Finding Bounds for the Number of Sudoku Squares

Beth Schaab

July 28, 2008

1 Introduction

This project will concentrate on looking at Sudoku squares from a graph theoretical and combinatorial standpoint. When given a particular Sudoku puzzle, a few questions might come to mind. For instance, "Is there only one solution to this puzzle?" or "Is there even a solution to this puzzle?" Mathematicians have also been looking at the minimum number of initial entries that need to be given in order for there to be a unique solution to a given puzzle.

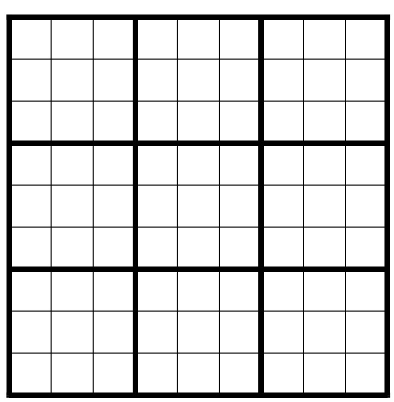
This paper will look at counting Sudoku solutions using tools from graph theory and combinatorics and the relationship between Latin Squares and Sudoku squares. With respect to the relationship between Latin Squares and Sudoku squares, specific tools will be used to look at Sudoku squares including some fundamental ideas of graph theory, which include proper colorings, chromatic numbers, regular graphs, and partial colorings. Rank 2 Sudoku squares will be useful in getting a general idea of how to count Sudoku squares. Permanents of matrices and systems of distinct representatives will be helpful in understanding and proving theorems. Hall's Marriage Theorem as well as Stirling's Formula will be utilized in order to prove two major theorems. The paper "Sudoku Squares and Chromatic Polynomials" by Agnes M. Herzberg and M. Ram Murty will be used as a guide for this paper. Proofs will be worked through and gaps filled in where necessary.

2 Definitions

It is beneficial to establish what a Sudoku square is and some basic properties that every Sudoku square has.

Definition 1 *A* ***Sudoku square of rank*** n *is an* $n^2 \times n^2$ *grid completed in a way such that the entries of the grid are the numbers* 1 *through* n^2*. A Sudoku square contains* n *horizontal bands made up of* n *rows and* n *vertical bands made up of* n *columns. The intersection of a horizontal band and a vertical band is a subsquare. It is required that the square is completed such that every row, column, and subsquare contains the numbers* 1 *to* n^2 *exactly once.*

Subsquares of a rank 3 Sudoku square are shown below.



A Latin Square of rank n is an $n \times n$ grid such that the entries of each column and row are the numbers 1 through n. Every Sudoku square of rank n is a Latin square of rank n^2.

A typical Sudoku square is a Latin square of rank 9. It contains 9 rows and 9 columns. The Sudoku square of rank 3 contains 3 horizontal bands as well as 3 vertical bands. The Sudoku square contains 9 subsquares.

Definition 2 *The **Sudoku graph of rank** n**,** X_n, is a graph on $(n^2)^2$ vertices. Two entries a_{ij} and $a_{i'j'}$ (where i, j, i', and j' are the numbers 1 through n^2) are adjacent if and only if $i = i'$, $j = j'$, or $\lceil \frac{i}{n} \rceil = \lceil \frac{i'}{n} \rceil$ **and** $\lceil \frac{j}{n} \rceil = \lceil \frac{j'}{n} \rceil$.*

Two vertices (i, j) and (i', j') are adjacent if they are located in the same row of the Sudoku square, $i = i'$, the same column of the Sudoku square, $j = j'$, or the same subsquare. In order to determine whether or not two vertices are in the same subsquare, it is necessary to take the greatest integer function of $\frac{i}{n}, \frac{i'}{n}, \frac{j}{n}$ and $\frac{j'}{n}$. Taking $\lceil \frac{i}{n} \rceil = \lceil \frac{i'}{n} \rceil$ will determine whether the two vertices are in the same horizontal band and taking $\lceil \frac{j}{n} \rceil = \lceil \frac{j'}{n} \rceil$ will determine whether or not the two vertices are in the same vertical band. When the two vertices are in the same horizontal band and vertical band, then they lie in the same subsquare.

A Sudoku graph can be thought of as a graph on $(n^2)^2$ vertices where each vertex represents an entry in our Sudoku square. The Sudoku graph of rank 3 will have $(3^2)^2 = 81$ vertices, each corresponding to a cell in the Sudoku puzzle. Two distinct vertices will be adjacent if and only if the corresponding cells in the grid are in the same row, column, or subsquare. It is convenient to convert a Sudoku puzzle into a graph so tools from graph theory can be utilized to analyze properties of Sudoku graphs.

Definition 3 *A **coloring** of a graph $G = \{V, E, \phi\}$ is an assignment of a color to each vertex in V.* [2]

Definition 4 *A coloring of a graph is a **proper coloring** if two vertices in every pair of distinct adjacent vertices are assigned different colors.* [2]

Using the concept of a proper coloring, if a_{ij} is the entry in a Sudoku graph that is located in the i^{th} row and j^{th} column, the entry must be assigned a different "color" (or numbers, in the case of Sudoku) than any entry that is located in the i^{th} row and j^{th} column, as well as any entry that is in the 3×3 subgraph where a_{ij} is located. In the case of a Sudoku graph, assigning each vertex a different "color" means that no number can be repeated within a row, column, or a subsquare. So, the entry a_{ij} must not be the same "color" or number as any of these adjacent entries. Figure 1 shows the entries that are adjacent to the a_{11} entry in a part of the Sudoku graph X_3.

Definition 5 *The minimum number of colors that are needed to properly color a graph, G, is called the **chromatic number** of G.* [3]

The Sudoku graph is properly colored when any two adjacent vertices in the Sudoku graph are assigned different numbers.

3 Graph Theory Results about Sudoku Puzzles

Definition 6 *A graph is called **regular** if the degrees of all of its vertices are the same.* [3]

Every vertex is adjacent to all vertices in the column that the initial vertex is located in, so there are $n^2 - 1$ vertices in the column that are adjacent to the initial vertex. The initial vertex must also be adjacent to every vertex in the row that the initial vertex is located in, so there are $n^2 - 1$ vertices in that row that are adjacent to the vertex. And the initial vertex must also be adjacent to every vertex that is located in the same subsquare, so there are $(n - 1)^2$ vertices in that subsquare which are adjacent to the vertex.

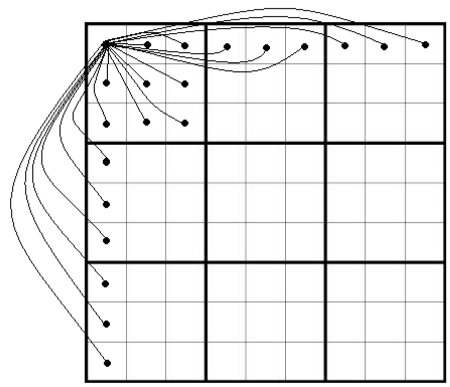


Figure 1: The entries adjacent to the a_{11}^{th} entry. This is a piece of the Sudoku graph showing the vertices that lie in the 1^{st} row and the 1^{st} column, as well as the subgrid in which a_{11} is located.

The subsquare has $(n-1)^2$ vertices because the uppermost row and left most column of the subsquare are accounted for, so $(n-1)^2$ vertices are left to examine. The degree of each vertex is

$$\begin{aligned} & (n^2-1)+(n^2-1)+(n-1)^2 \\ = \;& n^2-1+n^2-1+n^2-2n+1 \\ = \;& 3n^2-2n-1. \end{aligned}$$

Thus, each vertex of a Sudoku graph has a degree of $3n^2-2n-1$ or $(3n+1)(n-1)$.

The Sudoku graph is a regular graph since each vertex has to be adjacent to the same number of other vertices. The first theorem looks at partial colorings of a graph, G in relation to the number of ways of completing the coloring of the graph to obtain a proper coloring of the graph.

The following theorem, which will not be used, is an example of the type of results one can obtain. Its proof is beyond the scope of this project.

Theorem 7 (Theorem 1 from [3]) *Let G be a finite graph with v vertices. Let C be a partial proper coloring of t vertices of G using d_0 colors. Let $\rho_{G,C}(\lambda)$ be the number of ways of completing this coloring using λ colors to obtain a proper coloring of G. Then, $\rho_{G,C}(\lambda)$ is a monic polynomial (in λ) with integer coefficients of degree $v-t$ for $\lambda \geq d_o$.*

When a partial coloring of a Sudoku graph is presented, it is interesting to determine when a unique coloring can be obtained. The following theorem provides some insight.

Theorem 8 (Theorem 2 from [3]) *Let G be a graph with chromatic number $\chi(G)$ and let C be a partial coloring of G using only $\chi(G)-2$ colors. If the partial coloring can be completed to a total proper coloring of G, then there are at least two ways of extending the coloring.*

Proof. Two colors have not been used in the proper coloring of G. Thus, these two colors can be interchanged in completing the final coloring of G. ■

3.1 Explicit Coloring for X_n

Since a Sudoku graph is a graph, we can find the chromatic number of a Sudoku graph, X_n.

