Rose-Hulman Undergraduate Mathematics Journal

Volume 11
Issue 2

Article 6

Structure and Statistics of the Self-Power Map

Matthew Friedrichsen
St. Olaf College, friedric@stolaf.edu

Brian Larson
Wheaton College

Emily McDowell
University of Pennsylvania

Rose-
Hulman
Undergraduate
Mathematics
Journal

Structure and Statistics of the Self-Power Map

Matthew Friedrichsen[a] Brian Larson[b]
Emily McDowell[c]

Volume 11, No. 2, Fall, 2010

Sponsored by

Rose-Hulman Institute of Technology
Department of Mathematics
Terre Haute, IN 47803
Email: mathjournal@rose-hulman.edu
http://www.rose-hulman.edu/mathjournal

[a]St. Olaf College
[b]Wheaton College
[c]University of Pennsylvania

Rose-Hulman Undergraduate Mathematics Journal
Volume 11, No. 2, Fall, 2010

Structure and Statistics of the Self-Power Map

Matthew Friedrichsen Brian Larson Emily McDowell

Abstract. We investigate the structure of a function relevant to cryptography, given by $f : x \mapsto x^x \bmod p$, for p a prime. We call f the *self-power map*. Given x, it is easy to calculate $f(x) \equiv x^x \pmod{p}$. However, it is thought to be difficult to quickly calculate $f^{-1}(x^x)$. That is, given $x^x \equiv c \pmod{p}$, for a fixed c, it is difficult to quickly solve for x. We call the problem of finding the inverse of the self-power map the *Self-Power Problem*. As a variation of the Discrete Logarithm Problem, the Self-Power Problem is thought to be difficult to solve and therefore considered safe for use in some versions of the ElGamal Digital Signature Algorithm. Nonetheless, utilizing functional graphs to represent the map has revealed non-random structural properties, which we describe primarily through number theory and statistics.

Acknowledgements: This work was done at the Rose-Hulman Institute of Technology Number Theory REU 2010. We thank D. Cloutier, N. Lindle, and A. Hoffman for their code that allowed us to collect our data in a timely manner. We also thank the NSF for the grant that supported our REU program. Finally, we would like to extend our gratitude to Josh Holden for his invaluable guidance and unwavering support throughout this project.

1 Introduction

Modern cryptography is the study of transferring messages digitally between parties in a secure fashion. Cryptosystems considered "secure" are assumed to rest on problems that are difficult to solve. For example, RSA is associated with factoring, while protocols like Diffie-Hellman and ElGamal are associated with the Discrete Logarithm Problem (DLP). Both factoring and the DLP are thought to be very hard problems and are currently well-studied.

Of specific interest to our topic is the ElGamal Digital Signature Algorithm (DSA), a message-signing protocol. Assume Alice would like to send a message M to Bob, and that she needs to sign this message in a fashion that allows Bob to easily verify her identity. To execute this through the ElGamal DSA, Alice first chooses a large prime p and a private key $a \in \mathbb{Z}$, randomly selected from between 1 and $p-2$. She then releases the public key (p, α, y), where α is some primitive root modulo p and $y \equiv \alpha^a \pmod{p}$.

To sign M, Alice chooses a random k between 1 and $p-2$ that is relatively prime to $p-1$. Her signature (r, s) is given by $r \equiv \alpha^k \pmod{p-1}$ and $s \equiv k^{-1}(M - ar) \pmod{p}$.

Bob receives from Alice both the message M and the corresponding signature (r, s). He also knows Alice's public key (p, α, y). To verify that Alice is indeed the sender, Bob computes $v_1 \equiv y^r r^s \pmod{p}$ and $v_2 \equiv \alpha^M \pmod{p}$. If the verification equation holds, that is if $v_1 \equiv v_2 \pmod{p}$, then Bob deems the signature valid.

To forge a signature in Alice's name, Frank the forger must be able to construct a valid verification equation. He needs to find a v_1 and v_2 such that $v_1 \equiv y^r r^s \equiv \alpha^M \equiv v_2 \pmod{p}$, for whatever message M he would like Bob to read. He knows the public key (p, α, y), but without the private key a he can't compute a valid $s \equiv k^{-1}(M - ar) \pmod{p}$. He has the following options. Already knowing $y, \alpha,$ and M, Frank can fix r and rearrange the verification equation to solve the DLP for s:

$$r^s \equiv (y^r)^{-1}\alpha^M \pmod{p}.$$

Since solving the DLP is considered intractable, he then tries to fix s instead. This results in attempting to solve a similar problem to the DLP for r:

$$y^r r^s \equiv \alpha^M \pmod{p}.$$

There are two versions of the ElGamal DSA that rely on, among other things, the difficulty of computing the inverse map of

$$f : x \mapsto x^x \bmod p$$

where p is prime. We call this function f the self-power map. In the first variation of the ElGamal DSA, to take advantage of f, the verification equation is the following:

$$v_1 \equiv (y)^s r^r \equiv \alpha^M \equiv v_2 \pmod{p}.$$

Frank the forger has the following problems to overcome to successfully forge a signature. Still knowing $y, \alpha,$ and M, he can fix r and then rearrange to try to solve the DLP for s:

$$y^s \equiv \alpha^M (r^r)^{-1} \pmod{p}.$$

Else, he can fix s and solve the Self-Power Problem for r:

$$r^r \equiv \alpha^M (y^s)^{-1} \pmod{p}.$$

In the other variation of the ElGamal DSA that uses the Self-Power Problem, the verification equation is

$$v_1 \equiv (y)^M r^r \equiv \alpha^s \equiv v_2 \pmod{p}.$$

Forging a signature still presents similar problems. $y, \alpha,$ and M are still fixed. Frank can fix r and then try to solve the DLP for s:

$$(y)^M r^r \equiv \alpha^s \pmod{p}.$$

Or Frank can fix s and rearrange to solve the Self-Power Problem for r:

$$r^r \equiv (y^M)^{-1} \alpha^s \pmod{p}.$$

While the DLP has been studied extensively, the self-power map, to our knowledge, has seen little attention. Due to the feasibility of applying tools from DLP investigations directly to the Self-Power Problem, and the practical concern of whether or not the Self-Power Problem is in fact a difficult problem, we believe the self-power map merits specific attention. Overall, our theoretical and statistical methods exposed non-random structure within the self-power map that suggests there is more to be discovered and possibly exploited. For example, we can predict where some numbers map to based on their value, as well as rules that apply to large sets of numbers.

In this paper, we first discuss previous work that relates to our analysis of the self-power map. We then establish a few basic definitions, which leads to our investigation of the map using graph theory and number theory. Finally, we discuss the statistical behavior of various graph theoretical characteristics of the self-power map.

2 Previous Work

The self-power map has been previously studied in papers by Crocker [3, 4], and Balog, Broughan, and Shparlinski [1]. These papers investigate the number of distinct residues in the self-power map, and bounds on the number of solutions for x to $x^x \equiv a \mod p$, with a being a fixed residue modulo p.

In [3], Crocker looks specifically at solutions for x to $x^x \equiv 1 \pmod{p}$. He defines primes where only 1 and $p-1$ are solutions as *irreducible* primes. In [4] he gives both a lower and

an upper bound, respectively $[\sqrt{(p-1)/2}]$ and $p-4$, for the number of distinct residues of a given self-power map.

Functional graphs for DLP functions, which are related to the self-power map, have also been examined. Cloutier and Holden [2] described the method of constructing functional graphs for the discrete logarithm. They also established attributes of interest for our graphs, such as the number of components and number of terminal nodes (to be defined later).

Statistics was first applied to these parameters of DLP graphs by Lindle in [8] and later by Hoffman in [7]. None of Lindle's results relate specifically to the self-power map. Nonetheless, his idea of statistically comparing a given class of functional graphs with random functional graphs has served as inspiration for the statistical side of this paper.

