Quantum Algorithms 4/13/2023

Lecture 22: Eigenvalue Estimation – DRAFT

Instructor: Dieter van Melkebeek Scribe: Edward Barton, Zihao Zhu, David Zikel

Last lecture we introduced eigenvalue estimation as an application of phase estimation. This lecture we analyze two remarkable instantiations of eigenvalue estimation. First we analyze the unitary operator underlying amplitude amplification. This leads us to an interesting additional capability: estimating the success probability in amplitude amplification with a square root speed up in terms of the number of queries compared to the classical process. It also yields an alternate view of our earlier version of amplitude amplification to handle unknown success probability.

The second instantiation is order finding, the critical ingredient in the quantum algorithm for factoring integers. We cover the quantum part of the algorithm in full, and the classical part modulo continued fraction expansions, which we will cover next lecture.

1 Recap of phase estimation

We start with the problem statement.

Input: A pure state on of the form $|\psi\rangle = \frac{1}{\sqrt{M}}\sum_{x=0}^{M-1}\exp(2\pi i\omega x)\,|x\rangle$, where $\omega \in [0,1)$ is unknown. We refer to such a state as a harmonic state. Note that the input expression is very similar to the output of the Quantum Fourier Transform. The key difference is that ω can take any value between 0 and 1. The objective is to find ω (or a good approximation of it).

Classical output: ω, or a good approximation of the form y/M for $y \in \mathbb{Z}_M$ such that

$$|\omega - \frac{y}{M}|_{\mathbb{T}} \leq \delta, \tag{1}$$

where we can take δ to be another parameter and $\mathbb{T}$ refers to "modulo 1" (explained below).

Quantum output: Pure state $|\tilde{\omega}\rangle$ on n qubits with total weight of the good y's at least $1-\epsilon$, where good y's are the ones that satisfy (1).

To produce the quantum output, we apply the inverse Fourier transform F^{-1} over $\mathbb{Z}_{\mathbb{M}}, +$. For the problem with classical output, we output the result of measuring this state. The odds of observing y^* for $|\frac{y^*}{N} - \omega|_{\mathbb{T}}$ minimal is at least $\frac{4}{\pi^2}$, and the odds of observing some y such that $|\frac{y}{N} - \omega|_{\mathbb{T}} \geq \delta$ for some specified δ is at most $\frac{1}{2(\delta N-1)} = O(1/\delta N)$.

2 Eigenvalue Estimation

The problem statement is similar to phase estimation, except that we are given an eigenstate $|\phi\rangle$ rather than a harmonic state.

Input: A unitary operation U on n qubits, e.g., in the form of a unitary circuit, and an eigenstate $|\phi\rangle$ of U, i.e., a state such that $U|\phi\rangle = \lambda|\phi\rangle$, where $\lambda = \exp(2\pi i\omega)$ for some $\omega \in [0,1)$. Also an accuracy parameter $\delta > 0$.

Classical output: A good approximation to ω, namely $\tilde{\omega}$ such that $|\omega - \tilde{\omega}|_{\mathbb{T}} \leq \delta$.

Quantum output: Pure state $|\tilde{\omega}\rangle$ with "most" weight on good approximations to ω.

This is very similar to phase estimation, except that we are given an eigenstate $|\phi\rangle$ rather than a harmonic state.

2.1 Algorithm

1. Create an m-qubit harmonic state $|\psi\rangle$ with frequency ω. In order to do this:

   ○ Prepend to $|\phi\rangle$ an ancilla register initialized to $|0^m\rangle$: $|0^m\rangle\,|\phi\rangle$

   ○ Put that ancilla register in a uniform superposition: $\frac{1}{\sqrt{M}}\sum_{x=0}^{M-1}|x\rangle\,|\phi\rangle$

   ○ Use the first register as a control to apply U a number of times to the second register: $\frac{1}{\sqrt{M}}\sum_{x=0}^{M-1}|x\rangle\,U^x\,|\phi\rangle$

   In summary, we will have the following process:

   $$|0^m\rangle\,|\phi\rangle \rightarrow \frac{1}{\sqrt{M}}\sum_{x=0}^{M-1}|x\rangle\,|\phi\rangle \rightarrow \frac{1}{\sqrt{M}}\sum_{x=0}^{M-1}|x\rangle\,U^x\,|\phi\rangle$$

   For the rationale behind this, we observe that applying U x times to $|\phi\rangle$ is, as $|\phi\rangle$ is an eigenstate, equivalent to a phase shift of $\lambda^x = \exp(2\pi i\omega x)$. This gives us $\frac{1}{\sqrt{M}}\sum_{x=0}^{M-1}\exp(2\pi i\omega x)\,|x\rangle\,|\phi\rangle = |\psi\rangle\,|\phi\rangle$. Ignoring the (unchanged) $|\phi\rangle$ yields a harmonic state with frequency ω.

