

Spring 2018

## A MATHEMATICAL ANALYSIS OF THE GAME OF CHESS

John C. White  
*Southeastern University - Lakeland*

Follow this and additional works at: <https://firescholars.seu.edu/honors>



Part of the [Data Science Commons](#), [Mathematics Commons](#), and the [Numerical Analysis and Computation Commons](#)

---

### Recommended Citation

White, John C., "A MATHEMATICAL ANALYSIS OF THE GAME OF CHESS" (2018). *Selected Honors Theses*. 101.

<https://firescholars.seu.edu/honors/101>

This Thesis is brought to you for free and open access by FireScholars. It has been accepted for inclusion in Selected Honors Theses by an authorized administrator of FireScholars. For more information, please contact [firescholars@seu.edu](mailto:firescholars@seu.edu).

A MATHEMATICAL ANALYSIS OF THE GAME OF CHESS

by

John C White

Submitted to the Honors Program Committee

in partial fulfillment

of the requirements for University Honors Scholars

Southeastern University

2018

Copyright by John C White

2018

## Acknowledgements

I would like to thank my thesis advisor, Cody Tessler, for all that he did to help me complete my thesis. I am grateful his support and insight. I would also like to thank Dr. Gordon Miller for everything he does for the honors program. It has been a pleasure to be a part of the program throughout my years at Southeastern University. Lastly, I want to thank my family for their unwavering love and support.

## Abstract

This paper analyzes chess through the lens of mathematics. Chess is a complex yet easy to understand game. Can mathematics be used to perfect a player's skills? The work of Ernst Zermelo shows that one player should be able to force a win or force a draw. The work of Shannon and Hardy demonstrates the complexities of the game. Combinatorics, probability, and some chess puzzles are used to better understand the game. A computer program is used to test a hypothesis regarding chess strategy. Through the use of this program, we see that it is detrimental to be the first player to lose the queen. Ultimately, it is shown that mathematics exists inherently in chess. Therefore math can be used to improve, but not perfect, chess skills.

**KEY WORDS:** chess, mathematics, combinatorics, probability, Zermelo, Hardy, Shannon

## Table of Contents

Acknowledgements.....	iii
Abstract.....	iv
Introduction.....	1
Literature Review.....	5
Zermelo’s Work.....	5
Countability.....	6
Solvability.....	8
Data Analysis.....	10
Preliminaries.....	10
Probability and Statistics.....	12
Chess Puzzles.....	15
Methodology.....	19
Conclusion.....	21
Works Cited.....	23

## A Mathematical Analysis of the Game of Chess

### **Introduction**

The game of chess has been around for hundreds of years. It has played a prominent role in Western culture since the Middle Ages (Adams 2). Every day it is played by millions of people around the world. The game takes strategy, patience, and above all, problem solving. Chess is quite complex yet the objective is beautifully simple. This study will attempt to explain how mathematics, particularly probability, statistics, and combinatorics, fits into the world of chess.

The game is played on a square board made up of 64 smaller squares. Half of these small squares are black and the other half are white, making a checkerboard pattern. At the beginning of the game, there are 32 pieces, 16 black and 16 white. Each player has eight pawns, two castles, two knights, two bishops, one queen, and one king.

The pawn is often viewed as the weakest piece on the board. The pawn may move forward two squares on its first move, but may only move forward one square on all subsequent moves. The pawn can only move straight forward unless it is attacking an opposing piece. To attack, it must move to a square to its immediate forward diagonal. The pawn does have one very important characteristic. If a pawn makes it to the opposite end of the board, it is immediately “promoted” to become any other piece, except a king. It is also unable to remain a pawn (Let’s Play 2). Therefore, the weakest piece in the game can potentially become the most powerful piece.