Theorem 9 *The chromatic number of X_n is n^2.*

Proof. Looking at a rank 3 Sudoku square and finding a general way of filling in that square, it can be seen that certain measures have to be taken in order to comply with the rules of Sudoku. The first band can be completed in the following way:

1	2	3	4	5	6	7	8	9
4	5	6	7	8	9	1	2	3
7	8	9	1	2	3	4	5	6

The first row can be filled in quite simply. The second row has to follow the rule that no numbers can be repeated within a subsquare. Thus, the numbers in the second row are shifted three places to the left to follow that rule. The same logic applies to the third row, again shifting the numbers from the second row three places to the left. The first band is now filled in following the Sudoku rules. The focus shifts to the second band of the Sudoku puzzle. For this band, the rule that no number can be repeated in any column has to be taken into account. This can be achieved by taking the first row of the Sudoku puzzle and shifting each number to the left one place. Once this shift has taken place, the same technique from the first band of shifting the next two rows three places to the left can be used. The first two bands of the Sudoku puzzle now look like this:

1	2	3	4	5	6	7	8	9
4	5	6	7	8	9	1	2	3
7	8	9	1	2	3	4	5	6
2	3	4	5	6	7	8	9	1
5	6	7	8	9	1	2	3	4
8	9	1	2	3	4	5	6	7

Finally, the focus shifts to the third and final band in the Sudoku puzzle. The same logic from the first and second band can be applied. Again, the rule that no number can be repeated in any column has to be addressed. This can be achieved by looking at the first row of the second band (or the 4^{th} row), and again shifting the entries one place to the left. From this point on, the entries of the next two rows can be shifted three places to the left to create a valid Sudoku puzzle that looks similar to the following:

1	2	3	4	5	6	7	8	9
4	5	6	7	8	9	1	2	3
7	8	9	1	2	3	4	5	6
2	3	4	5	6	7	8	9	1
5	6	7	8	9	1	2	3	4
8	9	1	2	3	4	5	6	7
3	4	5	6	7	8	9	1	2
6	7	8	9	1	2	3	4	5
9	1	2	3	4	5	6	7	8

This technique works because it involves looking at the rules that make the Sudoku puzzle unique and addressing them. The process of shifting the entries one place to the left ensures that there are no entries repeated within a column. The process of shifting the entries three places to the left ensures that there are no entries repeated within a subsquare. Making sure that no entries are repeated within a row is a trivial process.

This same argument can be applied to a generic rank n Sudoku graph. The process for filling in the first row remains the same. The technique for the rest of the rows in the band changes slightly. Instead of shifting the entries three places to the left (as was done in the rank 3 case), they are shifted n places to the left. The technique for the first row in every band remains the same.

It also needs to be shown that the graph cannot be filled in with $n^2 - 1$ colors. In order to show this, consider a row in the rank 3 case. If there are only 8 colors in which to complete a row, then by the constraints of Sudoku, the puzzle can never be completed without using a number which has already been used in the row. Thus, there must be at least n^2 colors. $\blacksquare$

4 Counting Sudoku Squares of Rank 2

When considering the total number of Sudoku squares, one must consider the fact that two Sudoku puzzles might actually be the intimately related. This could be achieved by utilizing a number of valid symmetries. For example, when given a completed Sudoku solution, the process of relabeling numbers provides 9! new Sudoku solutions. There are several other symmetries to consider. Permutations of any of the 3 bands (vertically or horizontally) can occur, as well as any of the rows of the bands. To simplify this process, take into consideration the 4×4 Sudoku square shown below.

Without loss of generality, the first subsquare can be completed in the following way:

1	2		
3	4		

Interchanging the last two columns is a valid symmetry, so taking this into account, the first row can be completed as $(3,4)$ in this order. Similarly, the last two rows can be interchanged, thus the Sudoku grid can be filled in as follows:

1	2	3	4
3	4		
2			
4			

The only choices for entries in the diagonal position of $(3,3)$ are 1 and 4, since 2 has been used in the 3^{rd} row and 3 has been used in the 3^{rd} column. Some simple experimentation shows that the entry of 1 will not work.

1	2	3	4
3	4	2	1
2	3	1	
4	1		

If one were to fill in the Sudoku square using 1 in the $(3,3)$ position, 4 could not be anywhere in the 3^{rd} column. So, 4 must be the entry in $(3,3)$. The rank 2 Sudoku puzzle now looks like:

1	2	3	4
3	4		
2		4	
4			

Then the Sudoku square can take any of the following forms:

1	2	3	4
3	4		
2		4	
4			1

1	2	3	4
3	4		
2		4	
4			2

1	2	3	4
3	4		
2		4	
4			3

Any of these three possibilities can be finished according to the rules of Sudoku as follows:

1	2	3	4
3	4	1	2
2	1	4	3
4	3	2	1

1	2	3	4
3	4	2	1
2	1	4	3
4	3	1	2

1	2	3	4
3	4	1	2
2	3	4	1
4	1	2	3

However, it can been seen that the 2^{nd} and 3^{rd} completed Sudoku puzzle are related by a symmetry. This can be seen by taking the transpose of the 3^{rd} completed puzzle and permuting the numbers 2 and 3. So, there are really only 2 unique solutions for a rank 2 Sudoku puzzle, the 1^{st} and 2^{nd} completed Sudoku puzzles. So, considering that the 4 numbers in the puzzle can be interchanged, the last two rows can be interchanged, and the last two columns can be interchanged in the three different Sudoku puzzles, there are a total of $4! \times 2 \times 2 \times 3 = 288$ Sudoku squares of rank 2.

5 Counting Sudoku Squares

Considering that in the previous section, it was shown the number of valid Sudoku squares of rank 2 is 288, the focus shifts to finding the number of valid Sudoku squares of rank 3. In a paper by Felgenhauer and Jarvis [1], the number of Sudoku squares was found to be $6,670,903,752,021,072,936,960$. It is nearly impossible to count the different solutions by hand, so several tools will be used to help with the process.

The permanent of a matrix will be essential in determining a bound on the number.

Definition 10 *The **permanent** of an $n \times n$ matrix, A, is*

$$per(A) = \sum_{\sigma \in S_n} a_{1\sigma(1)} a_{2\sigma(2)} ... a_{n\sigma(n)} \tag{1}$$

where S_n denotes the symmetric group on n elements [6].

To determine the number of Latin squares of rank n, we will use Hall's Theorem. In order to understand Hall's Theorem, it is essential to first define what a system of distinct representatives is from [6].

Definition 11 *Suppose there are subsets $A_1, A_2, ..., A_n$ of the set $\{1, 2, ..., n\}$. A choice of $a_i \in A_i$ for each $i \in \{1, 2, ..., n\}$ such that $a_i \neq a_j$ when $i \neq j$ is a* ***system of distinct representatives*** [3].

Hall's Theorem is as follows (from [6]):

Theorem 12 *A collection of subsets $A_1, A_{2,}..., A_n$ has a system of distinct representatives if and only if for all $k = 1, ..., n$ and any choice of distinct indices $i_1, ..., i_k$ then every collection of k subsets $A_{i1}, A_{i2}, ..., A_{ik}$ satisfies $|A_{i1} \cup A_{i2} \cup ... \cup A_{ik}| \geq k$.*

5.1 Hall's Theorem Examples

To help unravel the conditions of Hall's Theorem, it is helpful to see examples of when Hall's Theorem does and does not find a system of distinct representatives. Examples follow.

5.1.1 Hall's Theorem guarantees a system of distinct representatives

Suppose there are 3 subsets of the set $\{1, 2, 3\}$, where

$$\begin{aligned} A_1 &= \{1, 2\} \\ A_2 &= \{2, 3\} \\ A_3 &= \{1, 3\}. \end{aligned}$$

It is necessary to check k values of $1, 2$, and 3. When $k = 1$,

$$\begin{aligned} i_1 &= 1 \\ A_1 &= \{1, 2\} \\ |A_1| &\geq 1 \\ 2 &\geq 1, \end{aligned}$$

or

$$\begin{aligned} i_1 &= 2 \\ A_2 &= \{2, 3\} \\ |A_2| &\geq 1 \\ 2 &\geq 1, \end{aligned}$$

or

$$\begin{aligned} i_1 &= 3 \\ A_3 &= \{1, 3\} \\ |A_3| &\geq 1 \\ 2 &\geq 1. \end{aligned}$$

When $k = 2$,

$$\begin{aligned} i_1 &= 1, \; i_2 = 2 \\ |A_1 \cup A_2| &\geq 2 \\ 3 &\geq 2, \end{aligned}$$

or

$$\begin{aligned} i_1 &= 2,\ i_2 = 3 \\ |A_2 \cup A_3| &\geq 2 \\ 3 &\geq 2, \end{aligned}$$

or

$$\begin{aligned} i_1 &= 1,\ i_2 = 3 \\ |A_1 \cup A_3| &\geq 2 \\ 3 &\geq 2. \end{aligned}$$

When $k = 3$,

$$\begin{aligned} i_1 &= 1,\ i_2 = 2,\ \ i_3 = 3 \\ |A_1 \cup A_2 \cup A_3| &\geq 3 \\ 3 &\geq 3. \end{aligned}$$

So, for this particular subset of numbers, Hall's theorem guarantees a system of distinct representatives. A system of distinct representatives for this set is choosing 1 from A_1, 2 from A_2, and 3 from A_3.