Hoffman's work is most closely related to ours. We have reproduced his methodology of data collection and subsequent statistical analysis of functional graph parameters. In fact, we gathered our data by executing a version of his code tailored specifically to the self-power map. We present these findings in the statistical section of our results.

3 Definitions and Methods

In this section we will establish the notation and definitions for terms that will be used throughout the paper. We will first discuss definitions pertaining to Graph Theory and Number Theory and then move on to discuss objects of statistical interest as well as the methods we will use to analyze them.

3.1 Functional Graphs

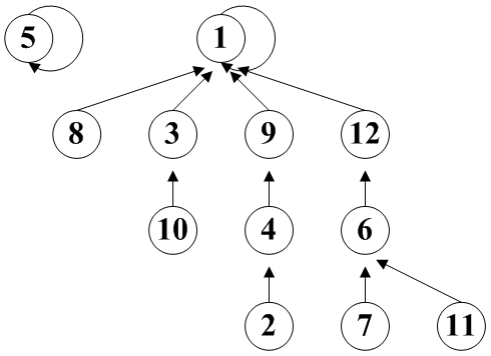
Functional Graph. A functional graph (FG) is a directed graph on the set $\{1, \ldots, n\}$ such that the out-degree of each node is one.

A functional graph allows us to represent a function on a mathematical object, such as $(\mathbb{Z}/p\mathbb{Z})^*$, where each node x represents an element in the domain, $\{1, \ldots, p-1\}$ and the arrow leaving each node points to $f(x)$ for the given function on $(\mathbb{Z}/p\mathbb{Z})^*$. Both the table and the graph representing the self-power function for $p = 13$ can be seen in Figure 1.

It is useful to model the self-power map as a functional graph. This is advantageous visually and mathematically because it allows us to investigate patterns in the graph. Some objects of interest in function graphs are components, cycles, image nodes, and terminal nodes.

Component. A component is a set of nodes that are connected. All components are pairwise disjoint and the components form a partition of the nodes.

In Figure 1, the graph is divided into two components. One component contains only one node, 5, and the other contains the remaining 11 nodes.



x	1	2	3	4	5	6	7	8	9	10	11	12
$f(x)$	1	4	1	9	5	12	6	1	1	3	6	1

Figure 1: $x \mapsto x^x \bmod 13$

Cycle. A cycle is a set of nodes within a component such that following the path from any starting node in the cycle will lead back to the starting node. In functional graphs, there is exactly one cycle within each component. We call cycles with n nodes n-cycles. We will also refer to cycles with one node as fixed points.

In Figure 1, the two cycles are $5 \mapsto 5$ and $1 \mapsto 1$. Both cycles are fixed points in this case.

Tail. A tail is a set of nodes in a component that are connected such that no node is connected to more than one other node and the final node is connected to the cycle of the component.

In Figure 1, the set of nodes $\{7, 6, 12\}$ form a tail because 7 maps to 6, 6 maps to 12, and 12 maps to a cycle. The set of nodes $\{7, 11, 6, 12\}$ do not form a tail because both 7 and 11 map to 6. Then there are five different tails in Figure 1, $\{8\}, \{3, 10\}, \{9, 4, 2\}, \{12, 6, 7\}$, and $\{12, 6, 11\}$.

Image Node. An image node is a node that has at least one incoming arrow. We call the nodes that map to any given image node the pre-images of the image node. We say that the in-degree of a node is the number of pre-images of that node.

In Figure 1, the image nodes are 1, 3, 4, 5, 6, 9, and 12.

Terminal Node. A terminal node is a node that has no incoming arrows, i.e., it is not an image node.

In Figure 1, the terminal nodes are 2, 7, 8, 10, and 11

3.2 Primitive Roots and Quadratic Residues

Quadratic Residue. If m is a positive integer, then the integer n is a quadratic residue of m if $\gcd(n, m) = 1$ and if there exists a solution to $x^2 \equiv n \pmod{m}$.

Primitive Root. Let ϕ be the Euler totient function. If g and m are relatively prime integers with $m > 0$ and if $\phi(m)$ is the smallest positive integer n such that $g^n \equiv 1 \pmod{m}$, then g is a primitive root modulo m.

With regards to number theory, we studied how quadratic residues and primitive roots behave within the graph. This is a natural thing to investigate because primitive roots and quadratic residues are large, disjoint sets of the multiplicative group $(\mathbb{Z}/p\mathbb{Z})^*$. In fact, of the $p-1$ elements of the group, there are $\frac{p-1}{2}$ quadratic residues and $\phi(p-1)$ primitive roots, which together make up a majority of the group.

3.3 Order of a Node

Order. Let m be a positive integer and let $n \in \{1, \ldots, m-1\}$. If $\gcd(n, m) = 1$, we define the order of $n \pmod{m}$, $\mathrm{ord}_m(n)$, as the smallest positive integer d such that $n^d \equiv 1 \pmod{m}$.

We know that for each positive $d \mid (p-1)$ there are exactly $\phi(d)$ many elements in $(\mathbb{Z}/p\mathbb{Z})^*$ of that order. For a prime p, we will define
$S_d = \{n \in (\mathbb{Z}/p\mathbb{Z})^* \mid \mathrm{ord}_p(n) = d\}$, where $d \mid (p-1)$. Because there are exactly $\phi(d)$ elements of order d in $(\mathbb{Z}/p\mathbb{Z})^*$, S_d has $\phi(d)$ elements. We will investigate how the order of a node affects its behavior in the self-power map.

3.4 Statistics

In general, our goal regarding functional graphs is to demonstrate that self-power functional graphs (SPFG's) do not look like random FG's. One efficient way to test this hypothesis is by applying statistical methods to properties of functional graphs that have been examined in both random graphs [5] and in DLP graphs [7, 8] but are not necessarily pliant to our theoretical tools.

From the collection of parameters described by Flajolet and Odlyzko in [5] for random functional graphs, we selected the following as pertinent to our investigations. We also define one additional parameter, denoted by $*$, that has not thus far been addressed in DLP-related literature.

Total Sums.

Number of components. Number of components in a given functional graph.

Number of cyclic nodes. Number of nodes that constitute the cycles.

Number of terminal nodes. Number of nodes without pre-images.

Number of n-cycles. Number of cycles comprising exactly n nodes.

Number of fixed points. Number of 1-cycles.

Total Sums As Seen From a Node.

Total cycle length. For each node in a graph, count the length of the cycle into which its directed path leads. Sum over these values for all $p-1$ nodes.

Total distance to a cycle. For each node in a graph, count the number of edges that must be traversed before reaching a cyclic node. (Let this value be 0 for cyclic nodes.) Sum over these values for all $p-1$ nodes.

Maximal Values.

Maximum cycle length. Number of nodes in largest cycle.

Maximum tail length. Number of nodes in longest tail.

Averages.

Average cycle length. Divide "Total cycle length" by $p-1$.

Average tail length. Divide "Total distance to a cycle" by $p-1$.

Average in-degree.* Expected number of pre-images for a random node, given by $\frac{p-1}{\text{no. image nodes}}$.

We modified code from previous work on the DLP [7] to generate data on the above parameters for the self-power map. Specifically, we calculated these values for maps corresponding to a class of 389 six-digit primes falling between 100,003 and 130,787, and again for a class of 701 larger seven-digit primes between 1,000,003 and 1,037,963. This data was then imported into Minitab for processing.

The primes above were taken in mostly consecutive order, with the exception of our data for *safe primes*.

Safe Prime. A prime p is called a safe prime when it is of the form $2q+1$, where q is also a prime.

Safe primes are significant because they are a popular choice for many security protocols. Because $p-1$ has one large factor q and only two factors total, $(\mathbb{Z}/p\mathbb{Z})^*$ has group structure that is cryptographically "nice". Specifically for protocols dependent upon difficulty of the DLP, it is important for the order of the generator α, namely $p-1$, to not factor into small primes [6]. SPFG's corresponding to safe primes also turned out to be relevant to our results. Because of their relatively sparse distribution within the primes as a whole, our data on safe primes were gathered partially outside of a consecutive range.

