Proofs of Parker's Square Sum Conjecture

Soren Johnson, Ella Rowley, and Elliott Stuenkel
Advisors: Dr. Mike Johnson and Dr. Tommy Occhipinti

July 19, 2024

Abstract

The square-sum problem askes if it is possible to take a list of integers from 1 to an arbitrary n and make a sequence where each consecutive pair of numbers adds to a perfect square. Matt Parker conjectured that this was possible for any $n > 24$. In this paper, we outline 3 proofs by induction of this conjecture. We also provide the initial theory of a potentially more beautiful proof involving hexagonal grids.

1 Introduction

The square sum problem regards the possibility of taking a list of integers from 1 to an arbitrary n and organizing them in such a way where each consecutive pair in the sequence of numbers add to a perfect square. An example of one such sequence can be seen below:

9 25 9 25 9 25 9 25
8 , 1 , 15 , 10 , 6 , 3 , 13 , 12 , 4 , 5 , 11 , 14 , 2 , 7 , 9 , 16
16 16 16 16 16 16 16

In this example, the black numbers are a sequence of numbers from 1 to 16 where each consecutive pair of numbers add to the perfect square in blue. The most common way to think about this problem is using graph theory. If each of the integers are thought to be vertices with an edge between two vertices if they add to a perfect square, you have a graph representative of the problem. The problem can then be boiled down to finding a Hamiltonian path on this graph, meaning you can use the edges to hit every point on the graph exactly once. This Hamiltonian path is a solution to the square sum problem.

In 2015, mathematician and internet personality Matt Parker released a book titled Things to Make and Do in the Fourth Dimension, in which he briefly discusses the square sum problem and mentions the existence of a solution had been verified up to $n = 89$. Later in 2018, Parker stars in series of "Numberphile", videos where he mentions his colleague Charlie Turner had verified for n up to 299 with solutions always existing after 24. Consequently, Parker conjectured that it is possible to make viable sequences for any $n > 24$. He also challenged viewers of the video to verfiy the conjecture beyond 299.

On the same day that the "Numberphile" videos were posted, user "henryzz" posted a link to the video on the internet forum www.mersenneforum.org , as well as the challenge to verify the conjecture beyond 299. By the end of the day, Robert Gerbicz had posted his verification up through $2^{20} = 1,048,576$. Six days later he posted the outlines of a proof for numbers of the form $n = \frac{71 \cdot 25^{k-1}}{2}$ for all $k \in \mathbb{Z}$ such that $k \geq 0$. Finally, four days after that Gerbicz had posted the outline for a proof of all $n > 24$. As far as we know, this work only existed publicly in the forum thread for 4.5 years. In July of 2022, "HexagonVideos" posted a video explaining Gerbicz proof in a more accessible way which brought his work to a larger audience.

Also in 2018, John Kelly, Samuel Moyer, and Laura Olson, three students at Luther college, were given this problem to work on. Although Gerbicz had already written the proof for all n, no one really knew about it so these students started from scratch. By the end of their summer research time frame, they were able to prove it was possible to make viable sequences for $n = 16 \cdot \frac{k(k+1)}{2}$, where $k \in \mathbb{Z}^+$. To construct this proof, they used a series of tables and hexagonal grids with a consistent pattern. A PDF of their paper is linked in our GitHub repository regarding this project.

Coming into this research with both of these prior advancements, we set out to understand and then extend some of the work Gerbicz and the past Luther students had done. In sections 3-6, we will focus on our extensions of Gerbicz method. Here, we begin by explaining Gerbicz proof in a way that is more comprehensible, especially to those who don't know C/C++. From there we explain a nicer proof of Parker's conjecture using the same framework, as well as a few generalizations. In section 7, you can see our extensions of Kelly, Moyer, and Olson's work with hexagonal grids.

2 Graph Theory

Before, explaining Gerbicz's methods it is important to provide some basic graph theory. Here we will provide formal graph theory definitions that are used throughout our paper. It will also include important notational information which will be essential to know when reading the rest of our paper.

Definition 2.1. A graph G is a finite set of V of vertices, and a set $E \subseteq \binom{V}{2}$ called edges.

Remark 2.1.1. Anytime we use the word "graph" in this paper, we are referring to a simple, undirected graph

Definition 2.2. $\boxed{n}$ is the graph on n vertices labeled with the integers 1 to n where two vertices are connected with an edge if they sum to a square number.

Definition 2.3. A walk on a graph $G = (V, E)$ from vertex $x \in V$ to vertex $y \in V$ is a sequence $x, v_1, v_2, ..., v_l, y$ of vertices of G such that $\{\{x, v_1\}, \{v_1, v_2\}, ..., \{v_l, y\}\} \in E$.

Definition 2.4. A path is a walk that does not repeat vertices except possibly when $x = y$.

Definition 2.5. A Hamiltonian path on a graph G ($\widetilde{G}$) is a path that contains every vertex. If the path begins and ends on two vertices that share an edge, then it is called a Hamiltonian cycle on a graph G ($\mathring{G}$).

Remark 2.5.1. The j^{th} term in these sequences will be denoted as $\widetilde{G}_j$ and $\mathring{G}_j$ respectively.

Definition 2.6. A bipartite graph is a graph in which the vertices can be divided into two disjoint sets, such that no two vertices within the same set are connected by an edge.

Definition 2.7. The degree of a vertex is the number of edges connected to that vertex.

3 Basic Mechanics of Gerbicz Proof

After establishing the the basics of graph theory, it is important to establish the basic sequence transformations needed for Gerbicz proof. There are three basic transformations needed to make sequences that can be strung together to make new sequences. These transformations are reverse, scale, and shift. Their definitions are below as well as important lemmas regarding the preservation of important properties for the square sum problem.

Definition 3.1. Reverse of a Sequence Q $(R(Q))$: A new sequence constructed by taking an original sequence Q, starting at the end point and working backwards so none of the consecutive pairs are different.

Lemma 3.2. The $R(\widetilde{\boxed{n}})$ is a sequence where each consecutive pair of numbers add to a perfect square and every number less than or equal to n is represented exactly once.

Proof. Let $\widetilde{\boxed{t}}$ be some sequence of t unique numbers arranged in a way where each consecutive pair sums to a perfect square. By definition 3.1, the $R(\widetilde{\boxed{t}})$ does not change any of the numbers used in the sequence so the t numbers are still unique. By definition 3.1, the $R(\widetilde{\boxed{t}})$ also does not change the consecutive pairs despite changing the order of them so they must add to perfect squares.

∎

Definition 3.3. To scale a sequence Q, you multiply every term in the sequence Q by some number k. This function is denoted by $S(k, Q)$)

Lemma 3.4. $S(k, \widetilde{\boxed{n}})$ for any perfect square k creates a new sequence such that each consecutive pair of numbers adds to a perfect square and there are no repeated numbers.

Proof. Let a, b be adjacent numbers in some $\widetilde{\boxed{n}}$ such that $a + b = m^2$ where m^2 is some perfect square. If we multiply a and b by p^2, this becomes $p^2(a) + p^2(b)$ which is equivalent to $p^2(a + b) = p^2(m^2) = (pm)^2$.

∎

Remark 3.4.1. A scaled sequence will not contain all integers.

Definition 3.5. To shift a sequence Q, you add or subtract some constant c in an alternating pattern so that each consecutive pair still adds to the same number as before the shift. The first number of the sequence has c added to it, and the second number will have c subtracted to it. This function is denoted at $Sh(Q)_c$.

Lemma 3.6. $Sh(\widetilde{\boxed{n}})_c$ maintains the property that each consecutive pair adds to a perfect square.