5.1.2 Hall's Theorem guarantees a lack of a system of distinct representatives

Suppose there are 3 subsets of the set $\{1, 2, 3\}$, where

$$\begin{aligned} A_1 &= \{2\} \\ A_2 &= \{2\} \\ A_3 &= \{1, 3\}. \end{aligned}$$

It is necessary to check k values of $1, 2$, and 3. There are 3 subsets of the set $\{1, 2, 3\}$, thus it is necessary to check the values $k = 1, 2, 3$. When $k = 2$,

$$\begin{aligned} i_1 &= 1 \\ i_1 &= 2 \\ |A_1 \cup A_2| &\geq 2 \\ 1 &\not\geq 2. \end{aligned}$$

When $k = 2$, it is evident that Hall's Theorem does not hold for this particular set. Thus, there is no system of distinct representatives for this particular set.

5.2 Bounding Sudoku Squares of Rank n

Definition 13 *A **doubly stochastic matrix** is a matrix in which the entries are nonnegative numbers and the row and column sums are equal to* 1.

Definition 14 *A **Hall matrix** for sets $A_1, ..., A_n \subseteq \{1, 2, ..., n\}$ is a $(0,1)$ matrix whose $(i,j)^{th}$ entry is 1 if and only if $i \in A_j$.*

The permanent of a Hall matrix will give the number of systems of distinct representatives.

In 1926, B.L. van der Waerden [5] determined that for any doubly stochastic matrix

$$per A \geq \frac{n!}{n^n}. \tag{2}$$

This is referred to as the van der Waerden Theorem.

In 1967, H. Minc conjectured (and it has since been proven, see [8]) that an upper bound for the permanent of a $(0,1)$ matrix, A, with row sums r_i is

$$perA \leq \prod_{i=1}^{n} r_i!^{\frac{1}{r_i}}. \tag{3}$$

Lemma 15 *The number of Latin squares of rank n is at least $\frac{n!^{2n}}{n^{n^2}}$.*

Proof. The first row of a Latin square of rank n can be filled in $n!$ ways because there are n choices for the first entry, $n-1$ choices for the second entry, and so on. Assume that k rows of the Latin square have been completed and the $(k+1)^{st}$ row needs to be filled in. Let A_i be the set of numbers not yet used in the ith column. A_i is the entire set of n numbers minus the k numbers that have already been used. The number of ways to fill in the $(k+1)$st row is the number of ways of choosing a system of distinct representatives of the sets $A_1, A_2, ..., A_n$. The number of ways in which this can be done is the permanent of the corresponding Hall Matrix, A. Looking at the rows of the Hall matrix, there will be $(n-k)$ 1's in each row because there are n possible choices for numbers and k of them have been used in filling in the first k rows. The row sum of the Hall Matrix is $n-k$. Similarly, the column sum of the Hall Matrix is also $n-k$; the columns represent whether or not a particular number can be entered in a particular row. Each row has to contain one and only one of the numbers $1, 2, ..., n$. Thus, when determining the number of ways of filling the k^{th} row, a particular number has been used in k out of n rows. Thus, the column sums are also $n-k$. To use the bounds stated previously, we need to make sure we have a doubly stochastic matrix. Currently, the row sums and column sums are greater than 1. To make the row and column sums equal 1, we will multiply matrix A by $\frac{1}{n-k}$. Multiplying matrix A by $\frac{1}{n-k}$ and thus taking the permanent leads to

$$per\left(\frac{1}{n-k} \cdot A\right).$$

The number of ways to compute $(k+1)$st row is $per(A)$. Using the definition of a permanent (Equation (1)), we have $\frac{1}{n-k}$ multiplied by itself n times, or $\left(\frac{1}{n-k}\right)^n$. The equation can be rewritten as

$$per\left(\frac{1}{n-k}A\right) = \left(\frac{1}{n-k}\right)^n per(A).$$

Using Equation (2), the upper bound can be applied to the permanent of A,

$$\left(\frac{1}{n-k}\right)^n per(A) \geq \frac{n!}{n^n}.$$

Solving for $per(A)$, yields

$$per(A) \geq \frac{(n-k)^n n!}{n!}.$$

The number of ways of filling in the first row through the $(n-1)^{st}$ row are multiplied, thus,

$$\begin{aligned}
\prod_{k=0}^{n-1} \frac{(n-k)^n n!}{n^n} &= \left(\frac{n!}{n^n}\right) \prod_{k=0}^{n-1} (n-k)^n \\
&= \frac{(n!)^n}{n^{n^2}} \cdot \prod_{k=0}^{n-1} (n-k)^n \\
&= \frac{(n!)^n}{n^{n^2}} \cdot (n!)^n \\
&= \frac{(n!)^n (n!)^n}{n^{n^2}} \\
&= \frac{n!^{2n}}{n^{n^2}}.
\end{aligned}$$

■

We will be discussing asymptotic bounds where we analyze behavior for large n. First we discuss "big oh" notation. As a reference for "big oh" notation we use [4].

Definition 16 *Let $f : \mathbb{R} \to \mathbb{R}$ and $g : \mathbb{R} \to \mathbb{R}$ be functions. We write*

$$f(n) \in O(g(n))$$

and say that $f(n)$ is of order at most $g(n)$ or $f(n)$ is big oh of $g(n)$ if there exists positive constants C and N such that

$$|f(n)| \leq C|g(n)| \forall n \geq N.$$

There are certain properties that hold with respect to "big oh" notation. The first lemma looks at what happens when the function is multiplied by a constant.

Lemma 17 *Let $f \in O(h)$. Then $\forall k \in \mathbb{R}$, $kf \in O(h)$.*

Proof. Let f be $O(h)$. By definition, f is $O(h)$ so $\exists C, N > 0$ such that

$$|f(x)| \leq C|(h(x))| \forall x > N.$$

Multiplying both sides by $|k|$,

$$|k||f(x)| \leq |k|C|(h(x))| \ \forall x > N.$$

C and k both represent constants, thus taking $|k| \cdot C$ results in another constant, C_1.

$$|kf(x)| \leq C_1|(h(x))| \ \forall x > N.$$

So, for all $k \in \mathbb{R}$, $kf \in O(h)$. ■

The next lemma looks at what happens when two functions are added together.

Lemma 18 *Let $f, g \in O(h)$. Then $(f + g) \in O(h)$.*

Proof. Let f and g be $O(h)$. By definition, f is $O(h)$ so $\exists C_1, N_1 > 0$ such that

$$|f(x)| \leq C_1|(h(x))| \forall x > N_1. \tag{4}$$

Also by definition, g is $O(h)$ so $\exists C_2, N_2 > 0$ such that

$$|g(x)| \leq C_2|(h(x))| \forall x > N_2. \tag{5}$$

Adding Equations (4) and (5),

$$\begin{aligned} |f(x)| + |g(x)| &\leq C_1|(h(x))| + C_2|(h(x))| \\ &\leq |h(x)|[C_1 + C_2] \forall x > \max\{N_1, N_2\}. \end{aligned}$$

Again, since C_1 and C_2 are constants, $C_1 + C_2$ will also be a constant, C_3. Thus,

$$|f(x) + g(x)| \leq C_3|(h(x))| \forall x > \max\{N_1, N_2\}.$$

■

Stirling's Formula will also be helpful in the following theorem. Stirling's formula provides a way to approximate logarithmic values of factorials for very large values of n.

Lemma 19 *Stirling's Formula by* [3] *is*

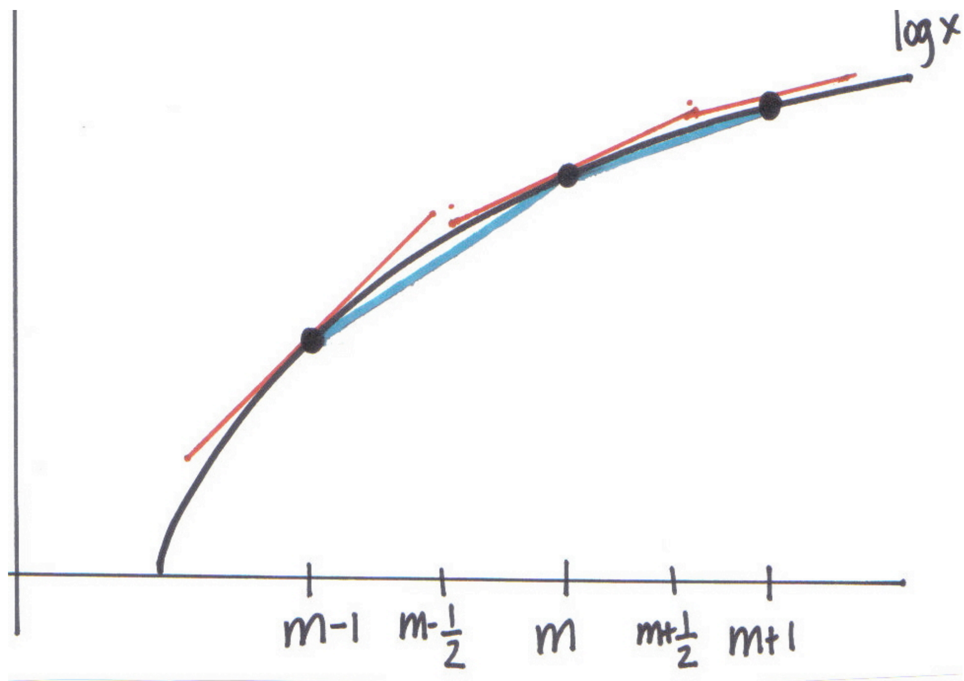


Figure 2: The graph of $f(x)$ (blue curve) and $g(x)$ (red curve) compared with $\log x$ (black curve).