	Six-digit	Seven-digit
No. Consecutive Primes	238	599
No. Safe Primes	180	132
Total Primes	**389**	**701**

Statistical Tests. In Minitab, there were three main tests we used to examine our data. We list these and give brief explanation of their function below.

Probability Plot. How is the data distributed for a specific parameter?

T-Test. Does the observed average value of a parameter differ significantly from the expected average value?

Analysis of Variance (ANOVA). Within a given parameter, are the average values significantly different between two or more categories of another variable?

Linear Regression. Does one parameter predict the value of another? How "good" is this prediction?

Now we are ready to investigate the different structural properties of the self-power map using theoretical tools. We begin by describing the basic behaviors of specific nodes in the graph. We then move on to discuss the behavior quadratic residues and primitive roots. Next, we examine the order of a node and its effect on cycles. Finally, we look specifically at fixed points in the self-power map and work towards establishing both upper and lower bounds for the number of fixed points in any particular graph.

4 Basic Behavior

It is evident upon inspection that 1 and $p-1$ always map to 1 for any prime p. Besides these observations about 1 and $p-1$, there exist patterns for the nodes $\frac{p-1}{2}$ and $\frac{p+1}{2}$ as well. Crocker proved some of the patterns for $\frac{p-1}{2}$ in [3], but we will include the full proof here.

Proposition 1. *Let p be prime. Let f denote the self-power map of p. If $p \equiv 1$ or $3 \pmod 8$, then $f(\frac{p-1}{2}) = 1$. If $p \equiv 5$ or $7 \pmod 8$, then $f(\frac{p-1}{2}) = -1$.*

Proof. Consider

$$\begin{aligned} f(\frac{p-1}{2}) &\equiv (\frac{p-1}{2})^{\frac{p-1}{2}} \\ &\equiv (-1)^{\frac{p-1}{2}}(2^{-1})^{\frac{p-1}{2}} \pmod p \end{aligned}$$

Case 1. Assume $p \equiv 1 \pmod 8$. Then $\frac{p-1}{2}$ is even and 2 is a quadratic residue modulo p. Therefore $(-1)^{\frac{p-1}{2}}(2^{-1})^{\frac{p-1}{2}} \equiv (1)(1)^{-1} \equiv 1 \pmod p$ by Euler's criterion.

Case 2. Assume $p \equiv 3 \pmod 8$. Then $\frac{p-1}{2}$ is odd and 2 is not a quadratic residue modulo p. Then $(-1)^{\frac{p-1}{2}}(2^{-1})^{\frac{p-1}{2}} \equiv (-1)(-1)^{-1} \equiv 1 \pmod p$ by Euler's criterion.

Case 3. Assume $p \equiv 5 \pmod 8$. Then $\frac{p-1}{2}$ is even and 2 is not a quadratic residue modulo p. Therefore $(-1)^{\frac{p-1}{2}}(2^{-1})^{\frac{p-1}{2}} \equiv (1)(-1)^{-1} \equiv -1 \pmod p$ by Euler's criterion.

Case 4. Assume $p \equiv 7 \pmod 8$. Then $\frac{p-1}{2}$ is odd and 2 is a quadratic residue modulo p. Therefore $(-1)^{\frac{p-1}{2}}(2^{-1})^{\frac{p-1}{2}} \equiv (-1)(1)^{-1} \equiv -1 \pmod p$ by Euler's criterion.

□

As an example, in Figure 1 we see that $\frac{p-1}{2} = 6$ maps to $-1 \equiv 12$, because $p \equiv 13 \equiv 5 \pmod 8$.

Proposition 2. *Let p be prime. If $p \equiv 1$ or $7 \pmod 8$, then $f(\frac{p+1}{2}) = \frac{p+1}{2}$. If $p \equiv 3$ or $5 \pmod 8$, then $f(\frac{p+1}{2}) = \frac{p-1}{2}$.*

Proof. Consider

$$\begin{aligned} f(\frac{p+1}{2}) &\equiv (\frac{p+1}{2})^{\frac{p+1}{2}} \\ &\equiv (\frac{p+1}{2})^{\frac{p-1}{2}+1} \\ &\equiv (\frac{p+1}{2})(\frac{p+1}{2})^{\frac{p-1}{2}} \\ &\equiv (\frac{p+1}{2})(1)(2^{-1})^{\frac{p-1}{2}} \pmod p \end{aligned}$$

Case 1. Assume $p \equiv 1 \pmod 8$. Then 2 is a quadratic residue modulo p. Hence $(\frac{p+1}{2})(2^{-1})^{\frac{p-1}{2}} \equiv (\frac{p+1}{2})(1)^{-1} \equiv \frac{p+1}{2}$ by Euler's Criterion.

Case 2. Assume $p \equiv 3 \pmod 8$. Then 2 is not a quadratic residue modulo p. Hence $(\frac{p+1}{2})(2^{-1})^{\frac{p-1}{2}} \equiv (\frac{p+1}{2})(-1)^{-1} \equiv \frac{p-1}{2}$ by Euler's Criterion.

Case 3. Assume $p \equiv 5 \pmod 8$. Then 2 is not a quadratic residue modulo p. Hence $(\frac{p+1}{2})(2^{-1})^{\frac{p-1}{2}} \equiv (\frac{p+1}{2})(-1)^{-1} \equiv \frac{p-1}{2}$ by Euler's Criterion.

Case 4. Assume $p \equiv 7 \pmod 8$. Then 2 is a quadratic residue modulo p. Hence $(\frac{p+1}{2})(2^{-1})^{\frac{p-1}{2}} \equiv (\frac{p+1}{2})(1)^{-1} \equiv \frac{p+1}{2}$ by Euler's Criterion.

□

Back to Figure 1 for an example, $\frac{p+1}{2} = 7$ maps to $\frac{p-1}{2} = 6$ because $p \equiv 13 \equiv 5 \pmod 8$.

5 Quadratic Residues and Primitive Roots in the Self-Power Map

Now that we have described patterns for $\frac{p+1}{2}$ and $\frac{p-1}{2}$, we will turn our attention to larger sets of nodes within the self-power map: quadratic residues and primitive roots. We start by looking at quadratic residues:

Proposition 3. *Let p be prime. Let $n \in (\mathbb{Z}/p\mathbb{Z})^*$ such that n is a quadratic residue. Then n^n is also a quadratic residue modulo p.*

Proof. Since n is a quadratic residue, then $x^2 \equiv n \pmod{p}$ for some $x \in (\mathbb{Z}/p\mathbb{Z})^*$. Since $n \equiv x^2 \pmod{p}$, then

$$n^n \equiv (x^2)^{x^2} \equiv (x^{x^2})^2 \pmod{p}$$

Hence $n^n \pmod{p}$ is a quadratic residue. □

From this fact it is clear that the existence of at least one quadratic residue in a component implies that the cycle of that component will consist completely of quadratic residues. For example, in Figure 2, 10 maps to 3 which maps to 1, and all are quadratic residues. The cycle of their component is a quadratic residue. The next observation gives another condition where $n^n \pmod{p}$ is a quadratic residue.

Proposition 4. *Let p be prime. Let $n \in (\mathbb{Z}/p\mathbb{Z})^*$ such that n is even. Then n^n is a quadratic residue modulo p.*

Proof. Since n is even, $n = 2a$ for some $a \in (\mathbb{Z}/p\mathbb{Z})^*$. Consider

$$n^n \equiv (2a)^{2a} \equiv ((2a)^a)^2 \pmod{p}$$

Hence $n^n \pmod{p}$ is a quadratic residue. □

In Figure 2, we see that all nodes that are even map to a quadratic residue.