Proof. Let a, b again be adjacent numbers in some $\widetilde{\boxed{n}}$ such that $a + b = m^2$. If we add our constant c to a, and subtract it from b, it looks like $(a+c)+(b-c) = a+c+b-c = a+b = m^2$ which maintains the property that the two numbers sum to a square.

∎

Lemma 3.7. $Sh(S(k, \widetilde{\boxed{n}}))_c$ maintains the property that there are no duplicate numbers in the sequence if $k \in \mathbb{Z}$ is odd and $|c| \leq \frac{k-1}{2}$.

Proof. Let a and $a + 1$ be two numbers that are one apart and placed in the sequence $\widetilde{\boxed{n}}$ such that a shift would add c to $k(a)$ and subtract c from $k(a+1)$. $k(a+1) - c = k(a) + k - c$, which cannot equal $k(a) + c$ unless $c = k - c$. This is the closest possible case where two numbers would potentially shift to the same number, so we can make sure that never happens for any case by keeping c such that $|c| \leq \frac{k-1}{2}$

∎

For this proof by induction to work, it is important to start with two sequences that solve the square sum problem. These sequences must have additional properties making them a "Nice Pair".

Definition 3.8. Two sequences $\widetilde{\boxed{n}}$ and $\widetilde{\boxed{n+1}}$ are called a nice pair if they maintain positional parity, meaning for all $i \le n$, if $i = \widetilde{\boxed{n}}_p = \widetilde{\boxed{n+1}}_q$, then $p - q$ is even. $\widetilde{\boxed{n}}$ will be referred to as the short nice sequence and $\widetilde{\boxed{n+1}}$ will be referred to as the long nice sequence.

Lemma 3.9. If $\widetilde{\boxed{n}}$ and $\widetilde{\boxed{n+1}}$ are a nice pair, the number $(n+1)$ in $\widetilde{\boxed{n+1}}$ will have positional parity equal to the parity of $(n+1)$

Example 3.9.1. Nice pair of length 41:

$\mathring{\boxed{41}} = 1, 8, 41, 40, 9, 27, 37, 12, 24, 25, 39, 10, 26, 38, 11, 5, 4, 32, 17, 19, 6, 30, 34, 15, 21, 28, 36, 13, 23, 2, 7, 18, 31, 33, 16, 20, 29, 35, 14, 22, 3$

$\mathring{\boxed{42}} = 1, 35, 29, 20, 16, 33, 3, 22, 14, 2, 23, 13, 36, 28, 21, 15, 34, 30, 6, 19, 17, 32, 4, 5, 31, 18, 7, 42, 39, 10, 26, 38, 11, 25, 24, 12, 37, 27, 9, 40, 41, 8$

The magenta numbers indicating odd positional parity are {1,3,4,6,7,9,11,14,16,17,21,23,24, 26,29,31,34,36,37,39,41}in both sequences. Since this is a list of 21 numbers the other numbers have to be even in both sequences as well maintaining positional parity.

Next, we provide a definition for a complete residue system, as well as provide an example of what a complete residue system will look like in the context of Gerbicz's proofs and their extensions.

Definition 3.10. A Complete Residue System is a set of m integers such that every integer is congruent modulo m to exactly one integer of the set $\mathbb{Z}/m\mathbb{Z}$. (We will denote this by $\mathbb{Z}_m$.)

Remark 3.10.1. res_k denotes an unspecified number in a complete residue system of $\mathbb{Z}_k$.

Remark 3.10.2. The complete residue system used in the proofs below will always be $\frac{k-1}{2} \le res_k \le k + \frac{k-1}{2} - 1$ where k is an odd perfect square and $res_k \in \mathbb{Z}$

Example 3.10.3. The set $\{24, 25, 26, ..., 60, 71, 72\}$ is a complete residue system for $\mathbb{Z}_{49}$.

Remark 3.10.4. The values used for c when using the shift function is a complete residue system of k. However, when referring to a complete residue system, this is not the one we are talking about.

4 Theorems and Definitions Regarding the Inductive Step

The last section before outlining the proof itself provides information regarding the inductive step of the Gerbicz proof. It includes a few important definitions including the definition of "glue" as well as a few algorithms and theorems regarding the way the sequences created using the basic mechanics are strung together to create new nice pairs. It is necessary to write the algorithm and theorems into the program that we use to find the "glue".

Definition 4.1. The term "Glue" refers to the formula used to string together the k sequences needed to create the nice pair of length $kn + res_k$ and the numbers between 1 and $kn + res_k$ that are not contained in the k sequences.

Definition 4.2. The term <u>"Glues"</u> refers to the numbers x such that $x \in \mathbb{Z}\ 1 \leq x \leq \frac{k-1}{2}$ and any additional numbers missing from the k sequences needed to create the nice pair of length $kn + res_k$ between 1 and $kn + res_k$. These numbers are used to "glue" the k sequences together giving them their name.

Algorithm 4.3. If $\widetilde{\boxed{n}}$ and $\widetilde{\boxed{n+1}}$ are a nice pair where the first j numbers of each nice sequence are exactly the same, then the following is a method using the nice pair to create all of the components needed to create k new nice pairs. These necessary components are k short sequences, k long sequences and the glue used to string the sequences together.

1. Check that for all $i \leq j$ and $i = n$, $\widetilde{\boxed{n}}_i$ fulfills one of the following requirements:

   (a) $1 \leq \widetilde{\boxed{n}}_i \leq \frac{k-1}{2}$

   (b) $\widetilde{\boxed{n}}_i = k \cdot \widetilde{\boxed{n}}_x + c$ where $\widetilde{\boxed{n}}_x$ is one of the first $j-1$ numbers in $\widetilde{\boxed{n}}$, $\widetilde{\boxed{n}}_x \neq \widetilde{\boxed{n}}_i$, and $|c| \leq \frac{k-1}{2}$

2. Check that $\widetilde{\boxed{n+1}}_{n+1}$ fulfills one of the following requirements:

   (a) $1 \leq \widetilde{\boxed{n+1}}_{n+1} \leq \frac{k-1}{2}$

   (b) $\widetilde{\boxed{n+1}}_{n+1} = k \cdot \widetilde{\boxed{n}}_x + c$ where $\widetilde{\boxed{n}}_x$ $\widetilde{\boxed{n}}_x$ is one of the first $j-1$ numbers in $\widetilde{\boxed{n}}$, and $|c| \leq \frac{k-1}{2}$

3. Create k unique short sequences $Sh(S(k, \widetilde{\boxed{n}}))_c$ such that $|c| \leq \frac{k-1}{2}$. The i^{th} term of $Sh(S(k, \widetilde{\boxed{n}}))_c$ is $k \cdot \widetilde{\boxed{n}}_i + c \cdot (-1)^{i+1}$

4. Create k unique long sequences $Sh(S(k, \widetilde{\boxed{n+1}}))_c$ such that $|c| \leq \frac{k-1}{2}$. The i^{th} term of $Sh(S(k, \widetilde{\boxed{n+1}}))_c$ is $k \cdot \widetilde{\boxed{n+1}}_i + c \cdot (-1)^{i+1}$

5. Let Q be a list of all the glues. These glues will be 1 to $\frac{k-1}{2}$, and the first $j-1$ entries of every $Sh(S(k, \widetilde{\boxed{n}}))_c$ such that $|c| \leq \frac{k-1}{2}$.

6. Let Tc where c is the shift, be a sub-sequence of $Sh(S(k, \widetilde{\boxed{n}}))_c$ such that $|c| \leq \frac{k-1}{2}$. This sub-sequence will be numbers in the j^{th} position to the n^{th} position of $Sh(S(k, \widetilde{\boxed{n}}))_c$.

