

# 量子计算(八)——振幅放大与振幅估计

## 一、振幅放大的问题背景

这个算法要解决的问题就是寻找符合要求的解。假设解空间可以被表示为二进制字符串，并且已知某种能够确定解空间中各个解的好坏的标准。

也就是给出布尔函数 $\mathcal{X}$ ，它将解空间中的解 $x$ 映射到 $\{0, 1\}$ ：

$$\mathcal{X}(x) = \begin{cases} 0 & , \quad \text{if } x \text{ is bad} \\ 1 & , \quad \text{if } x \text{ is good} \end{cases}$$

这个目标基本和Grover算法要解决的问题是一致的。事实上，Grover算法中的核心算法就是振幅放大，这篇论文就是对Grover算法的总结推广，使得任意无测量量子算法也可以使用。

### 1.量子化

显然，由于解空间被表示为二进制串，因而解空间可以作为一个希尔伯特空间，从而允许量子算法的运行。而解的好坏，则将其划分为两个子空间——称为好空间与坏空间。于是解空间的任意纯态 $|\psi\rangle$ 都可以被分解表示为：

$$|\psi\rangle = |\psi_0\rangle + |\psi_1\rangle$$

其中 $|\psi_0\rangle$ 表示落入坏空间的部分，相应地 $|\psi_1\rangle$ 表示落入好空间的部分。于是 $b_\psi = \langle\psi_0|\psi_0\rangle$ 表示了对这个纯态测量后得到坏结果的概率， $a_\psi = \langle\psi_1|\psi_1\rangle$ 则是得到好结果的概率，以后简记为 $a$ 。显然 $a_\psi + b_\psi = 1$ 。

到这一步，我们的目标就已经清晰明朗了：只要让好空间部分的 $|\psi_1\rangle$ 振幅变大，就提高了测到好结果的概率。之后只需要代入 $\mathcal{X}$ 判定其是否确实是好结果即可。

## 二、振幅放大的构建

### 1. 振幅放大算符Q

假定 $n$ 是解空间的二进制串的长度。假设无测量量子算法 $\mathcal{A}$ 是作用到解空间上的酉矩阵，并假设纯态 $|\Psi\rangle = \mathcal{A}|0^n\rangle$ 是由其作用到初始零态的结果。那么如下构建的算符即可实现振幅放大：

$$\mathbf{Q} = \mathbf{Q}(\mathcal{A}, \chi) = -\mathcal{A}\mathbf{S}_0\mathcal{A}^{-1}\mathbf{S}_\chi$$

其中， $\mathbf{S}$ 代表它会改变振幅的符号，而下标表示改变的条件：

$$\mathbf{S}_0|x\rangle = \begin{cases} -|x\rangle & , \quad \text{if } x = 0^n \\ |x\rangle & , \quad \text{if } x \neq 0^n \end{cases}$$

$$\mathbf{S}_\chi|x\rangle = \begin{cases} -|x\rangle & , \quad \text{if } \chi(x) = 1 \\ |x\rangle & , \quad \text{if } \chi(x) = 0 \end{cases}$$

显然可以将 $\mathbf{S}_0$ 写为 $I - 2|0^n\rangle\langle 0^n|$ ，于是：

$$\begin{aligned} \mathcal{A}\mathbf{S}_0\mathcal{A}^{-1} &= \mathcal{A}(I - 2|0^n\rangle\langle 0^n|)\mathcal{A}^{-1} \\ &= I - 2\mathcal{A}|0^n\rangle\langle 0^n|\mathcal{A}^{-1} \\ &= I - 2|\Psi\rangle\langle\Psi| \end{aligned}$$

现在我们来研究算符 $\mathbf{Q}$ 作用到任意态上会发生什么。由于在这个问题中纯态被分解为好纯态和坏纯态，因此研究此算符分别作用到好坏纯态上的结果。首先是坏纯态 $|\Psi_0\rangle$ ：

$$\begin{aligned} \mathbf{Q}|\Psi_0\rangle &= -\mathcal{A}\mathbf{S}_0\mathcal{A}^{-1}\mathbf{S}_\chi|\Psi_0\rangle \\ &= -\mathcal{A}\mathbf{S}_0\mathcal{A}^{-1}|\Psi_0\rangle \\ &= -(I - 2|\Psi\rangle\langle\Psi|)|\Psi_0\rangle \\ &= -|\Psi_0\rangle + 2(1 - a)|\Psi\rangle \\ &= (1 - 2a)|\Psi_0\rangle + 2(1 - a)|\Psi_1\rangle \end{aligned}$$