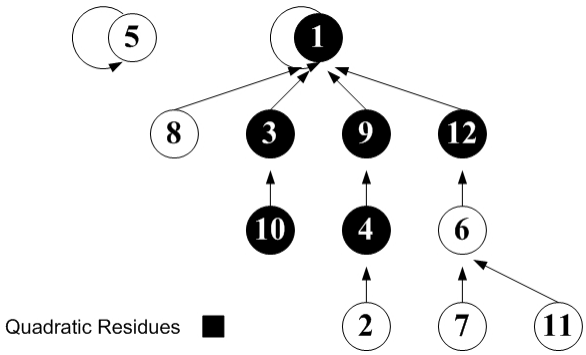


Figure 2: $x \mapsto x^x \bmod 13$

Quadratic residues are not the only nodes to follow these patterns. In fact, any k-th power residue behaves in an analogous fashion. A k-th power residue is a number $a \in (\mathbb{Z}/p\mathbb{Z})^*$ such that there exists $x \in (\mathbb{Z}/p\mathbb{Z})^*$ where $x^k \equiv a \pmod{p}$. For each $k \mid (p-1)$ there will be $\frac{p-1}{k}$ many k-th power residues.

Now we will examine the set of primitive roots within the self-power map. These nodes behave differently from quadratic residues and k-th power residues because primitive roots are not guaranteed to map to other primitive roots.

Proposition 5. *Let p be prime. A primitive root is an image node in the self-power map if and only if it is mapped to by a primitive root that is relatively prime to $p-1$.*

Proof. We first prove in the forward direction. Let g be a primitive root modulo p that is an image node. Then there exists $x \in (\mathbb{Z}/p\mathbb{Z})^*$ such that $x^x \equiv g \pmod{p}$. Since $x^x \equiv g \pmod{p}$, then

$$\operatorname{ord}_p(x^x) = \operatorname{ord}_p(g) = p-1.$$

But

$$\operatorname{ord}_p(x^x) = \frac{\operatorname{ord}_p(x)}{\gcd(\operatorname{ord}_p(x), x)}.$$

Then

$$\operatorname{ord}_p(x) = \operatorname{ord}_p(x^x)\gcd(\operatorname{ord}_p(x), x).$$

Since $\operatorname{ord}_p(x^x) = p-1$ and $p-1$ is the largest possible order for an element in $(\mathbb{Z}/p\mathbb{Z})^*$, then $\operatorname{ord}_p(x) = p-1$. Hence x is a primitive root. Since x is a primitive root and x^x is a primitive root, then $\gcd(x, p-1) = 1$ else x^x would not be a primitive root.

We now prove in the alternate direction. Let g be a primitive root modulo p such that $\gcd(g, p-1) = 1$. Since $\gcd(g, p-1) = 1$ and g is a primitive root, then g^g is a primitive root.

□

In Figure 3, 6 is the only image node that is also a primitive root. Its pre-images are 7 and 11, which are both relatively prime to $12 = p-1$.

Proposition 6. *Let $p > 3$ be prime. A primitive root can never be a fixed point in the self-power map for p.*

Proof. Let g be a primitive root modulo p. Assume $g^g \equiv g \pmod{p}$. Since $g^g \equiv g \pmod{p}$, then $g^{g-1} \equiv 1 \pmod{p}$. But this is a contradiction since $g-1 \neq p-1$ and g is a primitive root. Hence $g^g \not\equiv g \pmod{p}$. □

From these two propositions we gain the fact that a primitive root can only be either a terminal node or the image node of a different primitive root in the self-power map. Also, by a similar proof to the last proposition, we see that a primitive root will never be a pre-image of 1 in the self-power map. Looking at Figure 3, we can see that none of the primitive roots are fixed points or map to 1.

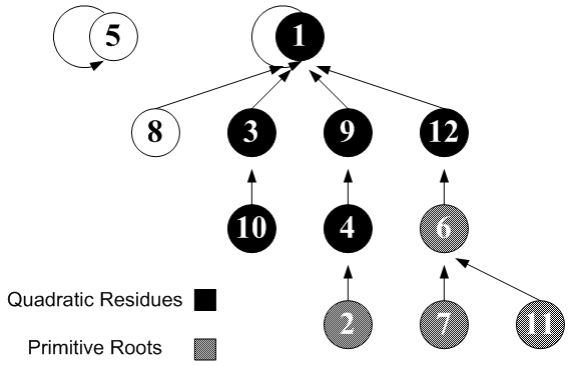


Figure 3: $x \mapsto x^x \bmod 13$

6 Effects of the Order of a Node

Now that we have looked at sets of nodes in the self-power map, we will look at how order affects the action of each separate node in the self-power map.

Proposition 7. *Let p be prime. A node n is a fixed point in the self-power map if and only if $\operatorname{ord}_p(n) \mid (n-1)$.*

Proof. We first prove in the forward direction. Let $n \in (\mathbb{Z}/p\mathbb{Z})^*$. Assume $n^n \equiv n \pmod{p}$. Then $n^{n-1} \equiv 1 \pmod{p}$. Hence $\operatorname{ord}_p(n) \mid (n-1)$.

Now we prove in the alternate direction. Let $n \in (\mathbb{Z}/p\mathbb{Z})^*$. Assume $\operatorname{ord}_p(n) \mid (n-1)$. Then $n^{n-1} \equiv 1 \pmod{p}$. Hence $n^n \equiv n \pmod{p}$.

□

Therefore, fixed points are determined by the orders of the nodes within the self-power functional graph. In Figure 4, the fixed points are 5 and 1. We see that $\operatorname{ord}_{13}(5) = 4$ divides 4, and also $\operatorname{ord}_{13}(1) = 1$ divides 1. Next, we show that the order of a node always divides the respective orders of its pre-images.

Proposition 8. *The order of a node m divides the respective orders of its pre-images in the self-power functional graph of p.*

Proof. Let p be prime. Let $n, m \in (\mathbb{Z}/p\mathbb{Z})^*$ be such that $n^n \equiv m \pmod{p}$. It follows that

$$\operatorname{ord}_p(m) = \frac{\operatorname{ord}_p(n)}{\gcd(n, \operatorname{ord}_p(n))}$$
$$\operatorname{ord}_p(m) \gcd(n, \operatorname{ord}_p(n)) = \operatorname{ord}_p(n).$$

Hence, $\operatorname{ord}_p(m) \mid \operatorname{ord}_p(n)$. □

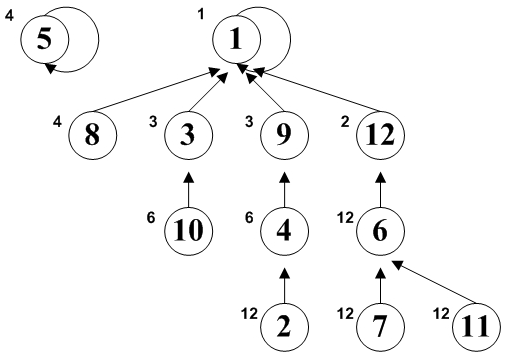


Figure 4: $x \mapsto x^x \bmod 13$, where superscripts denote the order of a node.

For example, in Figure 4 we see that $2 \mapsto 4 \mapsto 9$ and that $\mathrm{ord}_{13}(9)$ divides $\mathrm{ord}_{13}(4)$ which divides $\mathrm{ord}_{13}(2)$.

Because the order of x^x is

$$\frac{\mathrm{ord}_p(x)}{\gcd(\mathrm{ord}_p(x), x)},$$

we can calculate the order of the image of a node. This allows us to start with a terminal node and work our way inwards in a component of the self-power map, determining the orders of the image nodes as we go. This fact also shows that a node n with $\gcd(\mathrm{ord}_p(n), n) = 1$ will map to another node of the same order. The next corollary is a special case of this behavior.