2. Apply phase estimation to $|\psi\rangle$: $|\psi\rangle\,|\,|\phi\rangle \rightarrow |\tilde{\omega}\rangle\,|\phi\rangle$

Note that $|\phi\rangle$ has not been affected and works as a catalyst in this process. The final state we get is equivalent to the tensor product of $|\tilde{\omega}\rangle$ and $|\phi\rangle$. There is no entanglement between the two registers.

Here is the circuit obtained:

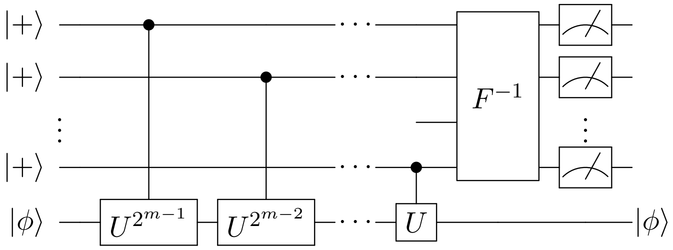

Figure 1: Eigenvalue estimation circuit. The i^{th} highest order bit controls $U^{2^{m-i}}$.

2.2 Analysis

In terms of accuracy:

○ The weight of the $y^* \in \mathbb{Z}_m$ with $|\omega - \frac{y^*}{M}|_{\mathbb{T}} \leq \frac{1}{2M}$ is at least $\frac{4}{\pi^2}$.

- The total weight of the $y \in \mathbb{Z}_m$ with $|\omega - \frac{y}{M}|_{\mathbb{T}} \geq \delta$ is $O(\frac{1}{\delta M})$. As such, we can get roughly m bits of accuracy with constant probability.

- If the confidence level is not enough, we can boost it classically. We can either increase M or run the process multiple times, and then select the result that appears most often or consider the median of different results.

- After measuring register $|\tilde{\omega}\rangle$, the register $|\phi\rangle$ remains intact, so we can immediately reuse it in subsequent runs for boosting the confidence..

In terms of efficiency, we can apply a Fourier transform in time $\operatorname{poly}\log(M)$ (i.e., polynomial in m). Applying high powers (up to M) of U can be more time consuming, and the efficiency of the whole process is typically dictated by this part (computing powers of U).

2.3 Process applied to an arbitrary state

In the problem of eigenvalue estimation, we are directly given an eigenstate of U, namely $|\phi\rangle$. What if we apply our algorithm to an arbitrary input state $|\varphi\rangle$? The analysis for an arbitrary input state $|\varphi\rangle$ goes like this:

1. Since U is a unary operator, it has a full orthonormal basis of eigenstates: $|\varphi\rangle = \sum_j \alpha_j |\phi_j\rangle$ where $U|\phi_j\rangle = \exp(2\pi i \omega_j)| |\phi_j\rangle$ for $\omega_j \in [0,1)$.

2. We can write $|\varphi\rangle$ as a linear combination of the eigenstates: $|\varphi\rangle = \sum_j \alpha_j |\phi_j\rangle$. Note that we may not know the coefficients α_j, but, from an analysis point of view, they must exist.

3. By linearity of the underlying unitary quantum circuit:

$$|0^m\rangle |\varphi\rangle = \sum_j \alpha_j |0^m\rangle |\phi_j\rangle \to \sum_j \alpha_j |\tilde{\omega}_j\rangle |\phi_j\rangle .$$

Measuring the first register of $\sum_j \alpha_j |\tilde{\omega}_j\rangle |\phi_j\rangle$ has the same distribution as picking j with probability $|\alpha_j|^2$ and then measuring $|\tilde{\omega}_j\rangle$.

3 Amplitude Amplification

We now develop the instantiation of eigenvalue estimation to the unitary operator underlying amplitude amplification. We start with a recap of the setting and algorithm for amplitude amplification.

3.1 Setting

- Black box access to $f : \{0,1\}^n \to \{0,1\}$. The predicate f indicates what states are good and what states are bad.

- Unitary circuit A on n qubits such that $A|0^n\rangle = \sum_z \alpha_z |z\rangle$ has non-zero amplitude α_z on some x with $f(z) = 1$. The state $A|0^n\rangle$ is the start state.