$$\log n! = n \log n - n + \frac{1}{2} \log n + O(1).$$

Proof. Rudin [6] outlines a simple computation which yields Sterling's Formula. For $m = 1, 2, 3, ...$, define

$$f(x) = (m + 1 - x) \log m + (x - m) \log(m + 1) \tag{6}$$

if $m \leq x \leq m + 1$, and define

$$g(x) = \frac{x}{m} - 1 + \log m \tag{7}$$

if $m - \frac{1}{2} \leq x < m + \frac{1}{2}$. Figure 2 shows the graphs of $f(x)$ (blue curve) and $g(x)$ (red curve) as well as the graph of $\log x$ (black curve). When investigating the graph of $f(x)$, it is important to note that $m \leq x \leq m + 1$. So, when $x = m$ and $x = m + 1$, $f(x)$ will be

x	**f(x)**
m	$\log m$
$m + 1$	$\log(m + 1)$

When looking at the graph of $g(x)$, it is important to see that $m - \frac{1}{2} \leq x < m + \frac{1}{2}$. So, when $x = m - \frac{1}{2}$ and $x = m + \frac{1}{2}$, $f(x)$ will be

x	**g(x)**
$m - \frac{1}{2}$	$-\frac{1}{2m} + \log m$
$m + \frac{1}{2}$	$\frac{1}{2m} + \log m$

Looking at Figure 2, $g(x)$ is the tangent line to $\log x$, and since $\log x$ is concave down, then the tangent line will always lie above a concave down curve, so $f(x) \leq \log x \leq g(x)$ if $x \geq 1$. Rudin suggests integrating $f(x)$ on the interval $[1, n]$. In order to integrate $f(x)$, we use the method of Riemann Sums. Since $f(x)$ connects the points on $\log(x)$, one can find the area under $f(x)$ by using the $\log(x)$ curve and right hand sums (Figure 3). Using Figure 3, it is easy to see that

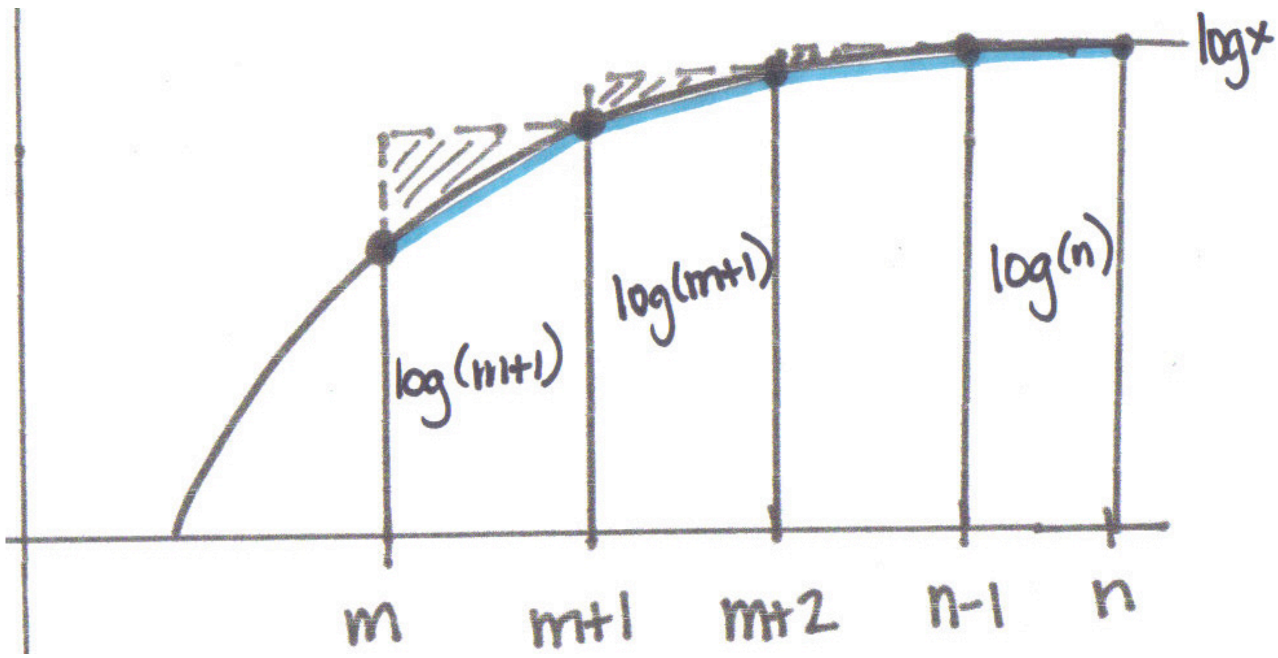


Figure 3: The graph of $f(x)$ with the corresponding right hand Reimann Sums.

$$\int_1^n f(x)dx = \log n + \log(n-1) + ... + \log 3 + \log 2 + \log 1 + p \tag{8}$$

where p represents the excess area of rectangles not under the curve of $\log x$. Using the method of sliding the excess triangle to the right, one can see that the excess part of the graph will be less than the area of the last rectangle, which is $\log n$. The question then becomes, how much error do we have? Using an interesting "sliding" argument, we can see that the error is definitely less than $\log n$. Looking at Figure 4, the portion of the rectangle that lies above the $\log x$ curve can be "slid" to the right and be contained in the rectangle immediately to the right. Continuing this process for all the Reimann Sum rectangles, it is easily seen that the error will be less than $\log n$. But is there a way to be more accurate? Consider the following: The portion of the rectangle that lies above the curve of $f(x)$ looks like the left picture in Figure 5. Taking half the width of the top leg of the triangle, and creating a rectangle with the two halves of the triangle, we get something that looks like the right picture in Figure 5. To find the area of the rectangle, we know that the vertical side of the rectangle stays the same, while the length of the horizontal leg is half of the original length. So, the area of the rectangle created by the excess triangle is half the original area. So, the area of the excess is $\frac{1}{2}\log n$, since the area of the original rectangle was $\log n$. So, now Equation (8) becomes

$$\int_1^n f(x)dx = \log(n!) + \frac{1}{2}\log n \tag{9}$$

We can look at the integral of $\log x$ on the interval $[1, n]$,

$$\begin{aligned} \int_1^n \log x dx &= (x\log x - x)|_1^n \\ &= (n\log n - n) - (1\log 1 - 1) \\ &= n\log n - n + 1. \end{aligned} \tag{10}$$

Combining Equations (9) and (10), since we know that $f(x) \leq \log x$, one can now see that

$$\int_1^n f(x)dx = \log(n!) + \frac{1}{2}\log n \leq n\log n - n + 1.$$

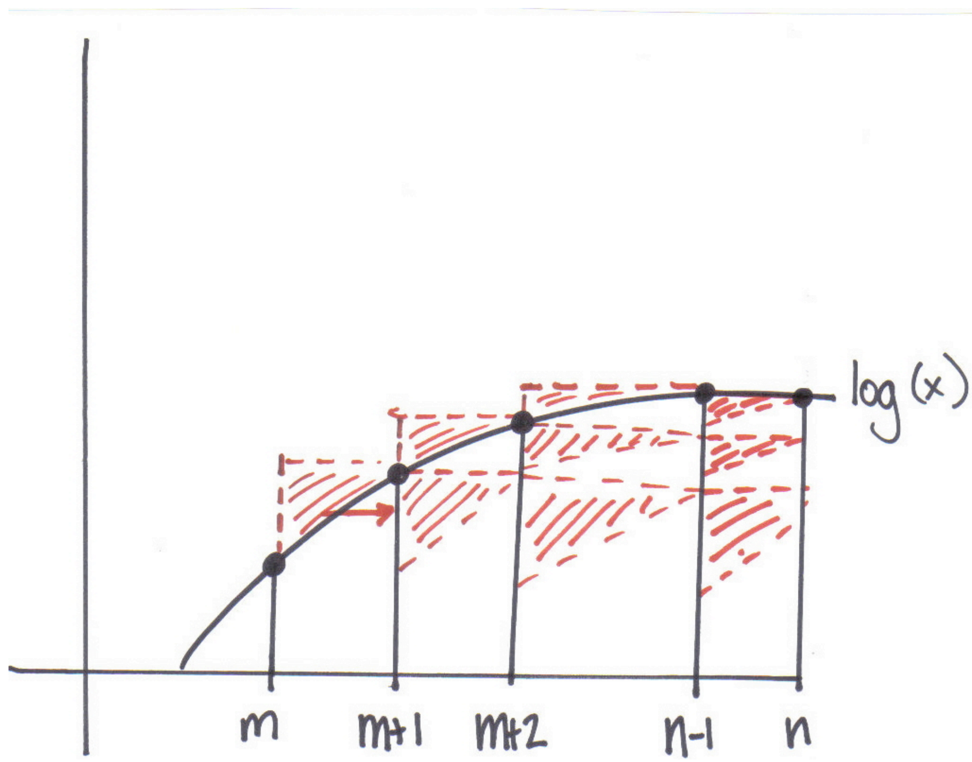


Figure 4: Sliding the excess areas of Reimann Sum triangles to fit within the area of the rectangle with height $\log n$.