Corollary 9. *Let p be prime and let $a_1, a_2, \ldots, a_n$ be the nodes of an n-cycle, where $f(a_1) = a_2, f(a_2) = a_3, \ldots, f(a_{n-1}) = a_n$, and $f(a_n) = a_1$. Then all the nodes of the n-cycle have the same order.*

Proof. By Proposition 8,

$$\mathrm{ord}_p(a_1) \mid \mathrm{ord}_p(a_n), \mathrm{ord}_p(a_n) \mid \mathrm{ord}_p(a_{n-1}), \ldots, \mathrm{ord}_p(a_2) \mid \mathrm{ord}_p(a_1).$$

By transitivity, $\mathrm{ord}_p(a_k) | \mathrm{ord}_p(a_i)$ and $\mathrm{ord}_p(a_i) | \mathrm{ord}_p(a_k)$. Therefore all the nodes in the cycle have the same order.

□

From this corollary and the proof of Proposition 8 we know all nodes in a cycle have the same order and are relatively prime to their order.

As an example we will consider the self-power map for $p = 47$. In this graph, the nodes $11, 39$, and 43 form a cycle, where $11 \mapsto 39 \mapsto 43$ and $43 \mapsto 11$. After computation we find that the order of each of these nodes is 46. Notice also that $11, 39$ and 43 are all relatively prime to 46.

Now we look at one more condition of nodes in a cycle:

Theorem 10. *Let p be prime. Let $a_1, a_2, \ldots, a_n$ be the nodes of an n-cycle in the self-power map for p. Then*

$$\prod_{k=1}^{n} a_k \equiv 1 \pmod{\operatorname{ord}_p[a_n]}.$$

The notation $[a_n]$ denotes the order of any node in the n-cycle since the order of all nodes in the same cycle is the same by Corollary 9.

Proof. Let $f(a_1) = a_2, f(a_2) = a_3, \ldots, f(a_{n-1}) = a_n$, and $f(a_n) = a_1$. Since $a_1^{a_1} \equiv a_2$ and $a_2^{a_2} \equiv a_3 \pmod{p}$, then

$$(a_1)^{a_1 a_2} \equiv a_3 \pmod{p}.$$

Continuing the argument, we get $(a_1)^{a_1 a_2 \ldots a_n} \equiv a_1 \pmod{p}$. Then

$$(a_1)^{a_1 a_2 \ldots a_n - 1} \equiv 1 \pmod{p}.$$

Thus, $\operatorname{ord}_p[a_n] \mid (a_1 a_2 \ldots a_n - 1)$, and $\prod_{k=1}^{n} a_k \equiv 1 \pmod{\operatorname{ord}_p[a_n]}$. □

For example, we consider again the self-power map for $p = 47$ and the cycle consisting of $11, 39, 43$. We know that the order of each of these nodes is 46. By multiplying, we get $11 \cdot 39 \cdot 43 = 18447 \equiv 1 \pmod{46}$.

Unfortunately, Theorem 10 does not work in the reverse direction. If you can find a set of nodes with the same order, that are relatively prime to their order, and their product is 1 modulo their order, the nodes are not guaranteed to form a cycle in the self-power map. On the other hand, given this information, it is possible to construct an upper bound on the number of components and the maximum cycle size in the self-power map for a given prime p.

From this knowledge we can also establish a condition under which nodes of certain orders will form a cycle in the self-power map. This is based on whether or not S_d contains any pre-images of 1.

Theorem 11. *Let d and p be primes such that $d \mid (p-1)$. If there does not exist $x \in S_d$ such that $x^x \equiv 1 \pmod{p}$, then there exists at least one cycle composed of nodes of order d in the self-power map for p.*

Proof. Assume there does not exist $x \in S_d$ such that $x^x \equiv 1 \pmod{p}$. Note that $\operatorname{ord}_p(1) = 1$, so if there exists $x \in S_d$ such that $x^x \equiv 1 \pmod{1}$, then

$$\frac{d}{\gcd(d, x)} = 1.$$

This would imply $\gcd(d, x) = d$. However, since d is prime and $x^x \not\equiv 1 \pmod{p}$ for all $x \in S_d$, then $\gcd(d, x) = 1$for all $x \in S_d$. This means that for any $x \in S_d$, $x^x \in S_d$ because

$$\operatorname{ord}_p(x^x) = \frac{\operatorname{ord}_p(x)}{\gcd(\operatorname{ord}_p(x), x)} = \frac{d}{\gcd(d, x)} = d.$$

We know that S_d is finite because it has exactly $\phi(d)$ elements. Since S_d is finite and for every $x \in S_d, x^x \in S_d$, then some of the nodes in S_d must form a cycle. □

As an example, let $p = 47$. The prime divisors of $p-1$ are 2 and 23. The only element of order 2 is $p-1$, which maps to 1. By calculation we see that $p-1$, with order 2, is the only pre-image of 1. Therefore, since there are no pre-images of 1 with order 23, there must be a cycle with nodes of order 23, and in fact there are two. 21 and 34 form a cycle of size 2, and their orders are 23. Also, 24 is a fixed point with order 23.

It should be noted that the presence of a pre-image of 1 in S_d does not prevent other nodes in S_d from forming a cycle. In the next section, we will see how knowing the pre-images of 1 and $p-1$ determine a certain type of cycle, the fixed point.

7 Investigations of Fixed Points and Pre-images of 1 and $p-1$

We are now going to focus on fixed points in the self-power map and examine how they are connected to pre-images of 1 and $p-1$.

Theorem 12. *Let p be prime and $n \in (\mathbb{Z}/p\mathbb{Z})^*$ be such that $n^n \equiv n \pmod{p}$. Then*

$$(p-n)^{(p-n)} \equiv (-1)^{n+1} \pmod{p}$$

Proof. Let $n \in (\mathbb{Z}/p\mathbb{Z})^*$ be such that $n^n \equiv n \pmod{p}$. Note that

$$\begin{aligned}(p-n)^{(p-n)} &\equiv (p-n)^{(p-1-n+1)} \\ &\equiv (p-n)^{p-1}(p-n)^{-n}(p-n) \\ &\equiv (p-n)^{-n}(p-n) \pmod{p}\end{aligned}$$

Case 1. Assume n is odd. Then $(p-n)^{-n} \equiv -(n^{-n}) \equiv -n^{-1} \pmod{p}$, because $n^n \equiv n \pmod{p}$. Thus,

$$(p-n)^{-n}(p-n) \equiv -n^{-1}(p-n) \equiv -n^{-1}p + 1 \equiv 1 \pmod{p}.$$

Case 2. Assume n is even. Then $(p-n)^{-n} \equiv n^{-n} \equiv n^{-1} \pmod{p}$, because $n^n \equiv n \pmod{p}$. Therefore,

$$(p-n)^{-n}(p-n) \equiv n^{-1}(p-n) \equiv n^{-1}p - 1 \equiv -1 \pmod{p}.$$

□

As an example, we consider the self-power map for $p = 41$. The fixed points of this map are $1, 9, 16, 21, 31$ and the additive inverses of these nodes are $40, 32, 25, 20, 10$ respectively. For 16, which is the only even fixed point, its additive inverse 25 is a pre-image of $p-1$. For

the rest of the fixed points, since they are odd, their additive inverses are all pre-images of 1.

Theorem 12 dictates that the additive inverses of fixed points be pre-images of either 1 or $p-1$. It is interesting to know how the additive inverses of pre-images of 1 and $p-1$ behave in general. Investigation reveals that these nodes also behave in a predictable pattern.