同理对于好纯态 $|\Psi_1\rangle$ ：

$$\begin{aligned} \mathbf{Q}|\Psi_1\rangle &= -\mathcal{A}\mathbf{S}_0\mathcal{A}^{-1}\mathbf{S}_\chi|\Psi_1\rangle \\ &= (I - 2|\Psi\rangle\langle\Psi|)|\Psi_1\rangle \\ &= |\Psi_1\rangle - 2a|\Psi\rangle \\ &= -2a|\Psi_0\rangle + (1 - 2a)|\Psi_1\rangle \end{aligned}$$

## 2. 振幅如何被放大

假设对 $|\Psi\rangle$ 施加 $k-1$ 次 $Q$ 算符后：

$$Q^{k-1}|\Psi\rangle = S_k|\Psi_0\rangle + T_k|\Psi_1\rangle$$

也就是好纯态的振幅变为 $T_k$ ，显然可以得到下列递推式：

$$\begin{cases} S_{k+1} = (1-2a)S_k - 2aT_k \\ T_{k+1} = (2-2a)S_k + (1-2a)T_k \end{cases}, \quad \begin{cases} S_1 = 1 \\ T_1 = 1 \end{cases}$$

解得：

为了解出上式，首先计算 $S_k$ 与 $T_k$ 的线性组合：

$$xS_{k+1} + yT_{k+1} = [x(1-2a) + y(2-2a)]S_k + [-2ax + y(1-2a)]T_k$$

上式能成为等比递推式的条件是：

$$\lambda = \frac{x(1-2a) + y(2-2a)}{x} = \frac{-2ax + y(1-2a)}{y}$$

显然可取 $x = \sqrt{a-1}$ ,  $y = \sqrt{a}$ 从而 $\lambda = 1 - 2a - 2\sqrt{a(a-1)}$ 。记 $u_k = \sqrt{a-1}S_k + \sqrt{a}T_k$ ，则 $u_1 = \sqrt{a-1} + \sqrt{a}$ ，则 $u_k = \lambda^{k-1}u_1$ 。于是：

$$S_k = \frac{\lambda^{k-1}u_1}{\sqrt{a-1}} - \sqrt{\frac{a}{a-1}}T_k$$

代入递推方程组之第二式得：

$$T_{k+1} = \left(1 - 2a + 2\sqrt{a(a-1)}\right) T_k - \frac{2\sqrt{a-1}u_1}{\lambda} \lambda^k$$

记 $A = 1 - 2a + 2\sqrt{a(a-1)}$ ,  $B = \frac{2\sqrt{a-1}u_1}{\lambda}$ ,  $C = \lambda = 1 - 2a - 2\sqrt{a(a-1)}$ 。已知递推形式 $a_{n+1} = Aa_n + BC^n$ 在 $A \neq C$ 时具有通项：

$$a_n = a_1 A^{n-1} + BC \frac{A^{n-1} - C^{n-1}}{A - C}$$

代入可得：

$$T_k = \frac{\sqrt{a} - \sqrt{a-1}}{2\sqrt{a}} \left(1 - 2a + 2\sqrt{a(a-1)}\right)^{k-1} + \frac{\sqrt{a} + \sqrt{a-1}}{2\sqrt{a}} \left(1 - 2a - 2\sqrt{a(a-1)}\right)^{k-1}$$

或者说，对于 $\mathbf{Q}^k|\Psi\rangle$ ，若记 $a = \sin^2\theta$ 且不等于0或1且限定 $\theta \in \left(0, \frac{\pi}{2}\right]$ ，那么其好纯态的振幅是：

