigma'}}}{|G|^2}\,\sigma(g)_{\ell m}\,\sigma'(g)^*_{\ell' n'}\,\sigma'(h)^*_{n'm'}\,|\sigma,\ell,m\rangle\langle\sigma',\ell',m'|$$

now use Schur orthogonality: $\frac{d_\sigma}{|G|}\sum_{g\in G}\sigma(g)_{\ell m}\,\sigma'(g)_{\ell' n'} = \delta_{\sigma,\sigma'}\,\delta_{\ell,\ell'}\,\delta_{m,n'}$

$$=\sum_{\sigma}\sum_{\ell,m,m'}\sum_{h}\frac{1}{|G|}\,\sigma(h)^*_{mm'}\,|\sigma,\ell,m\rangle\langle\sigma',\ell',m'|$$

$$=\frac{1}{|G|}\sum_{\sigma\in\hat{G}}|\sigma\rangle\langle\sigma|\otimes I_{d_\sigma}\otimes\sum_{h\in H}\sigma(h)^*$$

Comments:
- State is block diagonal, with blocks labeled by irreps $\sigma\in\hat{G}$. Can measure this WLOG. ("weak Fourier sampling") In general, not enough info here.
- Row state is maximally mixed. Discard it.
- Column state is basis-dependent. How to measure?

HSP as state distinguishability

We have a state identification problem: given ρ_H for some unknown H, determine H.
More generally, we can make $k = \mathrm{poly}(\log|G|)$ states $\rho_H^{\otimes k}$

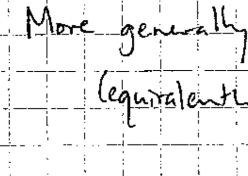
(equivalently, coset states $|g_1 H\rangle, |g_2 H\rangle, \ldots, |g_k H\rangle$ with each $g_i \in G$ independently uniformly random)

Good news: In principle, there is enough information in poly(log |G|) coset states to determine H, for any G [Ettinger–Høyer–Knill]

Bad news: Some groups require entangled measurements on $\Omega(\log|G|)$ coset states to determine H. [Moore–Russell–Schulman, Hallgren–Rötteler–Sen]

How can we identify measurements with nice properties that are likely to identify the states?
Idea: Try to understand the optimal measurement.
([Ip 03]: Shor implements the optimal measurement for the abelian HSP.)

Dihedral group



generic element: (a,b) $\quad a\in\mathbb{Z}_N,\ b\in\mathbb{Z}_2$ $\qquad G=\mathbb{Z}_N\rtimes\mathbb{Z}_2$

$$(a,b)(c,d) = (a+(-1)^b c,\ b+d)$$

Fact [Ettinger–Høyer]: To solve the HSP in the dihedral group, it is sufficient to be able to identify a hidden reflection, $H=\{(0,0),(a,1)\}$ (Proof later)

note $(a,1)^2 = (a-a,\ 1+1) = (0,0)$

Dihedral coset states and the subset sum problem

For any subgroup $\underbrace{\{(0,0),(a,1)\}}_{=:H_a}$, the elements $\{(a',0) : a' \in \mathbb{Z}_N\}$ form a complete set of coset reps.

Coset state: $|(a',0)H_a\rangle = \frac{1}{\sqrt{2}}(|a',0\rangle + |a'+a,1\rangle)$ $\quad a' \in \mathbb{Z}_N$ uniformly random

FT the 1st register over $\mathbb{Z}_N$: $\frac{1}{\sqrt{2N}} \sum_{x \in \mathbb{Z}_N} |x\rangle (\omega^{a'x}|0\rangle + \omega^{(a'+a)x}|1\rangle)$

now the mixed state is $\frac{1}{2N^2} \sum_{x,x',a'} \omega^{a'(x-x')} |x\rangle\langle x'| (|0\rangle + \omega^{ax}|1\rangle)(\langle 0| + \omega^{-ax'}\langle 1|)$

$$= \frac{1}{2N} \sum_x |x\rangle\langle x| (|0\rangle + \omega^{ax}|1\rangle)(\langle 0| + \omega^{-ax}\langle 1|)$$

which is block diagonal! (here x is basically the irrep name + row label)

so WLOG, we can measure x. discarding a global phase, we have

$\frac{1}{\sqrt{2}}(|0\rangle + \omega^{ax}|1\rangle)$ (note a' disappears, cf. abelian HSP)

goal: using k qubits like this, find a.

k copies: $\frac{1}{\sqrt{2}}(|0\rangle + \omega^{ax_1}|1\rangle) \otimes \cdots \otimes \frac{1}{\sqrt{2}}(|0\rangle + \omega^{ax_k}|1\rangle)$

$$= \frac{1}{\sqrt{2^k}} \sum_{b \in \mathbb{Z}_2^k} \omega^{a(b\cdot x)} |b\rangle$$

Now we would like to let $w = b\cdot x$ and do the sum over w instead of b.
Problem: given $x \in \mathbb{Z}_N^k$ and $w \in \mathbb{Z}_N$, what b's ($b \in \mathbb{Z}_2^k$) have $b \cdot x = w$?
This is the <u>subset sum problem</u>. (a classic NP-hard problem. but we have random x's, not worst case. more on this later.)