7. Let Lc where c is the shift, be a sub-sequence of $Sh(S(k, \widetilde{\boxed{n+1}}))_c$ such that $|c| \leq \frac{k-1}{2}$. This sub-sequence will be numbers in the j^{th} position to the $(n+1)^{th}$ position of $Sh(S(k, \widetilde{\boxed{n+1}}))_c$.

8. For each c value, chose either L or T. Chose long if:

   (a) n is odd and $c \geq k - (res_k(+1))$

   (b) n is even and $c \leq (res_k(+1)) - k$

   Otherwise, chose T.

Lemma 4.4. The largest number in $\widetilde{\boxed{n+1}}$ will always be contained in L

Proof. By algorithm 4.3, L contains every number in $\widetilde{\boxed{n+1}}$ not contained in Q. Also by algorithm 4.3, Q can only contain numbers in both $\widetilde{\boxed{n}}$ and $\widetilde{\boxed{n+1}}$. Finally by definition 2.2, $\widetilde{\boxed{n}}$ cannot contain $(n+1)$. Therefore $(n+1)$ must always be contained in L.

$\blacksquare$

Remark 4.4.1. If $j = 1$, $T = \widetilde{\boxed{n}}$ and $L = \widetilde{\boxed{n+1}}$. This will be the case for the first two proofs in the main results section. These proofs are largely credited to Robert Gerbicz.

Lemma 4.5. Given a nice pair $\widetilde{\boxed{n}}$ and $\widetilde{\boxed{n+1}}$, a constant $k \in \mathbb{Z}$, a constant $j \in \mathbb{Z}$, and a constant $c \in \mathbb{Z}$ such that $c \leq \frac{k-1}{2}$, the numbers contained in the sequence $Sh(S(k,T))_c$ are exactly the numbers contained in the sequence $Sh(S(k,L))_c$ except for $k \cdot (n+1) + c \cdot (-1)^n$ which will only be in the second sequence.

Proof. Consider the nice pair $\widetilde{\boxed{a}}$ and $\widetilde{\boxed{a+1}}$ where $a \in \mathbb{Z}$. By algorithm 4.3 the i^{th} term of a $\widetilde{\boxed{a}}$ is equal to $k \cdot \widetilde{\boxed{n}}_i + c \cdot (-1)^{i+1}$ and the i^{th} term of $\widetilde{\boxed{n+1}}$ is equal to $k \cdot \widetilde{\boxed{n}}_i + c \cdot (-1)^{j+1}$. We can ignore the existence of the term $k \cdot (a+1) + c \cdot (-1)^a$ which can never overlap with the other numbers by lemma 3.7. By definition 3.8 all the integers such that $1 \leq n \leq a$ have the same positional parity in both sequences. Thus the value of $c \cdot (-1)^{i+1} = c \cdot (-1)^{j+1}$ meaning that the numbers contained in the sequence $Sh(S(k, \widetilde{\boxed{n}}))_c$ are exactly the numbers contained in the sequence $Sh(S(k, \widetilde{\boxed{n+1}}))_c$ except for $k \cdot (n+1) + c \cdot (-1)^n$

$\blacksquare$

Lemma 4.6. Corollary of Lemma 4.5 If you compare any two sequences created by the same nice pair using Algorithm 4.3 with the same k value, the same j value, and distinct c values, they are completely disjoint.

Proof. By lemma 4.5 we know that either sequence with the same c value have the same numbers assuming constant k and j values. For two sequences of length n or two sequences of length $n+1$ It is obvious that different c values will create distinct sequences since if you add different constants to the same number, you are guaranteed to get different numbers. This means the possible problem arises if value $n+1$ in the sequence of length $n+1$ somehow overlapping with the value n in the sequence of length n. After scaling and shifting $n+1$ in the sequence of length $n+1$ the smallest number it can become is

$$k(n+1) - \frac{k-1}{2} = kn + \frac{k+1}{2}.$$

The largest number possible value of n in the sequence of length n is

$$kn + \frac{k-1}{2} < kn + \frac{k+1}{2}.$$

Since the largest value of n is strictly less than the smallest value of $n+1$ it is impossible for two sequences with different c values to contain any of the same numbers meaning they are completely disjoint.

$\blacksquare$

5 Gerbicz's Proofs and Their Direct Extensions

5.1 A General Summary of the Following Proofs

Since the following three proofs are extremely similar in nature, we outline the general structure here first. The goal overall goal is to prove that there is a sequence of length n that solves the square sum problem for every $n > 24$. The following proofs prove a stronger statement about the existence of nice pairs. The original conjecture can then be proved by also providing a solution to the square sum problem for every n less than or equal to the first base case in the proof that is also greater than 24. Each of the following proofs will rely on two text files. The first text file will contain all the glues found for that particular proof. The second file will contain the all of the nice pairs from where they are first found to exist to k times that number plus $\frac{k-1}{2}$. These nice pairs will all have specified start and end points where the end points change depending on the parity of the length of each sequence. We have broken each proof into two parts.

5.1.1 Base Cases

The proof is somewhat of a complex induction which will involve hundreds of base cases. For each proof below, we will require the base cases to be nice pairs and have one additional property we will call Ψ which will dictate the start and end of the nice pairs. We have included a table for Ψ below:

k value	Required Start	Required End
49	1	odd: 3
		even: 8
25	1	odd: 3
		even: 2
9	1,3,13,12,4,5	odd: 8
		even: 24

Figure 1: Assignment of Ψ

The following is a brief explanation of the computer algorithm we used to find these base cases. The first step is creating a graph in Sage that has the integers from 1 to n as vertices with edges between them if they sum to a square number. A starting sequence and an ending integer can be forced by creating a "start" and "end" vertex that only connects to the desired start and end points. Once a Hamiltonian path has found using a Hamiltonian path finder on Sage, create a dictionary that has the positional parity of each integer as well as what the parity of what $n+1$ will be. At this point, create a new graph that has the integers from 1 to $n+1$ as vertices with edges between them if they sum to a square and if they have opposite parity using the dictionary created. If a path cannot be found for the nice pair, find a new path to search for a nice pair.

5.1.2 Proof of the Inductive Step

After finding the base cases, it is important to prove the inductive step. The proof of the inductive step is a constructive proof by cases. There are two cases for each of the residues associated with the particular k value since the glue changes depending on the parity of n as well as the value of the residue. For the $k = 49$ and the $k = 25$ cases, the algorithm is slightly different because $j = 1$ making Q the empty set by algorithm 4.3. Therefore also by algorithm 4.3 the only "glues" that can be used in the new sequence are 1 to $\frac{k-1}{2}$. These proofs rely on algorithm 4.3 to create and determine which of the sequences need to be strung together for each new nice pair. The complete glue is found using a computer algorithm that makes sure that the sequences continue to maintain the same start and end points as well as to make sure positional parity is maintained in the new nice pairs.

1. Use the residue and length of n to determine when to use the short sequence and when to use the long sequence for each value of c. For each sequence, determine what the endpoints will be. Also determine the values in Q.

2. Remove the integers in Ψ from your set Q.

3. Represent each sequence as a triplet of vertices if the length of the sequence is odd, and four vertices if the length is even, one vertex each for the start and endpoint and extras to force a path between the start and endpoint of the sequence. This also forces the start and endpoint of each sequence to maintain proper parity. Make the designated endpoints of the final sequence into vertices, and make the integers from the previous step into vertices.

4. Connect all vertices(other than the middle vertices of each sequence) by drawing the edges between vertices that sum to a square.

5. Create new vertices for every integer in Ψ and connect them so they match the desired starting sequence. Finally, attach a vertex labeled "start" and "end" to the beginning and ending integers.

6. Run this graph through a Hamiltonian path finder function and once one is found, take the sequence, reverse it if it is backwards, and then remove the "start " and "end" vertices.