$$\begin{aligned} T_{k+1} &= \frac{\sqrt{a} - \sqrt{a-1}}{2\sqrt{a}} \left(1 - 2a + 2\sqrt{a(a-1)}\right)^k + \frac{\sqrt{a} + \sqrt{a-1}}{2\sqrt{a}} \left(1 - 2a - 2\sqrt{a(a-1)}\right)^k \\ &= \frac{\sin\theta - i\cos\theta}{2\sqrt{a}} (1 - 2\sin^2\theta + i2\sin\theta\cos\theta)^k + \frac{\sin\theta + i\cos\theta}{2\sqrt{a}} (1 - 2\sin^2\theta - i2\sin\theta\cos\theta)^k \\ &= \frac{-e^{i(\frac{\pi}{2}+\theta)}}{2\sqrt{a}} (\cos 2\theta + i\sin 2\theta)^k + \frac{e^{i(\frac{\pi}{2}-\theta)}}{2\sqrt{a}} (\cos 2\theta - i\sin 2\theta)^k \\ &= \frac{-e^{i(\frac{\pi}{2}+(2k+1)\theta)}}{2\sqrt{a}} + \frac{e^{i(\frac{\pi}{2}-(2k+1)\theta)}}{2\sqrt{a}} \\ &= \frac{1}{\sqrt{a}} \cdot \sin((2k+1)\theta) \end{aligned}$$

因此，作用 $\mathbf{Q}$ 算符 $k$ 次后，测得好纯态的概率就是 $\sin^2((2k+1)\theta)$ 。

### 3.算符 $\mathbf{Q}$ 作用的次数

显然我们希望测得好纯态的概率越大越好，这就要求 $\sin^2((2k+1)\theta) \rightarrow 1$ ，即 $(2k+1)\theta \rightarrow \frac{\pi}{2}$ ，即 $k = \frac{\pi}{4\theta} - \frac{1}{2} \rightarrow \left\lfloor \frac{\pi}{4\theta} \right\rfloor$ 。以后记 $\tilde{m} = \frac{\pi}{4\theta} - \frac{1}{2}$ ， $m = \left\lfloor \frac{\pi}{4\theta} \right\rfloor$ 。

此时重新计算测得好纯态的概率：

$$\begin{aligned}
\sin^2((2m+1)\theta) &\approx \sin^2\left(\left(\left\lfloor \frac{\pi}{2\theta} \right\rfloor + 1\right)\theta\right) \\
&\approx \sin^2\left(\left\lfloor \frac{\pi}{2} \right\rfloor + \theta\right) \\
&\geq \sin^2\left(\frac{\pi}{2} + \theta\right) \\
&= 1 - \sin^2\theta \\
&= 1 - a
\end{aligned}$$

虽然 $\sin^2((2m+1)\theta) \geq 1 - a$ 这一结论是正确的，但上式的证明是稍有问题的，即 $\sin^2(1 + \theta) \geq \sin^2\left(\frac{\pi}{2} + \theta\right)$ 这一步。但至少在 $\theta \in [0.2854, 1.8562]$ 范围内是没问题的。错误的根本原因在于向下取整内的分母不能直接约分。要想完美证明，接下来只需证 $\theta \in [0, 0.2854]$ 内成立即可，这里不再证明。论文中的证明是引用另一篇文章，请自行观阅。

也就是说测量到好纯态的概率是大于 $1 - a$ 的；同时由于本就是对好纯态振幅的放大，因此这个概率也大于 $a$ 。因此最后测到好纯态的概率至少是 $\max\{a, 1 - a\}$ 。

然而这引申出一个问题：之所以能要求让 $\mathbf{Q}$ 作用 $\left\lfloor \frac{\pi}{4\theta} \right\rfloor$ 次，是因为我们已知了 $\theta$ 值，即已知了 $a$ 值，即已知在最初的时候能够测得好纯态的概率，例如Grover算法中就是如此。但在实际问题中，很多时候这个值具有置信度，甚至完全未知，此时则不能确定 $\mathbf{Q}$ 作用的次数。因此，我们有必要从Grover搜索算法进一步推广。

## 4.算符 $\mathbf{Q}$ 的特征值与特征向量

除了数列递推外，我们也可以通过将 $\mathbf{Q}$ 对角化后，简便地得到 $\mathbf{Q}^k$ 。为此假设：

$$\mathbf{Q}(x|\Psi_0\rangle + y|\Psi_1\rangle) = \lambda(x|\Psi_0\rangle + y|\Psi_1\rangle)$$

显然有方程：

$$\lambda = \frac{x(1 - 2a) + -2ay}{x} = \frac{x(2 - 2a) + y(1 - 2a)}{y}$$

这与之前数列递推所得方程很像，但并没有必然联系。相似地步骤可得特征值与特征向量为：

$$\begin{cases} \lambda_{\pm} = 1 - 2a \pm 2\sqrt{a(a-1)} = e^{\pm i2\theta} \\ |\Psi_{\pm}\rangle = \frac{1}{\sqrt{2}} \left( \frac{i}{\sqrt{1-a}} |\Psi_0\rangle + \frac{1}{\sqrt{a}} |\Psi_1\rangle \right) \end{cases}$$