- Weight $p = \sum_{z:f(z)=1} |\alpha_z|^2 > 0$. This p may or may not be given.

The goal is to output $|G\rangle$.

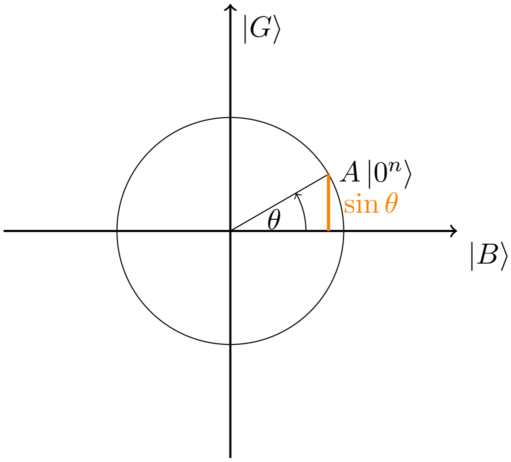

Figure 2: Two dimensional depiction

$|B\rangle$ is the part of the initial state supported on bad basis states, and $|G\rangle$ the part supported on good basis states. Notation:

- $A|0^n\rangle = \sqrt{1-p}\,|B\rangle + \sqrt{p}\,|G\rangle$
- $|B\rangle = \frac{1}{\sqrt{1-p}} \sum_{x:f(x)=0} \alpha_x |x\rangle$
- $|G\rangle = \frac{1}{\sqrt{p}} \sum_{x:f(x)=1} \alpha_x |x\rangle$
- $p = \sin^2\theta$

3.2 Algorithm

1. Start from $A|0^n\rangle$.
2. Repeatedly apply $U \doteq AR_{|0^n\rangle}A^{-1}U_f$

In the expression for U the part U_f represents a flip around the bad basis states, and the other part a flip around the start state, as we have seen before. We can analyze the effect of the operator U in a 2-dimensional diagram with with $|B\rangle$ on the x-axis and $|G\rangle$ on the y-axis. The combined effect of U is a rotation over 2θ counter clock-wise in $|B\rangle - |G\rangle$ plane, where θ is given by $p = \sin^2\theta$.

3.3 Eigenvalue estimation

We now apply eigenstate estimation on the operator U. Among other things, it will give us an efficient way to approximate success probability p.

The matrix representation of the rotation over 2θ counter-clockwise, U in the $|B\rangle - |G\rangle$ plane is given by:

$$\begin{bmatrix} \cos 2\theta & -\sin 2\theta \\ \sin 2\theta & \cos 2\theta \end{bmatrix}$$

The first column of U rotate $|B\rangle$ over 2θ counter clock-wise so that the x-axis is $\cos 2\theta$ and y-axis is $\sin 2\theta$. Similarity, the second column of U rotate $|G\rangle$ and the x and y axis are $-\sin 2\theta$ and $\cos 2\theta$ accordingly.

Now we want to know what the eigenstructure of U is (eigenvalues and eigenvectors) and how we can decompose the original state as a linear combination of these eigenstates. The structure of U is as follows:

- Eigenvalues: $\lambda_+ = \exp(2i\theta)$, $\lambda_- = \exp(-2i\theta)$
  One can verify by the trace of the matrix is the sum of the trace and the product of eigenvalues is the determinant of the matrix.

- Eigenvectors: $|\phi_+\rangle = \frac{1}{\sqrt{2}}\begin{bmatrix} i \\ 1 \end{bmatrix}$, $|\phi_-\rangle = \frac{1}{\sqrt{2}}\begin{bmatrix} -i \\ 1 \end{bmatrix}$

  One can verify the eigenvectors by applying eigenvectors with U. Applying U with $|\phi_+\rangle$

  $$U|\phi_+\rangle = \frac{1}{\sqrt{2}}\begin{bmatrix} i\cos 2\theta - \sin 2\theta \\ i\sin 2\theta + \cos 2\theta \end{bmatrix}$$

  Applying λ_+ with $|\phi_+\rangle$

  $$\lambda_+|\phi_+\rangle = \frac{1}{\sqrt{2}}\begin{bmatrix} i \cdot e^{2i\theta} \\ e^{2i\theta} \end{bmatrix}$$

  Similarly for $|\phi_-\rangle$ case.