Solutions: $S_w^x := \{b \in \mathbb{Z}_2^k : b\cdot x = w\}$
$\eta_w^x := |S_w^x|$
and also define $|S_w^x\rangle := \frac{1}{\sqrt{\eta_w^x}} \sum_{b \in S_w^x} |b\rangle$ (or 0 if $\eta_w^x = 0$)

Then the k-copy state is

$$\frac{1}{\sqrt{2^k}} \sum_{w \in \mathbb{Z}_N} \omega^{aw} \sqrt{\eta_w^x}\, |S_w^x\rangle$$

Here we see that if we could replace $|S_w^x\rangle$ by $|w\rangle$, and if the η's were close to uniform, then a FT would reveal a. This is exactly what the optimal measurement does!

Pretty good measurement

We will find the optimal measurement by looking at a particular explicit POVM and showing it is optimal.

States ρ_j, ~~[illegible]~~ a priori probabilities p_j.

Define $\Sigma := \sum_j p_j \rho_j$. Then let $E_j := \frac{1}{\sqrt{\Sigma}} p_j \rho_j \frac{1}{\sqrt{\Sigma}}$. Clearly $\sum_j E_j = 1$.

This is ~~[illegible]~~ usually pretty good at distinguishing (E.g. HSW decoder.) Sometimes it is optimal!

Optimal measurement

Theorem [Holevo, Yuen-Kennedy-Lax] For an ensemble of states ρ_j with prob.'s p_j, and a POVM E_j, define $R := \sum p_j \rho_j E_j$. Then the POVM maximizes the probability of successfully identifying the state $(\sum_j p_j \operatorname{tr}(\rho_j E_j))$ iff

(i) $R = R^\dagger$
(ii) $R \geq p_j \rho_j \quad \forall j$

Note: the problem of finding the (or an) optimal POVM is a semidefinite program!

Optimality of the PGM for the DHSP (uniform ensemble)

States: $\rho_a = \frac{1}{2^k} \sum_{w,v \in \mathbb{Z}_N} \omega^{a(w-v)} \sqrt{\eta^x_w \eta^x_v}\, |S^x_w\rangle\langle S^x_v| \qquad x \in \mathbb{Z}_N^k \quad a \in \mathbb{Z}_N$

$$\Sigma = \frac{1}{2^k N} \sum_{w,v \in \mathbb{Z}_N} \underbrace{\sum_{a \in \mathbb{Z}_N} \omega^{a(w-v)}}_{N\delta_{w,v}} \sqrt{\eta^x_w \eta^x_v}\, |S^x_w\rangle\langle S^x_v|$$

$$= \frac{1}{2^k} \sum_{w \in \mathbb{Z}_N} \eta^x_w |S^x_w\rangle\langle S^x_w|$$

so the PGM is $E_a = \frac{1}{\sqrt{\Sigma}} \frac{\rho_a}{N} \frac{1}{\sqrt{\Sigma}} = \frac{1}{N} \sum_{w,v \in \mathbb{Z}_N} \omega^{a(w-v)} |S^x_w\rangle\langle S^x_v|$

$= |E_a\rangle\langle E_a|$ where $|E_a\rangle = \frac{1}{\sqrt{N}} \sum_w \omega^{aw} |S^x_w\rangle$

(Note: The ρ_a's are nonorthogonal pure states: $\rho_a = |\rho_a\rangle\langle\rho_a|$ with $|\rho_a\rangle = \frac{1}{\sqrt{2^k}} \sum_w \omega^{aw} \sqrt{\eta^x_w}\, |S^x_w\rangle$.

$\langle\rho_a|\rho_b\rangle = \frac{1}{2^k} \sum_{w,v} \omega^{bv-aw} \sqrt{\eta^x_w \eta^x_v} \langle S^x_w|S^x_v\rangle = \frac{1}{2^k} \sum_w \omega^{(b-a)w} \eta^x_w \neq \delta_{a,b}$ in general.

But $\langle E_a|E_b\rangle = \frac{1}{N} \sum_{w,v} \omega^{bv-aw} \langle S^x_w|S^x_v\rangle = \frac{1}{N} \sum_w \omega^{(b-a)w} = \delta_{a,b}$

so the PGM is an orthogonal measurement.)

