ANALYSIS OF GENERALIZED SUDOKU PUZZLES: A MIXTURE OF DISCRETE TECHNIQUES

By

Ivan Wycliffe Haynes

Bachelor of Science
University of South Carolina, Columbia, South Carolina 1998

Submitted in Partial Fulfillment of the Requirements

for the Degree of Master of Science in

Mathematics

College of Arts and Sciences

University of South Carolina

2008

Accepted by:

Eva Czabarka, Director of Thesis

David Sumner, Second Reader

James Buggy, Dean of the Graduate School

Acknowledgments

First, I would like to thank my wife Joanna and my family. Your love and support has made this possible. For being the second reader of this thesis, I would like to thank Dr. Sumner. I also owe thanks to Dr. McNulty for his help with TEX issues. Finally, I would like to thank my thesis advisor, Eva Czabarka. Without her suggestions, explanations, and patience this thesis would not have been possible.

ABSTRACT

A $\mathrm{Dim}(n,m)$ Sudoku puzzle is an $nm \times nm$ grid with $n \times m$ subgrids. We interpret the $\mathrm{Dim}(n,m)$ Sudoku puzzle as a vertex coloring problem in graph theory. This provides a broad framework for investigation. We will also discuss the relationship between Latin squares and Sudoku puzzles and show that the set of $\mathrm{Dim}(n,m)$ Sudoku puzzles is substantially smaller than the set of rank nm Latin squares. Our work is a generalization of a paper that appeared in the "Notices" of the American Mathematical Society, June/July 2007, titled "Sudoku Squares and Chromatic Polynomials". [**8**]

CONTENTS

List of Figures

CHAPTER 1

INTRODUCTION

The Sudoku puzzle is a relatively new phenomenon in the United States that has become very popular. You will find them in many magazines and newspapers alongside the crossword puzzles.

Sudoku puzzles are related to Latin squares, which were developed by the 18$^{\text{th}}$ century Swiss mathematician Leonhard Euler. Latin squares are square-grids of size $n \times n$ where each of the numbers from 1 through n appear in every column and in every row precisely once. They are referred to as rank n Latin squares.

Magic squares are square grids that are filled with (not necessarily different) numbers such that the numbers in each row and column add up to the same sum. It is easy to see that Latin squares are also magic squares.

In the late 19$^{\text{th}}$ century a Paris-based daily newspaper, *Le Siecle* published a partially completed 9×9 magic square that had 3×3 subgrids. The object of the game was to fill out the magic square such that the numbers in the grids also sum to the same number as in the rows and columns.

The standard Sudoku puzzle consists of a partially filled out 9×9 grid in which some of the entries have a number from 1 to 9. We call this a $\mathrm{Dim}(3,3)$ puzzle because it is composed of subgrids of size 3×3. The challenge is to complete the grid in such a way that each row, column, and all nine 3×3 sub-grids contain each of the numbers from 1 to 9 exactly once. So it is easy to see that a standard Sudoku puzzle is actually a rank 9 Latin square. An example of a Sudoku puzzle and its completion is given in Figure 1.1.

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

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

FIGURE 1.1. A standard Sudoku puzzle and its solution

Soon after *Le Siecle*, the magazine *Le France* refined the puzzle to essentially the same format as the modern Sudoku; with the only exception that the puzzles were required to have the numbers 1 through 9 in both of the diagonals, to ensure a unique solution.

Dell Magazines began publishing Sudoku puzzles in the late 1970's. The puzzles most likely were developed by an independent puzzle maker and architect, Howard Garnes, and the newspaper called them Number Place.

While the name of the game is of Japanese origin ("SuDoku" means "single number"), it was not till 10 years later when the Japanese company Nikoli, Inc. started to publish a version of the Sudoku at the suggestion of its president, Mr. Maki Kaji. He gave the game its current name.

Almost two decades passed before (near the end of 2004) The Times newspaper in London has started to publish Sudoku as its daily puzzle due to the efforts of Wayne Gould, who has spent many years to develop a computer program that generates Sudoku puzzles.

By 2005 major newspapers in the US have begun publishing Sudoku puzzles and by now many new versions of the game can be found on the web. The reader is referred to more details on the history or the variations of Sudoku to the Wikipedia article [**1**] from which many of the above information were obtained.

1	2	3	4	5	6
4	5	6	1	2	3
3	4	5	2	6	1
2	6	1	3	4	5
6	3	2	5	1	4
5	1	4	6	3	2

FIGURE 1.2. A Dim$(2,3)$ Sudoku

The puzzles require logic, sometimes intricate, to solve but no formal mathematics is required. However, the puzzles lead naturally to certain mathematical questions. For example, how many Sudoku puzzles are there? How does the number of Dim(n,n) Sudoku puzzles compare to the number of rank n^2 Latin squares? Which puzzles have solutions and which do not? If a puzzle has a solution, is it unique? What is the minimum number of initial entries that need to be specified in order for a puzzle to have a unique solution? At this time, it is unknown if a puzzle beginning with 16 entries exists that has a unique solution. [**8**]

In the June/July issue of the American Mathematical Society's publication Notices, Agnes Herzberg and M. Ram Murty wrote an interesting article [**8**] on Dim(n,n)-puzzles such as the Dim$(3,3)$-puzzle shown in Figure 1.1.

In this article "Sudoku Squares and Chromatic Polynomials", the authors employed elements of graph theory, Chromatic polynomials, set theory, and the theory of permanents to prove some interesting things about Sudoku and also to arrive at an upper bound for the number of completed Dim(n,n) Sudoku puzzles. In particular, they show that the number of Dim(n,n) puzzles is much less than the number of rank n^2 Latin squares. So much so that as n tends to infinity, the probability that a randomly chosen rank n^2 Latin square is also a $\mathrm{Dim}(n,n)$ Sudoku puzzle goes to zero as n goes to infinity.

The organization of this thesis is as follows: In the first two chapters we will go through some standard definitions that are required for our results. For our generalized Sudoku puzzle we will define a graph such that a solution to a Sudoku puzzle corresponds to a proper coloring of this graph. We will then analyze this graph – much the same way as the Herzberg and Murty article does –, using results of Hall to bound the number of Latin squares from below and using matrix theory results to bound the number of Sudoku puzzles from above. This way we will obtain an upper bound on the fraction of Latin squares that are also Sudoku puzzles. The main result of this thesis is to generalize their work to the case of $\mathrm{Dim}(n,m)$ Sudoku puzzles such as the $\mathrm{Dim}(2,3)$ puzzle shown in Figure 1.2.

CHAPTER 2

GRAPH THEORY PRELIMINARIES

A Sudoku puzzle can easily be interpreted as a graph and then analyzed using concepts of graph theory. Graph colorings are of particular importance. All of the material in this chapter is standand, and can be found in textbooks such as [**4**] and [**13**].

2.1. DEFINITIONS

DEFINITION 2.1. A **simple graph** G is a set of elements called **vertices**, denoted V(G), together with a collection of unordered pairs of vertices called **edges**, denoted by E(G), that meets the following condition.

$$E(G) \subseteq \{\{u,v\} \mid u,v \in V(G), u \neq v\}.$$

For the remainder of this thesis, when referring to an edge, we will use the notation *uv*, or *vu* to mean the unordered pair {u,v}. We will also use the term **graph** as an abbreviation for simple graph.

A graph H is called a **subgraph** of a graph G if $V(H) \subseteq V(G)$ and $E(H) \subseteq E(G)$. The **order** of a graph is the number of vertices, denoted $V(G)$, and its **size** is the number of edges, denoted $E(G)$. Also, if u and v are two vertices of a graph and if the unordered pair {u,v} is an edge denoted by e, we say that e **joins** u and v or that it is an edge between u and v. In this case, the vertices are said to be **adjacent**, and both u and v are said to be **incident** upon e. A graph can be easily represented on paper using dots to represent the vertices and drawing a line (curved or straight) between unordered pairs of vertices to represent the edges. An example of such a depiction is in Figure 2.1.

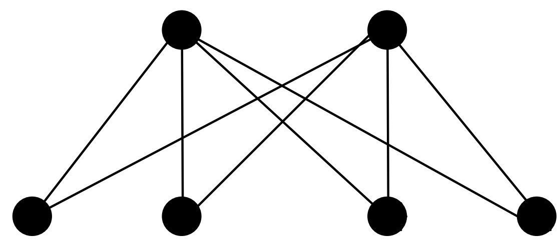


FIGURE 2.1. A visual illustration of a graph

DEFINITION 2.2. The **neighborhood** of a vertex v, denoted $N(v)$, is the collection of vertices which are adjacent to v. Formally, we write $N(v) = \{u \in V(G) : uv \in E(G)\}$.

The number of elements in $N(v)$ is referred to as the **degree** of vertex v. If all of the vertices in a graph have the same degree, then the graph is said to be **regular**.

DEFINITION 2.3. The **complete graph**, denoted K_n, is a graph with n vertices in which there is an edge joining each pair of vertices u,v for which $u \neq v$.

Note that K_n is a regular graph, the degree of each vertex is $n-1$, and the number of edges is $\binom{n}{2}$, since there is one edge for each pair of vertices.

Now, from a graph we can create new graphs by adding or subtracting edges, and also by identifying vertices. These kinds of modifications to a graph will be important so we define them precisely.

DEFINITION 2.4. Let $G = (V,E)$ be a graph and let $u, v \in V$, $u \neq v$. Then G_{+uv} is the graph with vertex set V and edge set $E' = E \bigcup uv$.

An example is in Figure 2.2. Note that if u and v are adjacent, then $G = G_{+uv}$.

DEFINITION 2.5. Let $G = (V,E)$ be a graph and let $u, v \in V$, $u \neq v$. Then G_{-uv} is the graph with vertex set V and edge set $E' = E \setminus uv$.

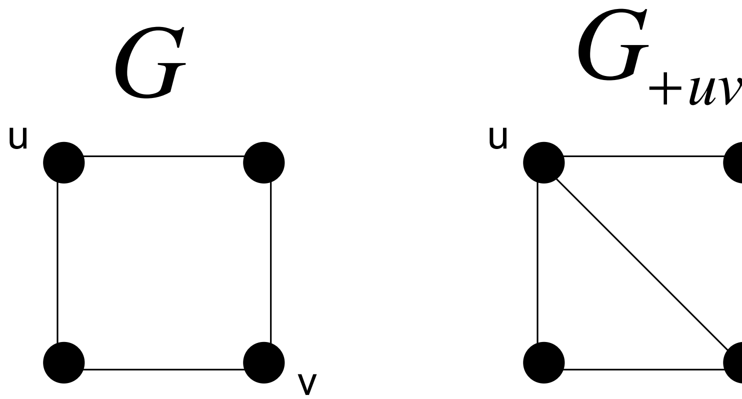


FIGURE 2.2. Edge addition

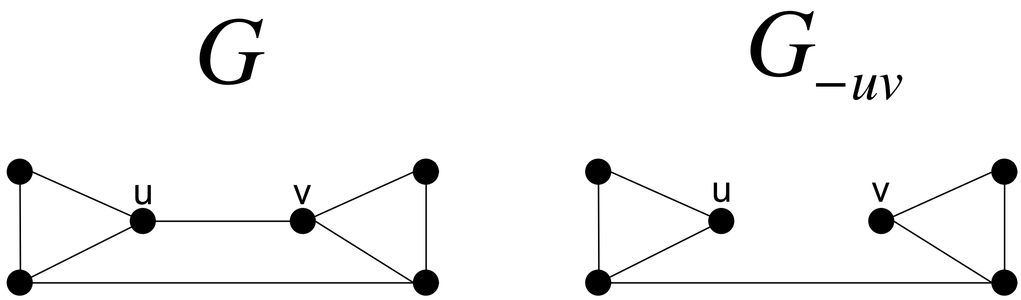


FIGURE 2.3. Edge removal

An example is in Figure 2.3.

Now what do we mean by identifying vertices? Below is the precise definition.

DEFINITION 2.6. Let $G = (V, E)$ be a graph and let $u, v \in V$, $w \notin V$. Then $G_{\cdot uv}$ is the graph with vertex set $V' = \{V \setminus \{u, v\}\} \cup \{w\}$ and edge set

$$\begin{aligned} E' &= \{E \setminus (\{xu \mid x \in N(u)\} \cup \{xv \mid x \in N(v)\})\} \\ &\quad \bigcup \{wx \mid x \in (N(u) \cup N(v)) \setminus \{u, v\})\}. \end{aligned}$$

In Figure 2.4 is a picture of a graph G and also $G_{\cdot uv}$.

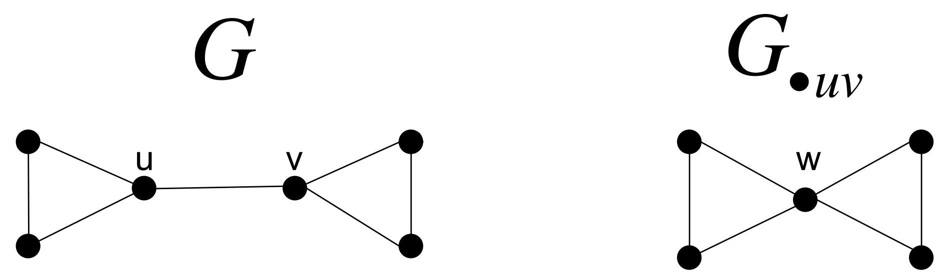


FIGURE 2.4. Vertex identification

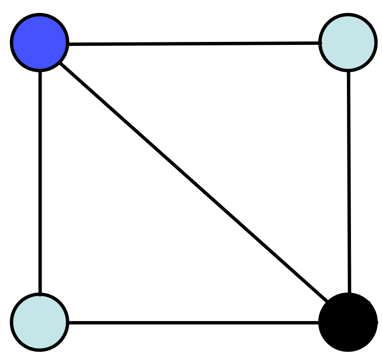


FIGURE 2.5. A depiction of a graph coloring

Next we define what we mean by a coloring of a graph. A λ **coloring** of a graph G is a function f from G to $\{1,2,...,\lambda\}$. We call this map a ***proper coloring*** if $f(x) \neq f(y)$ whenever x and y are adjacent in G. The minimal number of colors required to give the graph G a proper coloring is called the ***chromatic number*** of G and is denoted by $\chi(G)$. To make pictures easier to interpret, we replace the integers in the range of our function with actual colors. An example is given in Figure 2.5. Notice that no two adjacent vertices have the same color.

DEFINITION 2.7. The total number of ways one can properly color a graph G with λ colors is denoted $C_G(\lambda)$.

2.2. LEMMATA

Next we state a few important lemmata about graph colorings. It turns out that the number of ways to color a graph can be equal to coloring certain combinations of the same graph after it has been modified by adding or subtracting an edge, or identifying two vertices.

LEMMA 2.8. *Let G be a graph and let u and v be non-adjacent vertices in G. Then the number of proper λ-colorings of G that give u and v the same color is equal to$c_{G_{\cdot uv}}(\lambda)$.*

PROOF. Let $w \in G_{\cdot uv}$ be the vertex that results in the identification of u and v. Let $\mathscr{A}$ be the set of proper λ-colorings of G that give u and v the same color. Let $\mathscr{B}$ be the set of proper λ-colorings of $G_{\cdot uv}$.

Define $\alpha : \mathscr{A} \to \mathscr{B}$ by $\alpha(f) = f_\alpha$ where

$$f_\alpha = \begin{cases} f(x) & \text{if } x \in V(G_{\cdot uv}) \setminus w, \\ f(u) & \text{if } x = w \end{cases}$$

Clearly, $f_\alpha : V(G_{\cdot uv}) \to \{1, 2, \ldots, \lambda\}$. Moreover, if x, y are adjacent vertices of $G_{\cdot uv}$ and $w \notin \{x, y\}$, then they are adjacent vertices of G, and so $f_\alpha(x) = f(x) \neq f(y) = f_\alpha(y)$. Also, if w is adjacent to a vertex x, then x is adjacent to either u or v in G, which implies that $f_\alpha(x) = f(x) \neq f(u) = f(v) = f_\alpha(w)$. Thus, α is a well defined function from $\mathscr{A}$ to $\mathscr{B}$, since each coloring of G will determine a unique coloring of $G_{\cdot uv}$. We show that α is one-to-one and onto.