7. The nice pair for this path will have one more long sequence than the other, changing one sequence from T to L. Find this sequence using the last step of 4.3, and change the endpoint and number of empty vertices accordingly. Remember to change the target endpoint. Record the positional parity of each integer and start/endpoint vertex by using the path and going from left to right.

8. Repeat steps 1 through 6, but use a length of $n+1$ and connect integer vertices if they sum to a square and have opposite positional parity to find the nice pair. If a path is found, then proceed. If no path is found, start again at step 6 to get a new path. It is possible that glue for a nice pair does not exist. For that reason, this process should interrupted after enough time has passed.

9. Glue for a nice pair has been found and all that is left is to reformat the glue by removing the endpoint vertices from each sequence on each of the paths provided and return the finished glue.

Below we will show an example of glues for the nice pair of length $49n + 37$ when the initial nice pair has even length. As noted in table 1, for the 49 case both sequences must start with 1, the odd length sequence must end with 3 and the even length sequence must end with 8. Another thing to note is that j will always be one for this case.

49n+37 for even n:
[1,RT-7,7,18,RT10,5,20,RT12,T-10,RL-20,RT21,11,RT3,T20,L-21,RT-11,T13,21,T-6,RT6,9,16, RT8,24,RT16,RL-16,T-1,RT1,T22,L-19,RT-9,T11,19,RL-22,RT19,T4,RT-4,4,12,13,L-13,RT-3, L-14,RL-24,RT17,RL-17,RT24,8,RT0,T2,10,15,RT7,T-5,RT5,T18,L-23,T23,L-18,RT-8,23,RT15, RL-15,2,14,22,RT14,L-12,RT-2,17,RT9,6,3]
[1,RT-7,RL-20,RT21,11,L-11,L-16,T16,24,T8,16,9,T6,RT-6,21,RT13,RL-13,13,12,4,T-4,RT4, T19,L-22,19,RT11,T-9,RL-19,RT22,T1,RT-1,T24,L-17,T17,L-24,RL-14,14,22,RT14,L-12,T12, 20,5,T10,RT-10,10,15,RT7,T-5,RT5,T18,L-23,T23,L-18,RT-8,23,RT15,RL-15,2,7,18,RL-21, RT20,T3,RT-3,3,6,T9,17,T-2,RT2,T0,8]

This is the format found in the document for glues. Each piece of information is separated by a comma. If it is a number on it's own, it is just that number in the sequence. As defined in algorithm 4.3 "T" implies the use of the short sequence and "L" implies the use of the long sequence. As stated in definition 3.1, R implies the sequence is reversed. Finally, the constant after the "T" or "L" is the value of c. Reminder that since n is even and j is odd, "T" will be an even length sequence and "L" will be an odd length sequence.

Below we will demonstrate that how the formula works. First, we will replace all the sequences with their start and end point using the formula for the i^{th} term as provided in algorithm 4.3:

For example RT-7 first number is $49 \cdot 8 - 7 \cdot (-1) = 49 \cdot 8 + 7 = 399$ and the last number is $49 \cdot 1 - 7(-1)^2 = 49 - 7 = 42$.

It will only be important to know the start and end points of each sequence as well have proved previously that all the numbers inside the sequence add to a perfect square already.

Following the same process the entire sequence becomes:
[1,(399, ..., 42),7,18,(382, ..., 59),5,20,(380, ..., 61),(39, ..., 402),(127, ..., 29),(371, ..., 70),11, (389, ..., 52),(69, ..., 372),(28, ..., 126),(403, ..., 38),(62, ..., 379),21,(43, ..., 398), (386, ..., 55),9,16,(384, ..., 57),24,(376, ..., 65),(131, ..., 33),(48, ..., 393),(391, ..., 50), (71, ..., 370),(30, ..., 128),(401, ..., 40),(60, ..., 381),19,(125, ..., 27),(373, ..., 68), (53, ..., 388),(396, ..., 45),4,12,13,(36, ..., 134),(395, ..., 46),(35, ..., 133),(123, ..., 25), (375, ..., 66),(130, ..., 32),(368, ..., 73),8,(392, ..., 49),(51, ..., 390),10,15,(385, ..., 56), (44, ..., 397),(387, ..., 54),(67, ..., 374),(26, ..., 124),(72, ..., 369),(31, ..., 129),(400, ..., 41), 23,(377, ..., 64),(132, ..., 34),2,14,22,(378, ..., 63),(37, ..., 135),(394, ..., 47),17,(383, ..., 58),6,3]

[1,(399, ..., 42),(127, ..., 29),(371, ..., 70),11,(38, ..., 136),(33, ..., 131),(65, ..., 376),24, (57, ..., 384),16,9,(55, ..., 386),(398, ..., 43),21,(379, ..., 62),(134, ..., 36),13,12,4, (45, ..., 396),(388, ..., 53),(68, ..., 373),(27, ..., 125),19,(381, ..., 60),(40, ..., 401), (128, ..., 30),(370, ..., 71),(50, ..., 391),(393, ..., 48),(73, ..., 368),(32, ..., 130),(66, ..., 375), (25, ..., 123),(133, ..., 35),14,22,(378, ..., 63),(37, ..., 135),(61, ..., 380),20,5,(59, ..., 382), (402, ..., 39),10,15,(385, ..., 56),(44, ..., 397),(387, ..., 54),(67, ..., 374),(26, ..., 124), (72, ..., 369),(31, ..., 129),(400, ..., 41),23,(377, ..., 64),(132, ..., 34),2,7,18,(126, ..., 28), (372, ..., 69),(52, ..., 389),(395, ..., 46),3,6,(58, ..., 383),17,(47, ..., 394),(390, ..., 51), (49, ..., 392),8]

In both sequences the magenta numbers indicating odd positional parity are {1,2,3,4,10,11,13, 16,17,18,19,20,21,22,23,26,32,33,35,42,44,50,51,52,53,55,57,58,59,60,61,62,63,64,67,73,124,130, 131,133,369,370,371,372,373,375,376,385,387,392,393,394,395,396,398,400,401,402}. The shorter sequence also contains 403 which is indicative of the inclusion of "RT-11". Since the long nice sequence contains "L-11" instead it contains 136 as a replacement. These will be have the correct positional parity in both sequences because of definition 3.8. In both sequences the black numbers indicating even positional parity are {5,6,7,8,9,12,14,15,24,25,27, 28,29,30,31,34,36,37,38,39,40,41,43,45,46,47,48,49,54,56,65,66,68,69,70,71,72,123,125,126,127, 128,129,132,134,135,368,374,377,378,379,380,381,382,383,384,386,388,389,390,391,397,399}.

Hopefully this provided a basic understanding of how the glues work, and how they can be used inductively. Notice that we never used the value of n in our example and yet we are still certain that the resulting sequence is a solution to the square sum problem. If you would like to see more examples of the glue in action, we recommend looking at the files in our GitHub repository. In this folder, you will find the glues and base cases for the $k = 9$ and the $k = 25$ case, as well as code titled "squares_generate_sequences.sagews" which can be used to create sequences of any length using our recursion method.

5.2 Gerbicz's Proof with k=49

We will start by proving that nice pairs where both sequences start with 1, the odd length sequence ends with 3, and the even length sequence ends with 8, exist for all $n \geq 41$. This was the primary proof that Gerbicz proposed in his forum post. It appealed to him since it could be used to prove that there exists Hamiltonian cycles for all $n \geq 32$. He thought it was a much more exciting result than just proving the existence of Hamiltonian paths.