Now $R = \frac{1}{N} \sum_{a \in \mathbb{Z}_N} \rho_a E_a = \frac{1}{N^2 2^k} \sum_{a,w,v,r} \omega^{a(w-v)+a(v-r)} \sqrt{\eta^x_w \eta^x_v}\, |S^x_w\rangle\langle S^x_r|$

$= \frac{1}{N 2^k} \sum_{w,v} \sqrt{\eta^x_w \eta^x_v}\, |S^x_w\rangle\langle S^x_w| = R^\dagger$

and check $R \geq \frac{1}{N}\rho_a \quad \forall a \in \mathbb{Z}_N$: clearly $R \geq 0$, and

$\langle\rho_a|R|\rho_a\rangle = \frac{1}{N(2^k)^2} \sum_{w,v} \eta^x_w \sqrt{\eta^x_w \eta^x_v}$

$\frac{1}{N}\langle\rho_a|\rho_a\rangle = \frac{1}{N 2^k} \sum_w \eta^x_w = \frac{1}{N}$

now using (by Cauchy-Schwartz) $\sum a_j \cdot \sum b_j \geq (\sum \sqrt{a_j b_j})^2$, we have

$\frac{1}{2^k} \sum_w (\eta^x_w)^{3/2} \; \frac{1}{2^k} \sum_v \sqrt{\eta^x_v} \geq \left(\frac{1}{2^k} \sum_w \eta^x_w\right)^2 = 1$

which proves $R \geq \frac{1}{N}\rho_a$.

Success probability of the dihedral PGM

$\Pr(\text{success}) = \operatorname{tr} E_a \rho_a$ (independent of a)

$= |\langle E_a|\rho_a\rangle|^2$

$= \frac{1}{2^k N}\left(\sum_w \sqrt{\eta_w^x}\right)^2$

This is for fixed x; averaging over (uniformly random) $x \in \mathbb{Z}_N^k$, we have

$$\Pr(\text{success}) = \frac{1}{2^k N^{k+1}} \sum_{x\in\mathbb{Z}_N^k} \left(\sum_{w\in\mathbb{Z}_N} \sqrt{\eta_w^x}\right)^2.$$

This will be big when the η's are spread out

Consider subset sum problem: let $k = \nu \log_2 N$, ν = "density"

$\nu < 1$: few numbers. most subsets have a distinct sum
$\nu > 1$: many numbers. most sums are achieved a comparable # of times.
η's uniform $\Rightarrow$ Pr(success) close to 1.

Later we will see how to prove this.

Implementing the dihedral PGM

We want to project onto a basis $|E_a\rangle = \frac{1}{\sqrt{N}} \sum_w \omega^{aw} |S_w^x\rangle$

I.e., we want to do $|E_a\rangle \mapsto |a\rangle$ followed by a standard measurement.
This can be done as follows:

$$|E_a\rangle \xrightarrow{|S_w^x\rangle \mapsto |w\rangle} \frac{1}{\sqrt{N}} \sum_w \omega^{aw} |w\rangle \qquad (\text{FT of } |a\rangle)$$

$$\xrightarrow{FT^{-1}} |a\rangle$$

The key step here is the inverse of ~~the~~ quantum sampling from solutions to subset sum:

$|w\rangle \mapsto |S_w^x\rangle$ (an isometry): replace $|w\rangle$ by uniform superposition of solutions to subset sum (x, w)

Hardness of subset sum

As previously mentioned, subset sum is NP-hard. But random instances at fixed ν are easier.

Low density: for $k < c\sqrt{\log N}$, $\exists$ an efficient algorithm [Lagarias, Odlyzko 85]
but this is no good for DHSP: it has $\nu = o(1)$

High density: for $k > 2^{c\sqrt{\log N}}$, $\exists$ a poly(k) algorithm [Flaxman, Przydatek 05]
this is exactly where Kuperberg's algorithm works! closely related.
but even this is not good enough for optimal measurement: can only find 1 solution, weaker than quantum sampling

Part ②

Outline:
- Semidirect product groups; known algorithms
- Reduction to cyclic subgroups
- Coset states and the matrix sum problem
- Pretty good measurement (which is optimal)
- PGM success probability: general lower and upper bounds
- Implementation by quantum sampling: approximate q. sampling is good enough.

Semidirect product groups

$G = A \rtimes_\varphi B$ (mnemonics: $A \trianglelefteq G$; B acts on A ("arms"))

set of elements $= A \times B$; write (a, b) $a \in A$, $b \in B$
$\varphi: B \to \mathrm{Aut}\, A$ a homomorphism
$(a,b)(a',b') = (a + \varphi(b)(a'), b+b')$
$(a,b)^{-1} = (\varphi(-b)(-a), -b)$

Specialize to $B = \mathbb{Z}_p$, p prime. Then $\varphi(1)$ determines φ: $\varphi(b) = \underbrace{\varphi(1)\circ\cdots\circ\varphi(1)}_{b}$
Denote $\varphi^b := \varphi(b)$. Here $\varphi: A \to A$ is an automorphism of A.