To show that α is 1-1, let f_1 and f_2 be two different elements of $\mathscr{A}$. Then for some $x \in V(G)$, $f_1(x) \neq f_2(x)$. There are two cases.

Case 1: $x \neq u$ and $x \neq v$. Then $x \in V(G_{\cdot uv}) \setminus \{w\}$. Then

$$(f_1)_\alpha(x) = f_1(x) \neq f_2(x) = (f_2)_\alpha(x).$$

Hence

$$(f_1)_\alpha(x) = (f_2)_\alpha(x),$$

which implies that $(f_1)_\alpha \neq (f_2)_\alpha$.

Case 2: $x = u$ or $x = v$. Then $f_1(x) = f_1(u)$ and $f_2(x) = f_2(u)$. Now,

$$(f_1)_\alpha(w) = f_1(u) = f_1(x) \neq f_2(x) = f_2(u) = (f_2)_\alpha(w).$$

Hence

$$(f_1)_\alpha \neq (f_2)_\alpha$$

.

So α is 1-1 from $\mathscr{A}$ to $\mathscr{B}$.

To show that α is onto, let $g \in \mathscr{B}$. Define f by

$$f(x) = \begin{cases} g(x) & \text{if } x \in V(G) \setminus \{u,v\}, \\ g(w) & \text{if } x = u \text{ or } x = v \end{cases}$$

First we show that $f \in \mathscr{A}$. Clearly, $f : V(G) \to \{1,2,\ldots,\lambda\}$ and $f(u) = f(v)$, so we only need to show that f is a proper coloring.

Suppose $x,y \in V(G)$, and x,y are adjacent. Note that since u and v are non-adjacent, $\{u,v\} \neq \{x,y\}$. Now, there are two cases.

Case 1: $x,y \in V(G) \setminus \{u,v\}$. Then $f(x) = g(x) \neq g(y) = f(y)$. So x and y are given different colors by the function f.

Case 2: $\{x,y\} \cap \{u,v\} \neq \emptyset$. Then by our previous remark, only one of x or y is an element of $\{u,v\}$. WOLG, we let $x \in \{u,v\}$ and $y \in V(G) \setminus \{u,v\}$. Since x and y are adjacent in G, it must be that w and y are adjacent in $G_{\cdot uv}$. Hence $f(x) = g(w) \neq g(y) = f(y)$. So x and y are given different colors by the function f.

So f is a function that, using λ colors, properly colors vertices in G with the stipulation that u and v are given the same color. Hence $f \in \mathscr{A}$.

Now we show that $f_\alpha(x) = g(x)$. By definition,

$$f_\alpha(x) = \begin{cases} f(x) = g(x) & \text{if } x \in V(G_{\cdot uv}) \setminus w, \\ f(u) = g(w) & \text{if } x = w \end{cases}$$

So $f_\alpha(x) = g(x)$ for all $x \in V(G_{\cdot uv})$. Hence α maps $\mathscr{A}$ onto $\mathscr{B}$. Since α is both 1-1 and onto, $|\mathscr{A}| = |\mathscr{B}|$. $\square$

LEMMA 2.9. *Let G be a graph and let u and v be distinct vertices in G. Then $c_{G_{+uv}}(\lambda)$ is equal to the number of proper λ-colorings of G which give u and v different colors.*

PROOF. Let $\mathscr{A}$ be the set of proper λ-colorings of G such that u and v receive different colors. Let $\mathscr{B}$ be the set of proper λ-colorings of G_{+uv}.

We define $\alpha : \mathscr{A} \to \mathscr{B}$ by $\alpha(f) = f_\alpha$, where for each $x \in V(G_{+uv})$ we have $f_\alpha(x) = f(x)$. Then α is a function from $\mathscr{A}$ to $\mathscr{B}$, since each proper λ-coloring of G that assign different colors to u and v will determine a unique proper λ-coloring of G_{+uv}. We must show that α is 1-1 and onto.

For 1-1, let f_1 and f_2 be two separate elements of $\mathscr{A}$. Then for some $x \in V(G)$, $f_1(x) \neq f_2(x)$. But then $(f_1)_\alpha(x) = f_1(x) \neq f_2(x) = (f_2)_\alpha(x)$. Hence α is 1-1 from $\mathscr{A}$ to $\mathscr{B}$.

For onto, let $g \in \mathscr{B}$. Define f by $f(x) = g(x)$. Then $f_\alpha(x) = f(x) = g(x)$. So $(f(x)) = g(x)$ for all $x \in V(G_{+uv})$. Therefore α is onto. Since α is both 1-1 and onto, $|\mathscr{A}| = |\mathscr{B}|$. $\square$

LEMMA 2.10. *If u and v are non-adjacent vertices in a graph G, then*

$$C_G(\lambda) = C_{G_{+uv}}(\lambda) + C_{G_{\cdot uv}}(\lambda).$$

PROOF. In any proper coloring of the graph G that uses λ colors, there are two distinct possibilities. Either u and v will have the same color, or they will have different colors. By Lemma 2.8 the number of ways to color G giving u and v the same color is equal to $C_{G_{\cdot uv}}(\lambda)$. By Theorem 2.9 the number of ways to color G giving u and v different colors is equal to $C_{G_{+uv}}(\lambda)$. Hence $C_G(\lambda) = C_{G_{+uv}}(\lambda) + C_{G_{\cdot uv}}(\lambda)$. $\square$

LEMMA 2.11. *If u and v are adjacent vertices in a graph G, then*

$$C_G(\lambda) = C_{G_{-uv}}(\lambda) - C_{G_{\cdot uv}}(\lambda).$$

PROOF. Since u and v are adjacent, any coloring of G must assign different colors to u and v. Now, in any coloring of $C_{G_{-uv}}(\lambda)$, u and v may have different colors, or they may be the same. But by Lemma 2.8, $C_{(G_{-uv})_{\cdot uv}}(\lambda)$ is equal to the number of ways to properly color G_{-uv}, with the stipulation that u and v be given the same color. We must subtract these possibilities so $C_G(\lambda) = C_{G_{-uv}}(\lambda) - C_{(G_{-uv})_{\cdot uv}}(\lambda)$. Since $C_{(G_{-uv})_{\cdot uv}}(\lambda) = C_{G_{\cdot uv}}(\lambda)$ we have $C_G(\lambda) = C_{G_{-uv}}(\lambda) - C_{G_{\cdot uv}}(\lambda)$. $\square$

CHAPTER 3

POLYNOMIALS

As we have mentioned, but have not yet shown, the number of ways one can fill out a Sudoku puzzle is the same as the number of proper colorings of a corresponding graph. Hence we are interested in how to determine the number of ways to properly color a graph with λ colors, and hence the number of ways to fill out a Sudoku puzzle, is equal to a monic polynomial evaluated at λ. In this chapter we will develop and apply these ideas.

DEFINITION 3.1. A (complex or real) **polynomial** of x is a function of the form

$$p(x) = \sum_{i=1}^{\infty} a_i x^i,$$

where only finitely many of the a_i are nonzero (and each a_i is complex or real, alternatively). The a_i are called the **coefficients** of the polynomial.

Note that the above definition implies that a polynomial $p(x)$ can be written in the form $p(x) = a_n x^n + a_{n-1}x^{n-1} + \ldots + a_1 x + a_0$, which we will do from now on.

DEFINITION 3.2. A polynomial $p(x) = a_m x^m + a_{m-1}x^{m-1} + ... + a_1 x + a_0$ is the **zero polynomial**, if each of the a_i are zero; with other words $p(x) = 0$. If $p(x)$ is a nonzero polynomial, then its **degree** is n if $a_n \neq 0$ and $a_i = 0$ for all $i \geq n$.

We will need another idea:

DEFINITION 3.3. A polynomial with degree n is **monic** if and only if $a_n = 1$.

DEFINITION 3.4. Let $p(x)$ be a polynomial. The number x_0 is a **root** of $p(x)$ if $p(x_0) =$
0

THEOREM 3.5. *(Fundamental Theorem of Algebra) Let $p(x)$ be a non-zero polynomial of degree n with complex coefficients. Then $p(x)$ has n roots, when repeated roots are counted up to their multiplicity.* [**12**]

COROLLARY 3.6. *Let $P(x)$ and $Q(x)$ be two monic polynomials, and assume that there exists an integer m such that $P(\lambda) = Q(\lambda)$ for all integers λ with $\lambda \geq m$. Then $P(x) = Q(x)$.*

PROOF. Assume there exists $Q(x)$ which equals $P(x)$ for all $\lambda \geq m$. Assume that the maximum of the degrees of $P(x)$ and $Q(x)$ is n. Then Then $(P-Q)(x)$ is a polynomial of degree $\leq n$ with an infinite number of zero roots. This contradicts the Fundamental Theorem of Algebra. $\square$

Later we will make use of the following Lemma:

LEMMA 3.7. *Let $p(x)$ be a nonzero polynomial of degree n with integer coefficients and a be an integer root of $p(x)$. Then $p(x) = (x-a)q(x)$, where $q(x)$ is a polynomial of degree $n-1$ and has integer coefficients.*

PROOF. We will do this by induction on n, the degree of $p(x)$. We will assume that a_n is the leading coefficient of $p(x)$

If $n = 1$, then $\frac{1}{a_n}p(x)$ and $x-a$ are two monic polynomials with the same roots (since both have one root, and it must be n. By Corollary 3.6, $\frac{1}{a_n}p(x) = x-a$, so we may choose $q(x) = a_n$, which clearly satisfies the conditions.

Now let $n > 1$ and assume the statement is true for all polynomials with degree $n' < n$. Let $p'(x) = p(x) - a_n x^{n-1}(x-a)$. Since $a_n x^{n-1}(x-a)$ is a polynomial of degree n with a_n as its leading coefficient, and it has all integer coefficient, we have that $p'(x)$ has integer coefficients and the degree of $p'(x)$ is some n', where $n' < n$. Also, $p'(a) = p(a) - a_n a^{n-1} \cdot 0 = 0$. Therefore by the induction hypothesis there is a polynomial $q'(x)$ that has degree $n'-1 \leq n-2$ that has integer coefficients and $p'(x) = (x-a)q'(x)$. Therefore $p(x) = q'(x)(x-a) + a_n x^{n-1}(x-a) = (x-a)(a_n x^{n-1} + q'(x))$, and choosing $q(x) = a_n x^{n-1} + q'(x)$, $q(x)$ is a degree $n-1$ polynomial with all integer roots. hypothesis, there is $\square$

We begin with the complete graph on n vertices, and then work our way towards the general case of any graph on n vertices.

THEOREM 3.8. *Let G be the complete graph K_n. Then there exists a unique monic polynomial with integer coefficients of degree n, denoted $P_G(x)$, which equals $C_G(\lambda)$ for all nonnegative integers λ.*

PROOF. The uniqueness of such a monic polynomial follows from Corollary 3.6, so we only need to show the existence.

We will show that $P_G(x) = x(x-1)...(x-n+1)$. This is clearly is a monic polynomial of degree n with integer coefficients.

Let λ be a nonnegative integer.

Suppose first that $\lambda < n$. Clearly, $P_G(\lambda) = 0$. Now, in a proper coloring of K_n, any two vertices must have different colors. So any proper coloring of K_n must use n different colors. Hence $C_G(\lambda) = 0$ as well. So for each $\lambda < n$, $C_G(\lambda) = 0 = P_G(\lambda)$.

Now suppose that $\lambda \geq n$. Then $P_G(\lambda) = \lambda(\lambda-1)\cdots(\lambda-n+1)$. If we color G using λ colors, we may color the first vertex with λ colors, the second vertex with $\lambda - 1$ colors, etc.. Hence $C_G(\lambda) = (\lambda)(\lambda-1)...(\lambda-n+1)$. But this is equal to $P_G(\lambda)$.

So for any λ, $C_G(\lambda) = P_G(\lambda)$. □

THEOREM 3.9. *Let G be obtained from the complete graph K_n by removing one edge. Then there exists a unique monic polynomial of degree n with integer coefficients, denoted by $P_G(x)$, which equals $C_G(\lambda)$ for all nonnegative integers λ.*

PROOF. The uniqueness of such a monic polynomial follows from Corollary 3.6, so we only need to show the existence.

Suppose that u and v are the two non-adjacent vertices that result from removal of the single edge. By Theorem 2.9, we have $C_G(\lambda) = C_{G_{+uv}}(\lambda) + C_{G.uv}(\lambda)$. Now $C_{G_{+uv}}(\lambda)$ is a complete graph on n vertices and is equal to a monic polynomial of degree n with integer coefficients, $P_{G_{+uv}}(x)$ for all nonnegative integers λ by Theorem 3.8. But $G_{\cdot uv}$ is a complete graph on $n-1$ vertices and, also by Theorem 3.8, $c_{G_{\cdot uv}}(\lambda)$ is equal to a monic polynomial of degree $n-1$ with integer coefficients, $P_{G_{\cdot uv}}(x)$ for all nonnegative integers λ. Hence we may choose $P_G(x) = P_{G_{+uv}}(x) + P_{G_{\cdot uv}}(x)$, which is a monic polynomial of degree n and has integer coefficients. $\square$

THEOREM 3.10. *Let G be a graph on $n \geq 1$ vertices. Then there exists a unique monic polynomial of degree n with integer coefficients, denoted $P_G(x)$, which equals $C_G(\lambda)$ for all $\lambda \geq 0$.*

PROOF. We will use a double induction. Upward on the number of vertices, and then downward on the number of edges.

For the base step on the number of vertices, note that when $n = 1$, G consists of a single vertex, so we have that $C_G(\lambda) = \lambda$ for all nonnegative integers λ. We let $P_G(x) = x$, which is a monic polynomial of degree 1 and has integer coefficients. Then $C_G(\lambda) = P_G(\lambda) = \lambda$, and we are done.

So now let $n > 1$, and for the induction hypothesis on the number of vertices, assume that for any graph H on less than n vertices there exists a monic polynomial $P_H(x)$ of degree n with integer coefficients such that $P_H(\lambda) = C_H(\lambda)$ for all nonnegative integers λ.

Let G be a graph on n vertices. Let k be the number of edges in G. We need to show that there exists a monic polynomial $P_G(x)$ of degree n with integer coefficients, for which $P_G(\lambda) = C_G(\lambda)$ for all nonnegative integers λ. We will show this by using a downward induction on the possible values of k..

To begin the base case on the number of edges, assume that G has as many edges as possible. This means there is an edge between any pairs of two different vertices, so $G = K_n$, and $k = \binom{n}{2}$. Then by Theorem 3.8 we are done. Also, if $k = \binom{n}{2} - 1$ edges, then by Theorem 3.9 we are done.

So let $k \leq \binom{n}{2} - 2$. For the induction hypothesis on the number of edges, assume that for any graph H which has n vertices and k' edges, where $k < k' \leq \binom{n}{2}$, there exists a monic polynomial $P_H(x)$ of degree n with integer coefficients, for which $P_H(\lambda) = C_H(\lambda)$ for all nonnegative integers λ.

Since k, the number of edges in G, is less then $\binom{n}{2}$, there exist two vertices u, v which are non-adjacent. By theorem 2.9,

$$C_G(\lambda) = C_{G_{+uv}}(\lambda) + C_{G_{\cdot uv}}(\lambda).$$

But $C_{G_{+uv}}(\lambda)$ is the number of ways to color the graph G_{+uv} that has n vertices and $k+1$ edges. Since $k < k+1 \leq \binom{n}{2}$, by the induction hypothesis on the number of edges, there is a monic polynomial $P_{G_{+uv}}(x)$ of degree n with integer coefficients such that $P_{G_{+uv}}(\lambda) = c_{G_{+uv}}(\lambda)$ for all nonnegative integers λ. Also, $G_{\cdot uv}$ is a graph with $n-1$ points. By the induction hypothesis on the number of vertices, there is a monic polynomial $P_{G_{\cdot uv}}$ of degree $n-1$ with integer coefficients, such that $P_{G_{\cdot uv}}(\lambda) = C_{G_{\cdot uv}}(\lambda)$ for all nonnegative integers λ. Now choose $P_G(x) = P_{G_{+uv}}(x) + P_{H_{\cdot uv}}(x)$. Clearly, this is a monic polynomial of degree n that satisfies the required conditions. □

We have established that the number of ways to color a graph with λ colors is given by a unique monic polynomial. What can we say about a graph that is already partially colored? In particular can we represent the number of ways of extending a partial coloring to a complete coloring by a monic polynomial?