  This means: $|\phi_+\rangle = \frac{i}{\sqrt{2}}|B\rangle + \frac{1}{\sqrt{2}}|G\rangle$ and $|\phi_-\rangle = \frac{-i}{\sqrt{2}}|B\rangle + \frac{1}{\sqrt{2}}|G\rangle$. Note that the weights of $|B\rangle$ and $|G\rangle$ are the same both in $|\phi_+\rangle$ and $|\phi_-\rangle$, namely half of the total weight. So if we observe $|\phi_+\rangle$ or $|\phi_-\rangle$, we will get a good basis state with probability $\frac{1}{2}$.

- $A|0^n\rangle = \alpha_+|\phi_+\rangle + \alpha_-|\phi_-\rangle$ where $\alpha_\pm = \exp(\pm 2i\theta)/\sqrt{2}$. Verify the second component:

  $$\frac{1}{\sqrt{2}} \cdot \frac{1}{\sqrt{2}}e^{2i\theta} + \frac{1}{\sqrt{2}} \cdot \frac{1}{\sqrt{2}}e^{-2i\theta} = \frac{1}{2}(\cos 2\theta + i\sin 2\theta + \cos 2\theta - i\sin 2\theta) = \cos 2\theta$$

3.4 Analysis

- On input $|0^m\rangle A|0^n\rangle$, eigenvalue estimation yields $\alpha_+|\tilde{\omega}_+\rangle|\phi_+\rangle + \alpha_-|\tilde{\omega}_-\rangle|\phi_-\rangle$.

- Measure the first register. By the above exercise, this is equivalent to measuring $|\tilde{\omega}_+\rangle$ with probability 1/2, and measuring $|\tilde{\omega}_-\rangle$ with probability 1/2.

- Measuring $|\tilde{\omega}_\pm\rangle$ yields $y_\pm \in \mathbb{Z}_m$ such that $|\omega_\pm - \frac{y_\pm}{M}|_\mathbb{T} \le \delta$ with probability $1 - O(\frac{1}{\delta M})$ where $\omega_\pm = \pm\frac{2\theta}{2\pi} = \pm\frac{\theta}{\pi}$. A good approximation for θ gives us a good approximation for $p = \sin^2(\theta)$, as we will analyze next.

3.5 Estimating success probability

- We have $| \pm\theta - \tilde{\theta}_\pm|_\mathbb{T}| \le \delta$ with probability $1 - O(\frac{1}{\delta M})$ where $\tilde{\theta}_\pm = \pi\frac{y_\pm}{M}$. We don't know whether we get an approximation to $+\theta$ or to $-\theta$.

- $p = \sin^2(\pm\theta)$ and $\tilde{p} = \sin^2(\tilde{\theta}_\pm)$. Since we are squaring the the sine, it will not matter to us if the approximation we get is for $+\theta$ or to $-\theta$.

- By symmetry, it suffices to analyze $|p - \tilde{p}|$ for the the case of $+\theta$.
- $|p - \tilde{p}| = |\sin^2\theta - \sin^2\tilde{\theta}| = |\sin\theta - \sin\tilde{\theta}||\sin\theta + \sin\tilde{\theta}|$.
- Using Lipschitz continuity, we know $|\sin\theta - \sin\tilde{\theta}| \leq |\theta - \tilde{\theta}|$. So we have:
$$|p - \tilde{p}| \leq |\theta - \tilde{\theta}|(2\sin\theta + |\theta - \tilde{\theta}|).$$
- Using the guarantee for $\tilde{\theta}$ that we get from the analysis of eigenvalue estimation, we have:
$$|p - \tilde{p}| \leq \pi\delta(2\sin\theta + \pi\delta) = 2\pi\delta\sqrt{p} + \pi^2\delta^2.$$
- So now we can guarantee the error is bounded: $|p - \tilde{p}| \leq \eta \cdot p$ with probability at least $1 - \epsilon$ by setting $\delta = O(\eta\sqrt{p})$ and $M = O(\frac{1}{\delta\epsilon}) = O(\frac{1}{\eta\epsilon\sqrt{p}})$. Note that M is the maximum number of times we need to apply U.

This concludes a proof of the following result:

Theorem 1. *$O(\frac{1}{\sqrt{p}})$ iterations of A suffice to estimate the success probability p to within constant relative error with constant confidence.*

Classically, if we have a Bernoulli experiment with unknown probability of success p, and we we want to get a good estimate of p, we need to run this experiment $O(\frac{1}{p})$ times. So using the quantum algorithm we get a square root speed up.