Corollary 13. *Let p be prime and $n \in (\mathbb{Z}/p\mathbb{Z})^*$ be such that $n^n \equiv 1 \pmod{p}$. Then*

$$(p-n)^{(p-n)} \equiv n(-1)^{n+1} \pmod{p}.$$

Proof. Let $n \in (\mathbb{Z}/p\mathbb{Z})^*$ be such that $n^n \equiv 1 \pmod{p}$. As shown in Theorem 11,

$$(p-n)^{(p-n)} \equiv (p-n)^{-n}(p-n) \pmod{p}.$$

Case 1. Assume n is odd. Then $(p-n)^{-n} \equiv -(n^{-n}) \equiv -1 \pmod{p}$. Thus,

$$(p-n)^{-n}(p-n) \equiv -1(p-n) \equiv n-p \equiv n \pmod{p}.$$

Case 2. Assume n is even. Then $(p-n)^{-n} \equiv n^{-n} \equiv 1^{-1} \equiv 1 \pmod{p}$. Therefore,

$$(p-n)^{-n}(p-n) \equiv 1(p-n) \equiv p-n \equiv -n \pmod{p}.$$

□

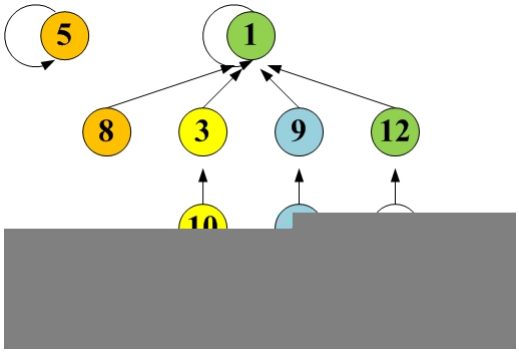


Figure 5: $x \mapsto x^x \bmod 13$, where pre-images of 1 and their additive inverses are highlighted.

Figure 5 illustrates Corollary 13. As an example, 8 is a pre-image of 1, and its additive inverse, 5, is a fixed point. This is because

$$(-8)^{-8} \equiv 8(-1)^9 \equiv 5 \pmod{13}.$$

Corollary 14. *Let p be a prime and $n \in (\mathbb{Z}/p\mathbb{Z})^*$ be such that $n^n \equiv -1 \pmod{p}$. Then*

$$(p-n)^{(p-n)} \equiv n(-1)^n \pmod{p}.$$

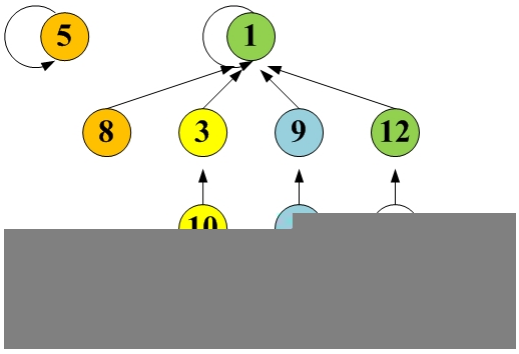


Figure 6: $x \mapsto x^x \bmod 13$, where the pre-image of -1 and its additive inverse are highlighted.

The proof of Corollary 14 is similar to the proof of Corollary 13. In Figure 6, we see that 6 is a pre-image of $-1 \equiv 12 \pmod{13}$. 7 maps to 6 because $(-6)^{-6} \equiv 6(-1)^6 \equiv 6 \pmod{13}$.

Using both of these corollaries and Theorem 12, it is possible to put bounds on either the number of components or the number of pre-images of 1 and $p-1$, depending on which information is known. With knowledge about the number of fixed points, one has a lower bound on both the total number of components and the total number of pre-images of 1 and $p-1$. With exact information on which nodes are fixed points, it is known which additive inverses will map to 1 and $p-1$. From the other direction, if the number of pre-images of 1 and $p-1$ is known, then an upper bound on the number of fixed points follows. And with exact information on which nodes are pre-images one can obtain an exact number of fixed points. Thus, significant information about the structure of the self-power functional graph is gained from solving $x^x \equiv x, x^x \equiv 1$ or $x^x \equiv p-1 \pmod{p}$ for x. If information is known about all the nodes and their orders, you will be able to pick out which nodes map to 1 and which map to $p-1$ with these facts:

Proposition 15. *Let p be prime. A node n maps to 1 in the self-power map of p if and only if $\mathrm{ord}_p(n) \mid n$.*

Proof. We first prove in the forward direction. Assume $n^n \equiv 1 \pmod{p}$. Then $\mathrm{ord}_p(n) \mid n$.

Now we prove in the alternate direction. Assume $\mathrm{ord}_p(n) \mid n$. Then $n^n \equiv 1 \pmod{1}$. □

In Figure 4, we see as an example of Proposition 15 that 9 maps to 1 and $\mathrm{ord}_{13}(9) \mid 9$.

Proposition 16. *Let p be prime. A node n maps to $p-1$ in the self-power map of p if and only if $\mathrm{ord}_p(n) \mid 2n$ and $\mathrm{ord}_p(n) \nmid n$.*

Proof. We first prove in the forward direction. Assume $n^n \equiv p-1 \pmod{p}$. Note that $\mathrm{ord}_p(n) \nmid n$ else $n^n \equiv 1 \pmod{p}$. Since $n^n \equiv p-1 \pmod{p}$, then $(n^n)^2 \equiv (p-1)^2 \pmod{p}$. But $(n^n)^2 = n^{2n}$ and $(p-1) \equiv 1 \pmod{p}$. Thus $n^{2n} \equiv 1 \pmod{p}$. Hence $\mathrm{ord}_p(n) \mid 2n$.

Now we prove in the alternate direction. Assume $\text{ord}_p(n) \mid 2n$ and $\text{ord}_p(n) \nmid n$. Since $\text{ord}_p(n) \mid 2n$, then $n^{2n} \equiv 1 \pmod{p}$. But $n^{2n} = (n^n)^2$. So $(n^n)^2 \equiv 1 \pmod{p}$. Then either $n^n \equiv 1 \pmod{p}$ or $n^n \equiv p-1 \pmod{p}$. But $\text{ord}_p(n) \nmid n$, so $n^n \not\equiv 1 \pmod{p}$. Hence $n^n \equiv p-1 \pmod{p}$.

□

Back to Figure 4, 6 is the only node to map to $p-1 \equiv 12 \pmod{13}$ because $\text{ord}_{13}(6) \mid 12$ and $\text{ord}_{13}(6) \nmid 6$.

With this knowledge one can determine which elements of $(\mathbb{Z}/p\mathbb{Z})^*$ are pre-images of 1 or $p-1$, and with the information about the factorization of $p-1$, it is possible to construct separate upper bounds for both the pre-images of 1 and the pre-images of $p-1$.

Proposition 17. *An upper bound on the number of pre-images of* 1 *is*

$$\sum_{d|(p-1)} \min\{\phi(d), \frac{p-1}{d} - 1\}, \text{ for positive } d \mid (p-1).$$

Proof. By Proposition 15, we know that the order of a pre-image of 1 must divide the value of the pre-image. Therefore, pre-images of 1 must be multiples of d for $d \mid (p-1)$. We know there are $\frac{p-1}{d}$ multiples of d. But we know there are only $\phi(d)$ elements of order d. Also, since $p-1$ always maps to 1 and is always a multiple of d, we take the minimun of $\phi(d)$ and $\frac{p-1}{d} - 1$. By summing $\min\{\phi(d), \frac{p-1}{d} - 1\}$ for all positive $d \mid (p-1)$ we get an upper bound on the number of pre-images of 1. □

Proposition 18. *An upper bound on the number of pre-images of* $p-1$ *is*

$$-1 + \sum_{d|(p-1)} \min\{\phi(d), \frac{p-1}{d}\}, \text{ for positive } d \mid (p-1), \text{ where } d \text{ is even.}$$