Cyclic subgroups $\langle (a,1)\rangle$:
$$(a,1)^2 = (a,1)(a,1) = (a+\varphi(a), 2)$$
$$(a,1)^3 = (a,1)(a+\varphi(a),2) = (a+\varphi(a)+\varphi^2(a), 3)$$
$$\vdots$$
$$(a,1)^b = (\Phi^{(b)}(a), b \bmod p) \qquad b \in \mathbb{N}$$
$$\text{where } \Phi^{(b)}(a) := \sum_{i=0}^{b-1} \varphi^i(a) \qquad (\Phi^{(b)}: A \to A)$$

Known algorithms

$\mathbb{Z}_2^n \wr \mathbb{Z}_2 = (\mathbb{Z}_2^n \times \mathbb{Z}_2^n) \rtimes \mathbb{Z}_2$	[Rötteler, Beth 98]
$\mathbb{Z}_{p^k}^n \rtimes \mathbb{Z}_2$ (p fixed)	[Friedl et al. 02]
$\mathbb{Z}_N \rtimes \mathbb{Z}_p$, $p = \phi(N)/\mathrm{poly}(\log N)$ prime	[Moore et al. 04]
$\mathbb{Z}_{p^k} \rtimes \mathbb{Z}_p$, $p>2$ prime	[Inui, Le Gall 04]

Reduction to cyclic subgroups

Lemma: To find an efficient algorithm for the HSP over $A \rtimes \mathbb{Z}_p$, it suffices to find an efficient algorithm for the HSP over $A_2 \rtimes \mathbb{Z}_p$ for any $A_2 \leq A$ with the promise that ~~[illegible]~~ $H = \langle (d,1)\rangle$ for some $d \in A_2$ with $|H| = p$.

Proof: Like Ettinger–Høyer for dihedral, with one additional possible complication.

Let $G_1 := A \times \{0\}$
$H_1 := H \cap G_1 =: A_1 \times \{0\}$
Since f restricted to G_1 hides H_1, and G_1 is abelian, we can efficiently find H_1.

Now is $H_1 \trianglelefteq G$? (If so, we'll factor it out.)
Let $g = (a,b) \in G$
$h_1 = (h, 0) \in H_1$
Then $g h_1 g^{-1} = (a,b)(h,0)(\varphi^{-b}(-a), -b) = (a+\varphi^b(h), b)(\varphi^{-b}(-a), -b) = (\varphi^b(h), 0)$
so $H_1 \trianglelefteq G$ iff $\varphi(H_1) = H_1$

We claim that this shows $H_1 \ntrianglelefteq G \Rightarrow H_1 = H$
i.e. $H_1 \neq H \Rightarrow H_1 \trianglelefteq G$.

If $H_1 \neq H$, then there is some $(d,1) \in H$
and by the previous calculation, $\forall h \in A_1$, $(d,1)(h,0)(d,1)^{-1} = (\varphi(h), 0) \in H$
so $\varphi(h) \in A_1$, i.e. $\varphi(H_1) = H_1 \Rightarrow H_1 \trianglelefteq G$

So we check whether $H_1 \trianglelefteq G$ (this can be done efficiently since G is solvable)
- if not, $H = H_1$ and we're done
- if so, we can factor it out

Let $G_2 := G/H_1 \cong A_2 \rtimes \mathbb{Z}_p$ where $A_2 := A/A_1$
$H_2 := H/H_1$

If $H = H_1$ then H_2 is trivial.

Otherwise, $H = \langle H_1, (a,1)\rangle$: we must have some $(a,1)$, and for any additional $(a',1)$,
$(a',1)(a,1)^{-1} = (a'-a, 0) \in H_1$, so $(a',1)$ is unnecessary.

so $H_2 = \langle (a,1)\rangle / \underbrace{(H_1 \cap \langle (a,1)\rangle)}_{\langle(\Phi^{(p)}(a), 0)\rangle} = \langle (d,1)\rangle$ for some $d \in A_2$ (clearly order p).

If we ~~can~~ identify d whenever H_2 is nontrivial, then this also handles the case where H_2 is trivial (just check). $\square$