DEFINITION 3.11. Let G be a graph on some n vertices and let t be a nonnegative integer, $t \leq n$. A **partial proper coloring** H of G on some t vertices is a function $H : B \to \{1, 2, \ldots, \lambda\}$, where $B \subseteq V(G)$, $|B| = t$, and if $u, v \in B$ are adjacent vertices of G, then $H(u) \neq H(v)$.

DEFINITION 3.12. Let n be a positive integer and t, d, λ be nonnegative integers such that $0 \leq t \leq n$. Let G be a finite graph on n vertices and H be a partial proper coloring of t vertices of G using some d colors. We denote by $C_{G,H}(\lambda)$ be the number of ways to extend H to a proper λ-coloring of G.

THEOREM 3.13. *Let n be a positive integer and t, d be nonnegative integers such that $0 \leq t \leq n$. Let G be a finite graph on n vertices and H be a partial proper coloring of t*

vertices of G using some d colors. Then there exists a unique monic polynomial of degree $n-t$, denoted $P_{G,H}(x)$, which equals $C_{G,H}(\lambda)$ for all integersλ for which $\lambda \geq d$.

PROOF. First note that for any n, if $t = n$, then or partial coloring already properly colors G, therefore we only have one way to extend it. So for any $\lambda \geq d$ we have $c_{G,H}(\lambda) = 1$, and we may choose $P_{G,H}(x) = 1$, which is clearly a monic polynomial of degree 0. Since $n-t = n-n = 0$, we are done.

Also, for any n, if $t = 0$, then $c_{G,H}(\lambda) = c_G(\lambda)$ and we are done by Theorem 3.10.

Therefore in the rest we will assume that $0 < t < n$.

We will use a double induction. First upward on the number of vertices n, and then upward on the number of edges of the graph. For the base step on the number of vertices, let $n = 1$. Then we are done since the only possible values of t are $t = 0$ or $t = 1 = n$.

So not let $n > 1$, and assume that the statement is true for any graph on less than n points. Let G be a graph on n points and k edges. We will proceed by induction on the number of edges in G.

For the base step on the number of edges, suppose G has zero edges. Let H be a partial proper coloring of G on t points and d colors, and let λ be an integer such that $\lambda \geq d$. Then we are free to color any of the remaining $n-t$ uncolored vertices with any of our λ colors. Hence $C_{G,H}(\lambda) = \lambda^{n-t}$, and we let $P_{G,H}(x) = x^{n-t}$, a monic polynomial of degree $n-t$.

So let $k > 1$, and suppose that the statement is true for any graph on n points and at most $k-1$ edges.

Let H be a partial proper coloring of G on t points and d colors, and let λ be an integer such that $\lambda \geq d$. There are two cases.

Case 1: Every edge of G connects two points that are colored by H: We may color the remaining vertices with any of our λ colors. So there are λ^{n-t} ways to extend the partial coloring to a complete coloring of G. Thus we let $P_{G,H}(x) = x^{n-t}$.

Case 2: There exists an edge, whose end vertices are u and v, with at least one end vertex in $G \setminus H$. Note that $C_{G_{-uv},H}(\lambda)$ is equal to the number of ways to extend the partial coloring of H to G if we allow u and v to have either the same or different colors. Let H'

be the coloring on $G_{\cdot uv}$ that agrees with H everywhere, except on u and v. If one of u or v is colored by H, then H' assigns this color to w, otherwise H' does not color w. Since only one of u, v is in H, it is clear that H and H' both color the same number of vertices, t and use the same number of colors, d.

Also, note that $C_{G_{\cdot uv}, H'}(\lambda)$ is equivalent to the number of ways to extend the partial coloring of H to G if we were to require that u and v are given the same color. Hence we have the equation:

$$C_{G,H}(\lambda) = C_{G_{-uv},H}(\lambda) - C_{G_{\cdot uv},H'}(\lambda)$$

By the induction hypothesis on the number of edges, there is a monic polynomial $P_{G_{-uv},H}(x)$ of degree $n-t$ such that $P_{G_{-uv},H}(\lambda) = C_{G_{-uv},H}(\lambda)$ for every integer $\lambda \geq d$.

By the induction hypothesis on the number of point, there is a monic polynomial $P_{G_{\cdot uv},H'}(x)$ of degree $n-1-t$ such that $C_{G_{\cdot uv},H'}(\lambda) = P_{G_{\cdot uv},H'}(\lambda)$ for all integers $\lambda \geq d$.

Set $P_{G,H}(x) = P_{G_{-uv},H}(x) - P_{G_{\cdot uv},H'}(x)$. This clearly is a monic polynomial of degree $n-t$ with $C_{G,H}(\lambda) = P_{G,H}(\lambda)$ for all integers $\lambda \geq d$. $\square$

CHAPTER 4

THE BIG-OH NOTATION

The big-oh notation was introduced by a German number theorist Paul Bachmann in his book *Analitische Zahlentheorie* in 1894 [**3**]. Though it can be extended to functions of real variables, we will only use it for functions of positive integers just as in D.E. Knuth's *The Art of Computer Programming* [**9**]. The interested reader is referred to more general definitions in standard textbooks.

The O-notation allows us to quantify the degree of accuracy in our approximation, for example in expressions like $f(n) = e^{n^2+O(n\ln(n))}$. In general, the notation $O(f(n))$ — or sometimes more precisely $O_n(f(n))$ — may be used whenever $f(n)$ is a function of the positive integer n; it roughly states that magnitude of the quantity for which we use $O(f(n))$ (while may not be explicitly known) is not too large.

DEFINITION 4.1. Suppose that $f(n)$ and $g(n)$ are two functions defined on the positive integers. We say that $f(n) = O(g(n))$ if and only if there exist integers n_o, M such that $\mid f(n) \mid \leq M \mid g(n) \mid$ for all $n \geq n_o$.

Note that in this context, the equality sign loses some of its usual conveniences. For example, from $f(n) = O(g(h))$ and $h(n) = O(g(n))$ we can not conclude that $f(n) = h(n)$.

DEFINITION 4.2. Suppose that $f(n)$ and $g(n)$ are two functions defined on the positive integers. We say that $f(n) \leq O(g(n))$ iff there exists a function $h(n)$ and an integer n_0 such that $h(n) = O(g(n))$ and $f(n) \leq h(n)$ for all $n \geq n_0$.

In order to be able to understand expressions like $f(n) = e^{n^2+O(n\ln(n))}$, we need another definition.

DEFINITION 4.3. Suppose that $f(n)$, $g(n)$ are functions on the positive integers, and $h(n,x)$ is an algebraic expression on the positive integers n and a variable x where the variable x appears once. We say that $f(n) = h(n, O(g(n)))$ iff $f(n) = h(n, \ell(n))$ for some function $\ell(n)$ where $\ell(n) = O(g(n))$.

In the following, $f(n)$ and $g(n)$ are functions and C is a constant. Here are some simple operations that we can do with the O-notation that follow fairly trivially from the definition:

$$\begin{aligned}
f(n) &= O(f(n)) \\
C \cdot O(f(n)) &= O(f(n)) \\
O(f(n)) + O(g(n)) &= O(f(n) + g(n)) \\
O(O(f(n)) &= O(f(n)) \\
O(f(n)) \cdot O(g(n)) &= O(f(n) \cdot g(n)) \\
f(n) \cdot O(g(n)) &= O(f(n) \cdot g(n))
\end{aligned}$$

CHAPTER 5

ANALYSIS OF THE SUDOKU GRAPH

In this chapter we use what we know about graph theory and polynomials to obtain some interesting results about the general $\mathrm{Dim}(n,m)$ Sudoku puzzle.

It is easy to see how a completed Sudoku puzzle is equivalent to a proper graph coloring. We will associate a graph X_{nm} with the $\mathrm{Dim}(n,m)$ Sudoku grid as follows. X_{nm} will have $(nm)^2$ vertices, each corresponding to a cell in the Sudoku grid. Two distinct vertices will be adjacent if and only if the corresponding cells in the grid are either in the same row, the same column, or the same sub-grid. This way each vertex will be given a color distinct from that of its neighbors. So for each completed Sudoku puzzle, there corresponds a proper coloring of the graph X_{nm}.

To put this in a more general and formal context, consider an $\mathrm{Dim}(n,m)$ grid. Each cell in the grid will be associated with a vertex in X_{nm} that is labeled (i,j) with $0 \leq i,j \leq nm-1$. We will consider (i,j) and (i',j') to be adjacent if either (1) $i=i'$ or $j=j'$ or (2) $\lfloor i/n \rfloor = \lfloor i'/n \rfloor$ and $\lfloor j/n \rfloor = \lfloor j'/n \rfloor$.

THEOREM 5.1. *X_{nm} is regular and the degree of each vertex is $3nm-(n+m)-1$.*

PROOF. Let v be and arbitrary vertex of X_{nm}. Then v is adjacent to $nm-1$ other vertices in its row, and $nm-1$ other vertices in its column. It is also adjacent to $nm-1$ others in the $n \times m$ subgrid it lies in, but $n-1$ of these were already counted in its column and $m-1$ of them were counted in its row. So v is adjacent to $(nm-1)+(nm-1)+[(nm-1)-(n-1)-(m-1)] = 3nm-(n+m)-1$ vertices. $\square$

To determine the chromatic number of X_{nm}, we will recall the following definitions:

DEFINITION 5.2. Let n be an integer, $n > 2$, and let k, m be integers. We will say that $k \equiv m \pmod{n}$ if n divides $m - k$.

DEFINITION 5.3. Let n be an integer, $n > 2$, and let k be an integers. k mod n denotes the unique integer m for which $k \equiv m \pmod{n}$ and $0 \leq m < n$.

Now we are ready to state and prove the following:

THEOREM 5.4. *The chromatic number of X_{nm} is nm.*

PROOF. Without loss of generality $m \geq n$. If $mn = 1$, then X_{nm} consists of a single point, and the statement is trivial. So assume that $mn > 1$.

First we show that X_{nm} cannot be properly colored with fewer than nm colors. Note that the vertices of X_{nm} which represent the cells in the upper $n \times m$ grid are all adjacent to each other. These vertices and the edges connecting them form the complete graph K_{nm}. Since we need at least nm colors to properly color K_{nm}, we need at least nm colors to properly color X_{nm}.

Next we show that nm colors are sufficient. To do this we will explicitly construct a proper coloring of X_{nm} using nm colors. First without loss of generality we assume that $n \leq m$. Consider the vertices (i, j) with $0 \leq i \leq nm - 1$ and $0 \leq j \leq nm - 1$. Now, using the division algorithm, we let $i = an + r$ and we let $j = bm + s$, where $a = \left\lfloor \frac{i'}{n} \right\rfloor$ and $b = \left\lfloor \frac{j}{m} \right\rfloor$. Hence,

$$\begin{aligned} 0 &\leq a < m \\ 0 &\leq r < n \\ 0 &\leq b < n \\ 0 &\leq s < m \end{aligned}$$

and the 4-tuple (a,r,b,s) is uniquely determines each (i,j). Now we will define a function f that will properly color each vertex of X_{nm} using nm colors as follows

$$f(i,j) = (rm + a + bm + s) \mod nm$$

Clearly, f only uses numbers from the set $\{0,1,2,...,nm-1\}$, which has order nm. Also, since (a,r,b,s) is unique for each (i,j), $f(i,j)$ assigns a unique color to each vertex (i,j). Hence f is a function that colors each vertex in the graph $X_{n,m}$ with nm colors (and note that by our previous remark, if f is a proper coloring it does indeed use all the colors). To show that the coloring will be proper, we need to show that if two vertices (i,j) and (i',j') receive the same color then they are not adjacent.

So assume that (i,j) and (i',j') receive the same color. Since if they are not in the same row, color or subgrid, they can not be adjacent, we need to examine three cases.

Case 1: Suppose that the two vertices (i,j) and (i',j') represent cells in the same $\text{Dim}(n,m)$ subgrid. By virtue of being in the same $n \times m$ subgrid, we have

$$\left\lfloor \frac{i}{n} \right\rfloor = \left\lfloor \frac{i'}{n} \right\rfloor \quad \text{and} \quad \left\lfloor \frac{j}{m} \right\rfloor = \left\lfloor \frac{j'}{m} \right\rfloor,$$

hence $a = a'$ and $b = b'$.

Now, since the vertices (i,j) and (i',j') have the same color,

$$\begin{aligned} f(i,j) = (rm + a + bm + s) \mod nm &= (r'm + a' + b'm + s') \mod nm = f(i',j') \\ (rm + a + bm + s) \mod nm &= (r'm + a + bm + s') \mod nm \\ (rm + s) \mod nm &= (r'm + s') \mod nm \end{aligned}$$

Notice that $rm + s \leq (n-1)m + s = nm + (s-m) < nm$ and similarly $r'm + s' < nm$. Hence $rm + s = r'm + s'$. This implies that $m \mid r - r' \mid = \mid s' - s \mid$. Since $0 \leq s, s' < m$, it must be that $\min(s,s') \geq 0$ and $\max(s,s') \leq m-1$. Therefore $\mid s' - s \mid s = \max(s,s') - \min(s,s') \leq m - 1 - 1 < m$ But $\mid s - s' \mid$ is an integer that is divisible by m. Therefore $\mid s - s' \mid = 0$, so $s = s'$. So $\mid r - r' \mid = \frac{|s'-s|}{m} = 0$ and $r = r'$.

We have established that $(a,r,b,s) = (a',r',b',s')$. Thus $i = an+r = a'm+r' = i'$ and $j = bm+s = b'm+s' = j'$. So whenever two vertices which represent cells in the same sub-grid have the same color, they are identical.

Case 2: Suppose that two vertices (i,j) and (i',j') represent cells in the same column. In this case $j = j'$, so $f(i,j) = f(i',j') = f(i',j)$. Then

$$\begin{aligned} f(i,j) = (rm+a+bm+s) \mod nm &= (r'm+a'+bm+s) \mod nm = f(i',j) \\ (rm+a) \mod nm &= (r'm+a') \mod nm \end{aligned}$$

Notice that $rm+a \leq (n-1)m+a < (n-1)m+m = nm$ and similarly $r'm+a' < nm$. Hence $rm+a = r'm+a'$. Since $0 \leq a,a' < m$, we reason as in case 1 to show that $r = r'$ and $a = a'$.

We have established that $(a,r) = (a',r')$. Thus $i = an+r = a'n+r' = i'$. So whenever two vertices which represent cells in the same column have the same color, they are identical.

Case 3: Suppose that two vertices(i,j) and (i',j') represent cells in the same row. In this case $i = i'$, so $f(i,j) = f(i',j') = f(i,j')$. Then

$$\begin{aligned} f(i,j) = (rm+a+bm+s) \mod nm &= (rm+a+b'm+s') \mod nm = f(i,j') \\ (bm+s) \mod nm &= (b'm+s') \mod nm \end{aligned}$$

By definition, $bm+s = j < nm$. Similarly $b'm+s' = j' < nm$. Hence $bm+s = b'm+s'$, therefore $j = j'$. So whenever two vertices which represent cells in the same row have the same color, they are identical.

Hence $f(i,j)$ properly colors X_{nm} using nm colors. Therefore the chromatic number of X_{nm} is nm. □

Suppose now that we have a $\mathrm{Dim}(n,m)$ Sudoku puzzle that is partially filled out using only $nm-2$ colors. Since two colors have not been used in the initial partial coloring, it is apparent that these two colors can be interchanged in a final coloring to get another coloring. Hence there will not be a unique solution to the Dim(n,m) Sudoku puzzle unless at least $nm-1$ colors are used in an initial, partial coloring. We make this rigorous in the next theorem.