3.6 Amplitude amplification with unknown probability of success

What if we don't know p then how we do amplitude amplification? Note that in the quantum algorithm for amplitude amplification, we need to specify the number of rounds to apply U, and if we run it more than necessary, it is going to bring down the weight of the good part again since it's a quasi-periodic function of p. In the lectures on amplitude amplification we developed a rather ad-hoc algorithm for this setting. We'll now see that essentially the same algorithm follows naturally from the eigenvalue estimation approach.

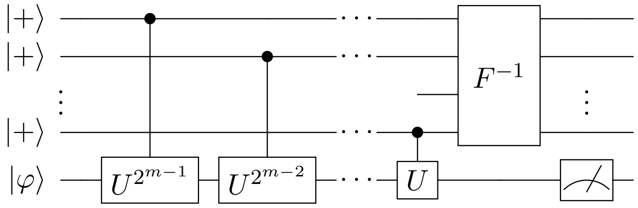

Figure 3: Amplitude amplification revisited

Thinking in terms of search of a good guy, we will measure the second register that contains the arguments. Thus, instead of measuring the first register as in Figure 1, we measure the second one with some probability measuring a good guy. The final state after unitary operations and without measurement is a superposition $\alpha_+ |\tilde{\omega}_+\rangle |\phi_+\rangle + \alpha_- |\tilde{\omega}_-\rangle |\phi_-\rangle$, where $|\alpha_+| = |\alpha_-| = \frac{1}{\sqrt{2}}$, $|\phi_\pm\rangle = \frac{1}{\sqrt{2}}(\pm i |B\rangle + |G\rangle)$. For the analysis, we assume the M is large enough then we have $M = \Omega(\frac{1}{\delta\epsilon})$, $\Pr[| \pm \theta - \tilde{\theta}_\pm|_{\mathbb{T}} \leq \pi\delta] \geq 1 - \epsilon$.

For $M = \Omega(\frac{1}{\sqrt{p}})$, consider in following two cases

1. **The argument based on the exercise:** There were no overlap at all, then the two states $|\tilde{\omega}_+\rangle$ and $|\tilde{\omega}_-\rangle$ are orthogonal. In that case, consider exercise**??** with the roles of with the roles of $|\tilde{\omega}_j\rangle$ and $|\phi_j\rangle$ reversed and only two values of j for which α_j is nonzero (namely $j = +$ and $j = -$). The exercise tells us that measuring the final state is equivalent to picking $j \in \{+, -\}$ with equal probability (since $|\alpha_+| = |\alpha_-|$) and then measuring $|\phi_j\rangle$. Since in both $|\phi_+\rangle$ and $|\phi_-\rangle$ the weight of $|G\rangle$ is 1/2, the probability of measuring a good state is 1/2.

2. **The argument based on the mixed states:** We have shown that most of the weight is on approximations that have a small relative error, where the approximant is either $+\theta$ or $-\theta$. This in particular means that most of the weight is on approximations that have the same sign as the approximant, and therefore that the overlap of $|\tilde{\omega}_+\rangle$ and $|\tilde{\omega}_-\rangle$ is small. Then we have approximately equal mixture $|\phi_+\rangle$ and $|\phi_-\rangle$, similarily, the weight of $|G\rangle$ is 1/2 for both cases. Thus the probability of measuring a good state is approximately 1/2.

Conclusion That one run of the circuit yields a good z with probablity about 1/2 provided $M = \Omega(1/\sqrt{p})$, because of we almost has equal mixture of $|\phi_+\rangle$ and $|\phi_-\rangle$. Both $|\phi_+\rangle$ and $|\phi_-\rangle$ have the weight $\frac{1}{2}$ on $|G\rangle$.

So if we set M large enough, we can get a good qubit with probability at least half. We do not know p, and thus do not know how to set M, but can try successive powers of two after boosting the success probability to at least 1/2, which is not enough. However, we can try twice to make the failure probability at most 1/4 to compensate this situation. This is what happens in the following algorithm, where m increases by 1 in every iteration, and thus $M \doteq 2^m$ doubles each time.

Algorithm for unknown success probability

1. Try $m = 1, 2, \ldots$ until first success.

2. Try twice for a given m to ensure each trial with $M = 2^m$ succeeds with probability at least about $\frac{3}{4}$ for $M \geq \Omega(1/\sqrt{p})$.