Proof. First, by Proposition 8, note that nodes with odd order cannot be pre-images of $p-1$ because 2, the order of $p-1$, does not divide odd numbers. By Proposition 16, we know that the order of a pre-image of $p-1$ must divide twice the value of the pre-image. Therefore, pre-images of $p-1$ are half of multiples of even d for $d \mid (p-1)$. We know that there are $\frac{p-1}{d}$ multiples of d. But we know there are only $\phi(d)$ elements of order d. Thus we take the minimum of $\phi(d)$ and $\frac{p-1}{d}$. By summing $\min\{\phi(d), \frac{p-1}{d}\}$ for all positive and even $d \mid (p-1)$ we get an upper bound on the number of pre-images of 1. We subtract 1 from this because $p-1$ is the only element of order 2 and $p-1$ always maps to 1. □

As an example, we will use these equations to find upper bounds on the pre-images of 1 and $p-1$ in the self-power map for 13. The divisors of 12 are $1, 2, 3, 4, 6, 12$. Summing

$$\sum_{d \mid 12} \min\{\phi(d), \frac{12}{d} - 1\}$$

for these divisors yields a value of 7 as an upper bound on the pre-images of 1. Likewise, summing

$$-1 + \sum_{d \mid 12} \min\{\phi(d), \frac{12}{d}\}$$

for the even divisors of 12 yields a value of 5 as an upper bound on the pre-images of $p-1$. Putting together the two results, we get 12 as a total upper bound on the number of pre-images of 1 and $p-1$. However there are only 12 nodes in the self-power map for 13 and we know that all 12 nodes cannot be either a pre-image of 1 or $p-1$ since we already know of at least one node, $\frac{p+1}{2}$, that is never a pre-image of 1 or $p-1$. Therefore, we would like to be able to construct a more accurate upper bound for the pre-images of 1 and $p-1$. [1] offers an upper bound on the number of pre-images for any image node in the self-power map but not specifically for 1 and $p-1$.

8 Statistical Analysis[9]

We now turn to statistically examining the structural parameters of self-power functional graphs. Each of the tests we utilized has a corresponding p-value, which indicates the likelihood that a positive finding is simply the result of chance. Note that for this paper, we will consider a finding significant when its corresponding p-value is 0.05 or less, meaning there is a 5% or lower probability that we are reporting a random instead of systematic result.

Given that cycles are one of our main theoretical discussion points, and that each component contains exactly one cycle, the number of components in a functional graph was a parameter of primary interest. We first examined the distribution of number of components for our group of 238 consecutive six-digit primes. Literature on random functional graphs led us to expect this would be a normal distribution [5]. Nonetheless, a probability plot showed that a normal distribution is a very poor fit, with $p < 0.005$. The data instead conforms to a lognormal distribution as seen in Figure 7, the probability plot reporting $p = 0.526$.

Furthermore, having observed a large number of fixed points in hand-drawn graphs, we suspected that fixed points account for a large number of the total cycles in SPFG's. To this end, we ran a linear regression that showed the number of fixed points is an excellent predictor of the number of components. The R-squared (R^2) value 92.8% means that approximately 93% of the variation in the data on number of components is accounted for by the number of fixed points, indicating a very strong correlation (Figure 8).

We also found that the average proportion of fixed points to total cycles is significantly larger for SPFG's than random FG's. From [5], the expected proportion for random functional graphs on n nodes is given by $\frac{1}{2}\log(n)$. We calculated this expected value by letting n be the average over our 238 six-digit primes. Our actual value was obtained by averaging the observed proportion of fixed points to cycles for these same primes. We then applied a

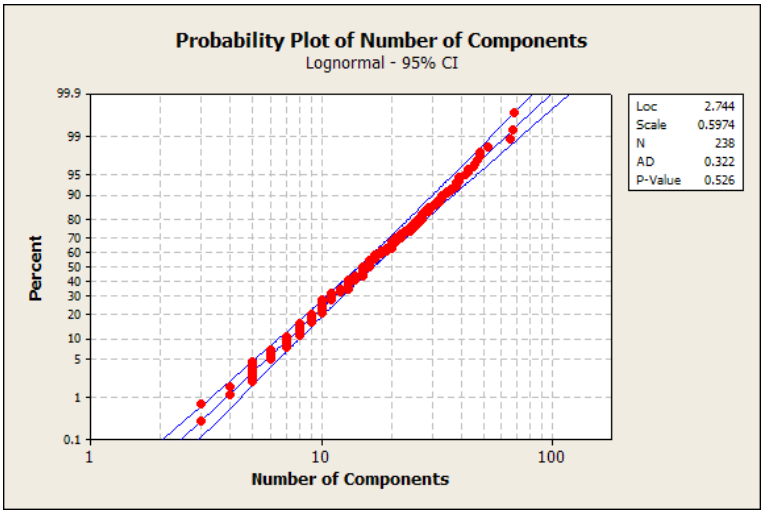


Figure 7: Lognormal Distribution of No. Components

t-test to compare the averages (Figure 9), which indicated a positive result ($p < 0.005$).

In an effort to confirm our findings thus far, we ran the above tests again for our class of 599 consecutive seven-digit primes. The lognormal probability plot was significantly poor ($p = 0.011$), especially compared to the six-digit primes (recall $p = 0.526$). Despite this, fixed points continued to be a good predictor of the number of components ($R^2 = 92.5\%$), and the average proportion of fixed points to total cycles was still significantly different from the average on random FG's ($p < 0.005$, Figure 9). This suggested the need to find another factor to account for this effect in the distribution.

Because of the cryptographic significance of safe primes mentioned earlier, we decided to investigate their role as possible effect contributors. Again considering the set of seven-digit primes, we split the data into safe prime and non-safe prime groups and repeated our distribution analysis on them separately. Once safe primes are removed from the set, the non-safe primes did follow a lognormal distribution more closely ($p = 0.051$).

The safe primes looked more like a normal than lognormal distribution ($p = 0.126$ and $p = 0.023$, respectively). There were, however, only 30 safe primes in this consecutively generated data set. If we consider a larger set of 132 seven-digit safe primes, neither the normal nor lognormal distributions are acceptable ($p < 0.005$ for both). The results are identical for normal and lognormal plots for our set of 180 six-digit safe primes. This supports the idea that safe primes are a more secure choice, since their corresponding graphs may be less predictable. For illustrative purposes, we present a lognormal probability plot for the seven-digit safe primes (Figure 10).

We note how choppy this plot looks, reminiscent of a bar chart, which suggests that

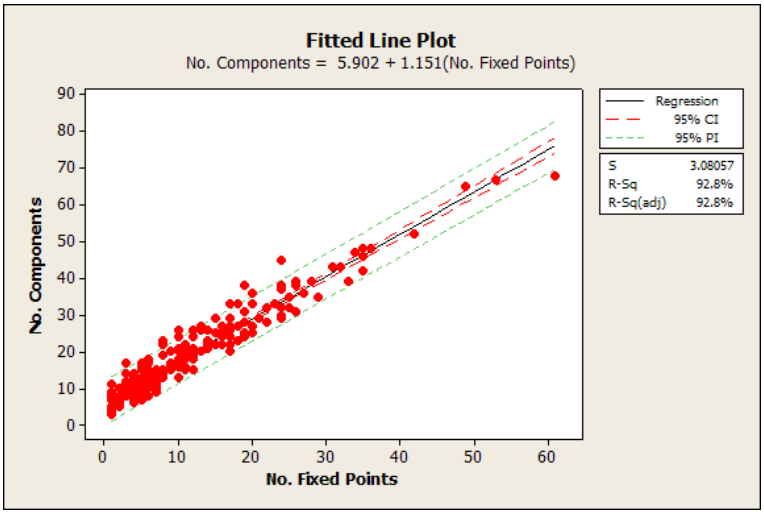


Figure 8: No. Components versus No. Fixed Points

	Expected Average	Observed Average	p-value
six-digit primes	0.1735	0.5248	$<$**0.005**
seven-digit primes	0.1447	0.4958	$<$**0.005**

Figure 9: t-Test: No. Fixed Points over No. Total Cycles

discrete distributions may be a better fit for safe primes. Preliminary testing on the set of six-digit safe primes shows promise in this area, with a χ^2-goodness-of-fit test for the Poisson distribution coming out much better than the continuous distributions ($p = 0.198$). But the test result drops precipitously in confidence when applied to our seven-digit safe primes ($p < 0.005$). Discrete distributions with respect to safe primes likely merit further investigation.