THEOREM 5.5. *Any Dim(n,m) solvable Sudoku puzzle will have a unique solution only if it begins with at least $nm-1$ colors.*

PROOF. Let H be the initial partial coloring of $0 \leq t < nm$ vertices of X_{nm}. Suppose that H uses only $d \leq nm-2$ colors. Then by theorem 3.13, there exists a unique monic polynomial of degree ≥ 2, denoted $P_{X_{nm},H}(x)$, which equals $C_{X_{nm},H}(\lambda)$ for all $\lambda \geq d$.

Since the chromatic number of X_{nm} is nm, we must have $C_{X_{nm},H}(\lambda) = 0$ for $\lambda \in \{d, d+1, ..., nm-1\}$. Therefore $P_{X_{nm},H}(x) = 0$ for $x \in \{d, d+1, ..., nm-1\}$. So we may write $P_{X_{nm},H}(x) = (x-d)(x-d-1)...(x-nm+1)q(x)$ for some monic polynomial $q(x)$. By repeated application of Lemma 3.7, we get that $q(x)$ has integer coefficients. For the case $x = nm$, we have $P_{X_{nm},H}(nm) = (nm-d)!q(nm)$.

Now we have assumed that $d \leq nm-2$, so certainly $(nm-d)! \geq 2$. Also, $q(nm)$ must be positive, else we would not have any ways to finish properly coloring X_{nm}. Since $q(x)$ has integer coefficients, $q(nm) \geq 1$. Finally, this implies that $C_{X_{nm},H}(nm) = P_{X_{nm},H}(nm) = (nm-d)!q(nm) \geq 2$, and so there is not a unique solution to the Dim(n,m) Sudoku puzzle when fewer than $nm-1$ colors are used. $\square$

CHAPTER 6

PERMANENTS AND SYSTEMS OF DISTINCT REPRESENTATIVES

In this chapter, we will develop the idea of a system of distinct representatives (**SDR**), and we also show how permanents of matrices can be used to count **SDR's**. This material was all developed in the 20$^{\text{th}}$ century.

6.1. SYSTEMS OF DISTINCT REPRESENTATIVES

We begin with the concept of an **SDR**. Basically this involves taking one unique element from a collection of non-empty sets.

DEFINITION 6.1. Let n be a positive integer. Let $\mathscr{S} = (S_1, S_2, ..., S_n)$ be an (ordered) collection of non-empty subsets of a set M. A **System of Distinct Representatives** (abbreviated **SDR**) is an n-tuple $X = (x_1, x_2, ..., x_n)$ of pairwise distinct elements of M, such that $x_i \in S_i$ for all $i \in \{1, 2, \ldots, n\}$.

A result known as Hall's marriage theorem tells us exactly when it is possible to have a system of distinct representatives. Let us work towards an understanding of this result.

DEFINITION 6.2. The finite collection $\mathscr{S}$ satisfies the **marriage condition** if for every $\Delta \subseteq \{0, 1, \ldots, n\}$, we have that

$$\mid \bigcup_{i \in \Delta} S_i \mid \geq \mid \Delta \mid .$$

(i.e. any k subsets taken together have at least k elements)

Note that the marriage condition is trivially satisfied for $\Delta = \emptyset$, so we do not need to check it for that case.

We first state a quick lemma that will help us prove the marriage theorem.

LEMMA 6.3. *If a collection $\mathscr{S} = (S_1, S_2, ..., S_n)$ of finite subsets of a set M satisfies the marriage condition, then each S_i is non-empty.*

PROOF. Assume that $\mathscr{S}$ satisfies the marriage condition, and fix an $i \in \{1, 2, \dots, n\}$ arbitrarily. Choose $\Delta = \{i\}$. The marriage condition implies that

$$1 = |\Delta| \leq |\bigcup_{\ell \in \Delta} S_\ell| = |S_i| .$$

Hence S_i contains at least one element and cannot be empty. □

Now we have another definition that will help us in the proof of our main theorem about **SDR**s

DEFINITION 6.4. Let $\mathscr{S} = (S_1, S_2, ..., S_n)$ be an (ordered) collection of non-empty subsets of a set M and let Δ be a nonempty proper subset of $\{1, 2 \dots, n\}$. Δ is **critical with respect to** $\mathscr{S}$, if

$$|\Delta| = |\bigcup_{i \in \Delta} S_i|$$

Now we are ready to state and prove the following:

THEOREM 6.5. *A collection $\mathscr{S} = \{S_1, S_2, ..., S_n\}$ of finite subsets of a set M has a* ***SDR*** *if and only if $\mathscr{S}$ satisfies the marriage condition.*

PROOF. Suppose first $(x_1, \dots, x_n)$ is an **SDR** for the collection $\mathscr{S}$. Let $\Delta \subseteq \{1, 2, \dots, n\}$. Define $X = \{x_i : i \in \Delta\}$. Then, since the x_i are part of an **SDR**, and consequently are all different, $|X| = |\Delta|$. But since $x_i \in S_i$ for each i, it must be that $X \subseteq \cup_{i \in \Delta} S_i$, therefore $|X| \leq |\cup_{i \in \Delta} S_i|$. Therefore $|\Delta| \leq |\cup_{i \in \Delta} S_i|$, and so $\mathscr{S}$ satisfies the marriage condition.

Now suppose that $\mathscr{S}$ satisfies the marriage condition. We will proceed by induction on n, the number of sets in $\mathscr{S}$. For the base case let $|\mathscr{S}| = 1$. Then $\mathscr{S} = (S_1)$. Choose $\Delta = \{1\}$. By Lemma 6.3 we have that $S_1 \neq \emptyset$, so we must have an $x_1 \in S_1$. But then (x_1) is an **SDR** for $\mathscr{S}$.

Now let $n > 1$ and assume the theorem is true for all $| \mathscr{S}' | < n$.

Since $\mathscr{S}$ satisfies the marriage condition, we have the following cases:

Case 1: $\mathscr{S}$ has no critical sets, with other words for every nonempty proper subsetsΔ of $\{1,2,\ldots,n\}$ we have that

$$| \Delta | < | \bigcup_{i\in\Delta} S_i |$$

By Theorem 6.3, each S_i is non empty. So we may pick an $x_n \in S_n$ arbitrarily. Define the ordered collection $\mathscr{S}' = (S'_1 \setminus \{x_n\},...,S'_{n-1})$ where $S'_i = S_i - \{x_n\}$. We will show that $\mathscr{S}'$ satisfies the marriage condition. Let $\Delta \subseteq \{1,2,3,\ldots,n-1\}$ be nonempty. Note that Δ is a nonempty proper subset of $\{1,2,\ldots,n\}$, therefore from our assumption we have that

$$\begin{aligned} | \Delta | &\leq | \bigcup_{i\in\Delta} S_i | - 1 = | \Big(\bigcup_{i\in\Delta} S_i\Big) \setminus \{x_n\} | \\ &= | \Big(\bigcup_{i\in\Delta} (S_i \setminus \{x_n\}\Big) | = | \bigcup_{i\in\Delta} S'_i, \end{aligned}$$

proving that $\mathscr{S}'$ does satisfy the marriage condition. Since $| \mathscr{S}' | < n$, an SDR $(x_1,\ldots,x_{n-1})$ exists for $\mathscr{S}'$ by the induction hypothesis. Clearly, $(x_1,\ldots,x_n)$ then is an **SDR** for $\mathscr{S}$

Case 2: $\mathscr{S}$ has a critical set $\Delta_0 = \{i_1,\ldots,i_k\}$. Clearly, $1 \leq k \leq n-1$. We will use $\Delta_1 = \{1,2,\ldots,n\} \setminus \Delta_0 = \{j_1,\ldots,j_{n-k}\}$, where $j_1 < j_2 < \ldots < j_{n-k}$. We will define two ordered collection of sets, $\mathscr{S}'$ and $\mathscr{S}$" as follows: $\mathscr{S}" = (S_{i_1}, S_{i_2},\ldots,S_{i_k})$ and $\mathscr{S}' = (S'_{j_1}, S'_{j_2},\ldots,S'_{j_{n-k}})$, where $S'_{j_i} = S_{j_i} - X$, where $X = \bigcup_{\ell=1}^{k} S_{i_\ell}$. Clearly, $\mathscr{S}$" satisfies the marriage condition, therefore by the induction hypothesis it has an SDR $(x_{i_1}, x_{i_2},\ldots,x_{i_k})$. Moreover, since Δ_0 is critical, we have that $X = \{x_{i_1}, x_{i_2},\ldots,x_{i_k}\}$. We will also show that $\mathscr{S}'$ satisfies the marriage condition. Let $\Delta' \subseteq \Delta_1$, and define $\Delta = \Delta' \cup \Delta_0$. Since Δ_0 and Δ_1 are disjoint

$$| \Delta' | + k = | \Delta' \cup \Delta_0 | = | \Delta |$$

Since Δ_0 is critical, we have that $|X| = k = |\Delta_0|$. Since $\Delta_0 \subseteq \Delta$ we have that

$$X = \bigcup_{\ell=1}^{k} S_{i_\ell} \subseteq \bigcup_{\ell\in\Delta} S_\ell,$$

which implies that

$$\bigcup_{\ell\in\Delta} S_\ell = \left(\Big(\bigcup_{\ell\in\Delta} S_\ell\Big) - X\right) \cup X = \Big(\bigcup_{\ell\in\Delta}(S_\ell - X)\Big) \cup X = \left(\bigcup_{\ell=1}^{n-k} S'_{j_\ell}\right) \cup X$$

Now each of the S'_{j_ℓ} are disjoint from X, therefore

$$\mid \bigcup_{\ell\in\Delta} S_\ell \mid = \mid \bigcup_{\ell=1}^{n-k} S'_{j_\ell} \mid + \mid X \mid = \mid \bigcup_{\ell=1}^{n-k} S'_{j_\ell} \mid + k$$

By the marriage condition on $\mathscr{S}$ we have that

$$k + \mid \Delta_1 \mid = \mid \Delta \mid \leq \mid \bigcup_{\ell\in\Delta} S_\ell \mid = \mid \left(\bigcup_{\ell=1}^{n-k} S'_{j_\ell}\right) \mid + k$$

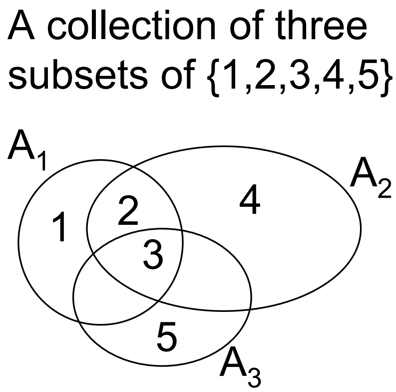
from which it follows that $\mathscr{S}'$ satisfies the marriage condition. Therefore it has an **SDR** $(x_{j_1}, x_{j_2}, \ldots, x_{j_{n-k}})$. Now for any $t \in \{1, 2, \ldots, n-k\}$ we have that $x_{j_t} \in S_{j_t} - X$, so $x_{j_t} \in S_{j_t}$ and $x_{j_t} \notin X$. But $x_{j_t} \notin X$ gives us that $x_{j_t} \neq x_{i_\ell}$ for any $\ell \in \{1, 2, \ldots, k\}$. This implies that $(x_1, \ldots, x_n)$ is an **SDR** for $\mathscr{S}$, completing the proof. □

6.2. PERMANENTS AND THE HALL MATRIX

We now know when it is possible to have an **SDR**. But how many **SDR's** does a collection of non-empty sets have? To count **SDR's** we will represent our sets as matrices. The matrix we will use is a special incidence matrix called the **Hall Matrix**. This matrix was named after Philip Hall, who originally proved the marriage theorem in 1935. An illustration of the Hall Matrix for three sets is given in Figure 6.1.

DEFINITION 6.6. Let $\mathscr{A} = (A_1, A_2, ...A_n)$ be a collection of finite subsets of a the set $A = \{1, 2, ..., n\}$. The **Hall Matrix**, associated with the collection $\mathscr{A}$ is the $n \times n$, $(0,1)$ matrix whose $(i, j) - th$ entry is 1 if and only if $i \in A_j$.

Additionally, we need to define the permanent of a matrix.



Hall Matrix for this collection

1	1	1	0	0
0	1	1	1	0
0	0	1	0	1

FIGURE 6.1. General Hall Matrix

DEFINITION 6.7. If A is an $n \times n$ matrix with the i, j entry given by a_{ij}, the **Permanent** of A, denoted $\text{per}(A)$, is

$$\sum_{\sigma \in \pi(n)} a_{1\sigma(1)} a_{2\sigma(2)} \ldots a_{n\sigma(n)}$$

where $\pi(n)$ denotes the symmetric group on the n symbols $\{1, 2, ..., n\}$.

The following is immediate from the definition:

LEMMA 6.8. *Let A be an $n \times n$ matrix and c be a constant. Then $per(cA) = c^n per(A)$*

PROOF. Let $A = (a_{ij})$ and $cA = (b_{ij})$. Then $b_{ij} = ca_{ij}$ and so for any $\sigma \in \pi(n)$ we have that

$$b_{1\sigma(1)} b_{2\sigma(2)} \ldots b_{n\sigma(n)} = ca_{1\sigma(1)} ca_{2\sigma(2)} \ldots ca_{n\sigma(n)} = c^n a_{1\sigma(1)} a_{2\sigma(2)} \ldots a_{n\sigma(n)}$$

Then

$$\begin{aligned} \text{per}(cA) &= \sum_{\sigma \in \pi(n)} b_{1\sigma(1)} b_{2\sigma(2)} \ldots b_{n\sigma(n)} = \sum_{\sigma \in \pi(n)} c^n a_{1\sigma(1)} a_{2\sigma(2)} \ldots a_{n\sigma(n)} \\ &= c^n \sum_{\sigma \in \pi(n)} a_{1\sigma(1)} a_{2\sigma(2)} \ldots a_{n\sigma(n)} = c^n \text{per}(A) \end{aligned}$$

□

Now we show that it is possible to count an **SDR's** by evaluating a permanent.

THEOREM 6.9. *Suppose $\mathscr{A} = (A_1, A_2, ...A_n)$ is a collection of finite subsets of a the set $\{1, 2, ..., n\}$. Then the number of ways to select an* **SDR** *for $\mathscr{A}$ is equal to the permanent of the Hall matrix, H, associated with $\mathscr{A}$.*

PROOF. In the evaluation of $\text{per}(H)$, the product corresponding to a particular permutation σ is 1 precisely when $i \in A_{\sigma(i)}$ for all i, otherwise the product equals to 0. So the permanent counts the number of permutations σ for which $i \in A_{\sigma(i)}$ for all i. But $i \in A_{\sigma(i)}$ for all i is equivalent with $\sigma^{-1}(i) \in A_i$, which means precisely that $(\sigma^{-1}(1), \sigma,^{-1}(2), ..., \sigma^{-1}(n))$ is an **SDR** for $(A_1, A_2, ...A_n)$. So each such permutation σ gives an **SDR** for $\mathscr{A}$. Thus the number of ways to select an **SDR** for $\mathscr{A}$ is at least equal to $\text{per}(H)$.

Now, any **SDR** arises from such a permutation, since its elements $x_1, \ldots, x_n$ are all different and they are in $\{1, 2, \ldots, n\}$. Hence $\text{per}(H)$ is at least equal to the number of ways to select an **SDR** for $\mathscr{A}$.

Hence the number of ways to select an **SDR** for $\mathscr{A}$ is equal to $\text{per}(H)$. □

6.3. TWO FAMOUS THEOREMS

So the number of **SDR**'s is equal to a permanent, but how do you evaluate a permanent? In 1926 B.L. van der Waerden suggested the problem of determining the minimal permanent among all $n \times n$ doubly stochastic matrices, which are matrices in which the row sums and the column sums are all equal to 1. He conjectured that for any doubly stochastic matrix A,

$$\text{per}(A) \geq n!/n^n$$

By 1981, D.I. Falikman and G.P. Egoritsjev had both provided different proofs of the conjecture. The theorem is known as the van der Waerden conjecture. [**7**] and [**6**]

Additionally, in 1967 H. Minc conjectured that if A is a $(0,1)$ matrix with row sums r_i, then

$$\operatorname{per}(A) \leq \prod_{i=1}^{n} r_i!^{\frac{1}{r_i}}$$

This was proved in 1973 by L.M. Bregman. [**5**]

CHAPTER 7

THE NUMBER OF LATIN SQUARES

In this chapter we will use a Hall Matrix to count the number of Latin squares. Recall that

DEFINITION 7.1. A **Latin square of rank** n is an $n \times n$ matrix where every row and every column contains precicely one of each of the numbers $1, 2, \ldots, n$.