Coset states

We're considering $G = A \rtimes \mathbb{Z}_p$
$H = \langle (a,1)\rangle$ of order p

can label left cosets by $(\ell, 0)$, $\ell \in A$: there are the right # of them, & they are distinct

coset state: $$|(\ell,0)H\rangle = \frac{1}{\sqrt{p}} \sum_{b \in \mathbb{Z}_p} |(\ell,0)(\Phi^{(b)}(a), b)\rangle = \frac{1}{\sqrt{p}} \sum_{b \in \mathbb{Z}_p} |\ell + \Phi^{(b)}(a), b\rangle$$

now FT the 1st register over A. $$\frac{1}{\sqrt{p|A|}} \sum_{x \in A} \sum_{b \in \mathbb{Z}_p} \chi_x(\ell + \Phi^{(b)}(a))\, |x, b\rangle$$

character of A: $\chi_x + \chi_{x'} = \chi_{x+x'}$; $\chi_x(y) = \chi_y(x)$
Ex: for $A = \mathbb{Z}_N$, $\chi_x(y) = \exp(2\pi i/N)$; for $A = \mathbb{Z}_p^r$, $\chi_x(y) = \exp(2\pi i (x\cdot y)/$

just as we argued for dihedral case,
the state mixed over $\ell \in A$ is block diagonal in x, so we can measure it. phase $\chi_x(\ell)$ disappears.

$$\rightarrow \frac{1}{\sqrt{p}} \sum_{b \in \mathbb{Z}_p} \chi_x(\Phi^{(b)}(a))\, |b\rangle$$

now one can show $\exists\, \hat{\Phi}^{(b)} : A \to A$ such that $\chi_x(\Phi^{(b)}(a)) = \chi_{\hat{\Phi}^{(b)}(x)}(a)$

(Ex: for $A = \mathbb{Z}_N$, $\varphi(a) = \mu a$ for $\mu \in \mathbb{Z}_N^\times$, $\hat{\Phi}^{(b)} = \Phi^{(b)}$
for $A = \mathbb{Z}_p^r$, $\varphi(a) = \mu a$ for $\mu \in GL_r(\mathbb{F}_p)$, $\hat{\Phi}^{(b)} = [\Phi^{(b)}]^T$)

so we have $$\frac{1}{\sqrt{p}} \sum_{b \in \mathbb{Z}_p} \chi_{\hat{\Phi}^{(b)}(x)}(a)\, |b\rangle$$

and for k copies, $$\frac{1}{\sqrt{p^k}} \sum_{b \in \mathbb{Z}_p^k} \chi_{\hat{\Phi}^{(b)}(x)}(a)\, |b\rangle$$
$\hookrightarrow \sum_i \hat{\Phi}^{(b_i)}(x_i)$

Matrix sum problem

$$S_w^x = \{ b \in \mathbb{Z}_p^k : \hat{\Phi}^{(b)}(x) = w \}$$ η_w^x, $|S_w^x\rangle$ as before

then the state is $$\frac{1}{\sqrt{p^k}} \sum_{w \in A} \chi_w(a) \sqrt{\eta_w^x}\, |S_w^x\rangle$$

PGM

The PGM calculations go through just as before. PGM is optimal, and is the projection onto the states $|E_a\rangle = \frac{1}{\sqrt{|A|}} \sum_{w \in A} \chi_w(a) |S_w^x\rangle$

Success probability

Again, by the same calculations as for the dihedral group,

$$\Pr(\text{success}) = \operatorname{tr} E_a \rho_a^{\otimes k} = \frac{1}{p^k |A|} \Big(\sum_w \sqrt{\eta_w^x} \Big)^2$$

and averaging over the uniformly random $x \in A^k$,

$$\Pr(\text{success}) = \frac{1}{p^k |A|^{k+1}} \sum_{x \in A^k} \Big(\sum_{w \in A} \sqrt{\eta_w^x} \Big)^2 = \frac{p}{|G|^{k+1}} \sum_{x \in A^k} \Big(\sum_{w \in A} \sqrt{\eta_w^x} \Big)^2$$

Lemma. (i) $\Pr(\text{success}) \leq p^k / |A|$
(ii) if $\Pr(\eta_w^x \geq \alpha) \geq \beta$, then $\Pr(\text{success}) \geq \alpha \beta^2 |A| / p^k$

Note: $\mathop{\mathbb{E}}_{x \in A^k, w \in A} \eta_w^x = \frac{1}{|A|^{k+1}} \sum_{x \in A^k} \sum_{w \in A} \eta_w^x = \frac{1}{|A|^{k+1}} \sum_{x \in A^k} p^k = \frac{p^k}{|A|}$, so $k = \log_p |A|$ is the expected critical value.

Proof: (i) $\Pr(\text{success}) \leq \frac{1}{p^k |A|^{k+1}} \sum_{x \in A^k} \Big(\sum_{w \in A} \eta_w^x \Big)^2 = \frac{1}{p^k |A|^{k+1}} \sum_{x \in A^k} p^{2k} = \frac{p^k}{|A|}$

(ii) using $\sum_{j=1}^N a_j^2 \geq \frac{1}{N} \Big(\sum_{j=1}^N a_j \Big)^2$,

$$\begin{aligned} \Pr(\text{success}) &\geq \frac{1}{p^k |A|^{2k+1}} \Big(\sum_{x \in A^k} \sum_{w \in A} \sqrt{\eta_w^x} \Big)^2 \\ &= \frac{|A|}{p^k} \Big(\underbrace{\frac{1}{|A|^{k+1}} \sum_x \sum_w \sqrt{\eta_w^x}}_{\geq \sqrt{\alpha} \Pr(\eta_w^x \geq \alpha)} \Big)^2 \\ &\geq \frac{|A|}{p^k} \alpha \beta^2 \end{aligned}$$