Analysis By an analysis similar as the one in the lecture on amplitude amplification, the expected number of iterations of A is $O(\frac{1}{\sqrt{p}})$. As the goal is to get a good state, the final state should be close to $\pm\frac{\pi}{2}$ as shown in fig. 2, which result in the optimal number of iteration should be $k^* = \frac{1}{2}(\frac{\pi}{2\theta_0} - 1)$. Thus $k^* = O(\frac{1}{\theta_0}) = O(\frac{1}{\sqrt{p}})$, with $p = \sin^2\theta_0$ and when θ_0 is small, $\sin\theta_0 \approx \theta_0$.

Note that we only observe the second register, and the Fourier operation on the first register has no effect on the measurements we get for second register. So we might as well not apply the Fourier transform. As we apply all $|+\rangle$, uniform distribution, there are 2^{m-1} bits considering as binary bit. Then when we do the measurement, it will collapse to one of the bits. The resulting circuit is equivalent to picking a number of iterations uniformly between 0 and $2^m - 1$ for a given value of m. Before, we would let the bound increase geometrically and then pick the actual number of iterations uniformly from 0 to that bound. This seemed like an ad-hoc trick, but now we see the same process follows naturally from the perspective of eigenvalue estimation.

4 Order Finding

We now develop the instantiation of eigenvalue estimation to order finding, i.e., finding the order of some integer modulo some other integer.

4.1 Problem

- **Input:** $a, \mu \in \mathbb{N}$
- **Output:** Smallest positive $r \in \mathbb{N}$ such that $a^r = 1 \bmod \mu$

Note that if a and μ have some factor in common, then $r \in \mathbb{N}$ does not exist. Thus, for r to exist, we require a and μ to be relatively prime, i.e., $\gcd(a, \mu) = 1$. Furthermore, $\gcd(a, \mu) = 1$ implies the existence of such an r that The classical algorithm has a time complexity of $O(\text{poly}(\mu))$. With the quantum algorithm, we will show that this problem can be solved in $O(\text{poly} \log(\mu))$.

4.2 Approach

Again, we view this as an application of eigenvalue estimation, so we will start with the unitary operator.

- First, we set up U such that eigenvalues of U have a strong connection with r and having powers that can be applied efficiently.
- Then, we apply eigenvalue estimation to U.
- Finally, we retrieve r from an approximate eigenvalue of U via classical post-processing.

4.3 Order Finding - Unitary Operator

In this part, we design a special unitary operator U such that its eigenvalues have a strong connection with r. We will also determine the state $|\zeta\rangle$ as required by the eigenvalue estimation.

The number of qubits, n is chosen to be the smallest integer n which satisfies $2^n \geq \mu - 1$. Thus we define the operator as one that maps the basis state $|x\rangle$ to $|ax \bmod \mu\rangle$ if $x \in [0, \mu)$ and leaves the state untouched if $x \in [\mu, N)$.

$$U : |x\rangle \mapsto \begin{cases} |ax \bmod \mu\rangle & \text{for } 0 \leq x < \mu \\ |x\rangle & \text{for } \mu \leq x < N \end{cases}$$

We verify this is a valid unitary operator for implementation in a quantum computer. Since $\gcd(a, \mu) = 1$, a has a multiplicative inverse modulo μ, and the mapping $x \mapsto ax \bmod \mu$ is one-to-one. Thus, U acts deterministically and reversibly, and is therefore a valid quantum operation. In addition, with this representation, we are able to efficiently apply U^k because U^k maps integers $x \in [0, \mu)$ to $a^k x \bmod \mu$, and we can compute $a^k \bmod \mu$ classically in time $\text{poly} \log(k\mu)$ using repeated squaring.

Analysis of Eigenstructure of U

- Notice that the excess states (states corresponding to $x > \mu$) are not modified by the operator. Thus, every basis state $|x\rangle$ for integral $x \in [\mu, N)$ is an eigenstate of U with eigenvalue 1. We are not interested in those eigenstates and their eigenvalues as they are uninformative.

- If we apply the operator r times, then we get the identity operator, i.e., $U^r = I$. This implies that the eigenvalues of U are r'th roots of unity. Meaning the solutions are elements of a regular r-gon inscribed in the unit circle in the complex plane with 1 as one of the points. Mathematically, the points can be expressed as:

$$\lambda_j = \exp(2\pi i \omega_j) \text{ with } \omega_j = j/r \text{ for } j \in \mathbb{Z}_r \tag{2}$$

  This shows that ω_j does have a strong connection with r. Thus, if we can approximate ω_j, then we can retrieve r from it, at least when j is relatively prime to r.