A partially filled rank 4 Latin Square and the corresponding Hall Matrix for it's third column is shown in Figure 7.1. The i^{th} row of the matrix tells us which numbers are allowable in the i^{th} cell in the third column of the Latin square.

THEOREM 7.2. *The number of rank s Latin squares is at least* $s!^{2s}/s^{(s^2)}$.

PROOF. For a rank s Latin square, the number of ways to fill in the first column is clearly $s!$. Suppose we have completed k columns of the Latin square. We now want to fill in the $k+1$-st column. For each cell $(i, k+1)$ of the $k+1^{st}$ column, we let A_i be the set of numbers not yet used in the i^{th} row. The size of A_i is therefore $s-k$. Now, filling in the $k+1^{st}$ column is equivalent to finding an **SDR** for the collection $(A_1, A_2, ..., A_s)$, since the number put in the i-th row has to be in A_i and the numbers in the column must all be different. The number of ways this can be done is equal to the permanent of the

1	4		
2	1		
3	2		
4	3		

0	1	1	0
0	0	1	1
1	0	0	1
1	1	0	0

FIGURE 7.1. A rank 4 Latin square and the Hall Matrix for its 3^{rd} column

corresponding Hall matrix for the collection $(A_1, A_2, \ldots, A_s)$. We denote this matrix H. Consider the matrix $(s-k)^{-1}H$. Each row has $s-k$ non-zero entries of size $1/(s-k)$. Hence the row sums are all equal to 1. Now consider the i^{th} column of the Hall matrix H. Since the number i was used exactly once in each of the k columns already filled in our $s \times s$ Latin square, there will be exactly $s-k$ 1's in the i^{th} column of H. But as this is true for each i, we can say that H has $s-k$ 1's in each column. Therefore the columns of $(s-k)^{-1}H$ all sum to 1. Hence $(s-k)^{-1}H$ is a doubly stochastic, $s \times s$ matrix.

By the van der Warden conjecture, $\operatorname{per}\Big((s-k)^{-1}H\Big) \geq s!/s^s$. Hence $\operatorname{per}(H) \geq (s-k)^s s!/s^s$. By Theorem 6.9, there are at least $(s-k)^s s!/s^s$ ways to fill in the $k+1$-st column, once the first k columns have been filled in. Now to obtain a lower bound on the number of Latin squares, we simply need to take the product over k ranging from 0 to $s-1$.

$$\prod_{k=0}^{s-1} \frac{(s-k)^s s!}{s^s} = \left(\frac{s!}{s^s}\right)^s \prod_{k=0}^{s-1} (s-k)^s = \left(\frac{s!}{s^s}\right)^s (s!)^s = \frac{s!^{2s}}{s^{(s^2)}}$$

Hence the number of rank s Latin squares is at least $s!^{2s}/s^{(s^2)}$. $\square$

At this point we would like to re-formulate the above result in terms of the exponential function. This will make it easier to compare to the corresponding formula for Sudoku puzzles. We will invoke Stirling's formula for factorials. There are many versions of this result. We will use the one used in [**8**]. The interested reader may also refer to [**11**] or [**14**] for details.

THEOREM 7.3.

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

THEOREM 7.4. *The number of rank s Latin squares is at least* $\frac{s!^{2s}}{s^{s^2}} = s^{s^2} e^{-2s^2 - O(s\ln s)}$.

PROOF. Let N be the number of rank s Latin squares. By Theorem 7.2, we see that

$$N \geq \frac{s!^{2s}}{s^{s^2}}$$

But, using Stirling's formula,

$$\begin{aligned}
\ln\left(\frac{s!^{2s}}{s^{s^2}}\right) &= 2s\ln(s!) - s^2\ln s \\
&= 2s[s\ln s - s + \frac{1}{2}\ln s + O(1)] - s^2\ln s \\
&= s^2\ln s - 2s^2 + s\ln s + 2sO(1) \\
&= \ln s^{s^2} + s\ln(s) - 2s^2 + 2sO(1)
\end{aligned}$$

Therefore,

$$\frac{s!^{2s}}{s^{s^2}} = s^{s^2}e^{-2s^2+s\ln(s)+2sO(1)}$$

Since for any constant C we have that if $s \geq e^{2C}$ then $\ln(s) \geq 2C$ and thus $s\ln(s) \geq 2sO(1)$, we get that $s\ln(s)+2sO(1) \leq 2s\ln(s)$. Also since $s\ln(s) \to \infty$ as $s \to \infty$, if s is big enough, $s\ln(s)+2sO(1) > 0$. Therefore $s\ln(s)+sO(1) = O(s\ln(s))$, and so

$$\frac{s!^{2s}}{s^{s^2}} = s^{s^2}e^{-2s^2+O(s\ln s)}$$

$\square$

From now on when we talk about n and m we will view m as some function $m(n)$ of n. This means that any function of m and n will ultimately become just a function of n. Thus, the O-notations we use will refer to functions of n only.

THEOREM 7.5. *Let $m = m(n)$ be a function of n such that $n \leq m(n)$. Then there is a positive constant C and a nummber n_0 such that fod all $n \geq n_0$ number of rank nm Latin squares is at least $(nm)^{(nm)^2}e^{-2(nm)^2-C(nm\ln m)}$.*

PROOF. Let $s = nm$. Then immediately from 7.4 we get that the number of Latin squares is at least

$$(nm)^{(nm)^2}e^{-2(nm)^2+O(nm\ln mn)}$$

This means that there is a function $g(n) = O(\ln(nm))$ such that the number of Latin squares is at least

$$(nm)^{(nm)^2} e^{-2(nm)^2 + g(n)}$$

Since $g(n) \in O(nm\ln(mn))$ measn that there is a positive constant C_1 and a num,ber n_0 such that

$$|g(n)| \leq Cmn\ln(nm) = Cmn(\ln(n) + \ln(m)) \leq 2C_1\ln(m),$$

we get that $-2C_1nm\ln(m) \leq g(n)$, and the result follows. $\square$

CHAPTER 8

TECHNICAL DETAILS

We use the same conceptual process to count the number of Sudoku puzzles as we used count the number of Latin squares. However, the algebra that results is much more difficult. In this section we present several lemmata which will eventually simplify our calculation at the end. These by themselves are just the algebraic details, their meaning will become clear in the later chapters.

Throughout this chapter, for each we will assume that $m = m(n)$ is an integer valued function of n, n, r are integers, where $2 \leq n \leq m$ and $1 \leq r \leq n$, we will use the following notation:

$$\alpha_r = \lceil \tfrac{(r-1)m}{n-1} \rceil \tag{1}$$

$$\beta = \frac{m}{n-1} \tag{2}$$

The goal of this chapter is to prove the following single equality, namely, that under appropriate conditions we have that

$$\sum_{r=1}^{n} \left\{ \sum_{k=1}^{\alpha_r} \ln(nm-(r-1)m-k+1) + \sum_{k=\alpha_r+1}^{m} \ln(nm-(k-1)n) \right\}$$
$$< nm\ln(nm) - 1.5nm$$

Note that since $m \geq n$, this gives us that $\beta > 1$. We now have the following inequalities (using also that $r \leq n$:

$$\begin{aligned}
(r-1)\beta &= \frac{(r-1)m}{n-1} \leq \alpha_r && (3)\\
(r-1)\beta + 1 &\geq \left\lfloor \frac{(r-1)m}{n-1} \right\rfloor + 1 \geq \alpha_r && (4)\\
\alpha_1 &= 0 && (5)\\
\alpha_r &\leq \left\lceil \frac{(n-1)m}{n-1} \right\rceil = m && (6)\\
\alpha_r &\geq \left\lceil \frac{m}{n-1} \right\rceil \geq 2 \text{ for all } r \geq 2 && (7)
\end{aligned}$$

Also,

$$\begin{aligned}
\sum_{r=0}^{n-1} \beta r &= \frac{\beta n(n-1)}{2} = \frac{mn}{2} && (8)\\
\sum_{r=0}^{n-2} \beta r &= \frac{mn}{2} - (n-1)\beta = \frac{mn}{2} - m && (9)
\end{aligned}$$

We will need the following Lemmata

LEMMA 8.1. *For each positive integers n,k, $\frac{k}{n+k} < ln(n+k) - ln(n) < \frac{k}{n}$*

PROOF. Since the derivative of $\ln(x)$ is $\frac{1}{x}$, and these two functions are continuous on $(0,\infty)$, by the Mean Value Theorem for derivatives [**2**] we have that there is a $\psi \in (n, n+k)$ such that

$$\frac{\ln(n+k) - \ln(n)}{k} = \frac{1}{\psi}$$

Since $\frac{1}{n+k} < \frac{1}{\psi} < \frac{1}{n}$, the statement follows. □

LEMMA 8.2. *For each integers ℓ, j where $2 \leq \ell \leq j$ we have that*

$$\sum_{i=\ell}^{j} \frac{1}{i} \leq \ln(j) - \ln(\ell) + \frac{1}{\ell}$$

PROOF. Since $\frac{1}{x}$ is a decreasing positive function on the interval $[1,\infty)$, we get, using the left-hand Riemann sums [**2**] of the corresponding intervals of length 1 that

$$\sum_{i=\ell}^{j} \frac{1}{x} \leq \int_{\ell-1}^{j} \frac{1}{x} dx$$

The statement follows from noting that

$$\int \frac{1}{x} dx = \ln(x) + C$$

and using Lemma 8.1 $\square$

LEMMA 8.3. *For each integers ℓ, j where $2 \leq \ell \leq j$ we have that*

$$\begin{aligned} j\ln(j) - (\ell-1)\ln(\ell) + \ell - j &\leq \sum_{i=\ell}^{j} \ln(i) \\ &\leq (j+1)\ln(j+1) - \ell\ln(\ell) + \ell - j - 1 \end{aligned}$$

PROOF. Since $\ln(x)$ is an increasing nonnegative function on the interval $[1,\infty)$, we get, using the left- and right-hand Riemann sums of the corresponding intervals of length 1 that

$$\int_{\ell-1}^{j} \ln(x) dx \leq \sum_{i=\ell}^{j+1} \ln(i) \leq \int_{\ell-1}^{j} \ln(x) dx$$

The statement follows from noting that

$$\int \ln(x) dx = x\ln(x) - x + C$$

$\square$

LEMMA 8.4. *For each integers ℓ, j where $2 \leq \ell \leq j$ we have that*

$$\begin{aligned} \frac{j^2}{2}\ln(j) - \frac{(\ell-1)^2}{2}\ln(\ell-1) &+ \frac{(\ell-1)^2}{4} - \frac{j^2}{4} \leq \sum_{i=\ell}^{j} i\ln(i) \\ &\leq \frac{(j+1)^2}{2}\ln(j+1) - \frac{\ell^2}{2}\ln(\ell) + \frac{\ell^2}{4} - \frac{(j+1)^2}{4} \end{aligned}$$

PROOF. Since $x\ln(x)$ is an increasing nonnegative function on the interval $[1,\infty)$, we get, using the left- and right-hand Riemann sums of the corresponding intervals of length 1 that

$$\int_{\ell-1}^{j} x\ln(x)dx \leq \sum_{i=\ell}^{j} i\ln(i) \leq \int_{\ell}^{j+1} x\ln(x)dx$$

Note that

$$\int x\ln(x)dx = \frac{x^2}{2}\ln(x) - \frac{x^2}{4} + C$$

The lemma follows. $\square$

LEMMA 8.5.

$$\sum_{r=1}^{n}\sum_{k=1}^{\alpha_r}\frac{1}{nm-(r-1)m-k+1} + \sum_{r=1}^{n}\sum_{k=\alpha_r+1}^{m}\frac{1}{nm-(k-1)n} \leq 3\ln m + 2$$

PROOF. We have from equation (5) that

$$\begin{aligned}\sum_{r=1}^{n}\sum_{k=1}^{\alpha_r}\frac{1}{nm-(r-1)m-k+1} &= \sum_{r=2}^{n}\sum_{k=1}^{\alpha_r}\frac{1}{nm-(r-1)m-k+1} \\ &= \sum_{r=2}^{n}\sum_{j=nm-(r-1)m-\alpha_r+1}^{nm-(r-1)m}\frac{1}{j}\end{aligned}$$

Using equation (6) we obtain

$$nm-(r-1)m-\alpha_r+1 \geq mn-(r-1)m-m+1 = mn-rm+1$$

Therefore

$$\sum_{r=2}^{n}\sum_{k=1}^{\alpha_r}\frac{1}{nm-(r-1)m-k+1} \leq \sum_{r=2}^{n}\sum_{j=nm-(r)m-1}^{nm-(r-1)m}\frac{1}{j}$$

because this one is summing more terms. Since reversing the sum over r nicely gives us

$$\begin{aligned}\sum_{r=2}^{n}\sum_{j=nm-rm+1}^{nm-rm+m}\frac{1}{j} &= \sum_{j=1}^{m}\frac{1}{j}+\sum_{j=2m+1}^{3m}\frac{1}{j}+\cdots+\sum_{j=(n-2)m+1}^{nm-m}\frac{1}{j}\\ &= \sum_{h=1}^{nm-m}\frac{1}{h}=1+\frac{1}{nm-m}+\sum_{h=1}^{nm-m-1}\frac{1}{h},\end{aligned}$$

we get from Lemma 8.2 and Lemma 8.1 that

$$\begin{aligned}\sum_{r=2}^{n}\sum_{k=1}^{\alpha_r}\frac{1}{nm-(r-1)m-k+1} &\leq 1+\frac{1}{nm-m}+\ln(nm-m)-\ln(2)\\ &\leq \frac{1}{m(n-1)}+\ln(nm)+\frac{1}{nm}\\ &\leq \ln(nm)+1\leq 2\ln(m)+1\end{aligned}$$

Also,

$$\begin{aligned}\sum_{r=1}^{n}\sum_{k=\alpha_r+1}^{m}\frac{1}{nm-(k-1)n} &< \sum_{r=1}^{n}\sum_{k=1}^{m}\frac{1}{nm-(k-1)n}=n\sum_{k=0}^{m-1}\frac{1}{nm-kn}=\sum_{k=0}^{m-1}\frac{1}{m-k}\\ &= \sum_{h=1}^{m}\frac{1}{h}\leq 1+\frac{1}{m}+\ln(m)-\ln(2)\leq \ln(m)+1\end{aligned}$$

and the statement follows. $\square$

LEMMA 8.6.

$$\begin{aligned}\sum_{r=1}^{n}\left\{\sum_{k=1}^{\alpha_r}\ln(nm-(r-1)m-k+1)\right\} &\leq mn\ln(nm)-m\ln(nm)+\ln(nm)-nm+\frac{m}{n}\\ &- \sum_{r=2}^{n}\left\{\sum_{k=\alpha_r+1}^{m}\ln\big(m(n-r+1)-k+1\big)\right\}\end{aligned}$$

PROOF. Clearly, if $r=1$ then $\alpha_r=0$. Therefore

$$\sum_{r=1}^{n}\left\{\sum_{k=1}^{\alpha_r}\ln(nm-(r-1)m-k+1)\right\}=\sum_{r=2}^{n}\left\{\sum_{k=1}^{\alpha_r}\ln(nm-(r-1)m-k+1)\right\}$$