The castle is a piece that can move forward, backward, right, or left. It can move as many squares as possible until it meets a square already occupied by another piece. The castle can attack from far away. Another name for the castle is the rook. For the







purposes of this paper, we will refer to castles as rooks from now on. The knight is an interesting piece. It may move two spaces forward or backward, then one space to the left or right. It can also move two spaces left or right, then one space forward or backward. The knight is the only piece that can “jump” over other pieces to reach its destination. The bishop may move any number of spaces diagonally. Therefore, the bishop that begins on a black square will never be on a white square for the duration of the game.

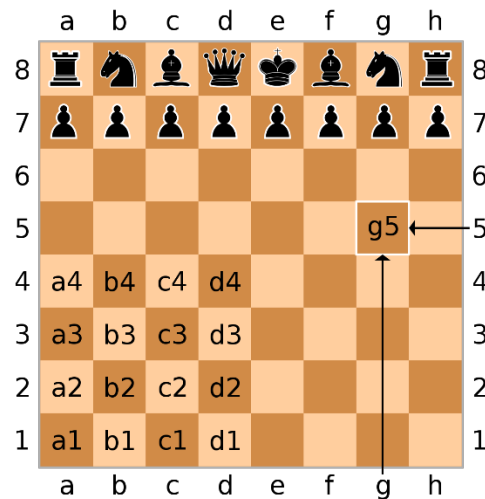
The queen is widely regarded as the most powerful piece on the board. It is essentially a combination of the rook and the bishop. The queen may move forward, backward, left, right, or diagonally as many squares as desired, until it meets a square already occupied by another piece. Finally, the king is the most important piece in the game. The king may move one square in any direction. Although the king is not overly powerful, it is vital because losing your king means losing the game.

The goal of chess is to capture the opposing player’s king. When a king is in danger of being captured by an opposing piece, that king is in “check”. If black makes a move to put the white king in check, white must, in its subsequent turn, move so that the king is no longer in check. This can be done by moving the king out of check, moving another piece to block the check, or capturing the piece that threatens the king. When a king is in check and no move can be made to get the king out of check, “checkmate” is declared and the game is over. Whichever player forces checkmate wins the game.

Notations are often used when writing about chess. Each piece is represented by a letter and each square on the board is represented by a combination of one letter and one number. These notations are demonstrated in the following table and image.