### 三、量子算法的去随机化

由上节我们已经知道，在已知 $a$ 的情况下，使用 $\mathbf{Q}^m \mathcal{A}|0^n\rangle$ 即可使我们得到好纯态的概率为 $\sin^2((2m+1)\theta) \geq \max\{a, 1-a\}$ 。然而我们依然有可能使这个概率为1，这就是量子算法的去随机化。论文中给出两种方法：

#### 1. $\theta$ 微调

当 $m = \tilde{m}$ 时，也就是 $\frac{\pi}{4\theta} - \frac{1}{2}$ 恰是整数，那么 $\mathbf{Q}^m \mathcal{A}|0^n\rangle$ 就自然完全得到好纯态。另一方面，若记 $\bar{m} = \lceil \tilde{m} \rceil$ ，那么 $\bar{m}$ 次迭代又稍微多了点，因此不妨使角度 $\theta$ 更小点，取 $\bar{\theta} = \frac{\pi}{4\bar{m} + 2}$ 。因此，只要调整算法的初始准确率至 $\bar{a} = \sin^2 \bar{\theta}$ ，那么 $\bar{m}$ 次迭代就恰恰好了。

于是问题转化为如何使算法 $\mathcal{A}$ 的初始准确率从 $a$ 变为 $\bar{a}$ 。这很简单：只要构建另一个算法 $\mathcal{B}$ 使得其作用在单个量子位上时：

$$\mathcal{B}|0\rangle = \sqrt{1 - \frac{\bar{a}}{a}}|0\rangle + \sqrt{\frac{\bar{a}}{a}}|1\rangle$$

(然后同时作用 $\mathcal{A}, \mathcal{B}$ ? 这里没太读懂)

#### 2. Q算符的改进(相位附加)

现在我们改进算符 $\mathbf{Q}$ ，使其在符合条件的情况下不是改变符号，而是附加相位：

$$\mathbf{Q} = \mathbf{Q}(\mathcal{A}, \mathcal{X}, \phi, \varphi) = -\mathcal{A}\mathbf{S}_0(\phi)\mathcal{A}^{-1}\mathbf{S}_{\mathcal{X}}(\varphi)$$

其中相位 $\phi, \varphi \in [0, 2\pi]$ ，并且：

$$\mathbf{S}_0(\phi)|x\rangle = \begin{cases} e^{i\phi}|x\rangle & , \quad \text{if } x = 0^n \\ |x\rangle & , \quad \text{if } x \neq 0^n \end{cases}$$

$$\mathbf{S}_{\mathcal{X}}(\varphi)|x\rangle = \begin{cases} e^{i\varphi}|x\rangle & , \quad \text{if } \mathcal{X}(x) = 1 \\ |x\rangle & , \quad \text{if } \mathcal{X}(x) = 0 \end{cases}$$

显然：

$$\begin{aligned}
\mathbf{S}_0(\phi) &= \sum |result_i\rangle\langle i| \\
&= e^{i\phi}|0^n\rangle\langle 0^n| + \sum_{i=1}^{2^n-1} |i\rangle\langle i| \\
&= \sum_{i=0}^{2^n-1} |i\rangle\langle i| + (e^{i\phi} - 1)|0^n\rangle\langle 0^n| \\
&= I + (e^{i\phi} - 1)|0^n\rangle\langle 0^n|
\end{aligned}$$

于是：

$$\begin{aligned}
\mathcal{A}\mathbf{S}_0(\phi)\mathcal{A}^{-1} &= \mathcal{A} \cdot (I + (e^{i\phi} - 1)|0^n\rangle\langle 0^n|) \cdot \mathcal{A}^{-1} \\
&= I + (e^{i\phi} - 1)\mathcal{A}|0^n\rangle\langle 0^n|\mathcal{A}^{-1} \\
&= I + (e^{i\phi} - 1)|\Psi\rangle\langle\Psi|
\end{aligned}$$

