Exercise 8: Continuous Fractions, Shor's Assumption and QFT

1. Here we will be proving some small claims you used in the lecture about **Continuous Fractions**.

   First recall the algorithm:

Data: $x \in \mathbb{R}, k \in \mathbb{N}$
Result: $[a_0, \dots, a_{k-1}, \lambda]$ at most $k+1$-long "almost" continuous-fractions representation of x: $a_i \in \mathbb{N}$ for $0 \leq i < k$, $\lambda \in \mathbb{R}_{>0}$, and $x = a_0 + \frac{1}{a_1 + \frac{1}{\cdots + \lambda}}$

```
a_0 ← ⌊x⌋;
x_0 ← x − ⌊x⌋;
for i=1...k do
  | if x_{i−1} == 0 then
  |   | return [a_0, ..., a_{i−1}]
  | end
  | a_i ← ⌊x_{i−1}^{−1}⌋;
  | x_i ← x_{i−1}^{−1} − ⌊x_{i−1}^{−1}⌋;
end
if x_{k−1} == 0 then
  | return [a_0, ..., a_{k−1}]
end
return [a_0, ..., a_{k−1}, x_{k−1}];
```

   (a) Show that for any rational number $x \in \mathbb{Q}$, the remainders are always rational $x_i = \frac{b_i}{c_i}$ for some $b_i c_i \in \mathbb{N}$. Prove that in this case the sequence of c_i is strictly decreasing.

   Conclude that any rational number has a finite representation in the continuous fractions representation $x = [a_0, a_1, \dots, a_n]$. Where do we need this conclusion in the analysis for Shor's usage of the continuous fractions algorithm? Can an irrational number $x \in \mathbb{R} \setminus \mathbb{Q}$ have such a finite continuous fractions representation?

(b) Let $x \in \mathbb{R}_{>0}$ and $[a_0, \dots, a_n]$ be the representation of x made by the continuous fractions algorithm as outlined in the lecture and included above. Show that for any $1 \leq i \leq n$ (note that 0 is not included), $a_i \geq 1$ (equivalently, it is not 0). Where (and for which i) did we need this claim in the analysis we saw in the lecture?

(c) For integers $a_0 \in \mathbb{N}, a_i \in \mathbb{N}_{\geq 1}$, denote by $p(a_0, \dots, a_k), q(a_0, \dots, a_k)$ the reduced numerator and denominator, respectively, of the fraction represented by $[a_0, \dots, a_k]$ (reduced meaning $GCD(p(a_0, \dots, a_k), q(a_0, \dots, a_k)) = 1$), i.e. $\frac{p(a_0, \dots, a_k)}{q(a_0, \dots, a_k)} = a_0 + \frac{1}{a_1 + \frac{1}{\dots + \frac{1}{a_k}}}$.

Prove that calculating the common denominators bottom-up results in a reduced fraction, and thus in calculating $p(a_0, \dots, a_k)$ and $q(a_0, \dots, a_k)$. Example: $[a_0, a_1, a_2] = a_0 + \frac{1}{a_1 + \frac{1}{a_2}} = a_0 + \frac{1}{\frac{a_1 a_2 + 1}{a_2}} = a_0 + \frac{a_2}{a_1 a_2 + 1} = \frac{(a_1 a_2 + 1) a_0 + a_2}{a_1 a_2 + 1} = \frac{a_0 + a_2 + a_0 a_1 a_2}{a_1 a_2 + 1}$.
hint: Use induction on the number of arguments $k+1$ of p, q. For the induction step look at $[a_1, \dots, a_k]$ and its relation to $[a_0, \dots, a_k]$.

(d) The previous subquestion allows us to understand $p(a_0, \dots, a_k), q(a_0, \dots, a_k)$: conclude from the previous subquestion for which $\ell \in \mathbb{N}$ and $\tilde{a}_i \in \mathbb{N}$ the relation $q(a_0, \dots, a_k) = p(\tilde{a}_0, \dots, \tilde{a}_\ell)$ holds (for $k \geq 1$)? Use this to prove the following recursive formula $p(a_0, \dots, a_k) = a_0 p(a_1, \cdots, a_k) + p(a_2, \dots, a_k)$.

hint: Use the relation you found to get an equivalence of two ways to write $[a_0, \dots, a_k]$.

(e) Use the previous subquestion together with subquestion (b) to show that q_k is a (weakly) increasing sequence. Where did we use this in the lecture?

(f) Use the recursive formula from subquestion (d) to prove by induction that

$$p(a_0, \dots, a_k) = \sum_{\ell=0}^{\lfloor \frac{k}{2} \rfloor} \prod_{\substack{i_1, \dots, i_\ell \in \{0, \dots, k-1\} \\ i_j + 2 \leq i_{j+1}}} a_0 a_1 \cdots a_{i_1 - 1} a_{i_1 + 2} \cdots a_{i_\ell - 1} a_{i_\ell + 2} \cdots a_k$$

is the sum of all the products of the numbers where we took away any number of consecutive pairs.

For example, we saw above that $p(a_0, a_1, a_2) = a_0 + a_2 + a_0 a_1 a_2$. The term $a_0 a_1 a_2$ is the product when no pairs were taken away, and a_0 and a_2 are each a product (of one number) where we took away one consecutive pair. Similarly we saw, $p(a_1, a_2) = 1 + a_1 a_2$ which is a product of no numbers, which is 1, for taking away the only two numbers a_1, a_2 and the term for not taking away any $a_1 a_2$.

(g) Conclude from the previous subquestion that $p(a_0, \dots, a_k) = p(a_k, a_{k-1}, \dots, a_0)$.