To improve on the idea of separating our data into safe prime and non-safe prime categories, we can consider the effect of the number of divisors of $p - 1$ on the number of components. This is a natural extension, since safe primes are considered good cryptographic choices for the fact that $p - 1$ has so few divisors. It is intuitive that there may be a graded effect on SPFG structure for the number of divisors in general.

We conducted an ANOVA test on our 599 consecutive seven-digit primes for number of components versus total divisors of $p - 1$ (we will call this number a "divisor class"). The ANOVA gives us two measurements. First, it gives us a p-value corresponding to whether or not the average values (means) for number of components are significantly different between divisor classes. Second, we get an R^2 value that states how good of a predictor the divisor classes are for number of components.

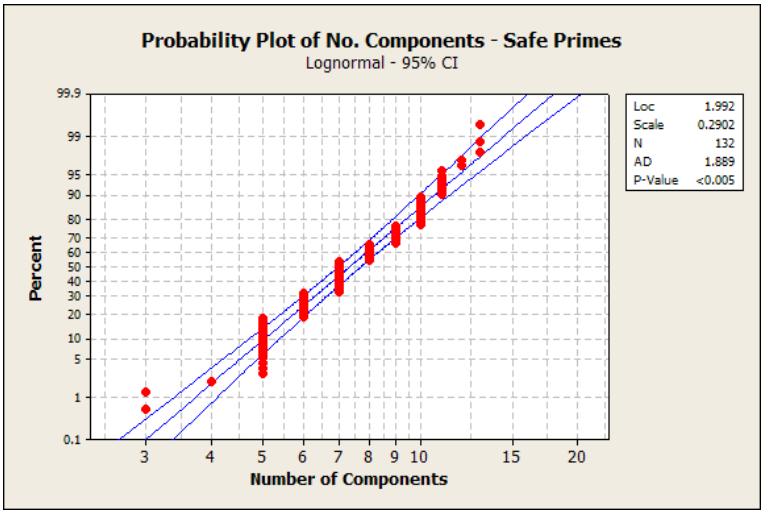


Figure 10: Distribution of No. Components in the Safe Primes

For number of components, the divisor classes had significantly different means ($p < 0.001$) and were highly successful predictors ($R^2 = 88.3\%$). This is especially significant for us because the number of divisors of $p-1$ can be explicitly calculated or well-approximated, unlike the number of fixed points, making it a much more useful predictor.

Having had success with number of components, we decided to run ANOVA tests against divisor classes for all of our parameters. The difference in means was significant for all of them, with $p < 0.005$ in all cases. The R^2 results are listed below.

SPFG Parameter	R^2-value (%)
Number of components.	88.30
Number of cyclic nodes.	55.81
Number of terminal nodes.	72.91
Number of fixed points.	92.45
Total cycle length.	45.87
Total distance to a cycle.	72.77
Maximum cycle length.	51.94
Maximum tail length.	79.10
Average cycle length.	45.87
Average tail length.	72.71
Average in-degree.	73.51

9 Conclusion

Our work has only begun to reveal the rich structure of the self-power functional graph. It was surprising to find that we can know where seemingly-uninteresting nodes such as $\frac{p-1}{2}$ and $\frac{p+1}{2}$ map to. A more exhaustive study of other nodes needs to be done to see if they are as predictable. It was also fruitful to examine well-studied sets of residues modulo p, such as quadratic residues and primitive roots. For both of them, we were able to glean information about what type of node their pre-images must be. Concerning primitive roots, we would expect that the self-power map would be more secure if $p-1$ were chosen such that it has less factors. That way, there would be less image nodes that are primitive roots, making it harder to predict the type of a node's pre-image. Future work on bounds for the number of primitive roots relatively prime to $p-1$ would be useful if this turns out to be valid.

Our investigations into fixed points led to results about the pre-images of 1 and $p-1$. Based on whether a node is a fixed point or a pre-image of 1 or $p-1$, we know exactly where its additive inverse will map to. This is doubly significant because fixed points are statistically more common in self-power functional graphs. Investigations to see if this knowledge can be expanded to the pre-images of other nodes would be helpful in cracking the Self-Power Problem. For cycles in general, we found that nodes in a cycle all have the same order. Since studying fixed points proved fruitful, we would like to find stronger conditions for a node to be in a cycle.

More future work lies in constructing a better bound on the number of pre-images of 1 and $p-1$, since much of the structure in the self-power functional graph was related to these nodes. Given the relationship between the pre-images of 1, $p-1$, and fixed points, this bound lends itself to another bound on the number of fixed points. This bound further contributes to a bound on the number of components and cycles, due to the high percentage of fixed points within the self-power functional graph. With current theoretical progress in mind, future work on the Self-Power Problem is promising.

On the statistical side, we have found a great deal of non-random structure in these maps. There are still many parameters to be examined which may reveal more predictable behavior. It would also be useful to consider more discrete distributions for safe prime data, since any predictable behavior in the safe primes would be notable. Perhaps most interesting for the future is the issue of the number of divisors of $p-1$. This number is a strong predictor of many of the parameters for which we have data, and perhaps an explanation for this could be found in theoretical results.

Other future work in a new direction lies in applying the methods used in this paper to the problem of solving $xg^x \equiv c \pmod{p}$ for x, where p is prime, g is a primitive root modulo p, and c is fixed. This congruence appears in the original version of the ElGamal Digital Signature Algorithm, when Frank the forger fixes s in the verification equation and attempts

to solve for r, and is a compounded version of the DLP.

In summary, our two-pronged approach with number theoretical and statistical tools has clearly demonstrated that self-power functional graphs look unlike random functional graphs, even without exploring the self-power map exhaustively. The current results and ideas presented here suggest that future progress on the Self-Power Problem is feasible and could potentially lead to practical information regarding cryptographic schemes.

References

[1] Antal Balog, Kevin A. Broughan, and Igor E. Shparlinski, *On the Number of Solutions of Exponential Congruences*, available at `http://arxiv.org/PS_cache/arxiv/pdf/1003/1003.1997v1.pdf`. Preprint.

[2] Daniel R. Cloutier and Joshua Holden, *Mapping the discrete logarithm* (2006), available at `http://xxx.lanl.gov/abs/math.NT/0605024`.

[3] Roger Crocker, *On a New Problem in Number Theory*, The American Mathematical Monthly **73** (1966), no. 4, 355–357.

[4] ———, *On Residues of n^n*, The American Mathematical Monthly **76** (1969), no. 9, 1028–1029.

[5] Philippe Flajolet and Andrew M. Odlyzko, *Random Mapping Statistics*, Advances in cryptology—EUROCRYPT '89 (Houthalen, 1989), Lecture Notes in Comput. Sci., vol. 434, Springer, Berlin, 1990, pp. 329-354.

[6] M. Friedl, N. Provos, and W. Simpson, *Diffie-Hellman Group Exchange for the Secure Shell (SSH) Transport Layer Protocol* (2006), available at `http://www.ietf.org/rfc/rfc4419.txt`.

[7] Andrew Hoffman, *Statistical Investigation of Structure in the Discrete Logarithm*, Rose-Hulman Undergraduate Mathematics Journal **10** (2009), no. 2.

[8] Nathan W. Lindle, *A Statistical Look at Maps of the Discrete Logarithm*, 2008. Senior thesis, Rose-Hulman Institute of Technology.

[9] Excel spreadsheets of the data we worked with in our Statistical Analysis section are online, available at `http://www.rose-hulman.edu/~holden/REU/`.