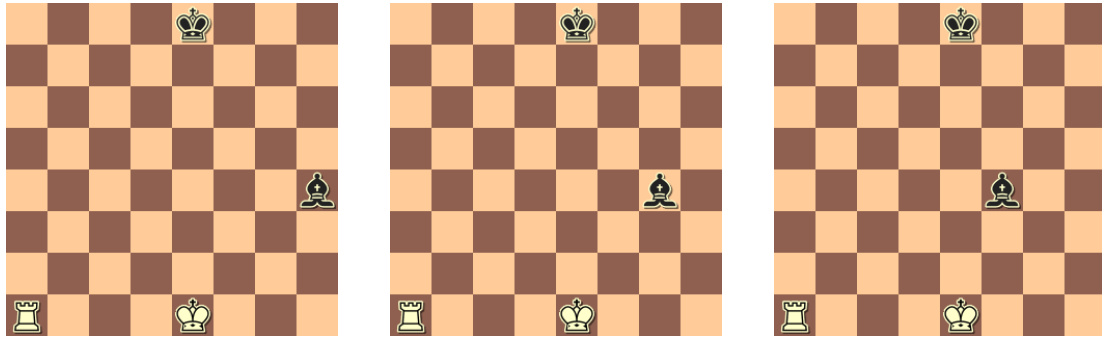


Piece	Notation	Symbol
King	K	
Queen	Q	
Rook	R	
Knight	N	
Bishop	B	
Pawn	P	

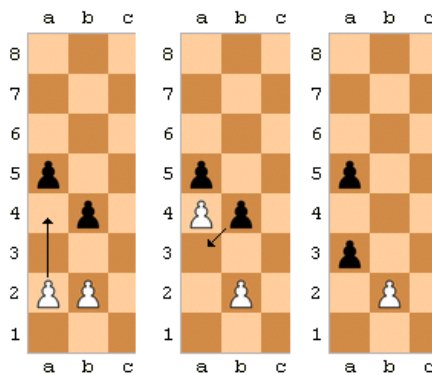


Chess contains a few special moves. The first and most prevalent special move is called castling. When a player castles, he essentially switches his king with one of his rooks. For example. If the white player has his king on e1 and a rook on h1 with no pieces in between the two, he may choose to castle. He would then move his king to g1 and his rook to f1, only using up one turn. If castling is done using the rook on a1, the king would move to c1 and the rook would move to d1. Castling is extremely useful when trying to protect the king. Many chess games will see at least one of the players castle at some point.

A player may castle only if neither his king nor his rook have moved from their starting position. A king cannot castle through check, out of check, or into check. So white would not be able to castle in any of the following situations.



Another special move, involving pawns, is called en passant. This much rarer move may only occur in specific circumstances. Let's say black moves a pawn to b4. If white moves his pawn two spaces to a4, black may then claim en passant, moving to a3 and capturing the white pawn on a4 in the process.



There is an abundant amount of unique strategies used by players around the world in their quest to capture the opponent's king. Some players come out with guns blazing in an all-out attack. Other players like to play it safe and wait for their opponent to make a mistake. With each piece having many different possible moves and the existence of countless strategies, it is obvious that a chess game can go a number of

different ways. There are in fact billions of different possible chess formations. The true extent of possible chess games will be explored later in this paper.

Does chess have an infinite amount of possible games? Does an unbeatable strategy exist? How far away is humanity from creating an unbeatable chess computer? These are some of the questions I will attempt to answer in this thesis.

This study will analyze the game of chess through the lens of statistics, probabilities, and combinatorics, which is the study of combinations and permutations of finite sets. The findings will hopefully be used to figure out how to use math to improve chess skills. The goal of this thesis is to ultimately present a way to understand the world's greatest board game via mathematics. Hopefully, the findings will answer the research question: Can math be used to perfect a player's chess skills?

## **Literature Review**

### *Zermelo's Work*

Ernst Zermelo was a German mathematician who received his doctorate from the University of Berlin in 1894 (O'Connor and Robertson). He pioneered research game theory and set theory. In 1913, he published an article titled, "On an Application of Set Theory to the Theory of the Game of Chess." In this article, Zermelo discusses two player games without chance moves where each player is trying to beat the other. He focuses on chess because it is the most well-known of these games. In a 1999 paper, Ulrich Schawlbe and Paul Walker discuss Zermelo's work and contribution to the discussion of chess solvability.

Zermelo presents two questions (Schwalbe and Walker 3). First, when is a player in a “winning” position and can this be defined in a mathematical sense? Second, if a player is in a winning position, can we determine the number of moves needed to force a win? To answer the first question, Zermelo states that there must exist a nonempty set that contains all sequences of moves that produce a victory for one player, say white, no matter how the opponent plays. If this set is empty, then the best white could do would be to force a draw. So Zermelo defined another set as the set containing all sequences of moves such that white can infinitely postpone a loss. However, a chess game cannot last forever. This is because a draw can be demanded if an exact position of the pieces is repeated three times. Therefore, this set must be empty. So white would only be able to postpone a loss for a finite amount of moves. This is equivalent to saying that black can force a victory. So theoretically, one player should be able to force a win or force a draw.

Now let's focus on the second question. If a player is in a winning position, can we determine the number of moves needed to force a win? Zermelo reasons that it would never take more moves than the amount of positions in the game. He proved this by contradiction: Assume that white can win in a number of moves greater than the possible positions. Then a winning position would have been repeated at least once. If white had of played the winning move the first time, he would have won in fewer moves than the amount of possible positions.