Implementation

Just as before, doing $|S_w^x\rangle \mapsto |w\rangle$ will implement the PGM. (Clear from form of $|E_a\rangle$)

In fact, it is good enough to do it approximately!

$$|w\rangle \mapsto \begin{cases} |S_w^x\rangle, & (x,w) \in Z_{\text{good}} \\ |\tilde{S}_w^x\rangle, & (x,w) \in Z_{\text{bad}} \end{cases} \quad \text{where} \quad \langle S_w^x | \tilde{S}_w^x \rangle = 0$$

(can be done if we can recognize bad instances)

then we have $\frac{1}{\sqrt{p^k}} \sum_{w \in A} \chi_w(a) \sqrt{\eta_w^x} |S_w^x\rangle$

$$\mapsto \frac{1}{\sqrt{p^k}} \sum_{w : (x,w) \in Z_{\text{good}}} \chi_w(a) \sqrt{\eta_w^x} |w\rangle + \frac{1}{\sqrt{p^k}} \sum_{w : (x,w) \in Z_{\text{bad}}} \chi_w(a) \sqrt{\eta_w^x} |v_w^x\rangle \quad \text{where} \quad \langle w | v_w^x \rangle = 0$$

fidelity with ideal state (averaged over x): $\frac{1}{(pN)^k} \sum_{(x,w) \in Z_{\text{good}}} \eta_w^x \geq \frac{|Z_{\text{good}}|}{p^k N^k}$ since $\eta_w^x > 1 \ \forall (x,w) \in Z_{\text{good}}$

and if a constant fraction of instances are good, $|Z_{\text{good}}| \geq \text{const.} \times N^{k+1}$, so fidelity $\geq \text{const.} \times \frac{N}{p^k}$.

Part ③: Examples

Dihedral group

Consider $G = A \rtimes_\varphi \mathbb{Z}_2$ with $\varphi(a) = -a$. In particular, $A = \mathbb{Z}_N$ is dihedral.

Here $\Phi^{(b)}(x) = \sum_{i=0}^{b-1} \varphi^i(x) = \begin{cases} 0 & b=0 \\ x & b=1 \end{cases} = b \cdot x$

so $S_w^x = \{ b \in \mathbb{Z}_2^k : \Phi^{(b)}(x) = w \} = \{ b \in \mathbb{Z}_2^k : b \cdot x = w \}$, the subset sum problem

can bound the success probability by calculating the mean & variance of η_w^x and using Chebyshev (see Regev)

Metacyclic groups

$G = \mathbb{Z}_N \rtimes_\varphi \mathbb{Z}_p$ with $\varphi(a) = \mu a$ for some $\mu \in \mathbb{Z}_N^\times$ with $\mu^p = 1 \bmod N$.

then $\Phi^{(b)}(x) = \sum_{i=0}^{b-1} \mu^i x$.

for simplicity, suppose $\mu - 1 \in \mathbb{Z}_N^\times$ (this is not necessary); then $\sum_{i=0}^{b-1} \mu^i = \frac{\mu^b - 1}{\mu - 1}$.

for $k=1$, we have $S_w^x = \{ b \in \mathbb{Z}_p : \frac{\mu^b - 1}{\mu - 1} x = w \}$.

$$(\mu^b - 1)x = (\mu - 1)w$$
$$\mu^b - 1 = (\mu - 1)w/x \quad \text{provided } x \in \mathbb{Z}_N^\times$$
$$\mu^b = 1 + (\mu - 1)w/x$$

(this is a discrete log problem!)

now for uniformly random $x \in \mathbb{Z}_N$, $\Pr(x \in \mathbb{Z}_N^\times) = \frac{\phi(N)}{N} = \Omega(1/\log\log N)$

~~so $\Pr(\eta_w^x \ge 1) \ge \Omega(1/\log\log N)$~~

and for uniformly random $w \in \mathbb{Z}_N$ and $x \in \mathbb{Z}_N^\times$, $\Pr(1 + (\mu-1)w/x = \mu^b) = p/N$

so $\Pr(\eta_w^x \ge 1) \ge p\,\phi(N)/N^2$, which is $1/\mathrm{poly}(\log N)$ provided $N/p = \mathrm{poly}(\log N)$.

