# Pollard's Rho Method

Justin Wyss-Gallifent

February 27, 2023



## 1 Introduction

John Pollard invented the Rho factorization algorithm in 1975. It does a fairly fast job for numbers with small prime factors, even if those numbers themselves are big, and it has a very small memory footprint, so it's a useful tool to do some initial probing.

## 2 Motivation

Given some $n \in \mathbb{Z}$ our goal is to find some factor of $n$. Let's assume that $n$ has a small prime factor called $p$. We won't necessarily find $p$ but it will help us find a factor of $n$.

Suppose we could somehow obtain two integers $x_i$ and $x_j$ with $x_i \equiv x_j \mod p$ but $x_i \not\equiv x_j \mod n$. If we did obtain two such integers then since $p \mid (x_j - x_i)$ and $p \mid n$ and since $n \nmid (x_j - x_i)$ then $p \leq \gcd(x_j - x_i, n) \leq n$ and so $\gcd(x_j - x_i, n)$ would be a nontrivial factor of $n$.

But how can we possibly find such $x_i$ and $x_j$?

Suppose we set $x_1 = 1$ and $p(x) = x^2 + 1$ and then we construct a sequence by reducing:

$$\begin{aligned} x_2 &\equiv p(x_1) \mod n \\ x_3 &\equiv p(x_2) \mod n \\ &\vdots \quad \vdots \end{aligned}$$

This generates a pseudorandom sequence mod $n$. But it repeats mod $p$, and this is key.

Since $p$ is assumed to be small it's likely that we'll quickly find $x_i$ and $x_j$ with $i < j$ such that $x_j - x_i \equiv 0 \mod p$ (because mod $p$ there are only $p$ distinct values) and $x_j - x_i \not\equiv 0 \mod n$. Of course we can't test for this because don't know $p$, but we can test $\gcd(x_j - x_i, n)$ and see if it's something other than 1 and $n$.

But suppose along the way there are in fact $x_i$ and $x_j$ satisfying $x_i \equiv x_j \mod p$, as is likely. Then after index $i$ they will repeat every $j - i$ indices. For example if $i = 22$ and $j = 27$ then they'll repeat every 5 indices from $i = 22$ onwards.

In light of this suppose $s$ is the smallest multiple of $j - i$ which is greater than or equal to $i$. Then since $2s - s = s$ is a multiple of $j - i$ it follows that $x_{2s} \equiv x_s \mod p$. With luck we'll also have $x_{2s} \not\equiv x_s \mod n$ and $x_{2s}$ and $x_s$ will do the job.

So how do we find $s$ if we don't know $i$ or $j$? Well, we just calculate our sequence but only test $x_{2s}$ and $x_s$ when possible.

Of course we can't actually test if $x_{2s} \equiv x_s \mod p$ since we don't know $p$ but we can examine $\gcd(x_{2s} - x_s, n)$ and see if it's other than 1 or $n$.

## 3 Pollard's Rho Method

Given an integer $n$ which we assume has a small factor we choose some $x_0$ (often $x_0 = 2$), and we choose $p(x) = x^2 + 1$. We generate $x_1 = p(x_0)$ reduced $\mod n$, $x_2 = p(x_1)$ reduced $\mod n$, and so on. At each even subscript $x_{2s}$ we calculate $\gcd(x_{2s} - x_s, n)$ and immediately upon obtaining a number greater than 1 we are done.

**Note 3.0.1.** The use of $p(x) = x^2 + 1$ is classically used as it generates good pseudorandom numbers when taken with many moduli.

**Note 3.0.2.** The gcd we find is not necessarily our hypothesized $p$, however $p$ is a divisor of it, and it is not uncommon to actually obtain a prime.

**Note 3.0.3.** It is possible (but in practice unlikely) that our repeats are in fact congruent mod $n$ as well as $p$ which yields $x_{2s} = x_x$ and a gcd of $n$. In such a case the algorithm fails. In such a case we try a different starting value or possibly a different polynomial.

**Example 3.1.** Let's factor $n = 1111$. We set $x_0 = 2$ and $p(x) = x^2 + 1$. We then calculate:

$$
\begin{aligned}
x_1 &\equiv 5 \mod 1111 & & \\
x_2 &\equiv 26 \mod 1111 & & \gcd(26-5, 1111) = 1 \\
x_3 &\equiv 677 \mod 1111 & & \\
x_4 &\equiv 598 \mod 1111 & & \gcd(598-26, 1111) = 11
\end{aligned}
$$

We know 11 is a factor and we're done.

**Example 3.2.** Let's factor $n = 1189$. We set $x_0 = 2$ and $p(x) = x^2 + 1$. We then calculate:

$$
\begin{aligned}
x_1 &\equiv 5 \mod 1189 & & \\
x_2 &\equiv 26 \mod 1189 & & \gcd(26-5, 1189) = 1 \\
x_3 &\equiv 677 \mod 1189 & & \\
x_4 &\equiv 565 \mod 1189 & & \gcd(565-26, 1189) = 1 \\
x_5 &\equiv 574 \mod 1189 & & \\
x_6 &\equiv 124 \mod 1189 & & \gcd(124-677, 1189) = 1 \\
x_7 &\equiv 1109 \mod 1189 & & \\
x_8 &\equiv 456 \mod 1189 & & \gcd(456-565, 1189) = 1 \\
x_9 &\equiv 1051 \mod 1189 & & \\
x_{10} &\equiv 21 \mod 1189 & & \gcd(21-574, 1189) = 1 \\
x_{11} &\equiv 442 \mod 1189 & & \\
x_{12} &\equiv 369 \mod 1189 & & \gcd(369-124, 1189) = 1 \\
x_{13} &\equiv 616 \mod 1189 & & \\
x_{14} &\equiv 166 \mod 1189 & & \gcd(166-1109, 1189) = 41
\end{aligned}
$$

We know 41 is a factor and we're done.

# 4 Nomenclature

The reason that this is called the Rho method is that when we obtain $x_{2s} \equiv x_s \mod p$ we have found $x_j \equiv x_i \mod p$ and we have a cycle. In the previous example $x_{14} \equiv x_7 \mod 41$ and hence because of the cyclic nature we have $x_{15} \equiv x_8 \mod 41$, $x_{16} \equiv x_9 \mod 41$ and so on. Our sequence of $x_i$, taken mod $p$, form the shape of the Greek letter $\rho$.