CPSC 502.2 Proposal: Experiments on the Densities of K-Parent Aliquot Numbers

Gavin Guinn

Supervisor: Michael J. Jacobson

September 2021

1 Introduction

Aliquot sequences are simple to define, take some whole number n and add together all numbers that evenly divide n (with the exclusion of n itself) to produce a new sum on which the formula is reapplied. This process of summation is referred to as the sum-of-proper-divisors function and is denoted $s(n)$, defined more formally:

$$s(n) = \sum_{d|n} d - n$$

Lets try this out on a concrete example and compute the aliquot sequence of 12:

$$s(12) = 1 + 2 + 3 + 4 + 6$$
$$s(16) = 1 + 2 + 4 + 8$$
$$s(15) = 1 + 3 + 5$$
$$s(9) = 1 + 3$$
$$s(4) = 1 + 2$$
$$s(3) = 1$$
$$s(1) = 0$$

In this example we can notice the sequence grows with $s(12) > 12$ and then shrinks till it terminates at 0. Numbers like 12 that have the property $s(n) > n$ are referred to as *abundant* and will grow a sequence, if the opposite is true and $s(n) < n$ then n is *deficient*. In this example we can also notice that the image of 12 under the sum-of-proper-divisors is 16. For our purposes it will be useful to formalize the inverse relationship as a set of pre-images:

$$s^{-1}(n) = \{k \in \mathbb{W} | s(k) = n\}$$

Numbers in the set $s^{-1}(n)$ are referred to as the aliquot parents of n so in our example we can refer to 12 as a parent of 16. A number can seemingly have any whole number of parents with 0-parent numbers referred to as non-aliquot. We will consider an sequence to terminate if it reaches 0 or a starts to repeat itself in a cycle of what are termed sociable numbers. The aliquot sequence of 12 is an example of a sequence that clearly terminates but that is not always the case, for instance take the sequence of 276:

$$s(276) \rightarrow s(396) \rightarrow s(696) \rightarrow s(1104) \rightarrow s(1872) \rightarrow ...$$

The aliquot sequence of 276 is the smallest example that has not been shown to terminate, [Zimmermann, 2016] has followed the it's growth over 2090 iterations to values with over 200 digits! The dichotomy between the seemingly unbounded growth of the 276 sequence and the clear termination of the 12 sequence gets to the most important open question relating to aliquot sequences:

Conjecture 1.1 (Catalan-Dickson)**.** All aliquot sequences are bounded and must terminate in a prime, perfect number, or a set of sociable numbers. [Catalan, 1888]

Conjecture 1.2 (Guy-Selfridge)**.** The Catalan-Dickson conjecture is false, some aliquot sequences are unbounded and never terminate. [Guy and Selfridge, 1975]

The computational exploration of these sequences that has taken place since Catalan originally proposed his conjecture in 1888 has certainly given credence to the Guy-Selfridge conjecture, the extent of the 276 sequence's momentous growth could not have been clear to Catalan at the time. However a trend observed in the data is far from a proof, under our current understanding is entirely possible that the 276 sequence could trip over a prime and crash down to 0 on some future iteration.

[Chum et al., 2018] employ a variety of techniques to study the average rate of aliquot sequence growth. If sequences tend to grow on average that would be evidence in support of the existence of unbounded sequences as conjectured by Guy-Selfridge however if sequences tend to shrink it would suggest Catalan-Dickson. Among other approaches the authors investigate the impact of k-parent aliquot numbers on the growth of sequences, when generating statistics the abundance or deficiency of numbers with many parents should be weighted higher than numbers with few parents. Continuing this line of inquiry Dr. Guy made the following observation which motivates my project:

> Think of a number!! Say 36%, which is nice and divisible. It appears that about 36% of the even numbers are "orphans".

> Divide by 1. For about 36% of the (even) values of n there is just one positive integer m such that $s(m) = n$. These values of n have just one "parent".

2

Divide by 2. About 18% of the even values of n have exactly two parents.

Divide by 3. About 6% of the even values of n have three parents.

Divide by 4. About 1.5% of the even values of n have just 4 parents.

This suggests that $1/(e \cdot p!)$ of the even numbers have p parents.

Experiments suggests that these values are a bit large for small values of p and a bit small for larger values of p. Can anything be proved?

2   Project / Previous Work

Since k-parent aliquot numbers are a generalization of non-aliquots the most effective strategy for quantifying the k-parent density is to generalize techniques used to study the density of non-aliquots. One method is the heuristic approach of [Pollack and Pomerance, 2016] who employ statistical methods to model the density of non-aliquots, this technique produces estimates which closely match the experimentally observed density. Another approach is that of [Chen and Zhao, 2011] who establish a provable lower bound for the density of non-aliquots, this provable lower bound is significantly lower than the observed density.

In my previous work as a USRI summer student I focused my attention on generalizing the heuristic model of Pollack and Pomerance:

Definition 2.1. Heuristic Model for the Density of Non-Aliquot Numbers

$$\Delta = \lim_{y \to \infty} \frac{1}{\log y} \sum_{\substack{a \leq y \\ 2|a}} \frac{1}{a} \mathrm{e}^{-a/s(a)}$$

During that work term I produced the following generalization of their model:

Definition 2.2. (tentative) Density of k-parent Aliquot Numbers

$$\Delta_k = \lim_{y \to \infty} \frac{1}{\log y} \sum_{a \leq y} \frac{a^{k-1}}{k! \cdot s(a)^k} \cdot \mathrm{e}^{-a/s(a)}$$

However due to time constraints some of the mathematical justification for my generalization of this model was left incomplete. A priority for this project is to return to this work to resolve the unjustified steps and polish my mathematical

writing.