(note this is exactly the condition from Moore et al.)

for $k=2$ (relevant to $N/p^2 = \mathrm{poly}(\log N)$), the matrix sum problem is

$$S_w^x = \{ b \in \mathbb{Z}_p^2 : \frac{\mu^{b_1} - 1}{\mu - 1} x_1 + \frac{\mu^{b_2} - 1}{\mu - 1} x_2 = w \}$$

$$\mu^{b_1} x_1 + \mu^{b_2} x_2 = x_1 + x_2 + (\mu - 1) w \qquad \text{how to solve?}$$

Stripped down version of the algorithm:

For hidden subgroup $H_a = \langle (a,1) \rangle$, the coset states are

$$|(\ell,0)H_a\rangle = \frac{1}{\sqrt{p}} \sum_{b \in \mathbb{Z}_p} |\ell + \Phi^{(b)}(a), b\rangle$$

Fourier transform 1^{st} register over $\mathbb{Z}_N$ $\quad (\omega := e^{2\pi i/N})$:

$$\frac{1}{\sqrt{Np}} \sum_{x \in \mathbb{Z}_N} \sum_{b \in \mathbb{Z}_p} \omega^{x(\ell + \Phi^{(b)}(a))} |x, b\rangle$$

measure x and postselect on $x \in \mathbb{Z}_N^\times$ $\quad$ (probability $\Omega(1/\log\log N)$)

compute $x\Phi^{(b)}(1)$:

$$\frac{1}{\sqrt{p}} \sum_{b \in \mathbb{Z}_p} \omega^{x\Phi^{(b)}(a)} |b, x\Phi^{(b)}(1)\rangle$$

note $\Phi^{(b)}$ can be computed efficiently since $\Phi^{(2b)} = (1+\mu^b)\Phi^{(b)}$

compute μ^b from $x\Phi^{(b)}(1) = x\frac{\mu^b - 1}{\mu - 1}$ $\quad$ (NB: x known)

use Shor to erase b (discrete log), then erase $x\Phi^{(b)}(1)$.

$$\mapsto \frac{1}{\sqrt{p}} \sum_{b \in \mathbb{Z}_p} \omega^{x\Phi^{(b)}(a)} |x\Phi^{(b)}(1)\rangle$$

now a Fourier transform gives a with probability $\geq p/N$.

Heisenberg group and friends

Heisenberg group:

(i) subgroup of $GL_3(\mathbb{F}_p)$: $\quad \left\{ \begin{pmatrix} 1 & 0 & 0 \\ b & 1 & 0 \\ a & c & 1 \end{pmatrix} : a,b,c \in \mathbb{F}_p \right\}$

(ii) group of $p \times p$ unitary matrices: $\langle X, Z \rangle = \{\omega^a X^b Z^c : a,b,c \in \mathbb{Z}_p\}$

$$X := \sum_{x \in \mathbb{Z}_p} |x+1\rangle\langle x| \qquad Z := \sum_{x \in \mathbb{Z}_p} \omega^x |x\rangle\langle x| \qquad \omega := e^{2\pi i/p}$$

(iii) semidirect product $\mathbb{Z}_p^2 \rtimes_\varphi \mathbb{Z}_p$ $\quad \varphi(c)(a,b) = (a+bc, b) = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}^c \begin{pmatrix} a \\ b \end{pmatrix}$

$$= \begin{pmatrix} 1 & c \\ 0 & 1 \end{pmatrix} \begin{pmatrix} a \\ b \end{pmatrix}$$

$$(a,b,c)(a',b',c') = (a+a'+b'c, b+b', c+c')$$

thus $\Phi^{(b)}\begin{pmatrix} x \\ y \end{pmatrix} = \sum_{i=0}^{b-1} \begin{pmatrix} 1 & i \\ 0 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} b & b(b-1)/2 \\ 0 & b \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix}$ $\quad$ MSP $\quad \sum_{j=1}^{k} \begin{pmatrix} b_j & b_j(b_j-1)/2 \\ 0 & b_j \end{pmatrix} \begin{pmatrix} x_j \\ y_j \end{pmatrix} = \begin{pmatrix} w \\ v \end{pmatrix}$

with $k=1$, probability of having a solution is $O(1/p)$ $\quad$ (p values of b, p^2 values of $\binom{w}{v}$)

but with $k=2$, probability is $\approx \frac{1}{2}$ that there are 2 solutions

$$\left. \begin{aligned} w &= b_1 x_1 + \tbinom{b_1}{2} y_1 + b_2 x_2 + \tbinom{b_2}{2} y_2 \\ v &= b_1 y_1 + b_2 y_2 \end{aligned} \right\} \Rightarrow \begin{aligned} b_1 &= [v y_1 + x_2 y_1 - x_1 y_2 \pm \sqrt{\Delta}] / y_1(y_1+y_2) \\ b_2 &= [v y_2 + x_1 y_2 - x_2 y_1 \mp \sqrt{\Delta}] / y_2(y_1+y_2) \end{aligned}$$