### *Countability*

For the very first move of the game, white has 20 options to choose from. White can move any of the eight pawns up one or two squares or he can move either knight up and to the left or up and to the right. Similarly, black has 20 moves to choose from for his

first turn. This leaves us with 400 (because  $20 \times 20 = 400$ ) possible variations after just the first two moves. These numbers keep getting larger as the game progresses. They beg the question, how many possible chess games are there?

In 1950, Claude E. Shannon wrote a paper titled “Programming a Computer for Playing Chess.” Shannon’s purpose of writing this paper was to describe how a computer could be programmed to analyze data and solve problems given by that data. He hoped this problem-solving technology could be used not just in chess, but in many areas such as telephone circuits, music, and language translation.

Shannon claims that it is feasible to play a perfect game or to build a machine to do so. From a given position this machine must consider every possible move, then evaluate all possible moves for the opponent, and so on until the end of the game. Since chess is not an infinite game, every game must end in either a win, loss, or draw. The machine would then work backward from the end to see if it can force a win with a particular set of moves (Shannon 4).

For this machine to work, it would have to be able to calculate every possible move. In chess, a move is defined by white taking its turn and black taking its turn. Each of these turns is called a ply. Shannon states that there are, on average, 30 legal plies that a player can make given a typical chess position. Then, a ply by white followed by a ply by black would give  $30 \times 30$  or roughly  $10^3$  possibilities. A standard game lasts about 40 moves (80 plies) before one of the players wins or resigns. If each move has  $10^3$  possibilities and a game lasts 40 moves, then there would be  $(10^3)^{40}$ , or  $10^{120}$ , possible games. This number is so massive that it is difficult to describe. According to Shannon,

even if a machine could calculate one variation per microsecond, it would take over  $10^{90}$  years to make the first move!

### *Solvability*

To better analyze the solvability of chess, we can look at games similar to chess that have been solved or nearly solved. Checkers is one such game. Both checkers and chess are games of perfect information. This means that at any point in the game, each player knows all the moves that were previously played. Also, in each game, turns are taken in an alternating fashion.

In a 1996 paper, “Solving the Game of Checkers,” Jonathan Schaeffer and Robert Lake ask the question, is it possible to program a computer to play checkers perfectly? If so, checkers would be a solved game. The paper states that there are three ways of solving a board game: publicly, practically, and provably. Publicly solving checkers means convincing the public that the game is solved. Practically solving means building a program powerful enough to consistently beat the best humans. Provably solving checkers means creating a perfect program that cannot lose. At the time this paper was written, checkers was in the process of being practically solved (Schaeffer and Lake 120).

In their paper, Schaeffer and Lake mainly focus on Chinook, a computer program created by Schaeffer and his team. This checkers playing program was the best in the world at the time. It was, however, not perfect. For the program to be unbeatable, it would require “a deeper search, an improving evaluation function, more endgame databases and a growing opening library” (Schaeffer and Lake 123). A program’s deep search helps it decide which positions to examine. Evaluation determines how favorable a position is. The endgame database stores perfect information for all positions with a

certain amount of pieces (or less) left on the board. In 1996, this was eight pieces or less. The opening library stores a vast amount of common openings to the game.

In 2007, Schaeffer announced that checkers was finally solved. From the starting position, if two players play a perfect game, making the optimal move on each turn, the result will be a draw. Schaeffer and his team proved this via their updated version of Chinook. This program utilizes a massive amount of computations that result in an explicit strategy that cannot lose (Schaeffer et al.). The program, playing as white or as black, would produce at least a draw against any opponent. This result was not a big surprise to the checkers community. Grandmasters had long assumed that two perfect players would play to a draw.

Chess is a much more complicated game than checkers. Each piece in checkers is essentially the same. A checkers piece can move forward diagonally, jump over opponent pieces, and promote to become a king on reaching the opposite side of the board. Chess contains six different types of pieces, most of which can move anywhere on the board. This begs the question, when, if ever, will chess be provably solved?