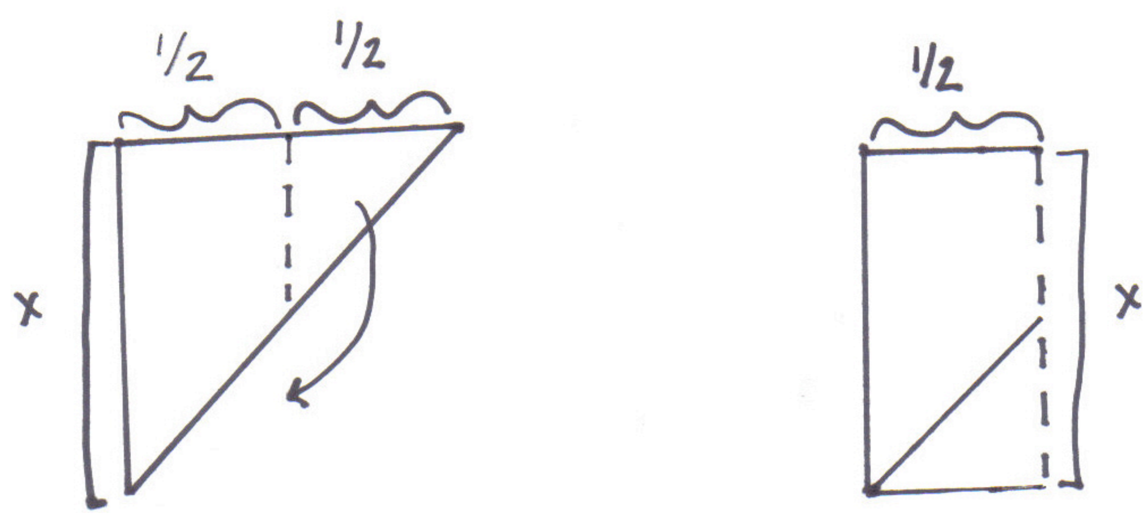


Figure 5: The portion of the rectangle that lies above $f(x)$.

So,

$$\log(n!) + \frac{1}{2}\log n - n\log n + n \leq 1.$$

To find a lower bound, one needs to determine the difference between $f(x)$ and $g(x)$, which are Equations (6) and (7),

$$f(x,m) - g(x,m) = -\left[\frac{[m-x][1+m\log m - m\log(m+1)]}{m}\right] \tag{11}$$

which is valid on $m \leq x \leq m+\frac{1}{2}$. So, the difference between $g(x)$ and $f(x)$ is represented in Equation (11). Integrating Equation (11) when x ranges from $[m, m+\frac{1}{2}]$,

$$-\int_m^{m+\frac{1}{2}} \left[\frac{[m-x][1+m\log m - m\log(m+1)]}{m}\right] dx = \frac{1}{8m} + \frac{\log m}{8} - \frac{1}{8}\log(1+m). \tag{12}$$

In order to integrate the difference of $g(x)$ and $f(x)$ when x ranges from $[m+\frac{1}{2}, m+1]$, we compute with $f(x,m) - g(x,m+1)$.

$$-\int_{m+\frac{1}{2}}^{m+1} \left[\frac{[1+m-x][1+(1+m)\log m - (1+m)\log(1+m)]}{1+m}\right] dx = \frac{1}{8}\left(-\frac{1}{1+m} - \log m + \log(m+1)\right). \tag{13}$$

Adding Equations (12) and (13) simplifies to

$$\frac{1}{8m+8m^2}.$$

Now, one has to compute the infinite series of

$$\begin{aligned}\sum_{m=1}^{\infty} \frac{1}{8m+8m^2} &= \frac{1}{8}\sum_{m=1}^{\infty}\frac{1}{m^2+m} \\ &= \frac{1}{8}\sum_{m=1}^{\infty}\frac{1}{m} - \frac{1}{m+1}\end{aligned}$$

which is actually a telescoping series. Working out the beginning of the series,

$$\frac{1}{8}\left[\left(\frac{1}{1}-\frac{1}{2}\right) + \left(\frac{1}{2}-\frac{1}{3}\right) + \left(\frac{1}{3}-\frac{1}{4}+ ...\right)\right]$$

one can see that all the terms except for when $m = 1$ will drop out, concluding that series sums to $\frac{1}{8}$. What has just been shown is that

$$\int_1^{\infty} [g(x) - f(x)]dx = \frac{1}{8}$$

but, if we only want to integrate from $[1, n]$, then it can be said that

$$\int_1^n [g(x) - f(x)]dx < \frac{1}{8}.$$

Rearranging the terms,

$$-\frac{1}{8} + \int_1^n g(x)dx < \int_1^n f(x)dx$$

and using Equation (9),

$$-\frac{1}{8} + \int_1^n g(x)dx < \log(n!) + \frac{1}{2}\log n.$$

Subtracting $\int_1^n \log x dx$ from both sides gives us

$$\begin{aligned}
\log(n!) + \frac{1}{2}\log n &> -\int_1^n \log x dx - \frac{1}{8} + \int_1^n g(x)dx \\
\log(n!) + \frac{1}{2}\log n - (n\log n - n + 1) &> -\frac{1}{8} + \int_1^n g(x)dx - \int_1^n \log x dx \\
\int_1^n f(x)dx - \int_1^n \log x dx &= \log(n!) + \frac{1}{2}\log n - n\log n + n - 1 > -\frac{1}{8} + \int_1^n g(x)dx - \int_1^n \log x dx.
\end{aligned}$$

Adding 1 to the equality gives us

$$\int_1^n f(x)dx - \int_1^n \log x dx + 1 > \log(n!) + \frac{1}{2}\log n - n\log n + n > \frac{7}{8} + \int_1^n g(x)dx - \int_1^n \log x dx$$

but also because

$$\int_1^n f(x)dx - \int_1^n \log x dx < 0$$

we can say

$$1 > \int_1^n f(x)dx - \int_1^n \log x dx + 1,$$

and similarly, since

$$\int_1^n g(x)dx - \int_1^n \log x dx > 0,$$

then

$$\frac{7}{8} + \int_1^n g(x)dx - \int_1^n \log x dx > \frac{7}{8}.$$

So, we can finally say that

$$1 > \log(n!) + \frac{1}{2}\log n - n\log n + n > \frac{7}{8}.$$

Since $\log(n!) + \frac{1}{2}\log n - n\log n + n$ is less than 1 , $\log(n!) + \frac{1}{2}\log n - n\log n + n \in O(1)$. ■

Theorem 20 (Corollary 5 from [3]) *The number of Latin squares of rank n^2 is at least*

$$n^{2n^4} e^{-2n^4 + O(n^2 \log n)}.$$

Proof. Lemma 13 states that the number of Latin squares of rank n is at least

$$\frac{n!^{2n}}{n^{n^2}}.$$

The number of Latin squares of order n^2 is found by replacing every n in Lemma 4 see with an n^2. Thus, we have

$$L = \frac{n^{2n^2}!}{n^{2n^4}}. \tag{14}$$

Using the properties of logarithms, taking the log of Equation (14) yields,

$$\begin{aligned}
\log\left(\frac{n^2!^{2n^2}}{n^{2n^4}}\right) &= \log(n^2!)^{2n^2} - \log n^{2n^4} \\
&= 2n^2 \log n^2! - \log n^{2n^4}
\end{aligned}$$

Now, using Stirling's Formula for $n^2!$,

$$\begin{aligned}
\log(L) &= 2n^2[n^2\log n^2 - n^2 + \frac{1}{2}\log n^2 + O(1)] - \log n^{2n^4} \\
&= 2n^4 \log n^2 - 2n^4 + n^2 \log n^2 + O(n^2) - \log n^{2n^4} \qquad (15)
\end{aligned}$$

Rearranging the terms of Equation (15) yields

$$\begin{aligned}\log(L) &= 2n^4 \log n^2 + n^2 \log n^2 - \log n^{2n^4} - 2n^4 + O(n^2) \\ &= 4n^4 \log n - 2n^4 + n^2 \log n^2 + O(n^2) - 2n^4 \log n\end{aligned}$$

Rearranging and applying some properties of logarithms yields,

$$\begin{aligned}\log(L) &= (4n^4 \log n - 2n^4 \log n) - 2n^4 + n^2 \log n^2 + O(n^2) \\ &= 2n^4 \log n - 2n^4 + n^2 \log n^2 + O(n^2)\end{aligned}$$

Let L be the number of Latin squares. Then,

$$\begin{aligned}L &= e^{\log(L)} = e^{(2n^4 \log n - 2n^4 + n^2 \log n^2 + O(n^2))} \\ &= e^{(\log n^{2n^4} - 2n^4 + 2n^2 \log n + O(n^2))} \\ &= n^{2n^4} e^{-2n^4 + O(n^2 \log n)}.\end{aligned}$$

■

5.2.1 Example of filling in a rank 3 Sudoku square

Before looking at the number of ways of filling in a general rank n Sudoku square, let's take a look at the number of ways of filling in a rank 3 Sudoku square.