More generally: recall $\varphi \in \mathrm{Aut}\, A$ with $\varphi^p = 1$. consider $A = \mathbb{Z}_p^r$.
now $\mathrm{Aut}\, \mathbb{Z}_p^r \cong GL_r(\mathbb{F}_p)$. defined by $\mu \in GL_r(\mathbb{F}_p)$ put in Jordan canonical form
$\mu^p = 1 \Rightarrow$ diagonal elements are 1
so μ is a direct sum of matrices $\begin{pmatrix} 1 & 1 & & 0 \\ & \ddots & \ddots & \\ & & \ddots & 1 \\ 0 & & & 1 \end{pmatrix}$

$$\text{and } \Phi^{(b)} = \begin{pmatrix} \binom{b}{1} & \binom{b}{2} & \binom{b}{3} & \cdots & \binom{b}{r} \\ & \binom{b}{1} & \binom{b}{2} & \ddots & \vdots \\ & & \ddots & \ddots & \binom{b}{3} \\ & & & \binom{b}{1} & \binom{b}{2} \\ & 0 & & & \binom{b}{1} \end{pmatrix}$$

MSP: r^{th} degree equations, r of them, in k variables
so we need to take $k = r$. then we expect $O(1)$ solutions.
can find them efficiently for r constant! (Buchberger's algorithm)

to bound success probability, need to show variance of # solutions is small.

$$\mathop{\mathbb{E}}_{x \in A^k, w \in A} \eta_w^x = \frac{1}{p^{r(k+1)}} \sum_x \sum_w \eta_w^x = \frac{1}{p^{r(k+1)}} \sum_x p^k = \frac{p^{rk}}{p^{r(k+1)}} p^k = p^{k-r}.$$

$$\mathop{\mathbb{E}}_{x,w} (\eta_w^x)^2 = \frac{1}{p^{r(k+1)}} \sum_{x,w} (\eta_w^x)^2$$

system of polynomial eqns

$$= \frac{1}{p^{r(k+1)}} \sum_{x,w} \Big(\sum_b \delta\big[\Xi_w^x(b) = 0\big] \Big) \Big(\sum_c \delta\big[\Xi_w^x(c) = 0\big] \Big)$$

$$= \frac{1}{p^{r(k+1)}} \sum_{x,w} \Big(\sum_b \delta\big[\Xi_w^x(b) = 0\big] + \sum_{b \neq c} \delta\big[\Xi_w^x(b) = \Xi_w^x(c) = 0\big] \Big)$$

linear in x

$$= \mathbb{E}\,\eta_w^x + \frac{1}{p^{r(k+1)}} \sum_{b \neq c} \sum_x \delta\big[\Xi_0^x(b) = \Xi_0^x(c)\big] \cancel{\sum_w \delta\big[\Xi_w^x(b) = 0\big]}\ 1$$

components of each x

$$= p^{k-r} + \frac{1}{p^{r(k+1)}} \sum_{b \neq c} \underbrace{\sum_x \delta\big[\text{some fixed value for one of the } \Xi\big]}_{p^{rk-r} \text{ (r components fixed)}}$$

$$= p^{k-r} + p^{2(k-r)} - p^{k-2r}$$

variance: subtract $p^{2(k-r)}$ get $p^{k-r}(1 - p^{-r}) =: \sigma^2$

for $k = r$, we have mean 1
variance $\sigma^2 = 1 - p^{-r} < 1$

Chebyshev: $\Pr(|\eta_w^x - \mathbb{E}\eta_w^x| \geq c) \leq \frac{\sigma^2}{c^2}$.

with $c = 2$, $\Pr(\eta_w^x \geq 3) \leq \frac{1}{4}$.

similar bound to show that 0 solutions is unlikely. then $\Pr(\eta_w^x = 1 \text{ or } 2) \geq \frac{1}{4}$.

Stripped down algorithm

$$\frac{1}{\sqrt{p}} \sum_{b \in \mathbb{Z}_p} \omega^{a_x x b + a_y [x\binom{b}{2} + yb]} |b\rangle$$

two copies: $$\frac{1}{p} \sum_{b_1, b_2 \in \mathbb{Z}_p} \omega^{a_x (x_1 b_1 + x_2 b_2) + a_y (x_1 \binom{b_1}{2} + y_1 b_1 + x_2 \binom{b_2}{2} + y_2 b_2)} |b_1, b_2\rangle$$

$$\mapsto \frac{1}{p} \sum_{b_1, b_2} \omega^{a_x v + a_y w} | b_1, b_2, x, w\rangle$$

now unitarily erase $\frac{1}{\sqrt{2}} (|b_{11}, b_{21}\rangle + |b_{12}, b_{22}\rangle)$ (2 solutions w.p. $\approx \frac{1}{2}$)

$$\mapsto \frac{1}{p} \sum_{v,w} \omega^{av + bw} |v, w\rangle ; \quad \text{inverse FT.}$$