- For eigenvectors $|\phi_j\rangle$, we would like to get values which have support for $x \in [0, \mu)$, i.e.,

$$|\phi_j\rangle = \sum_{s=0}^{r-1} \alpha_{j,s} |a^s \bmod \mu\rangle \tag{3}$$

  Because the excess states do not contain any information about r, we ignore them.

  Next we need to determine what our amplitudes, $\alpha_{j,s}$, need to be to satisfy our system. We start with applying U to our $|\phi_j\rangle$. This results in a linear combination of the operator applied to those states. That application raises a^s to a^{s+1}.

$$\begin{aligned} U|\phi_j\rangle &= \sum_{s=0}^{r-1} \alpha_{j,s} U |a^s \bmod \mu\rangle \\ &= \sum_{s=0}^{r-1} \alpha_{j,s} |a^{s+1} \bmod \mu\rangle \\ &= \sum_{s=1}^{r} \alpha_{j,s-1} |a^s \bmod \mu\rangle \\ &= \sum_{s=0}^{r-1} \lambda_j \alpha_{j,s} |a^s \bmod \mu\rangle \end{aligned} \tag{4}$$

  In the third step we re-index the summation so that we can maintain a^s as the basis state in our expression. The expression on the right-hand side of the fourth line is the expansion of $\lambda_j |\phi_j\rangle$, and the equality expresses that $|\phi_j\rangle$ is an eigenvector of U with eigenvalue λ_j. Equating the coefficients in the third and the fourth line, we can see for a fixed j, the current amplitude is equal to the previous divided by the eigenvalue for that fixed j. Or, the previous amplitude is equal to the product of the eigenvalue with the current amplitude. Note that the summation ranges don't quite match up, but the following still holds for every s as $a^r = a^0 \bmod \mu$.

$$\alpha_{j,s-1} = \lambda_j \alpha_{j,s}$$

- ∘ Next these amplitudes need to be normalized. Since $|\lambda_j| = 1$, the amplitudes $\alpha_{j,s}$ all have the same absolute value. As there are r of them, their absolute value should be $1/\sqrt{r}$.

- ∘ Finally, we can pick the power of λ_j for one component as we wish; we'll set the power to be zero when $s = 0$. Thus, we obtain:

  $$\alpha_{j,s} = \frac{1}{\sqrt{r}} \lambda_j^{-s} \tag{5}$$

  Note that we can express the resulting expression in terms of the character χ_ρ with $\rho = -s$ at point j.

  $$\alpha_{j,s} = \frac{1}{\sqrt{r}} \lambda_j^{-s} = \frac{1}{\sqrt{r}} \exp(-2\pi i s j/r) = \frac{1}{\sqrt{r}} \chi_{-s}(j). \tag{6}$$

  There are other possible combinations to represent this same expression, this is just the one that was chosen for the purposes of lecture.

- ∘ If we could generate an eigenvector $|\phi_j\rangle$ for some j that is relatively prime to r, say for $j = 1$, we could use eigenvalue estimation to approximate j/r and extract r from it. Unfortunately, we don't know how to generate these eigenvectors because they require knowledge of r. However, we know we can express the basis states a^s as linear combinations of the eigenvectors $|\phi_j\rangle$. We can then apply the eigenvalue estimation to that linear combination. In particular, we have the following decomposition:

  $$\frac{1}{\sqrt{r}} \sum_{s=0}^{r-1} |\phi_j\rangle = \frac{1}{r} \sum_{s=0}^{r-1} \sum_{j=0}^{r-1} \chi_{-s}(j) \, |a^s \bmod \mu\rangle = |1\rangle \tag{7}$$

  We use our previous decomposition of the eigenvectors and the corresponding amplitudes, so we can express the amplitudes of the basis states $|a^s \bmod \mu\rangle$ as sums of character values. Recall that $\sum_j \chi_{-s}(j) = 0$ for all s except $s = 0$. This causes the outer sum to reduce since the only s that contributes is $s = 0$. For $s = 0$ the inner sum of character values equals r, and the corresponding basis state is $|a^0 \bmod \mu\rangle = |1\rangle$, yielding the right-hand side $|1\rangle$ in (7). We conclude that we can easily generate the superposition of eigenvectors on the left-hand side of (7), namely as the basis state $|1\rangle$. This is the superposition we will use to run eigenvalue estimation on.