The number of ways of filling in the first row of the Sudoku puzzle is rather trivial. There are 9! ways of filling in the first row. Now, in order to determine the number of ways of filling in the second row, one must look at the Hall matrix which corresponds to row 2. The row sum of the Hall matrix is 6 because three numbers have been used in the first subsquare, so the permanent of this Hall matrix will be the number of ways of filling in the second row. From Equation (3), the permanent is bounded by

$$per(A) \leq \prod_{i=1}^{9} 6!^{\frac{1}{6}}.$$

In order to determine the number of ways of filling in the third row, one must look at the Hall matrix which corresponds to row 3. The row sum of the Hall matrix is 3 because six numbers have been used already in the first subsquare, so the permanent is bounded by

$$per(A) \leq \prod_{i=1}^{9} 3!^{\frac{1}{3}}.$$

Now, the number of ways of completing the first band of the Sudoku graph have been calculated. The focus now shifts to the number of ways of filling in the second band. The first entry of the first row of the second band has six choices for entries. The entries that have already been used in the first column need to be excluded, so there are 6 choices for the first entry in the first row of the second band. Thus, the row sum of the corresponding Hall matrix will be 6 , so the permanent is bounded by

$$per(A) \leq \prod_{i=1}^{9} 6!^{\frac{1}{6}}.$$

The number of ways of filling in the second row of the second band is slightly more complicated. One must take into consideration two things, the numbers that have been used in the first column as well as the numbers that have been used in the subsquare in which the entry is located. For the second row of the second band, there have been 4 numbers used in the column, while only 3 numbers have been used in the subsquare. It is important to exclude the most potential entries, so the numbers used in the column need to be excluded. So, there are up to 5 choices for the first entry of the second row of the second band. The row sum of the corresponding Hall matrix is 5, and the permanent is bounded by

$$per(A) \leq \prod_{i=1}^{9} 5!^{\frac{1}{5}}.$$

The number of ways of filling in the last row of the second band again uses the same method. Looking at the column, 5 numbers have previously been used, while 6 numbers have been used in the subsquare. Thus, it makes sense that there are only 3 choices for the first entry of the last row of the second band. So, the row sum of the corresponding Hall matrix is 3, and the permanent is bounded by

$$per(A) \leq \prod_{i=1}^{9} 3!^{\frac{1}{3}}.$$

The process for filling in the last band proceeds in the same way. In filling in the first row of the first band, 6 numbers have already been used in the column, so there are 3 choices for the first entry of the first row of the third band. So, the row sum of the corresponding Hall matrix is 3, and the permanent is bounded by

$$per(A) \leq \prod_{i=1}^{9} 3!^{\frac{1}{3}}.$$

The number of ways of filling in the second row of the third band takes into consideration that 7 colors have been used in the column, so there are only 2 choices for first entry of the second row of the third band. The row sum of the corresponding Hall matrix is 2, and the permanent is bounded by

$$per(A) \leq \prod_{i=1}^{9} 2!^{\frac{1}{2}}.$$

The number of ways of filling in the last row of the third band is somewhat trivial. Since 8 numbers have been used in the column already, there is only 1 choice for the first entry of the last row of the third column. Thus, the row sum of the corresponding Hall matrix is 1, and the permanent is bounded by

$$per(A) \leq \prod_{i=1}^{9} 1!^{\frac{1}{1}}.$$

Thus, the number of ways of filling in a rank 3 Sudoku square is the product of the number of ways of filling in each of the rows. The approximate number of ways of filling in a rank 3 Sudoku square is 1.70719×10^{26}. This is a horrible bound, since the actual number is known and it is no where close!

There are approximately 1.70719×10^{26} ways of filling in a rank 3 Sudoku square. However, if one wants to look at the general rank n case Sudoku square, it would be nice to have a way of determining the number of Sudoku squares possible. The following theorem considers such a instance.

Theorem 21 (Theorem 6 **from** [3]**)** *The number of Sudoku squares of rank n is bounded above by*

$$n^{2n^4} e^{-2.5n^4 + O(n^3 \log n)}.$$

Proof. A rank n Sudoku square is composed of n^2 subsquares of size $n \times n$. There are n bands in the Sudoku grid. The ways of completing the first band can be estimated in the following way: There are $n^2!$ choices for the first row. In order to find the ways of completing the second row, evaluate the permanent of the $n^2 \times n^2$ matrix whose rows correspond to the second row of our Sudoku grid. For each cell, there are $n^2 - n$ possibilities because we have already used n numbers in a corresponding $n \times n$ subsquare. So, the number of ways of filling the second row can be found by determining the number of ways of choosing a system of distinct representatives from the list of possible entries. The permanent is bounded by (Equation (3))

$$perA \leq \prod_{i=1}^{n} r_i!^{(1/r_i)}, \tag{16}$$

where r_i represents the row sum of the row of the matrix. In this case, the row sum of this matrix is (n^2-n). Now, substitute the known row sum into Equation (16) to get

$$\begin{aligned} per(A) &\leq \prod_{i=1}^{n^2}(n^2-n)!^{\frac{1}{n^2-n}} \\ &= (n^2-n)!^{\frac{n^2}{n^2-n}} \\ &= (n^2-n)!^{\frac{n}{n-1}} \end{aligned}$$

Using the same type of argument for the third row, the row sum is n^2-2n. Plugging the row sum into Equation 16 the number of ways of filling in the third row are

$$(n^2-2n)!^{\frac{n}{n-2}}.$$

So, using this same type of argument again, the final estimate of the number of ways of filling in the first band of a Sudoku square of rank n is

$$\prod_{k=0}^{n-1}(n^2-kn)!^{\frac{n}{n-k}}.$$

Now, assume that $(i-1)$ of n bands have been completely filled in. How many ways can we fill in the i^{th} band? To do this, find out how many possible entries there are for the first cell of the ith band. There are $[n^2-(i-1)n]$ entries for the first cell of the band because there are n^2 possible entries and $(i-1)n$ of these entries have been used in the column. The number of ways of completing the 2^{nd} row of the i^{th} band is at most

$$\begin{aligned} &(n^2-((i-1)n+1)!^{\frac{n^2}{n^2-((i-1)n+1)}} \\ =\ &(n^2-(in-n+1))!^{\frac{n^2}{n^2-(in-n+1)}}. \end{aligned}$$

Using the same process to determine the number of ways of filling in the i^{th} row of the i^{th} band, to get an estimate of

$$(n^2-(in-n+i))!^{\frac{n^2}{n^2-((i-1)n+i)}}.$$

The goal of this entire process is to get the best estimate for the number of ways of filling in the n^{th} row of the i^{th} band. There are two methods that can be used to find the best estimate for filling in the n^{th} row of the i^{th} band. Up until now, using the method of looking at the entries in the column that the entry is in has provided the best estimate for the number of ways of completing that entry. For the $(i-1)^{st}$ row, it becomes a better estimate to look within the subsquare that the entry is located in and exclude all of the numbers that have been used in the subsquare, rather than in the column that the entry belongs.

For the $(i-1)^{st}$ row, switch techniques and exclude the numbers that are already entered in the subsquare in which the cell being looked at belongs. For example, when filling in the first cell of the $(i+1)^{th}$ row, n^2+in entries have been used. These entries have already been entered in the subsquare and excluding these gives an estimate of

$$(n^2-in)!^{\frac{n^2}{n^2-in}}.$$

for the number of ways of filling in the i^{th} row of the i^{th} band. Use the same strategy to get that the number

S_n of Sudoku squares of rank n is bounded by

$$
\begin{aligned}
S_n &\leq \prod_{i=1}^{n}\left[\prod_{j=0}^{i-1} n^2-[(i-1)n+j]!^{\frac{n^2}{n^2-[(i-1)n+j]}}\right]\left[\prod_{k=i}^{n-1}(n^2-kn)!^{\frac{n^2}{n^2-kn}}\right] \\
\log S_n &\leq \sum_{i=1}^{n}\left(\sum_{j=0}^{i-1}\log(n^2-[(i-1)n+j])!^{\frac{n^2}{n^2-[(i-1)n+j]}}+\sum_{k=i}^{n-1}\log[n^2-kn]!^{\frac{n^2}{n^2-kn}}\right) \\
\log S_n &\leq \sum_{i=1}^{n}\left(\sum_{j=0}^{i-1}\frac{n^2}{n^2-[(i-1)n+j]}\cdot\log(n^2-[(i-1)n+j])!+\sum_{k=i}^{n-1}\frac{n^2}{n^2-kn}\log[n^2-kn]!\right) \\
\log S_n &\leq n^2\cdot\sum_{i=1}^{n}\left(\sum_{j=0}^{i-1}\frac{\log(n^2-[(i-1)n+j])!}{n^2-[(i-1)n+j]}+\sum_{k=i}^{n-1}\frac{\log[n^2-kn]!}{n^2-kn}\right) \\
\frac{\log S_n}{n^2} &\leq \sum_{i=1}^{n}\left(\sum_{j=0}^{i-1}\frac{\log(n^2-[(i-1)n+j])!}{n^2-[(i-1)n+j]}+\sum_{k=i}^{n-1}\frac{\log[n^2-kn]!}{n^2-kn}\right)
\end{aligned}
$$