Since the 1950s, programmers have attempted to build machines that are unbeatable at chess. As computers became more and more competent, so did their chess playing abilities. Garry Kasparov, arguably the greatest chess player of all time, discusses his experiences with chess computers in a 2010 article, "The Chess Master and the Computer." In 1985, Kasparov played against 32 different computers at once. He walked from one machine to the next, making his move. These computers came from the top manufacturers in the world. Kasparov won every game with a perfect score of 32-0. This

was not a big surprise at the time. Most people expected the grandmaster to easily take down the machines (Kasparov).

In 1996, Kasparov faced off against IBM's supercomputer Deep Blue. He barely beat the machine. A year later, after IBM doubled Deep Blue's processing power, Kasparov lost the rematch. In 2003, he played a few matches against two commercially available programs. Both of these matches ended in a draw. Today, anyone can purchase a PC program that would defeat most grandmasters (Kasparov). By our definition, then, chess is practically solved.

Will chess ever be provably solved? Kasparov doesn't think so. "Chess is far too complex to be definitively solved with any technology we can conceive of today" (Kasparov). Unfortunately, many chess computing projects around the world lost funding and shut down after Kasparov's defeat in 1997. However, with the constant improvement of technology, we may someday witness a program powerful enough to provably solve the game.

## **Data Analysis**

### *Preliminaries*

Combinatorics is essentially the mathematics of counting. It often deals with questions that ask "how many?" (Mazur 1). In this paper, we will be exploring chess and all the different ways a chess game can play out. We will be asking a lot of "how many" questions, so a basic understanding of combinatorics will be necessary.

There are two basic counting principles that are vital to the mathematics of counting. The first of these, the Addition Principle, states:



Assume there are  $n_1$  ways for event  $E_1$  to occur,  $n_2$  ways for event  $E_2$  to occur, ..., and  $n_k$  ways for event  $E_k$  to occur. If these ways for the different events to occur are pairwise disjoint, then the number of ways for at least one of the events  $E_1, E_2, \dots, \text{ or } E_k$  to occur is given by:

$$n_1 + n_2 + \dots + n_k = \sum_{i=1}^k n_k$$

To demonstrate this, let's imagine two cities that exist a large distance away from each other. We will call the first city, A and the second city, B. There are four ways that one can drive from A to B. There are two ways one can fly from A to B. And there are three ways one can sail from A to B. So the number of ways that someone can drive, fly, or sail from A to B is  $4 + 2 + 3$ , which equals 9. In this example, driving from A to B is  $E_1$ , flying from A to B is  $E_2$ , and sailing from A to B is  $E_3$ . The Addition Principle might seem somewhat elementary but it is an imperative building block of combinatorics.

The next basic principle, called the Multiplication Principle, states: Assume that an event E can be decomposed into r ordered events  $E_1, E_2, \dots, E_r$ , and that there are  $n_1$  ways for the event  $E_1$  to occur,  $n_2$  ways for the event  $E_2$  to occur, ...,  $n_r$  ways for the event  $E_r$  to occur. Then the total number of ways for the event E to occur is given by:

$$n_1 \times n_2 \times \dots \times n_r = \prod_{i=1}^r n_i$$

To demonstrate this, we will again imagine cities. This time we will have cities W, X, Y, and Z. Our goal is to travel from W to Z. The only way to do this, is by first passing through X and then Y. There are two ways to get from W to X, three ways to get from X to Y, and 5 ways to get from Y to Z. So the total number of ways to get from W to Z is  $2 \times 3 \times 5$ , which equals 30.

The next two vital combinatorics operations are permutations and combinations.

A permutation is a way of ordering distinct objects from a given set. Let set  $A = \{a, b, c, d\}$ . Then finding the 3-permutations of  $A$  would mean finding all the ways we can arrange three elements from set  $A$ . Let  $n$  be the number of distinct objects in a set  $B$ .