5.2.1 Proof of Inductive Step with k=49

Theorem 5.1. For any nice pair $\overset{\circ}{\boxed{n}}$ and $\overset{\circ}{\boxed{n+1}}$ where both sequences start with 1,the odd length sequence ends with 3, and the even length sequence ends with 8, you can always generate 49 new nice pairs of the form $\overset{\circ}{\boxed{49n + res_{49}}}$ and $\overset{\circ}{\boxed{49n + res_{49} + 1}}$ which also both start with 1 and end with 3 or 8 depending on the parity of the sequence length. [Gerbicz, 2018]

Proof. This is a proof by cases with 98 cases. There are two cases for each residue. The first case is for when n is an even number and the second case is for when n is an odd number. Note that since these nice pairs only start with an identical sequence of length 1, $j = 1$. Using algorithm 4.3, when $j = 1$ Q is the empty set, T will just be the entirety of $\boxed{n}$, and L will be the entirety of $\boxed{n+1}$. Using algorithm 4.3 we can create 98 sequences for each original nice pair. We can also determine that our only "glues" are 1-24 since Q is the empty set in this case. Then using 4.3 step 8, we can determine which of the sequences need to be strung together to create each sequence in the new nice pairs. A computer program was written to find how to arrange the 49 sequences and the numbers 1 through 24 that are needed for each sequence in each new pair. The resulting pairs are contained in a link below.

Here is a link to the formulas for all 98 nice pairs.
Case 1: res_{49}=24, n is even
Case 2: res_{49}=24, n is odd

•

•

•

Case 97: res_{49}=72, n is even
Case 98: res_{49}=72, n is odd

∎

Remark 5.1.1. This is just a proof of the inductive step of the main proof.

5.2.2 Proof by Induction with k=49

Theorem 5.2. There exists a nice pair where both sequences start with 1, the odd length sequence ends with 3, and the even length sequence ends with 8 for every $n \geq 41$. [Gerbicz, 2018]

Proof. Define S_n as follows:

$$S_1 = 41$$
$$S_2 = 2033$$
$$S_n = 49 \cdot S_{n-1} + 24 \text{ for } n \geq 3$$

We will prove by induction that for every $k \geq 1$, there exists a nice pair where both sequences start with 1, the odd length sequence ends with 3, and the even length sequence ends with 8 for $S_k \leq a < S_{k+1}$.

Base Case(s): Consider $k = 1$

For $S_1 \leq a < S_2$, we need to show that there exists a nice pair of sequences for all $41 \leq a < 2033$. By enumeration[1] or an external reference, it can be verified that such pairs exist for this range. The list provided contains all the nice pairs for these values.

Inductive Step:

Assume that for some $k \geq 1$, there exists a nice pair where both sequences start with 1, the odd length sequence ends with 3, and the even length sequence ends with 8 for $S_k \leq a < S_{k+1}$.
We need to show that there exists a nice pair for $S_{k+1} \leq a < S_{k+2}$.
Given the inductive hypothesis, we have nice pairs for all a in the interval $S_k \leq a < S_{k+1}$. By theorem 5.1 we know that we can create 49 new nice pairs of this form from each initial nice pair. By algorithm 4.3 and remark 3.10.2, we know the nice pairs created are of the form $\boxed{49n + \mathring{res}_{49}}$ and $\boxed{49n + \mathring{res}_{49} + 1}$ and the smallest value of $res_{49} = 24$. Therefore, given nice pairs from $S_k \leq a < S_{k+1}$ we should be able to make nice pairs from $49 \cdot S_k + 24 \leq a < 49 \cdot S_{k+1} + 24$. By definition, that is equivalent to $S_{k+1} \leq a < S_{k+2}$.

Therefore, by induction, there must be a nice pair where both sequence start with 1, the odd length sequence ends with 3, and the even length sequence ends with 8 for every $n \geq 41$.

∎

Remark 5.2.1. This proof is credited to Robert Gerbicz who posted all the evidence that this proof existed on 01/21/2018 in a forum post. We have attempted to make his work more clear by outlining the theorems, lemmas, and algorithms he relied on in his proof. Here is the link to the forum post where his work was originally published. A PDF version of this thread has also been uploaded to our project's Github Repository if you are unable to access the thread through the forum.

5.3 Gerbicz's Proof with k=25

5.3.1 Proof of Inductive Step with k=25

Theorem 5.3. For any nice pair $\widetilde{\boxed{n}}$ and $\widetilde{\boxed{n+1}}$ where both sequences start with 1, odd length sequences end with 3, and even length sequences end with 2, you can always generate 25 new nice pairs of the form $\boxed{25n + res_{25}}$ and $\boxed{25n + res_{25} + 1}$ which also start with 1 and end with 2 or 3 depending on the parity of the sequence length.

This proof is structured in exactly the same manner as the previous proof. For this reason we have chosen not to rewrite the proof. All the base cases for this proof can be found here. The "glue" for the 25 case can be found here.

Remark 5.3.1. This proof was also proposed to be true by Robert Gerbicz on January 21, 2018. He mentioned that his method used in the proof above should also work for paths using 25 as k and nice pairs that start with 1 and end with 2 or 3 depending on the parity of n. However, unlike the 49 case, he did not propose any glue or base cases. The appeal of this proof is that it requires significantly less base cases and glue formulas. Here is the link to the forum post. A PDF version of this thread has also been uploaded to our project's Github Repository if you are unable to access the thread through the forum.

5.4 Proof to be Used for the Existence of Cycles k=9

The proof regarding the existence of cycle with $k = 9$ was slightly more complicated. This proof is a direct extension of Gerbicz's proof and functions in a very similar way. However, his original method made the proof impossible, because there is not enough vertices on the glue graph to make a path. We believed we could make the $k- = 9$ work, and subsequently create a proof with far fewer base cases.

We needed to come up with a way to make more "glues" to connect the sequences together. This also needed to be done in a way that does not repeat numbers when making new sequences. We decided that the easiest way to do this was to remove some numbers from the nice pairs before creating the $2k$ sequences that needed to be strung together. After some thought, we decided the easiest way to make this possible was to require a starting sequence for each of the nice pairs. This way, we could still ensure the sequences had the same starting point, and were missing the same numbers when we removed the first few numbers from the front to be used for glue. The idea was that even if our base cases for nice pairs needed to start a little bit higher there would still be far less base cases because of the significantly smaller k value meaning ultimately we would still have a proof with less base cases.

The next problem we ran into was attempting to determine how long the starting sequence would have to be as well as what numbers had the highest probability of success. We knew that smaller numbers typically had higher degrees because they could be used to add to a larger number of perfect squares. We also quickly discovered the formula given in step 2 and 3 of algorithm 4.3 must be true. The starting sequence, and the end points of the two sequences had to be included in the glue, or else it was impossible to force it to be the starting sequence in the glues. This would mean that the conjecture would be unable to be proved inductively. After a lot of trial and error, the best we were able to find was nice pairs that had a starting sequence of $(1, 3, 13, 12, 4, 5, 11)$ and ended with either 8 or 24 depending on the parity of the length of each sequence in the nice pair.

5.4.1 Proof of Inductive Step with k=9

Theorem 5.4. For any nice path pair $\overset{\circ}{\boxed{n}}$ and $\overset{\circ}{\boxed{n+1}}$ where both sequences start with $(1, 3, 13, 12, 4, 5, 11)$, the odd length sequence ends with 8, and the even length sequence ends with 24 you can always generate 9 new nice pairs of the form $\overset{\circ}{\boxed{9n + res_9}}$ and $\overset{\circ}{\boxed{9n + res_9 + 1}}$ where these sequences also start with $(1, 3, 13, 12, 4, 5, 11)$ and end with either 8 or 24 depending on the parity of the length of the sequence.

