HOMEWORK #3 – MATH 5405
SPRING 2016

DUE: TUESDAY 3/22/2016

(5) Let's describe one more factorization algorithm (you can even implement it before class on March 10th). This is called Pollard's Rho. If p is a factor of n then the expectation is that this will usually find a factor of n in some constant times $\sqrt{p}$ steps.

(a) Lookup the birthday problem on the internet and find an answer the following question. How many randomly chosen numbers x_i in the range $0, \ldots, n-1$ are required so that there is at least a 50% probability that $x_i = x_j$ for some $i \neq j$. You don't need to prove that your answer is right.

Solution: See for instance `https://en.wikipedia.org/wiki/Birthday_problem` .

(b) The idea for us now we create a list of "random" numbers less than n. We do this by picking a polynomial $g(x)$ (usually $g(x) = x^2 + 1$), and then setting

$$x_1 = 2, x_2 = g(x_1) \bmod n, \ldots, x_{i+1} = g(x_i) \bmod n, \ldots$$

We want to quickly find i and j where $x_i \equiv_p x_j$ so that $d = \gcd(x_i - x_j, n) > 1$.

First define a new sequence y_i as follows.

$$y_1 = g(x_1), y_2 = g(g(y_1)) \bmod n, y_3 = g(g(y_2)) \bmod n, \ldots, y_{i+1} = g(g(y_i)) \bmod n, \ldots$$

Prove the following.

Claim: *If t and l are the smallest integers such that $x_t \equiv_p x_{t+l}$ (and hence $x_t = x_{t+2l} = x_{t+3l} = \ldots$), then $x_i \equiv_p y_i$ in at most $t + l$ steps.*

Hint: What is y_i in terms of x_i. Try setting $i = t + l - (t \bmod l)$ and showing that $x_i = x_{2i}$.

Solution: First since $x_t = x_{t+l} = x_{t+2l} = ... = x_{t+ml}$ we see that $x_{t+j} = x_{t+ml+j}$ for any integers m, j. We next observe that $y_1 = x_2$, $y_2 = g(g(y_1)) = g(g(x_2)) = g(x_3) = x_4$, $y_3 = g(g(y_2)) = g(g(x_4)) = g(x_5) = x_6$, $y_4 = g(g(y_3)) = g(g(x_6)) = x_8$. etc. In general $y_i = x_{2i}$. Write $t = ql + r$ with $r = (t \bmod l)$. If we set $i = t + l - r = t + l - t + ql = (q+1)l$, then, if we write

$$2i = 2(q+1)l = (q+1)l + (q+1)l = i + ml.$$

Hence $x_i = x_{i+ml} = x_{2i} = y_i$. This proves the claim.

(c) Now define the sequences x_i and y_i as in part (b). If $x_i \equiv_p y_i$ how could you use that to find find a factor p of n?

Solution: Just compute $\gcd(x_i - y_i, n)$ as usual.

(d) Write pseudo-code for that uses the above method to find factors of n.

Solution: See below

```
Define g(x) = x^2 + 1
x = 2
y = g(x)
d = 1
While d == 1:
    x = g(x)
    y = g(g(y))
    d = gcd(x-y, n)
```