于是：

$$\begin{aligned}
\mathbf{Q}|\Psi_0\rangle &= -\mathcal{A}\mathbf{S}_0(\phi)\mathcal{A}^{-1}\mathbf{S}_{\mathcal{X}}(\varphi)|\Psi_0\rangle \\
&= -(I + (e^{i\phi} - 1)|\Psi\rangle\langle\Psi|)|\Psi_0\rangle \\
&= -|\Psi_0\rangle + (1 - e^{i\phi})(1 - a)|\Psi\rangle \\
&= (a(e^{i\phi} - 1) - e^{i\phi})|\Psi_0\rangle + (1 - e^{i\phi})(1 - a)|\Psi_1\rangle \\
\\
\mathbf{Q}|\Psi_1\rangle &= -e^{i\varphi}(I + (e^{i\phi} - 1)|\Psi\rangle\langle\Psi|)|\Psi_1\rangle \\
&= -e^{i\varphi}|\Psi_1\rangle - e^{i\varphi}(e^{i\phi} - 1)a|\Psi\rangle \\
&= e^{i\varphi}(1 - e^{i\phi})a|\Psi_0\rangle + e^{i\varphi}((1 - e^{i\phi})a - 1)|\Psi_1\rangle
\end{aligned}$$

于是在作用 $\lfloor \tilde{m} \rfloor$ 次 $\mathbf{Q}(\mathcal{A}, \mathcal{X}, \pi, \pi)$ 后，系统处于叠加态：

$$\frac{1}{\sqrt{1-a}}\cos((2\lfloor \tilde{m} \rfloor + 1)\theta)|\Psi_0\rangle + \frac{1}{\sqrt{a}}\sin((2\lfloor \tilde{m} \rfloor + 1)\theta)|\Psi_1\rangle$$

之后再作用一次 $\mathbf{Q}(\mathcal{A}, \mathcal{X}, \phi, \varphi)$ ，但要求相位 $\phi, \varphi$ 符合下式：

$$\frac{1}{\sqrt{1-a}}\cos(\dots)(a(e^{i\phi} - 1) - e^{i\phi}) + \sqrt{a}\sin(\dots)e^{i\varphi}(1 - e^{i\phi}) = 0$$

也就是在这次作用之后， $|\Psi_0\rangle$ 的振幅为0，那么 $|\Psi_1\rangle$ 的振幅自然就为1了。于是问题转换为最后一次作用时应当如何选取 $\phi, \varphi$ 。如何选取并不是本论文的重要内容，但其存在性是可以保证的。

## 四、QSearch算法

在 $a$ 未知的情况下，Grover算法不能发挥太大的作用，因为不知道算符 $Q$ 应当作用的次数，使得测得好纯态的概率偏离最优值。为此，我们推广其至QSearch算法。

### 1. 算法流程

1. `int l = 1;`     `float c = random(start=1, end=2);`  $c$ 不等于1或2。
2. `l++;`     `int M = (int)pow(c, l) + 1;`
3. 将 $\mathcal{A}$ 作用于初始零态 $|0^n\rangle$ ，并测量得到 $|x\rangle$ 。如果是好结果，即 $\mathcal{X}(x) = 1$ ，那么 `return x`;
4. 否则，记录 $|\Psi\rangle = \mathcal{A}|0^n\rangle$ 。
5. `int j = (int)random(1, M);`
6. 将算符 $Q$ 作用到 $|\Psi\rangle$ 上 $j$ 次，即 $Q^j|\Psi\rangle$ 。
7. 测量结果得到 $|x\rangle$ 。如果是好结果，那么 `return x`; 否则， `goto step2`;

尽管这个算法看起来比已知 $a$ 的搜索算法的时间规模更大，但可以证明其依然是 $O(\sqrt{N})$ 的。

### 2. 启发式经典算法的植入

然而很多启发式经典算法也可以做到 $O(\sqrt{N})$ ，这导致我们之前讨论的量子算法看起来没多少优势。但是如果考虑将这些启发式算法改造为量子算法，那不就可以在 $\sqrt{N}$ 的基础上再取一次根号么！可以证明，如果经典算法解决问题的期望时间是 $T$ ，那么将其改造为量子算法后再使用QSearch算法，规模是 $O(\sqrt{T})$ 的。