Proof. This is a proof by cases with 18 cases. There are two cases for each residue. The first case is for when n is an odd number and the second case is for when n is an even number. Using algorithm 4.3, we can define Q to be $(1, 3, 13, 12, 4, 5)$. This means that the sequences T and L will have a starting point of 11. These shorter sequences will be used to create the 18 sequences to be strung together as defined in algorithm 4.3. Also by algorithm 4.3 we know the "glues" are now $\{1, 2, 3, 4, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, 104, 105, 106, 107, 108, 109, 110, 111, 112, 113, 114, 115, 116, 117, 118, 119, 120, 121\}$. Then using step 8 of 4.3, we can determine which of the sequences need to be strung together to create each sequence in the new nice pairs. A computer program was written to find how to arrange the 9 sequences and the "glues" that are needed for each sequence in each new pair. The resulting pairs are contained in a link below.

Here is a link to the formulas for all 18 nice pairs.
Case 1: res_9=4, n is odd
Case 2: res_9=4, n is even

•

•

•

Case 17: res_9=12, n is odd
Case 18: res_9=12, n is even

■

5.4.2 Induction proof with k=9

Theorem 5.5. There exists a nice pair that starts with $\{1, 3, 13, 12, 4, 5, 11\}$ and ends with 8 for the odd length sequence and 24 for the even length sequence for every $n \geq 53$.

Proof. Define S_n as follows:

$$S_1 = 53$$
$$S_2 = 481$$
$$S_n = 9 \cdot S_{n-1} + 4 \text{ for } n \geq 3$$

We will prove by induction that for every $k \geq 1$, there exists a nice pair where both sequences start with $(1, 3, 13, 12, 4, 5)$, the odd length sequence ends with 8, and the even length sequence ends with 24 for $S_k \leq a < S_{k+1}$.

Base Case(s): Consider $k = 1$

For $S_1 \leq a < S_2$, we need to show that there exists a nice pair of sequences of the form specified above for all $53 \leq a < 481$. By external reference, it can be verified that such pairs exist for this range. The list provided contains all the nice pairs for these values that meet the necessary criteria.

Inductive Step:

Assume that for some $k \geq 1$, there exists a nice pair where both sequences start with $(1, 3, 13, 12, 4, 5, 11)$, the odd length sequence ends with 8, and the even length sequence ends with 24 for $S_k \leq a < S_{k+1}$.

We need to show that there exists a nice pair of this form for $S_{k+1} \leq a < S_{k+2}$.

Given the inductive hypothesis, we have nice pairs for all a in the interval $S_k \leq a < S_{k+1}$. By theorem 5.5 we know that we can create 9 new nice pairs of this form from each initial nice pair. By algorithm 4.3 and remark 3.10.2, we know the nice pairs created are of the form $\boxed{9n + res_9}$ and $\boxed{9n + res_9 + 1}$ and the smallest value of $res_9 = 4$. Therefore, given nice pairs from $S_k \leq a < S_{k+1}$ we should be able to make nice pairs from $9 \cdot S_k + 4 \leq a < 9 \cdot S_{k+1} + 4$. By definition, that is equivalent to $S_{k+1} \leq a < S_{k+2}$.

Therefore, by induction, there must be a nice pair where both sequences start with $(1, 3, 13, 12, 4, 5, 11)$, the odd length sequence ends with 8, and the even length sequence ends with 24 for every $n \geq 53$.

■

Remark 5.5.1. The proof using the 9 case, we expanded upon Gerbiczs process, finding additional glues to solve the problem that made it seem impossible.

6 Generalizations

This method can be generalized to other sets of numbers beyond squares, but it requires several conditions to met. Firstly, we must be able to multiply any number in the set by some odd constant and receive another number in that set. Secondly, the graph must always be connected after some point. If both conditions are met, then this method has a possibility of working.

6.1 Cubes

The most natural generalization of this conjecture is to have the integers sum to cubes instead of squares. This is certainly more difficult as cubes are more spread apart than squares so paths are likely to appear later than with squares. The conjecture that there exists a point where there is always a path after that point is likely true, but proving it would require a high of computing power. For this reason, a proof that infinitely many of such sequences exists is outlined below that uses a similar method but does not rely on nice pairs. Instead of multiplying the sequences by an odd square, they are multiplied by an odd cube to preserve the cube sum property.

Theorem 6.1. There are infinitely many values of n that have the property that the integers from 1 to n can be arranged in a sequence where each consecutive pair sums to a cube.

Proof. An inductive argument can be made by taking a valid sequence and then finding glue that only relies on that sequence to build a larger sequence that has the same endpoints as before. For the base case, a viable sequence that starts with 1, 7 and ends with 2 is provided in the repository. For the inductive step, a glue that takes a sequence of length n that starts with 1, 7, and ends with 2, and creates a sequence of length $125n + 62$ that also starts with 1, 7 and ends with 2 is also provided in the repository. By taking this base case and using the same glue on it repeatedly, it is possible to construct infinitely many viable sequences. ■

6.2 Even Squares and 9

Theorem 6.2. $\boxed{n}$ forms three disjoint graphs when only using even squares

Proof. All odd numbers squared are odd as $(2m + 1)^2 = 4m^2 + 4m + 1 = 2(2m^2 + 2m) + 1$ and even numbers squared are divisible by 4 as $(2m)^2 = 4k^2$, meaning that all even squares are divisible by 4. This implies that when only using even squares in $\boxed{n}$, integers of the form $4m$ can only share an edge with other integers of the form $4m$, creating our first graph. Likewise, anything of the form $4m + 2$ can only share an edge with other integers of the form $4m + 2$, making the second graph. Finally, integers of the form of $4m + 1$ can only share an edge with integers of the form $4m + 3$, creating a bipartite graph. ■

Another generalization is to try to find a path on $\boxed{n}$ that only uses even squares, but 6.2 makes it clear that since $\boxed{n}$ forms three disjoint graphs when using only even squares, there will only be a path in the trivial case of $n = 1$. This means that at least one odd square must be used in order to create a path. Fortunately, only one odd square is needed and that odd square can be as low as 9.

Lemma 6.3. It is always possible arrange integers of the form $4m$ from 4 to $4m$ in sequence that starts with 4, ends with 8, and has the property that each consecutive pair sums an even square for $n \geq 49$.

Proof. The proof is similar to the 9-case outlined earlier in the paper. The materials to prove that such sequences exist that start with 4 and 12, end in 8 if n is odd, and end in 16 if n is even are provided in the repository with base cases and glue starting at $n \geq 49$. The materials to prove that such sequences exist that start with 4 and 12, end in 16 if n is odd, and end in 8 if n is even, switching the endpoints are provided in the repository with base cases and glue starting at $n \geq 46$ as well. This guarantees that there will always be a sequence that starts with 4 and ends with 8, regardless of the parity of n for $n \geq 49$. ■

Lemma 6.4. It is always possible arrange integers of the form $4m + 2$ from 2 to $4m + 2$ in sequence that starts with 2, ends with 6, and has the property that each consecutive pair sums an even square for $n \geq 43$.

Proof. The proof is similar to the 9-case outlined earlier in the paper. The materials to prove that such sequences exist that start with 2 and 14, end in 6 if n is odd, and end in 10 if n is even are provided in the repository with base cases and glue starting at $n \geq 43$. The materials to prove that such sequences exist that start with 2 and 14, end in 10 if n is odd, and end in 6 if n is even, switching the endpoints are provided in the repository with base cases and glue starting at $n \geq 41$ as well. This guarantees that there will always be a sequence that starts with 2 and ends with 6, regardless of the parity of n for $n \geq 43$.

■

Lemma 6.5. It is always possible arrange the first n odd numbers in a sequence such that the sequence starts with 1 and 3, ends in 5 if n is odd, ends in 7 if n is even, and has the property that each consecutive pair sums an even square for $n \geq 61$.