4.4 Order Finding - Quantum Part

The quantum part consists of applying the eigenvalue estimation with:

$$U : |x\rangle \mapsto \begin{cases} |ax \bmod \mu\rangle & \text{for } 0 \leq x < \mu \\ |x\rangle & \text{for } \mu \leq x < N \end{cases}$$

with initial state $|\zeta\rangle$ as:

$$|\zeta\rangle = |1\rangle = \frac{1}{\sqrt{r}} \sum_{j=0}^{r-1} |\phi_j\rangle \,,$$

where $U|\phi_j\rangle = \exp(2\pi i \omega_j)|\phi_j\rangle$ and $\omega_j = j/r$. This gives us, for some $j \in \mathbb{Z}_r$ chosen randomly with probability $1/r$, a value $y \in \mathbb{Z}_m$ such that:

$$\Pr\left[|\frac{j}{r} - \frac{y}{M}|_{\mathbb{T}} \leq \delta\right] \geq 1 - O(1/(\delta M))$$

For a uniformly chosen and unknown j, eigenvalue estimation yields y/M, which is an approximation of j/r, or ω_j, such that the absolute error is not larger than δ with high probability.

4.5 Order Finding - Classical Part

Given such an approximation, the classical part extracts r from y and M.

Starting Point

- For some $j \in \mathbb{Z}_r$ chosen randomly with probability $1/r$, a value $y \in \mathbb{Z}_m$ such that

  $$|\frac{j}{r} - \frac{y}{M}|_{\mathbb{T}} \leq \delta$$

  The quality of an approximation is given by δ. We still need to decide what δ to use and make sure the probability is high enough to achieve that.

Retrieving r

- We first note that from a single run, the best we can do is to retrieve j and r in reduced form as j' and r' where $j' \doteq j/\gcd(j,r)$ and $r' \doteq r/\gcd(j,r)$.
- For now, we assume that we can retrieve j' and r' efficiently from y and M.
- To get a good estimate of r, we leverage the fact that if j and r are relatively prime, then $r' = r$. In order to achieve this condition, we can run the estimation multiple times and return the largest r' obtained. It is important to note here that this is not a guarantee, all we can say is: if we run this enough times, with high probability we can claim that $\max(r') = r$. For this, $O(\log N)$ runs suffice as $\Pr[\gcd(j,r) = 1] = \Omega(1/\log(r))$. Note r can be up to 2^k when working with k bits, but we only need $\text{poly}(k)$ runs to have a high probability of success.
- We can improve this by noting that every r' will be a divisor or r. So, we can run multiple times and return the least common multiple (lcm) of all r' obtained. Now we only need a constant number of runs to guarantee a probability of success. This is because, run twice, we get the following:

  $$\Pr[\text{lcm}(r_1', r_2') = r] = \Pr[\gcd(j_1, j_2) = 1] \geq 1 - \sum_{p \text{ prime}} \frac{1}{p^2} \geq 54\% \tag{8}$$

Next, we discuss how we can efficiently retrieve j' and r' from y and M.

Retrieving j' and r'

Given

- ◦ $y \in \mathbb{Z}_M$ such that $|\frac{j'}{r'} - \frac{y}{M}|_{\mathbb{T}} \leq \delta$ for some $j \in \mathbb{Z}_r$ chosen uniformly at random.
- ◦ $j' \doteq j/\gcd(j,r)$ and $r' \doteq r/\gcd(j,r)$.

Approach

- ◦ We first observe that if $\delta < \frac{1}{2N^2}$, then $\frac{j'}{r'}$ is the unique rational with denominator at most N such that $|\frac{j'}{r'} - \frac{y}{M}|_{\mathbb{T}} \leq \delta$.

  Proof. Suppose $|\frac{j_1}{r_1} - \frac{y}{M}|_{\mathbb{T}} \leq \delta$ and $|\frac{j_2}{r_2} - \frac{y}{M}|_{\mathbb{T}} \leq \delta$.
  Then, via the triangle inequality, $|\frac{j_1}{r_1} - \frac{j_2}{r_2}|_{\mathbb{T}} \leq 2\delta < \frac{1}{N^2}$, so $|j_1 r_2 - j_2 r_1|_{\mathbb{T}} < \frac{r_1 r_2}{N^2} \leq 1$.
  Because $j_1 r_2$ and $j_2 r_1$ are integers, $|j_1 r_2 - j_2 r_1|$ is a nonnegative integer less than 1, and therefore equals zero. So, $\frac{j_1}{r_1} = \frac{j_2}{r_1}$. □

- ◦ If $\delta < \frac{1}{2N^2}$, continued fraction expansion of $\frac{y}{m}$ allows efficient retrieval of r' and j'.