# A statistical look at maps of the discrete logarithm


**Dr. Joshua Holden and Nathan Lindle**

Mathematics and Computer Science, Rose-Hulman Institute of Technology, Terre Haute, Indiana 47803


## Definitions

**Functional Graph** – A directed graph where the edges are determined by a transition function. In this case the function is

$$\varphi : x \to g^x \bmod p$$

**Binary Functional Graph** – A functional graph where the in-degree of each node is either 0 or 2. All the graphs studied were binary functional graphs.

**Component** - A connected set of nodes. The average number of components is measured for each prime modulus (e.g. 1.75 for p = 11)

**Cyclic Nodes** – Nodes that are part of a cycle, including nodes which loop back on themselves. The average cyclic nodes are measured for each prime (e.g. 3.25 for p = 11)

**Average Cycle** – The average cycle size as seen from a random node in a functional graph divided by the number of nodes in all the functional graphs for a given prime (e.g. 2.05 for p = 11)

**Average Tail** – The average distance to the cycle as seen from a random node in the graph. Cyclic nodes have a distance of 0. Computation is similar to that of the average cycle (e.g. 1.225 for p = 11)

**Max Cycle** – The largest cycle in a graph. The average is taken over all bases for a given p (e.g. 2.5 for p = 11)

**Max Tail** – The longest distance from a node to its cycle in a graph. Similar to max cycle (e.g. 2.75 for p = 11)

## Methods

Generating functions:

$$\text{Binary Functional Graphs} = f(z) = e^{c(z)}$$

$$\text{Components} = c(z) = \ln\left(\frac{1}{1 - zb(z)}\right)$$

$$\text{Binary Trees} = b(z) = z + \frac{1}{2} z b(z)^2$$

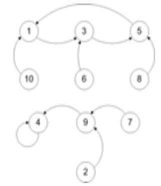
Marked generating function for total number of components:

Generating function fc $e^{-\frac{1}{2}u\ln(1-2z^2)}$ of components:

$$\text{Total Components} = \frac{\partial}{\partial u} e^{u\,c(z)}\Big|_{u=1} = -\frac{1}{2}\frac{\ln(1-2z^2)}{\sqrt{1-2z^2}}$$

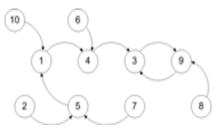
## Example Graphs (mod 11)

g = 3



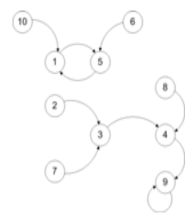

Components: 2
Cyclic nodes: 4
Average cycle: 2.2
Average tail: 0.8
Max cycle: 3
Max tail: 2

g = 4



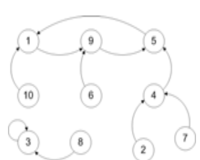

Components: 1
Cyclic nodes: 2
Average cycle: 2
Average tail: 2
Max cycle: 2
Max tail: 4

g = 5



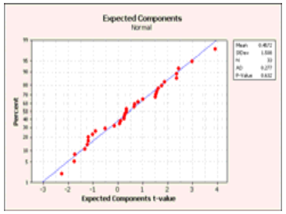

Components: 2
Cyclic nodes: 3
Average cycle: 1.4
Average tail: 1.3
Max cycle: 2
Max tail: 3

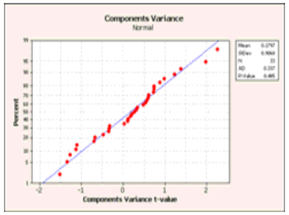
g = 9

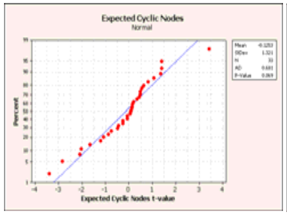


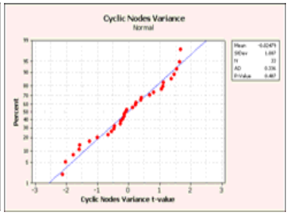

Components: 2
Cyclic nodes: 4
Average cycle: 2.6
Average tail: 0.8
Max cycle: 3
Max tail: 2