Proof. The proof is similar to the 9-case outlined earlier in the paper. The materials to prove that such sequences exist are provided in the repository with base cases and glue.

■

Theorem 6.6. For $n \geq 123$, it always possible to arrange the integers from 1 to n in a sequence such that each consecutive pair of integers sums to an even square number or 9.

Proof. By 6.3, 6.4, and 6.5 it is possible to find paths through integers of the form $4m$, $4m+2$, and $2m+1$ individually using only even squares after a certain point and the last thing to do is to link them together using 9. This can be divided into two cases.

Case 1: There are an even number of odd integers from 1 to n

This means that the path from 6.5 starts with 1 and ends with 7. By taking the path from 6.3 which starts with 4 and ends with 8, we then attach the 1 to the 8. Similarly, by taking the path path from 6.4 which starts with 2 and ends with 6, we can attach the 2 to the 7. This completes a path that only uses even squares and 9.

Case 2: There are an odd number of odd integers from 1 to n

This means that the path from 6.5 starts with 1 and 3 and ends with 5. The first step is take off the 1 from the sequence so it starts with 3 and ends with 5 and then reattach it later. Take the path from 6.4 which starts with 2 and ends with 6 and attach it to the 3. Next, take the path from 6.3 which starts with 4 and ends with 8 and attach the the 4 to the 5. Finally, place the 1 after the 8 at the end. This completes a path that only uses even squares and 9.

This process can start as early as $n = 196$ as this process is limited by 6.3 as 196 is the 49 multiple of 4 when starting at 4. All cases where this can be done from 123 to 195 are provided in a text file in the provided repository.

■

6.3 Powers of 3

Theorem 6.7. A graph with vertices consisting of the integers from 1 to n with edges between them if they sum to a power of 3 has no Hamiltonian path for all $n \geq 3$

Proof. The only power of 3 that is not divisible by 3 is $3^0 = 1$, but no two integers in the set of integers from 1 to n could possibly sum to 1. All other powers of 3 are divisible by 3 and so multiples of 3 can only sum with other multiples of 3 to get to a power of 3. This means that the graph must be disconnected as soon as 3 appears and at all time after that. Since this graph is disconnected, as soon as 3 appears, it has no Hamiltonian path. ■

Since powers are significantly more spread apart than squares, it may be tempting to attempt to use this to prove that there exists a point such that for all n after that point, the integers from 1 to n can be arranged in a sequence such that each consecutive pair of integers sums to a power of 3. By 6.7, however, it is clear that such a point will never exist. By letting the integers sum to multiples of powers of 3 using a finite set of multiples, there is a possibility of guaranteeing a path after a certain point. This is possible by utilizing the fact that 1 is a power of 3 and using our set of multiples to overcome any problems caused by modular arithmetic.

Lemma 6.8. If a and b sum to some constant c times a power of 3, then $3^m \cdot a$ and $3^m \cdot b$ sum to c times a power of 3.

Proof. Since a and b sum to c times a power of 3, $a+b = c \cdot 3^d$. It then follows that $3^m \cdot a + 3^m \cdot b = 3^m(a+b) = 3^m \cdot c \cdot 3^d = c \cdot 3^{m+d}$.

■

Theorem 6.9. It is possible to arrange the integer from 1 to n in a sequence such that each consecutive pair sums to an integer in the set $\{c \cdot 3^m | c \in \{1, 5, 7, 13, 17\}, m \in \mathbb{N}\}$ for all n.

Proof. By 6.8, it is possible to multiply existing sequences by a power of 3, which is an odd multiple, and still have pairs sum to integers in the desired set, making Gerbicz's method still applicable in this context. Glues that involve multiplying existing sequences by 3, base cases that start at 8 and all sequences before 8 are all provided in the repository to prove that there will always be a viable sequence.

■

Remark 6.9.1. While several multiples must be used to ensure that there is always a path, it likely can be done using fewer than 5 multiples as shown here. The minimum is believed to be 4.

Remark 6.9.2. A similar statement can almost assuredly be proven using odd powers other than 3, but the larger the power, the more multiples will have to be used.

6.4 Other Generalizations

Several generalizations of this method have been given in the previous section. For other generalizations with a different set of numbers, there are two necessary conditions that must be met. First, there must be a point in which the graph is always connected. If there are problems related to modular arithmetic or the density of the numbers is too low, this method will never work as there will never be a point where there is always a path in the first place. Similarly, it must be a set of numbers where it is possible to multiply everything in the set by an odd number and get a another number in that set. This condition is not necessary for the existence of a path, but it is necessary for this proof method to work.

6.5 Factorials

Theorem 6.10. For any finite set of positive integers, C, there exists an arbitrarily large n such that the integers from a 1 to n cannot be arranged in a sequence where each consecutive pair of numbers sums to an integer in the set of factorial multiples defined as $\{c \cdot m! | c \in C, m \in \mathbb{N}\}$.

Proof. Let d denote the largest element in C and $m \geq 2d - 1$. From there, it follows that

$$\begin{aligned} m + 1 &\geq 2d \\ (m+1) \cdot m! &\geq 2d \cdot m! \\ (m+1)! &\geq 2d \cdot m!. \end{aligned}$$

Given that d is defined as the as the largest integer in C, the next smallest integer that could possibly be in the set of factorial multiples after $d \cdot m!$ is $(m+1)!$. Since $(m+1)!$ is at least twice as large as $d \cdot m!$, if $n = d \cdot m!$, then n will have nothing that it can sum with that is smaller than itself to be in the set of factorial multiples, implying that the integers from 1 to n will be unable to form a viable sequence. Since m is able to get arbitrarily large, so too will n and so there exists an arbitrarily large n that has no viable sequence.

■

A similar conjecture cannot be made about factorials as 6.10 suggests, regardless of how many multiples are used. The density of a set of integers is important as to whether or not a path can always exist or not and powers appear to be the limit. Factorials are generally considered to be the next step up from powers and while it is possible to prove that there is always a path through powers of 3 using only a few multiples, it is also possible to prove that there cannot always be a path through through selected multiples of factorials after a certain point, no matter how many multiples are chosen.

7 A New Result

7.1 An Interesting Pattern

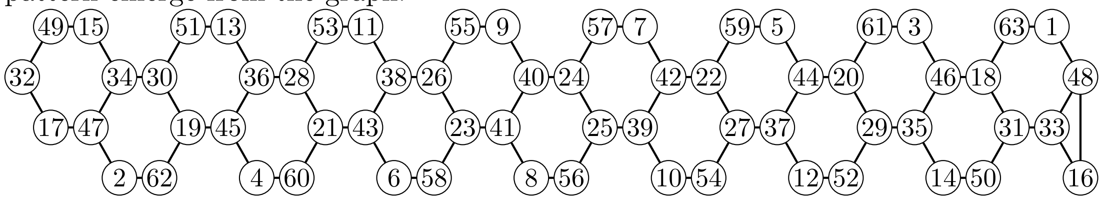
As cool as Gerbicz's Proof is, the question of whether or not one could achieve a similar result via a different method is still worth asking. The results of John Kelly, Samuel Moyer, and Laura Olson, the previous group of Luther students who worked on this project, indicate a number of underlying patterns, some of which seem promising in the pursuit of a nicer proof. Let's take another look at the first example presented in this paper:

9 25 9 25 9 25 9 25
8,1 , 15 , 10 , 6,3 , 13 , 12 , 4,5 , 11 , 14 , 2,7 , 9 , 16
16 16 16 16 16 16 16

In blue above and below the main sequence are the squares summed to by each pair of the sequence. There are two important patterns in these squares. The first is the alternation between 16 and the odd squares on either side of it. We looked into this first, to see if we could make paths for n bigger than 16 where every other edge taken is along the middle square. We will be back to this idea later. The second pattern is the use of only 3 unique perfect squares in the creation of this path. For $n = 16$ it is clear to see why this is, as the only connection between any numbers in the set that isn't taken in this path is the connection between 1 and 3 that sums to 4. In addition to this there are no possible pairs in the integers 1 through 16 that sum to any square other than 4, 9, 16, or 25, so it makes sense to only use three squares. However, we wondered if paths for n other than 15, 16, and 17 could be found using only the edges drawn between pairs that sum to only 3 specific squares. A quick search by a computer program indicated all n of the form $n = q^2 - 1$, where q is an integer, seemed to always have a valid path using only 3 squares. The 3 squares used in that path are $(q+1)^2$, q^2, and $(q-1)^2$.