Assume we fix some sequence of a_i's and denote by $p_k := p(a_0, \dots, a_k), q_k := q(a_0, \dots, a_k)$. Use the symmetry above and the recursive formula you've proven to show that the following recursion formulas hold: $p_k = a_k p_{k-1} + p_{k-2}$ and $q_k = a_k q_{k-1} + q_{k-2}$.
Where did we use these in the lecture?

(h) Denote $\Delta_k := p_k q_{k-1} - p_{k-1} q_k$. Prove by induction on k that $\Delta_k = (-1)^{k-1}$. (Use $q_0 = 1$ as $[a_0] = \frac{a_0}{1}$)
Where did we use this in the lecture?

2. Here we will prove an upper bound for the probability that the "bad" event happens in Shor's algorithm:

Claim 8.1

Let $N = p_1^{\alpha_1} \cdots p_m^{\alpha_m}$ for different odd primes $p_i \in \mathbb{N}_{\geq 3}$ and natural numbers $\alpha_i \in \mathbb{N}_{\geq 1}$. For a uniform distribution $A \sim \mathbb{Z}_N^*$, where $o_N(A)$ denotes the order of A in $\mathbb{Z}_N^*$ (or just $o(A)$ where the N is clear)[1],

$$\Pr_{A \sim \mathbb{Z}_N^*}[o(A) \textit{ is odd } \vee A^{o(A)/2} \equiv -1 \pmod N] \leq \frac{1}{2^{m-1}}$$

(a) Why is it enough for Shor's factorisation algorithm to consider N's without 2 in their prime factorisation?

(b) Show that for $A \in \mathbb{Z}_N^*$, if $A^r \equiv 1 \pmod N$ then $o_N(A) \mid r$. Conclude that for every $i \in [m]$, $o_{p_i^{\alpha_i}}(A) \mid o_N(A)$.

(c) Let $A \in \mathbb{Z}_N^*$ be such that either $o_N(A)$ is odd or $A^{o_N(A)/2} \equiv -1 \pmod N$. Define $d := \max\{c : 2^c \mid o_N(A)\}$ to be the power of 2 in the factorisation of A's order in $\mathbb{Z}_N^*$ (equivalently, the maximal power of 2 that divides it), and similarly for $i \in [m]$, $d_i := \max\{c : 2^c \mid o_{p_i^{\alpha_i}}(A)\}$ the power of 2 in the prime factorisation of the order of A in $\mathbb{Z}_{p_i^{\alpha_i}}^*$.

Show that $\forall i \in [m] : d_i = d$.

hint: For both cases use the previous subquestion. For the case where $o_N(A)$ is even show how $B \equiv k \pmod N$ implies knowledge about $B \pmod{p_i^{\alpha_i}}$ and think whether an $r \in \mathbb{N}$ for which $A^r \not\equiv 1 \pmod M$ can have $r \mid o_M(A)$?

(d) The Chinese remainder theorem says that the function $\varphi : \mathbb{Z}_N^* \to \mathbb{Z}_{p_1^{\alpha_1}}^* \times \cdots \times \mathbb{Z}_{p_m^{\alpha_m}}^*$ defined as $\varphi(A) := (A \pmod{p_1^{\alpha_1}}, \dots, A \pmod{p_m^{\alpha_m}})$ is an isomorphism, and specifically it is a bijection. Thus, drawing uniformly at random $A \sim \mathbb{Z}_N^*$ is equivalent to drawing uniformly and independently at random its modulo remainders $(A_i \sim \mathbb{Z}_{p_i^{\alpha_i}}^*)_{i \in [m]}$.

From the previous subquestion, in order to prove the claim we need to show that the probability of drawing all of them with the same

[1] i.e. $o(A) \in \mathbb{N}$ s.t. $A^{o_N(A)} \equiv 1 \pmod N$ while $A^r \not\equiv 1 \pmod N$ for any $1 \leq r < o_N(A)$.

d_i is upper-bounded by $\frac{1}{2^{m-1}}$, which is equivalent to bounding the probability to draw $m-1$ of them with a specific d (after drawing the first one which "decides" it $d = d_1$).

Note the following two facts: a consequence of the Chinese remainder theorem is $\text{lcm}\left(o_{p_1^{\alpha_1}}(A), \cdots, o_{p_m^{\alpha_m}}(A)\right) = o_N(A)$ and the fact that $\mathbb{Z}^*_{p_i^{\alpha_i}}$ is cyclic. Use these (without proving them) to show that the probability of a uniform draw $B \sim \mathbb{Z}^*_{p_i^{\alpha_i}}$ to have $d_i = \max\{c : 2^c \mid |\mathbb{Z}^*_{p_i^{\alpha_i}}|\}$ (the power of 2 in the prime factorisation of $|\mathbb{Z}^*_{p_i^{\alpha_i}}|$[2]) is $\frac{1}{2}$. Conclude the claim.

hint: Take an element $g^k \in \mathbb{Z}^*_{p_i^{\alpha_i}}$*. Divide the elements with odd* k *and even* k *and use subquestion (b). For the odd case look at their order. For the even case use the fact that for any* $B \in \mathbb{Z}^*_n$, $B^{|\mathbb{Z}^*_n|} \equiv 1 \pmod{n}$.

3. The QFT circuit we saw in class uses 2-qubit gates. Show that if we want to measure the output of the QFT in the computational basis then we can modify the circuit to use only 1-qubit gates (that are classically controlled). Can we use your modified circuit in Shor's algorithm?

[2]Recall: $|\mathbb{Z}^*_{p_i^{\alpha_i}}| = p_i^{\alpha_i} - p_i^{\alpha_i - 1} = p_i^{\alpha_i - 1}(1 - \frac{1}{p_i})$.