Now, Stirling's formula can be used to approximate the values of $\log(n^2-[(i-1)n+j])!$ and $\log[n^2-kn]!$,

$$
\begin{aligned}
\frac{\log S_n}{n^2} &\leq \sum_{i=1}^{n}\left[\begin{array}{c}\sum_{j=0}^{i-1}\frac{n^2-[(i-1)n+j]\log(n^2-[(i-1)n+j])-(n^2-[(i-1)n+j])+\frac{1}{2}\log(n^2-[(i-1)n+j])+O(1)}{n^2-[(i-1)n+j]}+ \\ \sum_{k=1}^{n-1}\frac{(n^2-kn)\log(n^2-kn)-(n^2-kn)+\frac{1}{2}\log(n^2-kn)+O(1)}{n^2-kn}\end{array}\right] \\
&\leq \sum_{i=1}^{n}\left[\begin{array}{c}\sum_{j=0}^{i-1}\log(n^2-[(i-1)n+j])-\frac{(n^2-\{(i-1)n+j])}{(n^2-\{(i-1)n+j])}+\frac{1\log(n^2-[(i-1)n+j)}{2(n^2-[(i-1)n+j]}+\frac{O(1)}{(n^2-[(i-1)n+j])} \\ +\sum_{k=i}^{n-1}\log(n^2-kn)-\frac{(n^2-kn)}{(n^2-kn)}+\frac{1\log(n^2-kn)}{2(n^2-kn)}+\frac{O(1)}{(n^2-kn)}\end{array}\right] \\
&\leq \sum_{i=1}^{n}\left[\begin{array}{c}\sum_{j=0}^{i-1}\log(n^2-[(i-1)n+j])-1+\frac{1\log(n^2-[(i-1)n+j)}{2(n^2-[(i-1)n+j]}+\frac{O(1)}{(n^2-[(i-1)n+j])}+ \\ \sum_{k=i}^{n-1}\log(n^2-kn)-1+\frac{1\log(n^2-kn)}{2(n^2-kn)}+\frac{O(1)}{(n^2-kn)}\end{array}\right] \qquad (17)
\end{aligned}
$$

$\sum\limits_{i-1}^{n}\left[\sum\limits_{j=0}^{i-1}-1\right]$ reduces to $\frac{-n(n+1)}{2}$ and similarly $\sum\limits_{i-1}^{n}\left[\sum\limits_{k=i}^{n-1}-1\right]=-\left[n^2-\frac{n(n+1)}{2}\right]$. Now, Equation (17) further simplifies to

$$
\begin{aligned}
&\leq -\frac{n(n+1)}{2}-\left[n^2-\frac{n(n+1)}{2}\right]+\sum_{i=1}^{n}\left[\begin{array}{c}\sum\limits_{j=0}^{i-1}\log(n^2-[(i-1)n+j])+\frac{\log(n^2[(i-1)n+j])}{2(n^2-[(i-1)n+j])}+\frac{O(1)}{(n^2-[(i-1)n+j])} \\ +\sum\limits_{k=i}^{n-1}\log(n^2-kn)+\frac{\log(n^2-kn)}{2(n^2-kn)}+\frac{O(1)}{(n^2-kn)}\end{array}\right] \\
&\leq -n^2+\sum_{i=1}^{n}\left[\sum_{j=0}^{i-1}\log(n^2-[(i-1)n+j])+\frac{\log(n^2-[(i-1)n+j])}{2(n^2-[(i-1)n+j])}+\frac{O(1)}{(n^2-[(i-1)n+j])}+\sum_{k=i}^{n-1}\left(\begin{array}{c}\log(n^2-kn) \\ +\frac{\log(n^2-kn)}{2(n^2-kn)}+\frac{O(1}{(n^2-k}\end{array}\right.\right.
\end{aligned}
$$

Concentrating on the $\frac{\log(n^2[(i-1)n+j])}{2(n^2-[(i-1)n+j])}$ term, and working it out in full detail, it can be seen that

$$
\begin{aligned}
\sum_{i=1}^{n}\sum_{j=0}^{i-1}\frac{\log(n^2[(i-1)n+j])}{2(n^2-[(i-1)n+j])} &= \left[\begin{array}{c}\frac{\log(n^2)}{2(n^2)}+\frac{\log(n^2-n)}{2(n^2-n)}+\frac{\log(n^2-n-1)}{2(n^2-n-1)}+\frac{\log(n^2-2n)}{2(n^2-2n)} \\ +\frac{\log(n^2-2n-1)}{2(n^2-2n-1)}+\frac{\log(n^2-2n-2)}{2(n^2-2n-2)}+\ldots+\frac{\log(n)}{2(n)}+\frac{\log(n-1)}{2(n-1)}+\ldots+\frac{\log(1)}{2(1)}\end{array}\right] \\
&\leq \sum_{k=1}^{n^2}\frac{\log(k)}{k}.
\end{aligned}
$$

We approximate the series by interpreting it as a Reimann Sum of $\frac{\log k}{k}$. The function reaches a maximum at $k=e$ and then decreases, which means that it's slope is always negative, as it approaches ∞. Taking the integral of the function will give the total area under the curve,

$$\sum_{k=1}^{n^2} \frac{\log(k)}{k} = \frac{\log(2)}{2} + \frac{\log(3)}{3} + \sum_{k=4}^{n^2} \frac{\log(k)}{k}$$

So, one can see, that

$$\sum_{i=1}^{n}\sum_{j=0}^{i-1} \frac{\log(n^2[(i-1)n+j])}{2(n^2-[(i-1)n+j])} \leq \frac{\log(2)}{2} + \frac{\log(3)}{3} + \int_3^{n^2} \frac{\log(x)}{x} dx.$$

In order to compute this integral, it is useful to use u substitution. Letting

$$\begin{aligned} u &= \log x \\ du &= \frac{1}{x} dx \end{aligned}$$

yields

$$\begin{aligned} \int u\ du &= \frac{u^2}{2} + C \\ &= \frac{(\log x)^2}{2} + C. \end{aligned}$$

So,

$$\begin{aligned} \sum_{k=1}^{n^2} \frac{\log(k)}{k} &\leq \int_3^{n^2} \frac{\log(x)}{x} dx + \frac{\log(2)}{2} + \frac{\log(3)}{3} \\ &= \frac{1}{2}(\log x)^2 |_3^{n^2} + \frac{\log(2)}{2} + \frac{\log(3)}{3} \\ &= \frac{1}{2}(\log n^2)^2 - \frac{1}{2}(\log(3))^2 + \frac{\log(2)}{2} + \frac{\log(3)}{3} \\ &= \frac{1}{2} \cdot 2(\log n)^2 - \frac{1}{2}(\log(3))^2 + \frac{\log(2)}{2} + \frac{\log(3)}{3} \\ &\in O(\log^2 n). \end{aligned}$$

It has just been shown that

$$\sum_{i=1}^{n}\sum_{j=0}^{i-1} \frac{\log(n^2[(i-1)n+j])}{2(n^2-[(i-1)n+j])} \in O(\log^2 n).$$

In a similar fashion,

$$\sum_{i=1}^{n}\sum_{k=i}^{n-1} \frac{\log(n^2-kn)}{2(n^2-kn)} \leq \int_3^{n^2} \frac{\ln(x)}{x} dx$$

which leads us to the conclusion that

$$\sum_{i=1}^{n}\sum_{k=i}^{n-1} \frac{\log(n^2-kn)}{2(n^2-kn)} \in O(\log^2 n).$$

Equation (18) now looks like:

$$\begin{aligned} \frac{\log S_n}{n^2} \leq & \sum_{i=1}^{n} \left[\sum_{j=0}^{i-1} \log(n^2-[(i-1)n+j]) + \frac{O(1)}{(n^2-[(i-1)n+j])} + \sum_{k=i}^{n-1} \log(n^2-kn) + \frac{O(1)}{(n^2-kn)} \right] \\ & -n^2 + O(\log^2 n). \end{aligned} \tag{19}$$

One can say that

$$\frac{1}{(n^2-[(i-1)n+j])} < \frac{\log(n^2-[(i-1)n+j])}{(n^2-[(i-1)n+j])}$$

so,

$$\sum_{i=1}^{n}\sum_{j=0}^{i-1}\frac{O(1)}{(n^2-[(i-1)n+j])} \in O\left(\sum_{i=1}^{n}\sum_{j=0}^{i-1}\frac{O(\log(n^2-[(i-1)n+j]))}{(n^2-[(i-1)n+j])}\right).$$

But

$$O\left(\sum_{i=1}^{n}\sum_{j=0}^{i-1}\frac{O(\log(n^2-[(i-1)n+j]))}{(n^2-[(i-1)n+j])}\right) \in O(\log^2 n)$$