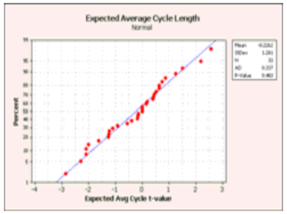
## Results

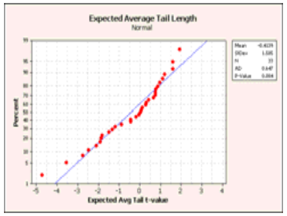


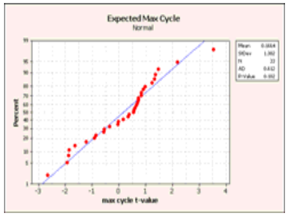




| Sample Maximum Tail Statistics | | | | |
|---|---|---|---|---|
| | Mean | | | Std. Dev. |
| p | Predicted | Observed | P-value | Observed |
| 99923 | 547.605802 | 543.281073 | 0 | 163.809 |
| 99961 | 547.710225 | 541.005022 | 0 | 163.494 |
| 99971 | 547.737702 | 544.967041 | 0.002 | 165.249 |
| 99989 | 547.787156 | 542.47563 | 0 | 163.809 |
| 99991 | 547.792651 | 541.265167 | 0 | 163.805 |
| 100003 | 547.825617 | 543.876996 | 0 | 163.295 |
| 100019 | 547.86957 | 542.008421 | 0 | 163.79 |
| 100049 | 547.951971 | 544.38604 | 0.002 | 165.651 |
| 100069 | 548.006899 | 549.379291 | 0.318 | 165.926 |
| 100103 | 548.100263 | 540.966673 | 0 | 164.496 |

| Sample Average Cycle Std. Dev. | | |
|---|---|---|
| p | Predicted | Observed |
| 99923 | 164.857 | 143.637 |
| 99961 | 164.888 | 140.91 |
| 99971 | 164.897 | 143.442 |
| 99989 | 164.911 | 144.856 |
| 99991 | 164.913 | 143.678 |
| 100003 | 164.923 | 143.686 |
| 100019 | 164.936 | 143.632 |
| 100049 | 164.961 | 142.121 |
| 100069 | 164.978 | 143.612 |
| 100103 | 165.006 | 142.79 |

| Sample Average Tail Std. Dev. | | |
|---|---|---|
| p | Predicted | Observed |
| 99923 | 164.858 | 86.025 |
| 99961 | 164.89 | 85.319 |
| 99971 | 164.898 | 86.914 |
| 99989 | 164.913 | 85.661 |
| 99991 | 164.914 | 85.227 |
| 100003 | 164.924 | 85.261 |
| 100019 | 164.938 | 86.029 |
| 100049 | 164.962 | 86.398 |
| 100069 | 164.979 | 87.876 |
| 100103 | 165.007 | 86.234 |

| Sample Max Cycle Std. Dev. | |
|---|---|
| p | Observed |
| 99923 | 154.949 |
| 99961 | 152.417 |
| 99971 | 154.931 |
| 99989 | 156.06 |
| 99991 | 154.534 |
| 100003 | 154.874 |
| 100019 | 154.535 |
| 100049 | 153.792 |
| 100069 | 155.018 |
| 100103 | 154.1 |

## Summary

- Many of the statistics gathered do not provide sufficient evidence to question the theory that modular exponentiation graphs are similar to random functional graphs.
- The observed variance in the average cycle and the average tail were significantly lower than the expected variance for a random binary functional graph.
- A few tests had surprisingly low p-values, but the normality tests indicate that these were just outliers.

## Future Work

- Get theoretical values for maximum tail and maximum cycle variance.
- Analyze lower variances in average cycle length and average tail length to try and come up with a reason.
- Find an explanation for the lower maximum tail.