A complementary approach to heuristic modeling is direct computation of the densities of k-parent aliquot numbers. If the experimentally computed densities closely match the predictions of the heuristic model we would have strong evidence for the model's correctness. Like the theoretical work the most effective approach to this problem is to generalize techniques used to compute the density of non-aliquots, in this case the work of [Chum et al., 2018]. The authors leverage an algorithm of [Pomerance and Yang, 2014] (P-Y algorithm) to efficiently calculate all aliquot numbers within a range, the density of non-aliquots simply being the percentage of numbers in this range that are not aliquot. In my previous work as a USRI summer student I was provided the multi-threaded C code used by [Chum et al., 2018] which I modified to calculate the densities of k-parent aliquot numbers. When I deployed my implementation the university's High Performance Cluster (HPC) I found that my counts of non-aliquots matched those of the original authors for all bounds up to 2^{40}, at this bound the two programs produce results that differ ± 1.

In this project I will resolve this discrepancy and take all possible measures to ensure the accuracy of my k-parent aliquot number counts. To achieve this end I will apply software engineering techniques to refactor and carefully test the implementations component functions. As part of this testing I will compare the results of my implementation of P-Y to a much simpler program that uses trial division to enumerate k-parent aliquots. If the bug cannot be isolated in this manner I will rerun the code of [Chum et al., 2018] and compare the outputs of the programs directly. Once I am confident in the accuracy of my program I will run it to the highest bound possible, at least 2^{40} which took about 12 hours of compute on the HPC. If I see any opportunity during the refactoring to optimize my implementation to run this program to higher bounds I will do so.

If there is time left in this project after these goals are accomplished I will generalize the provable lower bound of [Chen and Zhao, 2011] for the density of non-aliquots to establish provable lower bounds for the densities of k-parent aliquot numbers.

3   Timeline

(Nov 9, 2021) **Related work list and critiques**
Survey reading will focus on the wider literature surrounding aliquot sequences and elementary number theory. A wide perspective on this topic may elucidate additional connections between my work on k-parent aliquot numbers and the Guy-Selfridge conjecture. I will also look for resource on software engineering techniques multi-threaded mathematical computation.

(Dec 1, 2021) **P-Y implementation refactored and tested**
For this deliverable I will refactor my previous implementation of the P-Y algorithm to produce testable code. Once this is complete I will write a battery of unit tests to run against my implementation to establish the source of the discrepancy between my work and [Chum et al., 2018]. I will set up a build and test pipeline to ensure that bugs do not get introduced during the development process.

(Dec 6, 2021) **Interm report**
This report will document the software engineering techniques I have employed to ensure accuracy in my implementation of P-Y. If the discrepancy is resolved at this point the report will include computed statistics of k-parent aliquot numbers to at least 2^{40} and compare those results to those of my k-parent density model. If the discrepancy is not resolved the report will outline the next steps in establishing the accuracy which may include a careful analysis of [Chum et al., 2018] implementation of P-Y.

(Feb 15, 2022) **Justification of conjectural k-parent density result complete**
For this deliverable I will carefully review my previous work in generalizing the heuristic model of [Pollack and Pomerance, 2016] to estimate the density of k-parent aliquot numbers. This will include proving intermediate steps in the justification that were left unjustified on my last work term due to lack of time.

(Mar 24, 2022) **Final presentation**
Engaging presentation that will combine my heuristic model with the computationally produced statistics on the densities of k-parent aliquot numbers.

(Apr 12, 2022) **Final report**
Publication ready report that will combine my heuristic model with the computationally produced statistics on the densities of k-parent aliquot numbers.

(TBD) **Generalize non-aliquot density lower bounds to k-parent aliquot numbers**
If time permits I will generalize the provable lower bound of [Chen and Zhao, 2011] for the density of non-aliquots to establish provable lower bounds for the densities of k-parent aliquot numbers.

References

[Catalan, 1888] Catalan, E. (1888). Propositions et questions diverses. *Bulletin de la Société mathématique de France*, 2:128–129.

[Chen and Zhao, 2011] Chen, Y.-G. and Zhao, Q.-Q. (2011). Nonaliquot numbers. *Publicationes Mathematicae Debrecen*, 78(2):439–442.

[Chum et al., 2018] Chum, K., Guy, R. K., Jacobson, J. M. J., and Mosunov, A. S. (2018). Numerical and statistical analysis of aliquot sequences. *Experimental Mathematics*, 29(4):414–425.

[Guy and Selfridge, 1975] Guy, R. K. and Selfridge, J. L. (1975). What drives an aliquot sequence? *Mathematics of Computation*, 29(129):101–101.

[Pollack and Pomerance, 2016] Pollack, P. and Pomerance, C. (2016). Some problems of erdős on the sum-of-divisors function. *Transactions of the American Mathematical Society, Series B*, 3:1–26.

[Pomerance and Yang, 2014] Pomerance, C. and Yang, H.-S. (2014). Variant of a theorem of erdös on the sum-of-proper-divisors function. *Mathematics of Computation*, 83(288):1903–1913.

[Zimmermann, 2016] Zimmermann, P. (2016). Aliquot sequences 276, 552, 564, 660, 996, 1074 and 1134.