From equations (6) and (7) we get $1 \leq \alpha_r \leq m$ for $r > 1$. Also note that for any k such that $1 \leq k \leq m$, we have that

$$m(n-r)+1 \leq nm-(r-1)m-k+1 \leq nm-(r-1)m \tag{10}$$

Now,

$$\begin{aligned}
E &= \sum_{r=2}^{n}\left\{\sum_{k=1}^{\alpha_r}\ln(nm-(r-1)m-k+1)\right\} \\
&= \sum_{r=2}^{n}\left\{\sum_{k=1}^{\alpha_r}\ln\big(m(n-r+1)-k+1\big)\right\} \\
&= \sum_{r=2}^{n}\left\{\sum_{j=(n-r+1)m-\alpha_r+1}^{(n-r+1)m}\ln(j)\right\} = \sum_{r=2}^{n}\left\{\sum_{j=(n-r)m+1+m-\alpha_r}^{(n-r+1)m}\ln(j)\right\} \\
&= \sum_{r=2}^{n}\left\{\sum_{j=(n-r)m+1}^{(n-r+1)m}\ln(j) - \sum_{j=(n-r)m+1}^{(n-r)m+m-\alpha_r}\ln(j)\right\} \\
&= \sum_{r=2}^{n}\left\{\sum_{j=(n-r)m+1}^{(n-r+1)m}\ln(j) - \sum_{j=(n-r)m+1}^{(n-r+1)m-\alpha_r}\ln(j)\right\} \\
&= \sum_{r=2}^{n}\left\{\sum_{j=(n-r)m+1}^{(n-r+1)m}\ln(j) - \sum_{k=\alpha_r+1}^{m}\ln\big(m(n-r+1)-k+1\big)\right\}
\end{aligned}$$

It is easy to see using Lemma 8.3 that

$$\begin{aligned}
\sum_{r=2}^{n}\sum_{j=(n-r)m+1}^{(n-r+1)m}\ln(j) &= \sum_{s=1}^{(n-1)m}\ln(s) = \sum_{s=2}^{(n-1)m}\ln(s) \\
&\leq ((n-1)m+1)\ln((n-1)m+1) - 2\ln(2) - (n-1)m+1
\end{aligned}$$

Thus, using Lemma 8.1 we get that

$$\sum_{r=2}^{n}\sum_{j=(n-r)m+1}^{(n-r+1)m}\ln(j) \leq ((n-1)m+1)\left(\ln(nm) - \frac{m-1}{mn}\right) - (n-1)m - 1.8$$

Now,

$$\begin{aligned}\frac{(m-1)(n-1)m+m-1}{nm} &= \frac{m^2n-mn-m^2+2m-1}{mn}\\ &= m-1-\frac{m}{n}+\frac{2}{n}+\frac{1}{nm}\\ &\geq m-1-\frac{m}{n}\end{aligned}$$

therefore

$$\begin{aligned}\sum_{r=2}^{n}\sum_{j=(n-r)m+1}^{(n-r+1)m}\ln(j) &\leq ((n-1)m+1)\ln(nm)-nm-1+\frac{m}{n}\\ &= mn\ln(nm)-m\ln(nm)+\ln(nm)-nm+\frac{m}{n}\end{aligned}$$

Hence,

$$\begin{aligned}\sum_{r=2}^{n}\left\{\sum_{k=1}^{\alpha_r}\ln(nm-(r-1)m-k+1)\right\} &\leq mn\ln(nm)-m\ln(nm)+\ln(nm)-nm+\frac{m}{n}\\ &- \sum_{r=2}^{n}\left\{\sum_{k=\alpha_r+1}^{m}\ln\Big(m(n-r+1)-k+1\Big)\right\}\end{aligned}$$

□

LEMMA 8.7. *We have that*

$$\sum_{r=2}^{n-2}\Big(m(n-r+1)-(r-1)\beta-1\Big)\ln\Big(m(n-r+1)-(r-1)\beta-1\Big)\geq$$
$$\Big(\frac{mn^2}{2}-n-\frac{3mn}{2}+2\Big)\ln(nm)+\frac{m}{2}-\frac{n^2}{2}-\frac{1}{2}-\frac{mn(n-1)}{4}+\frac{m+\beta}{2}\ln(n)$$

PROOF.

$$\begin{aligned}C_2 &:= \sum_{r=2}^{n-2}\Big(m(n-r+1)-(r-1)\beta-1\Big)\ln\Big(m(n-r+1)-(r-1)\beta-1\Big)\\ &= \sum_{r=1}^{n-3}\Big(mn-r(\beta+m)-1\Big)\ln\Big(mn-r(\beta+m)-1\Big)\end{aligned}$$

Clearly, $(n-2)(\beta+m) = (n-1)\beta - \beta + (n-2)m = nm - m - \beta$, so using appropriate Riemann sums on intervals of length $\beta + m$ we obtain that

$$
\begin{aligned}
(\beta+m)C_2 &\geq \int\limits_{m+\beta-1}^{mn-\beta-m-1} x\ln(x)dx \\
&\geq \frac{(mn-\beta-m-1)^2}{2}\ln(mn-\beta-m+1) \\
&\quad -\frac{(mn-\beta-m-1)^2-(m+\beta-1)^2}{4} \\
&\quad -\frac{(m+\beta-1)^2}{2}\ln(m+\beta-1) \\
&= \frac{(mn-\beta-m-1)^2}{2}\ln(mn-\beta-m-1) \\
&\quad -\frac{m^2n^2}{4}+\frac{mn(\beta+m-1)}{2}-\frac{(m+\beta-1)^2}{2}\ln(m+\beta-1),
\end{aligned}
$$

therefore

$$
\begin{aligned}
C_2 &> \frac{(mn-\beta-m-1)^2}{2(m+\beta)}\ln(mn-\beta-m-1)-\frac{m^2n^2}{4(m+\beta)} \\
&\quad +\frac{mn}{2}-\frac{mn}{2(m+\beta)}-\frac{(m+\beta-1)^2}{2(m+\beta)}\ln(m+\beta-1) \\
&> \frac{(mn-\beta-m-1)^2}{2(m+\beta)}\ln(mn-\beta-m-1)-\frac{m^2n^2}{4(m+\beta)} \\
&\quad +\frac{mn}{2}-\frac{mn}{2(m+\beta)}-\frac{m+\beta}{2}\ln(m+\beta)
\end{aligned}
$$

Now, it is easy to see that

$$
\frac{\left(mn-(\beta+m)-1\right)^2}{2(\beta+m)} = \frac{m^2n^2}{2(m+\beta)}-\frac{mn}{(m+\beta)}-mn+\frac{\beta+m}{2}+1+\frac{1}{2(\beta+m)}
$$

so using the fact that $m+\beta = \frac{mn}{n-1}$, we obtain that

$$
\begin{aligned}
\frac{\left(mn-(\beta+m)-1\right)^2}{2(\beta+m)} &= \frac{mn(n-1)}{2}-n+1-mn+\frac{\beta+m}{2}+1+\frac{1}{2(\beta+m)} \\
&> \frac{mn^2}{2}-n-\frac{3mn}{2}+\frac{\beta+m}{2}+2
\end{aligned}
$$

Also by Lemma 8.1,

$$\ln\Big(mn-\beta-m-1\Big) \;\; \geq \;\; \ln(mn)-\frac{\beta+m+1}{mn-\beta-m-1}$$

Now,

$$\begin{aligned}\frac{\big(mn-(\beta+m)-1\big)^2}{2(\beta+m)} \quad \cdot \quad \frac{(\beta+m+1)}{mn-\beta-m-1} &= \frac{mn-(\beta+m)-1}{2} \\ &\quad +\frac{(mn-(\beta+m)-1)}{2(\beta+m)} \\ &< \frac{mn-(\beta+m)}{2}-1+\frac{n(n-1)}{2}\end{aligned}$$

Putting all these together, we obtain that

$$\begin{aligned}&\frac{\big(mn-(\beta+m)-1\big)^2}{2(\beta+m)}\cdot\ln\Big(mn-\beta-m-1\Big) \\ &> \Big(\frac{mn^2}{2}-n-\frac{3mn}{2}+\frac{\beta+m}{2}+2\Big)\ln(nm)-\frac{mn-(\beta+m)}{2}-\frac{n^2}{2}+\frac{n}{2}-1\end{aligned}$$

Now,

$$\frac{m+\beta}{2}\ln(m+\beta) \;\; < \;\; \frac{m+\beta}{2}\Big(\ln(m)+\frac{\beta}{m+\beta}\Big) = \frac{m+\beta}{2}\ln(m)+\frac{\beta}{2}$$

Combining all these, we get

$$\begin{aligned}C_2 \;\; > \;\; & \Big(\frac{mn^2}{2}-n-\frac{3mn}{2}+\frac{\beta+m}{2}+2\Big)\ln(nm)-\frac{mn-(\beta+m)}{2}-\frac{n^2}{2}+\frac{n}{2}-1 \\ & -\frac{m^2n^2}{4(m+\beta)}+\frac{mn}{2}-\frac{mn}{2(m+\beta)}-\frac{m+\beta}{2}\ln(m)-\frac{\beta}{2} \\ \geq \;\; & \Big(\frac{mn^2}{2}-n-\frac{3mn}{2}+\frac{\beta+m}{2}+2\Big)\ln(nm)+\frac{m}{2}-\frac{n^2}{2}-\frac{1}{2}-\frac{mn(n-1)}{4} \\ & -\frac{m+\beta}{2}\ln(m)\end{aligned}$$

□

LEMMA 8.8.

$$\sum_{r=2}^{n-1} m(n-r)\ln(m(n-r)) \;\leq\; \ln(m)\Big(\frac{mn^2}{2}-\frac{3mn}{2}+m\Big) + \ln(n)\Big(\frac{mn^2}{2}-mn+m\Big)-\frac{mn^2}{4}-0.8m$$

PROOF.

$$\begin{aligned}
\sum_{r=2}^{n-1} m(n-r)\ln(m(n-r)) &= \sum_{r=1}^{n-2} mr\ln(mr) = m\ln(m)\left\{\sum_{r=1}^{n-2} r\right\}+m\left\{\sum_{r=1}^{n-2} r\ln(r)\right\}\\
&= m\ln(m)\frac{(n-1)(n-2)}{2}\\
&\quad +m\left\{-(n-1)\ln(n-1)+\sum_{r=2}^{n-1} r\ln(r)\right\},
\end{aligned}$$

where

$$\begin{aligned}
\sum_{r=2}^{n-1} r\ln(r) &\leq \frac{n^2}{2}\ln(n)-\frac{n^2}{4}-2\ln(2)+1,\\
-(n-1)\ln(n-1) &\leq -(n-1)\Big(\ln(n)-\frac{1}{n-1}\Big) = -(n-1)\ln(n)+1.
\end{aligned}$$

Therefore

$$\begin{aligned}
\sum_{r=2}^{n-1} m(n-r)\ln(m(n-r)) &\leq m\ln(m)\frac{(n-1)(n-2)}{2}\\
&\quad +m\left\{-(n-1)\ln(n)+\frac{n^2}{2}\ln(n)-\frac{n^2}{4}-1.8\right\}\\
&= \ln(m)\Big(\frac{mn^2}{2}-\frac{3mn}{2}+m\Big)\\
&\quad +\ln(n)\Big(\frac{mn^2}{2}-mn+m\Big)-\frac{mn^2}{4}-0.8m
\end{aligned}$$

$\square$

LEMMA 8.9.

$$\sum_{r=2}^{n}\left\{\sum_{k=\alpha_r+1}^{m}\ln\big(m(n-r+1)-k+1\big)\right\} >$$

$$2.3m-\frac{n^2}{2}-\frac{1}{2}-\frac{mn}{4}-\ln(m)\Big(m+n-2\Big)+\ln(n)\Big(\frac{5mn}{2}-\frac{m}{2}-n+2+\frac{\beta}{2}\Big)$$

PROOF. Using equation $\alpha_n = m$, we obtain

$$B_2 = \sum_{r=2}^{n}\sum_{k=\alpha_r+1}^{m}\ln\Bigg(m(n-r+1)-k+1\Bigg)=\sum_{r=2}^{n-1}\sum_{k=m(n-r)+1}^{m(n-r+1)-\alpha_r}\ln(k)$$

Now,

$$\begin{aligned}\sum_{k=m(n-r)+1}^{m(n-r+1)-\alpha_r}\ln(k) &\geq \int_{m(n-r)}^{m(n-r+1)-\alpha_r}\ln(x)dx\\ &= (m(n-r+1)-\alpha_r)\ln(m(n-r+1)-\alpha_r)-m(n-r+1)+\alpha_r\\ &\quad -m(n-r)\ln(m(n-r))+m(n-r)\\ &= (m(n-r+1)-\alpha_r)\ln(m(n-r+1)-\alpha_r)-m+\alpha_r\\ &\quad -m(n-r)\ln(m(n-r)),\end{aligned}$$

So

$$\begin{aligned}B_2 \geq \sum_{r=2}^{n-1}\Bigg\{&(m(n-r+1)-\alpha_r)\ln(m(n-r+1)-\alpha_r)-m+\alpha_r\\ &-m(n-r)\ln(m(n-r)),\Bigg\}.\end{aligned}$$

Using Lemma 8.8 we can see that

$$\begin{aligned}
B_2 &\geq (2m-\alpha_{n-1})\ln(2m-\alpha_{n-1}) \\
&\quad +\sum_{r=2}^{n-2}(m(n-r+1)-\alpha_r)\ln(m(n-r+1)-\alpha_r) \\
&\quad +\left\{\sum_{r=2}^{n-1}(-m+\alpha_r)\right\}+\ln(m)\Big(-\frac{mn^2}{2}+\frac{3mn}{2}-m\Big) \\
&\quad +\ln(n)\Big(-\frac{mn^2}{2}+mn-m\Big)+\frac{mn^2}{4}+0.8m
\end{aligned}$$

Now, $2m-\alpha_{n-1}=2m-\lceil(n-2)\beta\rceil\geq 2m-(n-1)\beta=m$, so

$$\begin{aligned}
B_2 &\geq \sum_{r=2}^{n-2}(m(n-r+1)-\alpha_r)\ln(m(n-r+1)-\alpha_r) \\
&\quad +\left\{\sum_{r=2}^{n-1}(-m+\alpha_r)\right\}+\ln(m)\Big(-\frac{mn^2}{2}+\frac{3mn}{2}\Big) \\
&\quad +\ln(n)\Big(-\frac{mn^2}{2}+mn-m\Big)+\frac{mn^2}{4}+0.8m
\end{aligned}$$

By equations (3) and (4), if we substitute $-(r-1)\beta-1$ for $-\alpha_r$ and $(r-1)\beta$ for α_r, we obtain

$$\begin{aligned}
B_2 &\geq \sum_{r=2}^{n-2}(m(n-r+1)-(r-1)\beta-1)\ln(m(n-r+1)-(r-1)\beta-1) \\
&\quad +\left\{\sum_{r=2}^{n-1}(-m+(r-1)\beta)\right\}+\ln(m)\Big(-\frac{mn^2}{2}+\frac{3mn}{2}-m\Big) \\
&\quad +\ln(n)\Big(-\frac{mn^2}{2}+mn-m\Big)+\frac{mn^2}{4}+0.8m
\end{aligned}$$

Now

$$\begin{aligned}
\sum_{r=2}^{n-1}\Big(-m+(r-1)\beta\Big) &= -m(n-2)+\frac{\beta(n-1)(n-2)}{2} \\
&= -mn+2m+\frac{m(n-2)}{2}=-\frac{mn}{2}+m.
\end{aligned}$$

Using this and Lemma 8.7 we obtain

$$
\begin{aligned}
B_2 \;\geq\; & \Big(\frac{mn^2}{2} - n - \frac{3mn}{2} + 2\Big)\ln(nm) \\
& + \frac{m}{2} - \frac{n^2}{2} - \frac{1}{2} - \frac{mn(n-1)}{4} + \frac{m+\beta}{2}\ln(n) \\
& + \ln(m)\Big(-\frac{mn^2}{2} + \frac{3mn}{2} - m\Big) + \ln(n)\Big(-\frac{mn^2}{2} + mn - m\Big) \\
& + \frac{mn^2}{4} + 0.8m - \frac{mn}{2} + m \\
= \;& 2.3m - \frac{n^2}{2} - \frac{1}{2} - \frac{mn}{4} \\
& - \ln(m)\Big(m + n - 2\Big) + \ln(n)\Big(\frac{5mn}{2} - \frac{m}{2} - n + 2 + \frac{\beta}{2}\Big)
\end{aligned}
$$

□

LEMMA 8.10.