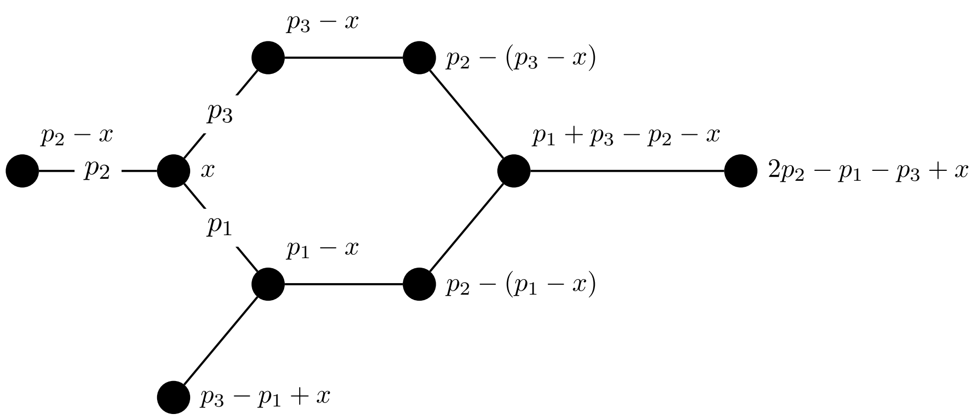
7.2 Visualizations

Since any vertex on a graph using only 3 squares can have at most degree 3, drawing a meaningful graph of this type was much less daunting than a graph with all possible edges for a similar n. Upon drawing the graph for several smaller q, we can see a relatively consistent hexagonal pattern emerge from the graph.



As we can see in this graph of $n = 63$, the majority of the numbers fit nicely into one of many hexagons formed, with the exception of 16, which makes a triangle with 33 and 48 on the right hand side of the graph. There are several patterns contained within this graph and other graphs of this form for different $n = q^2 - 1$. But first lets look at why this form makes hexagons at all.

To make this kind of graph, we can set some parameters as to how we draw it. By starting with some vertex x, we draw the edges to the other vertices its connected to, following a horizontal line (-) if the two numbers sum to q^2, going along / if the two numbers sum to $(q+1)^2$, and \ if they sum to $(q-1)^2$. If we name these moves, we can use them to explain some of the patterns seen in the graphs of this type. We'll call the movement along the $(q-1)^2$ edge p_1, the movement along the q^2 edge p_2, and the movement along the $(q+1)^2$ edge p_3. Then we can draw a more general version of the graph.



Using a general phrasing, we can examine several of the patterns from the hexagonal graphs. The first and most important is the relationship between some number x, and the number represented by $2p_2 - p_1 - p_3 + x$ in our general graph. If we look at vertices in the same relative position in our graph of $n = 63$, we can see that those numbers have a relationship where the second number is the first number minus two. To prove that this is the case, we can look at the relationship described by $2p_2 - p_1 - p_3$ in terms of $(q-1)^2, q^2$, and $(q+1)^2$.

$$\begin{aligned} &2q^2 - (q-1)^2 - (q+1)^2 \\ &= 2q^2 - q^2 + 2q - 1 - q^2 - 2q - 1 \\ &= -2 \end{aligned}$$

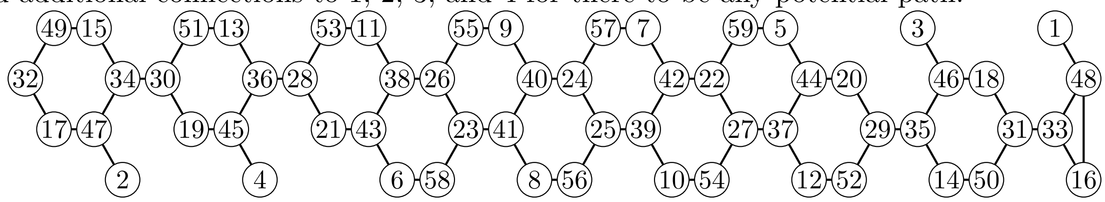
Using this framework, the authors have come up with the following conjectures.

Conjecture 7.1. For all n of the form $n = q^2 - 1$, as long as $n \geq 63$ a path can be found using only the edges $(q+1)^2$, q^2, and $(q-1)^2$.

Conjecture 7.1 was checked for $n = q^2 - 1$ up too 440 by a computer finding paths using a modified version of previously discussed code. In addition to this, graphs were constructed for each n up to 1155. These graphs followed the hexagonal pattern of the shown 63 graph, and seemed to have one of four predictable shapes based on whether q was 0mod4, 1mod4, 2mod4, or 3mod4. The authors believe that this fact could be used to prove Conjecture 7.1.

Conjecture 7.2. For all $n \geq 55$, a path can be found using $(q+1)^2$, q^2, $(q-1)^2$, and $(q-a)^2$, where q^2 is the next perfect square above n, and a is some integer $2 \leq a \leq q$.

Conjecture 7.2 is the generalization of Conjecture 7.1, trying to solve for any n using one additional group of edges. If we look at the graph from earlier, but remove a couple of the points, we can see that $(q+1)^2$, q^2, and $(q-1)^2$ provide a very strong structure, but we must add additional connections to 1, 2, 3, and 4 for there to be any potential path.



Much of the structure still remains from the $n = q^2 - 1$ case, and it seems very reasonable that there is some way to connect this graph in a way that lets us follow a predictable path. It would seem that the most obvious step is to use $a = 2$ instead of some vague $(q-a)^2$, but when checked it isn't consistently able to find paths at any point. Other constant values for a were checked, but it seems like you need to increase the value of a as n increases, based on preliminary searches. Despite the failure to find an easy option for what a 4th square should be, we can still be hopeful that it exists. A proof of this manner would be really beneficial in understanding an underlying structure of the problem, if one exists.

8 Closing Remarks

Even though there is a solution to the square sum problem for all n, it still stands to reason that there is very likely a far nicer proof than the proofs outlined in sections 3-6. The proof we now have is very nice when it comes to generating these paths using a computer, but it is hard to understand and the method only works for perfect squares and very similar sets of numbers, with little other possible application elsewhere. Another proof of this conjecture based on existing patterns within the graph, such as our hexagon lattice, could help lead to new insights in related areas. The general drawing of a hexagonal lattice works for numbers summing to any p_1, p_2, and p_3, though the connectedness of the graph depends on the 3 numbers chosen. In addition to that, in drawing graphs of this type, we glimpsed a great deal of interesting patterns that weren't apparently relevant and were difficult to describe. Though solved, there is still a lot to understand about this problem.

9 Acknowledgements

Any work that we referenced in this paper can be found in our GitHub repository. The README file should do a good job to explaining what every file contains.

This research was supported by Luther College, the Tomson Family Faculty Fellowship in Mathematics Fund*, and the Luther College Actuarial Science Endowment Fund.