And similarly,

$$\sum_{i=1}^{n}\sum_{k=i}^{n-1}\frac{O(1)}{(n^2-kn)} \leq \sum_{i=1}^{n}\sum_{k=i}^{n-1}\frac{\log(n^2-kn)}{2(n^2-kn)}$$

so,

$$\sum_{i=1}^{n}\sum_{k=i}^{n-1}\frac{O(1)}{(n^2-kn)} \in O(\log^2 n).$$

Equation (19) now looks like

$$\frac{\log S_n}{n^2} \leq \sum_{i=1}^{n}\left[\sum_{j=0}^{i-1}\log(n^2-[(i-1)n+j]) + \sum_{k=i}^{n-1}\log(n^2-kn)\right] - n^2 + O(\log^2 n). \tag{20}$$

Working with the term

$$\sum_{j=0}^{i-1}\log(n^2-[(i-1)n+j])$$

it is easily seen that the term is biggest when j is 0, thus,

$$\sum_{j=0}^{i-1}\log(n^2-[(i-1)n+j]) \leq \sum_{j=0}^{i-1}\log(n^2-(i-1)n) \leq i\log(n^2-(i-1)n).$$

Equation (20) now looks like

$$\frac{\log S_n}{n^2} \leq \sum_{i=1}^{n}\left[i\log(n^2-(i-1)n) + \sum_{k=i}^{n-1}\log(n^2-kn)\right] - n^2 + O(\log^2 n). \tag{21}$$

Adding $n^2 - O(n\log n)$ to both sides of Equation (21) yields

$$\frac{\log S_n}{n^2} + n^2 + O(n\log n) \leq \sum_{i=1}^{n}\left[i\log(n^2-(i-1)n) + \sum_{k=i}^{n-1}\log(n^2-kn)\right] - n^2 + O(\log^2 n) + n^2 + O(n\log n). \tag{22}$$

Combining like terms in Equation (22) yields

$$\frac{\log S_n}{n^2} + n^2 \leq \sum_{i=1}^{n}\left[i\log(n^2-(i-1)n) + \sum_{k=i}^{n-1}\log(n^2-kn)\right] + O(\log^2 n) + O(n\log n). \tag{23}$$

Using the properties of big oh notation, $O(\log^2 n) \in O(n\log n)$ because

$$\begin{aligned} \log n &\leq n \ \forall n \geq 1 \\ (\log n)(\log n) &\leq n(\log n) \\ \log^2 n &\leq n\log n, \end{aligned}$$

so equation (23) is now

$$\begin{aligned}
\frac{\log S_n}{n^2} + n^2 + O(n \log n) &\leq \sum_{i=1}^{n} \left[i \log(n^2 - (i-1)n) + \sum_{k=i}^{n-1} \log(n^2 - kn) \right] \\
&\leq \sum_{i=1}^{n} \left[i \log(n^2 - in + n) + \sum_{k=i}^{n-1} \log(n^2 - kn) \right] \\
&\leq \sum_{i=1}^{n} \left[i \log n(n-i+1) + \sum_{k=i}^{n-1} \log n(n-k) \right] \\
&\leq \sum_{i=1}^{n} \left[i \log n + i \log(n-i+1) + \sum_{k=i}^{n-1} [\log n + \log(n-k)] \right]. \qquad (24)
\end{aligned}$$

But,

$$\begin{aligned}
\sum_{k=i}^{n-1} \log n &= \log n \sum_{k=i}^{n-1} 1 \\
&= \log n(n-1-i+1) \\
&= n \log n - i \log n \\
&= (n-i) \log n. \qquad (25)
\end{aligned}$$

Combining Equations (24) and (25) yields

$$\frac{\log S_n}{n^2} + n^2 + O(n \log n) \leq \sum_{i=1}^{n} \left[i \log n + i \log(n-i+1) + (n-i) \log n + \sum_{k=i}^{n-1} \log(n-k) \right]. \qquad (26)$$

Now,

$$\begin{aligned}
\sum_{i=1}^{n} i \log n &= \log n \sum_{i=1}^{n} i \\
&= \frac{n(n+1)}{2} \log n \\
&= \frac{n^2}{2} \log n + \frac{n}{2} \log n \\
&= \frac{n^2}{2} \log n + O(n \log n). \qquad (27)
\end{aligned}$$

Combining (26) and (27),

$$\frac{\log S_n}{n^2} + n^2 + O(n \log n) \leq \frac{n^2}{2} \log n + \sum_{i=1}^{n} \left[i \log(n-i+1) + (n-i) \log n + \sum_{k=i}^{n-1} \log(n-k) \right]. \qquad (28)$$

The summation over k is $\log(n-i)!$ because

$$\begin{aligned}
\sum_{k=i}^{n-1} \log(n-k) &= \log(n-i) + \log(n-(i+1)) + \log(n-(i+2)) + \ldots \log(2) + \log(1) \\
&= \log((n-i)!).
\end{aligned}$$

Using Sterling's Formula on $\log(n-i)!$ yields

$$\log(n-i)! = (n-i) \log(n-i) - (n-i) + \frac{1}{2} \log(n-i) + O(1).$$

Combining the approximation for $\log(n-i)!$ and Equation (28) yields

$$\frac{\log S_n}{n^2} + n^2 + O(n\log n) \leq \frac{n^2}{2}\log n + \sum_{i=1}^{n}\left[\begin{array}{c} i\log(n-i+1) + (n-i)\log n + \\ (n-i)\log(n-i) - (n-i) + O(\log n) \end{array}\right]. \tag{29}$$

Some of the summations are equivalent to:

$$\begin{aligned} \sum_{i=1}^{n}(n-i)\log n &= \frac{(n^2-n)}{2}\log n = \frac{n^2}{2}\log n - \frac{n}{2}\log n & (30)\\ \sum_{i=1}^{n}(n-i) &= \frac{(n^2-n)}{2} = \frac{n^2}{2} - \frac{n}{2}. & (31) \end{aligned}$$

Adding

$$\sum_{i=1}^{n} i\log n(n-i+1) + \sum_{i=1}^{n}(n-i)\log(n-i) = (n+1)(\log(n!)) \tag{32}$$

and also,

$$\sum_{i=1}^{n} O(\log n) \tag{33}$$

is $O(n\log n)$, which is already in Equation (29). Combining Equations (29), (30), (31), and (32), yields

$$\begin{aligned}
\frac{\log S_n}{n^2} + n^2 + O(n\log n) &\leq \frac{n^2}{2}\log n + (n+1)(\log(n!)) + \frac{n^2}{2}\log n - \frac{n}{2}\log n - \frac{n^2}{2} + \frac{n}{2}\\
\frac{\log S_n}{n^2} + O(n\log n) &\leq -n^2 + \frac{n^2}{2}\log n + (n+1)(\log(n!)) + \frac{n^2}{2}\log n - \frac{n}{2}\log n - \frac{n^2}{2} + \frac{n}{2}\\
\frac{\log S_n}{n^2} + O(n\log n) &\leq -n^2 + \frac{n^2}{2}\log n + (n+1)\log(n!) + \frac{n^2}{2}\log n - \frac{n}{2}\log n - \frac{n^2}{2} + \frac{n}{2}\\
\frac{\log S_n}{n^2} + O(n\log n) &\leq -n^2 + \frac{n^2}{2}\log n + n^2\log n - n^2 + \frac{n}{2}\log n + n\log n - n + \frac{1}{2}\log n\\
&\quad + \frac{n^2}{2}\log n - \frac{n}{2}\log n - \frac{n^2}{2} + \frac{n}{2}\\
\frac{\log S_n}{n^2} &\leq -2.5n^2 + 2n^2\log n + O(n\log n)\\
\log S_n &\leq -2.5n^4 + 2n^4\log n + O(n^3\log n)\\
e^{\log S_n} &\leq e^{-2.5n^4} + e^{2n^4\log n} + e^{O(n^3\log n)}\\
S_n &\leq n^{2n^4} + e^{-2.5n^4 + O(n^3\log n)}.
\end{aligned}$$

■

6 Conclusion

It is interesting to see how a seemingly simple Sudoku puzzle can have ties so deeply rooted in mathematics. This paper has provided upper and lower bounds for the number of Sudoku puzzles of rank n, but presently, no precise number is known. It is also known that in order for a unique solution to exist for a Sudoku puzzle, there must be at least 8 different numbers presented initially.

7 References

[1] B. Felgenhauer and A. F. Jarvis. Mathematics of Sudoku I. *Mathematical Spectrum* **39**(2006), 15 − 22.

[2] Gossett, E. (2003). Discrete Mathematics with Proof. New Jersey: Prentice Hall.

[3] Herzberg, A. M., and Murty, M. R. Sudoku squares and chromatic polynomial. *Notices of the American Mathematical Society,* **54**(2007) no. 6, $708-717$.

[4] Johnsonbaugh, R. (2005). Discrete Mathematics (6^{th} ed.). Upper Saddle River, NJ: Prentice Hall.

[5] Knuth, D. A permanent inequality. *American Mathematics Monthly*, **88**(1981) no. 10, $731-740$.

[6] Rosen, K. H. (2002). Handbook of Discrete and Combinatorial Mathematics. New York, NY: CRC Press.

[7] Rudin, W. (1976). Principles of Mathematical Analysis (3^{rd} ed.). New York, NY: McGraw-Hill.

[8] van Lint, J.H. and Wilson, R.M. *A Course in Combinatorics*, Cambridge University Press, 1992.