$$
\begin{aligned}
\sum_{r=1}^{n-2}\Big((m - r\beta)\ln(m - r\beta - 1) \;\leq\; & \ln(m)\Big(\frac{mn}{2} - \frac{n\beta}{2} - 1\Big) - (\beta - 3)\ln(\beta - 1) \\
& - \frac{mn}{4} - \frac{3m}{4} + -1 + \frac{3\beta}{2(n-1)}
\end{aligned}
$$

PROOF. Since $(n-2)\beta = m - \beta$,

$$
\begin{aligned}
\sum_{r=1}^{n-3}\ln(m - r\beta - 1) \;=\; & \frac{1}{\beta}\sum_{r=1}^{n-3}\beta\ln(m - r\beta - 1) \\
\leq\; & \frac{1}{\beta}\int_{\beta-1}^{m-\beta-1}\ln(x)dx \\
\leq\; & (m - \beta - 1)\ln(m - \beta - 1) - (\beta - 1)\ln(\beta - 1) - m - 2
\end{aligned}
$$

Now,

$$
\begin{aligned}
(m - \beta - 1)\ln(m - \beta - 1) \;\leq\; & (m - \beta - 1)(\ln(m) - \frac{\beta + 1}{m}) \\
=\; & (m - \beta - 1)\ln(m) - \beta - 1 + \frac{(\beta + 1)^2}{m}
\end{aligned}
$$

Now, if n (and therefore m) is big enough, then

$$\begin{aligned}
\frac{(\beta+1)^2}{m} &= \frac{\beta^2}{m}+\frac{2\beta}{m}+\frac{1}{m} \\
&= \frac{\beta}{n-1}+\frac{2}{n-1}+\frac{1}{m} \\
&< \frac{\beta}{n-1}+1,
\end{aligned}$$

from which

$$(m-\beta-1)\ln(m-\beta-1) \leq (m-\beta-1)\ln(m)-\beta+\frac{\beta}{n-1}.$$

Thus,

$$\begin{aligned}
\sum_{r=1}^{n-3}\ln(m-r\beta-1)+\beta\ln(\beta-1) &\leq (m-\beta-1)\ln(m)-\beta+\frac{\beta}{n-1} \\
&\quad +\ln(\beta-1)-m-2
\end{aligned}$$

Therefore,

$$\begin{aligned}
\sum_{r=1}^{n-2}(m-r\beta)\ln(m-r\beta-1) &= \beta\ln(\beta-1)+\sum_{r=1}^{n-3}(m-r\beta)\ln(m-r\beta-1) \\
&\leq (m-\beta-1)\ln(m)-\beta+\frac{\beta}{n-1} \\
&\quad +\ln(\beta-1)-m-2 \\
&\quad +\sum_{r=1}^{n-3}(m-r\beta-1)\ln(m-r\beta-1)
\end{aligned}$$

Now, since $(n-2)\beta = m-\beta$,

$$\begin{aligned}
D &= \frac{1}{\beta}\sum_{r=1}^{n-3}\beta(m-r\beta-1)\ln(m-r\beta-1) \\
&\leq \frac{1}{\beta}\int_{\beta-1}^{m-\beta-1} x\ln(x)dx \\
&\leq \frac{(m-\beta)^2}{2\beta}\ln(m-\beta)-\frac{(m-\beta-1)^2-(\beta-1)^2}{4\beta}-\frac{(\beta-1)^2}{\beta}\ln(\beta-1)
\end{aligned}$$

Now,

$$\ln(m-\beta) \leq \ln(m) - \frac{\beta}{m}$$

we have, using that $m-\beta = (n-2)\beta$, that

$$\frac{(m-\beta)^2}{2\beta}\ln(m-\beta) \leq \frac{(m-\beta)(n-2)}{2}\ln(m) - \frac{(m-\beta)^2}{2m}.$$

Also, an easy computation gives that

$$\begin{aligned}
\frac{(m-\beta-1)^2-(\beta-1)^2}{4\beta} &= \frac{m^2-2m\beta+2\beta-1}{4\beta} > \frac{m(n-1)}{4} - \frac{m}{2} + \frac{1}{4} \\
&> \frac{mn}{4} - \frac{3m}{4} + \frac{1}{4}
\end{aligned}$$

Therefore, we get that

$$\begin{aligned}
D &< \frac{(m-\beta)(n-2)}{2}\ln(m) - \frac{(m-\beta)^2}{2m} \\
&\quad -\frac{mn}{4} + \frac{3m}{4} - \frac{1}{4} - \beta\ln(\beta-1) + 2\ln(\beta-1) \\
&< \frac{(m-\beta)(n-2)}{2}\ln(m) - \beta\ln(\beta-1) + 2\ln(\beta-1) \\
&\quad -\frac{m}{2} + \beta - \frac{mn}{4} + \frac{3m}{4} + \frac{\beta}{2(n-1)}
\end{aligned}$$

Therefore,

$$\begin{aligned}
\sum_{r=1}^{n-2}(m-r\beta)\ln(m-r\beta-1) &\leq \frac{(m-\beta)(n-2)}{2}\ln(m) - \beta\ln(\beta-1) + 2\ln(\beta-1) \\
&\quad -\frac{m}{2} + \beta - \frac{mn}{4} + \frac{3m}{4} + \frac{\beta}{2(n-1)} \\
&\quad +(m-\beta-1)\ln(m) - \beta + \frac{\beta}{n-1} \\
&\quad +\ln(\beta-1) - m - 2 \\
&= \ln(m)\left(\frac{mn}{2} - \frac{n\beta}{2} - 1\right) - (\beta-3)\ln(\beta-1) \\
&\quad -\frac{mn}{4} - \frac{3m}{4} + -1 + \frac{3\beta}{2(n-1)}
\end{aligned}$$

Now,

$\square$

LEMMA 8.11. *For big enough n, we have*

$$\sum_{r=2}^{n}\left\{\sum_{k=\alpha_r+1}^{m}\ln(nm-(k-1)n)\right\}\leq \ln(n)\Big(\frac{mn}{2}-3m-n+\beta-1\Big)$$
$$+\ln(m)\Big(\frac{mn}{2}-\frac{n\beta}{2}+2-\beta\Big)-\frac{3mn}{4}+3m+n-\beta$$

PROOF. Using $\alpha_n=m$, we obtain that

$$\begin{aligned}
B_1 &= \sum_{r=2}^{n}\sum_{k=\alpha_r+1}^{m}\ln\Big(nm-(k-1)n\Big)=\sum_{r=2}^{n-1}\sum_{k=\alpha_r+1}^{m}\ln\Big(nm-(k-1)n\Big)\\
&= \sum_{r=2}^{n-1}\sum_{k=\alpha_r+1}^{m}\ln\Big(n(m-k+1)\Big)=\sum_{r=2}^{n-1}\sum_{k=1}^{m-\alpha_r}\Big(\ln(n)+\ln(k)\Big)
\end{aligned}$$

and consequently, using equation (3)

$$\begin{aligned}
B_1 &= \ln(n)\left\{\sum_{r=2}^{n-1}(m-\alpha_r)\right\}+\sum_{r=2}^{n-1}\sum_{k=1}^{m-\alpha_r}\ln(k)\\
&\leq \ln(n)\left\{\sum_{r=1}^{n-2}(m-r\beta-1)\right\}\\
&\quad+\sum_{r=2}^{n-1}\Big(\ln(m-\alpha_r)+(m-\alpha_r)\ln(m-\alpha_r)-m+\alpha_r)\Big)\\
&\leq \ln(n)\Big((m-1)(n-2)-\frac{\beta(n-1)(n-2)}{2}\Big)\\
&\quad+\sum_{r=1}^{n-2}\Big((m-r\beta)\ln(m-r\beta-1)-m+r\beta)\Big)\\
&= \ln(n)\Big((m-1)(n-2)-\frac{m(n-2)}{2}\Big)+\sum_{r=1}^{n-2}\Big((m-r\beta)\ln(m-r\beta-1)\Big)\\
&\quad-m(n-2)+\frac{m(n-2)}{2}\\
&= \ln(n)\Big(\frac{mn}{2}-3m-n+2\Big)-\frac{mn}{2}+3m+n-2\\
&\quad+\sum_{r=1}^{n-2}\Big((m-r\beta)\ln(m-r\beta-1)\Big)
\end{aligned}$$

Now using Lemma 8.10 we obtain

$$\begin{aligned}
B_1 \;\leq\;& \ln(n)\Big(\frac{mn}{2} - 3m - n + 2\Big) - \frac{mn}{2} + 3m + n - 2 \\
& + \ln(m)\Big(\frac{mn}{2} - \frac{n\beta}{2} - 1\Big) - (\beta - 3)\ln(\beta - 1) \\
& - \frac{mn}{4} - \frac{3m}{4} + -1 + \frac{3\beta}{2(n-1)} \\
=\;& \ln(n)\Big(\frac{mn}{2} - 3m - n + 2\Big) + \ln(m)\Big(\frac{mn}{2} - \frac{n\beta}{2} - 1\Big) \\
& - \frac{3mn}{4} + 3m + n - 3 - (\beta - 3)\ln(\beta - 1)
\end{aligned}$$

Now,

$$\ln(\beta - 1) \leq \ln(\beta) - \frac{1}{\beta},$$

so

$$\Big(3 - \beta\Big)\ln(\beta - 1) < \Big(3 - \beta\Big)\ln(\beta) + 1$$

Using that $\ln(\beta) = \ln(m) - \ln(n-1) \leq \ln(m) - \ln(n) + \frac{1}{n-1}$, we get that

$$\Big(3 - \beta\Big)\ln(\beta - 1) \leq \Big(3 - \beta\Big)(\ln(m) - \ln(n)) + 3 - \beta$$

This means

$$\begin{aligned}
B_1 \;\leq\;& \ln(n)\Big(\frac{mn}{2} - 3m - n + 2\Big) + \ln(m)\Big(\frac{mn}{2} - \frac{n\beta}{2} - 1\Big) \\
& - \frac{3mn}{4} + 3m + n - 3 + \Big(3 - \beta\Big)(\ln(m) - \ln(n)) + 3 - \beta \\
=\;& \ln(n)\Big(\frac{mn}{2} - 3m - n + \beta - 1\Big) + \ln(m)\Big(\frac{mn}{2} - \frac{n\beta}{2} + 2 - \beta\Big) \\
& - \frac{3mn}{4} + 3m + n - \beta
\end{aligned}$$

$\square$

LEMMA 8.12. *If for some δ such that $0 < \delta \leq 3$ we have that whenever $m \geq n$ and $m \in O(n^{4-\delta})$, then there is an N such that for all $n \geq N$ we have*

$$\sum_{r=1}^{n} \left\{ \sum_{k=1}^{\alpha_r} \ln(nm-(r-1)m-k+1) + \sum_{k=\alpha_r+1}^{m} \ln(nm-(k-1)n) \right\} < nm\ln(nm) - 1.5nm$$

PROOF. By Lemma 8.6 we have that

$$\begin{aligned} \sum_{r=1}^{n} \left\{ \sum_{k=1}^{\alpha_r} \ln(nm-(r-1)m-k+1) \right\} \leq \; & mn\ln(nm) - m\ln(nm) + \ln(nm) - nm + \frac{m}{n} \\ & - \sum_{r=2}^{n} \left\{ \sum_{k=\alpha_r+1}^{m} \ln\big(m(n-r+1)-k+1\big) \right\} \end{aligned}$$

By Lemma 8.9 we have

$$\sum_{r=2}^{n} \left\{ \sum_{k=\alpha_r+1}^{m} \ln\big(m(n-r+1)-k+1\big) \right\} >$$
$$2.3m - \frac{n^2}{2} - \frac{1}{2} - \frac{mn}{4} - \ln(m)\Big(m+n-2\Big) + \ln(n)\Big(\frac{5mn}{2} - \frac{m}{2} - n + 2 + \frac{\beta}{2}\Big)$$

By Lemma 8.11 we have

$$\begin{aligned} \sum_{r=2}^{n} \left\{ \sum_{k=\alpha_r+1}^{m} \ln(nm-(k-1)n) \right\} \leq \; & \ln(n)\Big(\frac{mn}{2} - 3m - n + \beta - 1\Big) \\ & + \ln(m)\Big(\frac{mn}{2} - \frac{n\beta}{2} + 2 - \beta\Big) - \frac{3mn}{4} + 3m + n - \beta \end{aligned}$$

Combining these three Lemmata we get that the expression F we want to estimate in the Lemma is

$$\begin{aligned} F &\leq mn\ln(nm) - m\ln(nm) + \ln(nm) - nm + \frac{m}{n} \\ &-2.3m + \frac{n^2}{2} + \frac{1}{2} + \frac{mn}{4} + \ln(m)\Big(m+n-2\Big) \\ &+\ln(n)\Big(-\frac{5mn}{2} + \frac{m}{2} + n - 2 - \frac{\beta}{2}\Big) + \ln(n)\Big(\frac{mn}{2} - 3m - n + \beta - 1\Big) \\ &+\ln(m)\Big(\frac{mn}{2} - \frac{n\beta}{2} + 2 - \beta\Big) - \frac{3mn}{4} + 3m + n - \beta \end{aligned}$$

Thus,

$$\begin{aligned} F &\leq mn\ln(nm) - 1.5nm + \ln(m)\Big(n+1+\frac{mn}{2} - \frac{n\beta}{2} - \beta\Big) \\ &+\ln(n)\Big(-2mn - \frac{m}{2} - 2 + \frac{\beta}{2} - 3m\Big) + 01.2 + \frac{n^2}{2} + n - \beta + \frac{m}{n} \end{aligned}$$

Now, using

$$\begin{aligned} G &= \ln(m)\Big(n + \frac{mn}{2} - \frac{n\beta}{2} - \beta\Big) + \ln(n)\Big(-2mn - \frac{m}{2} - 2 + \frac{\beta}{2} - 3m\Big) \\ &+1.2 + \frac{n^2}{2} + n - \beta + \frac{m}{n} \end{aligned}$$

we get that

$$\begin{aligned} \frac{G}{mn} &= \ln(m)\Big(\frac{1}{m} + \frac{1}{mn} + \frac{1}{2} - \frac{1}{2(n-1)} - \frac{1}{n(n-1)}\Big) \\ &+\ln(n)\Big(-2 - \frac{1}{2n} - \frac{2}{mn} + \frac{1}{n(n-1)} - \frac{3}{n}\Big) \\ &+\frac{1.2}{mn} + \frac{n}{2m} + \frac{1}{m} - \frac{1}{n(n-1)} + \frac{1}{n^2} \end{aligned}$$

Clearly, for a fixed $\varepsilon > 0$, we get that if $n \geq n_0$, then

$$\begin{aligned}\frac{G}{mn} &= \ln(m)\Big(\frac{1}{m}+\frac{1}{mn}+\frac{1}{2}-\frac{1}{2(n-1)}-\frac{1}{n(n-1)}\Big)\\ &\quad +\ln(n)\Big(-2-\frac{1}{2n}-\frac{2}{mn}+\frac{1}{n(n-1)}-\frac{3}{n}\Big)+\varepsilon\end{aligned}$$

Now, since $m \geq n$, if $n-1 < m$, so$\frac{1}{mn} < \frac{1}{n(n_1)}$. Thus,

$$\frac{1}{m} + \frac{1}{mn}-\frac{1}{2(n-1)}-\frac{1}{n(n-1)} < \frac{1}{m}-\frac{1}{2(n-1)}$$

So if $m \geq 2n-2$, then we have that

$$\frac{G}{mn} < \frac{1}{2}\ln(m)-2\ln(n)+\frac{1.2}{mn}+\frac{n}{2m}+\frac{1}{m}-\frac{1}{n(n-1)}+\frac{1}{n^2}$$

and if also $m = O(n^{4-\delta})$, then, using that if $n \geq n_0$ then $m \leq Cn^{4-\delta}$, so $\ln(m) \leq \ln(C) + (4-\delta)\ln(n)$, we get

$$\begin{aligned}\frac{G}{mn} &\leq 2\ln(n)-\frac{\delta}{2}\ln(n)+\frac{1}{2}\ln(C)+\varepsilon\\ &= -\frac{\delta}{2}\ln(n)+\varepsilon,\end{aligned}$$

and, since $\ln(n) \to \infty$, we get that $\frac{G}{mn} < 0$, thus, $G < 0$. If $m \leq 2n-2$, then

$$\ln(n) \leq \ln(2n-2) = 2\ln(n-1) \leq 2\ln(n) - \frac{n-1}{,}$$

and we get similarly that if n is big enough, then for some $\varepsilon_1 > 0$ we have that

$$\frac{G}{mn} \leq -2\ln(n)+\ln(n)+\varepsilon_1 < -\ln(n)+\varepsilon_2,$$

and as before, we get that $G < 0$. Therefore,

$$F \leq mn\ln(nm)-1.5nm+G < mn\ln(mn)-1.5mn$$

$\square$

CHAPTER 9

COUNTING SUDOKU PUZZLES

We now turn our attention to calculating an upper bound for the number of Sudoku puzzles. Again, we will do this by calculating the sum of the permanents of the Hall matrices for each column of the Sudoku puzzle. However, we will formulate the Hall Matrix one way for some of the columns, and another way for the others. We will extensively use results from the previous chapter

A picture of a Hall matrix, formulated the first way, is given in Figure 9.1. This Matrix represents the third column of the $\mathrm{Dim}(2,3)$ puzzle to its left. As with the Latin square, the i^{th} row of the matrix tells us which numbers are allowable in the i^{th} cell in the third column of the Sudoku puzzle. For this Hall Matrix, we only allow numbers for a cell which have not been used in the sub-grid the cell lies in.