## 五、振幅估计

### 1. 振幅估计的问题引入

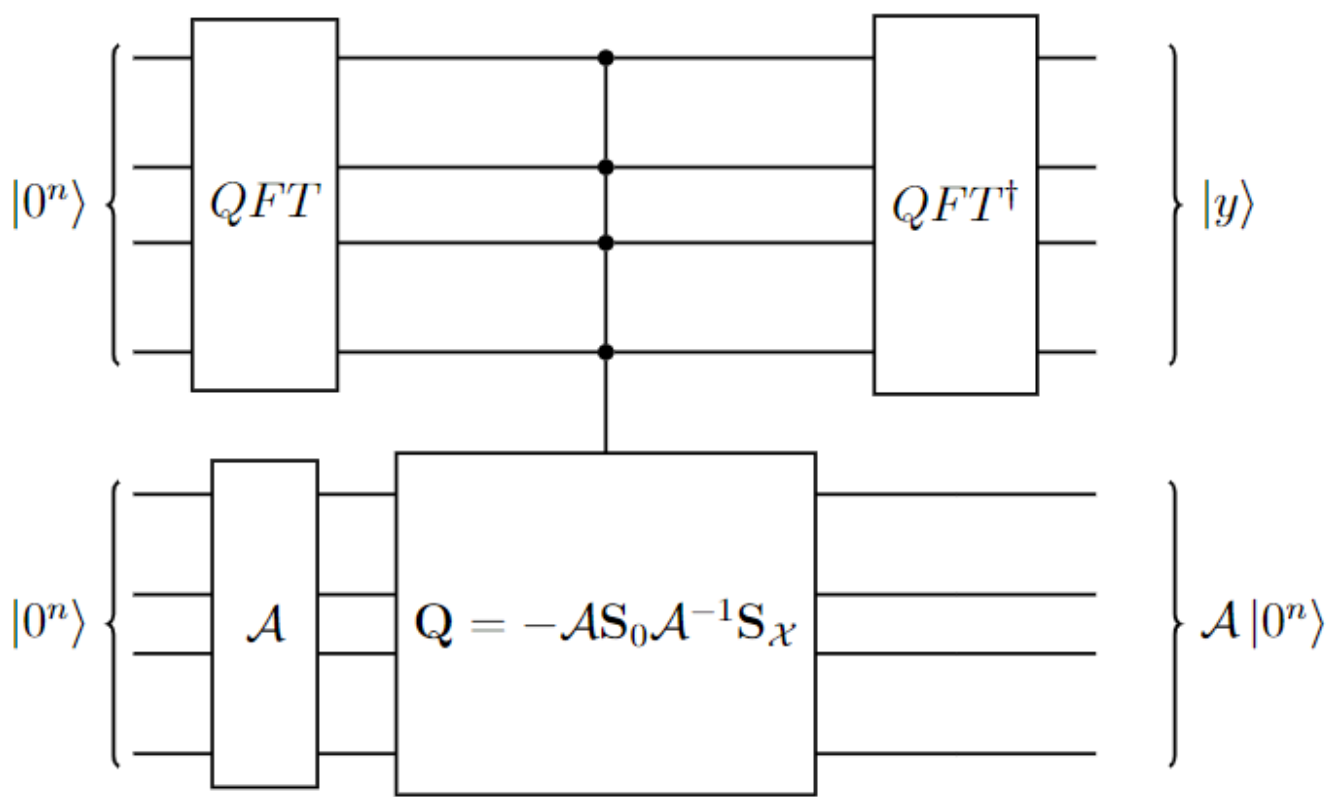
振幅放大是为了寻找解，但这却不能告诉我们解空间中有多少解是好的。振幅估计就是为了解决这个问题。记 $t = |\{x | \mathcal{X}(x) = 1\}|$ ，由于 $a$ 是首次运行 $\mathcal{A}$ 后测量得到好纯态的概率，因此 $a = \frac{t}{N}$ 。因此，只要我们能估计振幅 $a$ ，就能估计好解的数量。

注意到 $a = \sin^2\theta$ ，现在问题转化为估计 $\theta$ 。注意到算符 $Q$ 及其作用次数使得好纯态与坏纯态的振幅形成了三角函数，因此估计这些三角函数的最大周期 $\frac{\pi}{\theta}$ 亦可。而利用量子傅里叶变换的周期查找则是非常成熟的量子算法。当然另一方面，注意到算符 $Q$ 的特征根是 $e^{\pm i2\theta}$ ，利用相位估计来获得 $Q$ 的特征根也是一种可行的方法。



## 2.利用相位估计的振幅估计

将相位估计中的 $U$ 矩阵替换为 $Q$ 即可，如下图所示：



Phase Estimation with  $Q$

于是 $y$ 就代表了 $Q$ 特征值中的相位，从而可以得到 $a$ 的一个估计值：

$$\tilde{a} = \sin^2 \left( \pi \cdot \frac{y}{n} \right)$$