Again we will assume that $m = m(n)$ is some function of n such that $m \geq n$. Note that when $n = 1$, then a $\mathrm{Dim}(n,m)$ Sudoku puzzle just becomes a rank m Latin square: the size of a subgrid is just $1 \times m$, and the condition that we do not have repeating numbers in a sub-grid is the same as the condition that we do not have repeating numbers in a row.

1	2				
4	5				
3	4				
2	6				
6	1				
5	3				

0	0	1	0	0	1
0	0	1	0	0	1
1	0	0	0	1	0
1	0	0	0	1	0
0	1	0	1	0	0
0	1	0	1	0	0

FIGURE 9.1. A $\mathrm{Dim}(2,3)$ Sudoku puzzle and the Hall Matrix for its 3^{th} column

1	2	3	6		
4	5	6	2		
3	4	1	5		
2	6	5	3		
6	1	2	4		
5	3	4	1		

0	0	0	1	1	0
1	0	1	0	0	0
0	1	0	0	0	1
1	0	0	1	0	0
0	0	1	0	1	0
0	1	0	0	0	1

FIGURE 9.2. A Dim$(2,3)$ Sudoku puzzle and the Hall Matrix for its 5^{th} column

THEOREM 9.1. *If there is a δ such that $0 < \delta < 3$ and the function $m = m(n)$ satisfies $n \leq m$ and $m = O(n^{4-\delta})$, then there is a positive constant C and a number n_0 such that if $n \geq n_0$ then the number of Dim(n,m) Sudoku puzzles is at most*

$$(nm)^{(nm)^2} e^{-2.5(nm)^2 + C(nm\ln(m))^2}$$

PROOF. Note that since the result is valid for large n only, we will assume that $n \geq 2$.

Recall that a band is a column of sub-grids. Suppose we have filled in the first $r-1$ bands. Now consider the r^{th} band, and suppose we already have filled the first $k-1$ column How many ways can we fill in the k^{th} column of this band? For each cell $(i,(r-1)m+k)$, we let A_i be the set of numbers available for that cell. In the i^{th} row, there are $(r-1)m$ options already taken; and in the first $(k-1)$ columns of this sub-grid, there are $(k-1)n$ options already taken. The options taken in the row and taken in the sub-grid might overlap, so we can only be sure that we have used up at least

$$\max\{(r-1)m+k-1,(k-1)n\}$$

options that we can not use for this cell. There are also possibly some options already used in the column we are in.

The Hall matrix for this situation will look like the one in Figure 9.2

So the size of A_i is at most $nm - \max\{(r-1)m+k-1,(k-1)n\}$. Hence we have at most

$$\min\{nm-[(r-1)m+k-1], nm-(k-1)n\}$$

many numbers available to use. Since we must use each number from 1 to nm exactly once in this column, filling this column is equivalent to finding an **SDR** for the collection $\{A_1, A_2, ..., A_{nm}\}$. By Theorem 6.9 the number of ways this can be done is equal to the permanent of the corresponding Hall matrix for the collection $\{A_1, A_2, ..., A_{nm}\}$. Denote this matrix H. Each row i of H will have sum $s_i \le \min\{nm-[(r-1)m+k-1], nm-(k-1)n\}$.

We will now use Minc's conjecture. Notice that

$$nm-(k-1)n \le nm-(r-1)m$$

implies that

$$k \ge \tfrac{(r-1)m}{n-1}+1$$

Hence when $k \in \{1,2,...,\lceil \frac{(r-1)m}{n-1} \rceil\}$ the number of ways to fill in the k^{th} column is at most

$$(nm-(r-1)m-k+1)!^{\frac{nm}{nm-(r-1)m-k+1}}$$

When $k \in \{\lceil \frac{(r-1)m}{n-1} \rceil + 1, ..., m\}$ the number of ways to fill in the k^{th} column is at most

$$[nm-(k-1)(n)]!^{\frac{nm}{nm-(k-1)(n)}}$$

We must take the product over all columns to estimate the number of ways to fill in this band.

Hence the number of ways to fill in the r^{th} band is at most

$$\prod_{k=1}^{\alpha_r} [nm-(r-1)m-k+1]!^{\frac{nm}{nm-(r-1)m-k+1}} \times \prod_{k=\alpha_r+1}^{m} [nm-(k-1)(n)]!^{\frac{nm}{nm-(k-1)(n)}}$$

So if $S_{n,m}$ is defined to be the number of ways to fill in the entire Sudoku puzzle, then

$$S_{n,m} \leq \prod_{r=1}^{n} \left\{ \prod_{k=1}^{\alpha_r} [nm-(r-1)m-k+1]!^{\overline{nm-(r-1)m-k+1}^{\,nm}} \times \right.$$
$$\left. \times \prod_{k=\alpha_r+1}^{m} [nm-(k-1)(n)]!^{\frac{nm}{nm-(k-1)(n)}} \right\} \tag{11}$$

In order to prove our theorem, it is enough to show that there is a constant C and a number n_0 such that for all $n \geq n_0$ we have

$$S_{n,m} \leq (nm)^{(nm)^2} e^{-2.5(nm)^2+C(\ln(m))^2)}$$

or, alternatively, that

$$\ln(S_{n,m}) \leq (nm)^2 \ln(nm) - 2.5(nm)^2 + C(nm\ln(m))^2$$

This is equivalent with showing that

$$\frac{\ln(S_{n,m})}{nm} + nm \leq nm\ln(nm) - 1.5nm + C(\ln(m))^2 \tag{12}$$

Now from the expression we have in equiation (11)

$$\ln(S_{n,m}) \leq \sum_{r=1}^{n} \left\{ \sum_{k=1}^{\alpha_r} \ln\left\{ [nm-(r-1)m-k+1]!^{\frac{nm}{nm-(r-1)m-k+1}} \right\} + \right.$$
$$\left. + \sum_{k=\alpha_r+1}^{m} \ln\left\{ [nm-(k-1)(n)]!^{\frac{nm}{nm-(k-1)(n)}} \right\} \right\}$$

And so

$$\frac{\ln(S_{n,m})}{nm} \leq \sum_{r=1}^{n} \left\{ \sum_{k=1}^{\alpha_r} \frac{\ln[nm-(r-1)m-k+1]!}{nm-(r-1)m-k+1} + \sum_{k=\alpha_r+1}^{m} \frac{\ln[nm-(k-1)(n)]!}{nm-(k-1)n} \right\}$$

applying Sterlings formula to the factorials shows that there is a constant $C_1 > 0$ and a number n_1 such that whenever $n \geq n_1$, the expression on the right side is at most

$$\sum_{r=1}^{n}\sum_{k=1}^{\alpha_r}\left\{\ln(nm-(r-1)m-k+1)-1+\frac{(\frac{1}{2})\ln(nm-(r-1)m-k+1)+C_1}{nm-(r-1)m-k+1}\right\}$$
$$+\sum_{r=1}^{n}\sum_{k=\alpha_r+1}^{m}\left\{\ln(nm-(k-1)n)-1+\frac{(\frac{1}{2})\ln(nm-(k-1)n)+C_1}{nm-(k-1)n}\right\}$$

Notice that -1 is summed $n\alpha_r$ times in the first line, and then $n(m-\alpha_r)$ times in the second line for a total of nm times. Pulling this out and re-arranging terms gives us

$$\frac{\ln(S_{n,m})}{nm}+nm \leq$$
$$\sum_{r=1}^{n}\left\{\sum_{k=1}^{\alpha_r}\ln(nm-(r-1)m-k+1)+\sum_{k=\alpha_r+1}^{m}\ln(nm-(k-1)n)\right\} \tag{13}$$
$$+\sum_{r=1}^{n}\left\{\sum_{k=1}^{\alpha_r}\frac{(\frac{1}{2})\ln(nm-(r-1)m-k+1)}{nm-(r-1)m-k+1}+\sum_{k=\alpha_r+1}^{m}\frac{(\frac{1}{2})\ln(nm-(k-1)n)}{nm-(k-1)n}\right\} \tag{14}$$
$$+\sum_{r=1}^{n}\left\{\sum_{k=1}^{\alpha_r}\frac{C_1}{nm-(r-1)m-k+1}+\sum_{k=\alpha_r+1}^{m}\frac{C_1}{nm-(k-1)n}\right\} \tag{15}$$

Now, by Lemma 8.12 we have

$$(\ 13) \leq nm\ln(nm)-1.5nm \tag{16}$$

So in order to show equation(12), it is enough to show that for n big enough, there is a constant C such that (14)+(15) is at most $C(\log(m))^2$.

In (14), the numerator of each expression is at most $\frac{1}{2}\ln nm \leq \ln m$. Hence

$$(\ 14) \leq \ln m\sum_{r=1}^{n}\left\{\sum_{k=1}^{\alpha_r}\frac{1}{nm-(r-1)m-k+1}+\sum_{k=\alpha_r+1}^{m}\frac{1}{nm-(k-1)n}\right\}$$

So by lemma 8.5, we have,

$$(\ 14) \leq \ln m \times (3\ln m+2) = 3(\ln m)^2+2\ln(m)$$

In (15), we again use lemma 8.5 and immediately see that for some constant C

$$(15) \leq C \times 3 \ln m + 2$$

Therefore, (14) and (15) together are at most $3((\ln(m))^2 + 5\ln(m) + 2$, which is certainly smaller than $4(\ln(m))^2$ if n (and therefore m) are big enough. $\square$

CHAPTER 10

THE PROPORTION OF LATIN SQUARES THAT ARE ALSO SUDOKU PUZZLES

In this chapter we estimate the fraction of rank nm Latin squares that are also $\mathrm{Dim}(n,m)$ Sudoku puzzle, with other words, we estimate the probability that if we randomly select a rank nm Latin square such that every such square is selected with the same probability, then this randomly selected Latin square is also a $\mathrm{Dim}(n,m)$ Sudoku puzzle.

THEOREM 10.1. *Let p_{nm} be the probability that a randomly chosen rank nm Latin square is also a Sudoku puzzle. Then there is a positive constant C and a number n_0 such that if $n \geq n_0$ then*

$$p_{nm} \leq e^{-(nm)^2 + C(mn(\ln(m))^2)}$$

In particular, $p_{nm} \to 0$ as n tends to infinity.

PROOF. By Theorem 9.1, there is a positive constant C_1 and a number n_1 such that if $n \geq n_1$, then the number of Sudoku puzzles of Dim (n,m) is at most

$$(nm)^{(nm)^2} e^{-2.5(nm)^2 + C_1(nm(\ln m)^2)}.$$

By Theorem 7.5, there is a negative constant C_2 and a number n_2 such that if $n \geq n_2$ number of Latin squares of rank nm is at least

$$(nm)^{(nm)^2} e^{-2(nm)^2 + C_2(nm \ln m)}.$$

Select $n_3 = \max(n_1, n_2)$. If $n \geq n_3$, then the probability that a random Latin square of rank nm is also a $\mathrm{Dim}(m,n)$ Sudoku puzzle is at most

$$\frac{(nm)^{(nm)^2} e^{-2.5(nm)^2 + C_1(nm(\ln m)^2)}}{nm^{(nm)^2} e^{-2(nm)^2 + C_2(nm \ln m)}} \leq e^{-0.5(nm)^2 + C_1 nm(\ln m)^2) - C_2 nm \ln(m)}$$

Now,

$$C_1 nm(\ln m)^2) - C_2 nm \ln(m) \quad = \quad C_1 nm(\ln(m)) \Big(1 - C_2 \frac{1}{\ln(m)}\Big)$$

But clearly, if $m \geq e^{|C_2|}$, then this is at most $2C_2 nm(\ln(m))^2$. Since $m \geq n$, if $n \geq e^{|C_2|}$, this is achieved. So selecting $C_0 = 2C_1$ and $n_0 = \max(n_2, e^{|C_2|})$ will be enough for the first part of the claim. Since a probability is never negative, we only need to show that

$$\lim_{n \to \infty} \left(e^{-(nm)^2 + C(mn(\ln(m))^2})\right) = 0$$

Since

$$e^{-0.5(nm)^2 + C(mn(\ln(m))^2} = \frac{1}{e^{0.5(nm)^2 - C(mn(\ln(m))^2}} = \frac{1}{\left(e^{0.5(nm) - C(\ln(m))^2}\right)^{nm}},$$

it is enough to show that

$$\lim_{n \to \infty} (e^{0.5(nm) - C(\ln(m))^2}) = \infty,$$

or, alternatively, that

$$\lim_{n \to \infty} (0.5(nm) - C(\ln(m))^2) = \infty,$$

But this follows from using the L'Hopital Rule [**2**] to obtain

$$\lim_{m \to \infty} \frac{m}{\ln^2(m)} = \lim_{n \to \infty} \frac{m}{2\ln(m)} = \lim_{n \to \infty} \frac{m}{2} = \infty$$

and using that

$$0.5(nm) - C(\ln(m))^2 = 0.5n(\ln(m))^2 \left(\frac{m}{(\ln(m))^2} - \frac{C}{2n} \right) \to \infty$$

□

BIBLIOGRAPHY

1. `http://en.wikipedia.org/wiki/Sudoku`

2. H. Anton, I. Bivins, S. Davis *Calculus, eighth edition*, John Wiley and Sons, 2005

3. P. Bachmann *Analitische Zahlentheorie*, University of Michigan Libraries, 2006

4. R. Balakrishnan and K. Ranganathan, *A Textbook of Graph Theory*, Springer-Verlag, 1999

5. L.M. Bregman, *Some properties of non-negative matrices and their permanents*, Soviet Math.Dokl.14(4)(1973), pp.945-949

6. G.P. Egoritsjev, *Solution of van der Waerdens permanent conjecture*, Advances in Math. 42 (1981), pp.731-740

7. *Proof of the van der Waerden conjecture on the permanent of a doubly stochastic matrix*, Mat. Zametki 29 (1981) pp.931-938

8. Agnes M. Herzberg and M. Ram Murty, *Sudoku Squares and Chromatic Polynomials*, Notices of the American Mathematical Society, 54 (2007), pp.708-717.

9. D.E. Knuth, *The Art of Computer Programming: Volume I, Fundamental Algorithms* Addison-Wesley Professional, 1998

10. J.H. Van Lint and R.M. Wilson, *A Course in Combinatorics*, Cambridge University Press, 1992

11. R. B. Paris and D. Kaminsky, *Asymptotics and the Mellin-Barnes Integrals*, Cambridge University Press, 2001

12. Richard A. Silverman, *Introductory Complex Analysis*, Dover Publications, 1972

13. W.T. Tutte, *Graph Theory*, Addison-Wesley, 1984

14. E. T. Whittaker and G. N. Watson, *Course in Modern Analysis, fourth edition* Cambridge University Press, 1963