Then the way of arranging any  $r$  objects of  $B$  would be given by:

$${}_n P_r = \frac{n!}{(n-r)!} \quad (\text{where } n! = n(n-1)(n-2)\dots(2)(1))$$

Combinations differ from permutations in that the order does not matter in a combination. A combination of a set  $Q$  is simply a subset of  $Q$ . Let  $Q = \{w, x, y, z\}$ . Finding the 3-combinations of  $Q$  would mean finding all the ways we can combine three elements from  $Q$ . Let  $n$  be the number of distinct objects in a set  $B$ . Then the way of combining any  $r$  objects of  $B$  would be given by:

$${}_n C_r = \frac{n!}{r!(n-r)!}$$

The concepts of permutations and combinations will be applied to chess later in this paper. Further details and applications of these formulas can be found in “Principles and Techniques of Combinatorics” by Chen Chuan-Chong and Koh Khee-Meng.

### *Probability and Statistics*

When observing the rules of chess, one might make some assumptions about the outcome of a game. For instance, one might say that white has a clear advantage since the white player moves first. Another assumption might be that some pieces have more value than others. In his paper, “Evaluation of Material Imbalances,” Larry Kaufman assigns a number value to each chess piece. He gives the pawn a value of 1 and rates the other pieces accordingly.

Piece	Value
Pawn	1
Rook	5
Knight	$3 \frac{1}{4}$
Bishop	$3 \frac{1}{4}$
Queen	$9 \frac{3}{4}$

By looking at the table above, we might assume that the queen is the most powerful piece in the game. We can test this assumption by observing a collection of games that have been played. Obviously, a bigger sample size yields a more accurate estimate. It would be very tedious to compile a list of games and analyze them one by one. Fortunately, we can use computer programming to do the job for us.

My hypothesis was that the first player to lose a queen will lose the game a majority of the time. To test this hypothesis, my mathematics professor and I wrote a program that analyzes chess games, tells which queen was taken first if any, and tells which player won the game. With this information, we were able to test my hypothesis. The details of this program will be explained in the methodology chapter.

For every game in which a queen was captured, there are four different possibilities:

1. The white queen was captured first and the winner was white (w\_white)
2. The white queen was captured first and the winner was black (w\_black)
3. The black queen was captured first and the winner was white (b\_white)
4. The black queen was captured first and the winner was black (b\_black)

The programs tallies up the amount of each of the four outcomes and we are left with our results. The data set consisted of 20,058 games. The results are given below:

- w\_white: 1926
- w\_black: 3535
- b\_white: 3822
- b\_black: 1917

This gives us a total of 11,200 games in which a queen was taken. First let us try to figure out if white has a major advantage over black. Out of these 11,200 games, white won 5748. If we take 5,748 divided by the total number of games, we get  $5,748/11,200 \approx 0.5132 = 51.32\%$ . So for this data set, white won just over 51% of the games. From this, we cannot say for sure that white has a clear advantage since black won nearly the same amount of games as white.

Now we will look at how losing a queen affects the chances of winning. Adding together w\_white and w\_black, we get the amount of games where the white queen was taken first: 5,461. Out of these games, white only won 1,926. So we divide 1,926 by 5,461 to get approximately 0.3527, or 35.27%. As we can see, white's probability of winning drops from 51.32% to 35.27% just from losing the queen.

Now let's look at the results when the black queen is taken first. By adding b\_white and b\_black, we get 5,739. Out of these games, black won 1,917 times. By dividing 1,917 by 5,739, we get approximately 0.3340 or 33.4%. So black's probability of winning drops from 48.68% to 33.4% from losing the queen.

From our analysis we can deduce that a player's odds of winning are dramatically reduced upon losing a queen. This analysis is fairly simple, but it demonstrates how math

can be used to influence chess strategy. Understanding how losing certain pieces affects the chances of winning can certainly help to improve a player's abilities.

### *Chess Puzzles*

An easy way to observe the workings of math in chess is to look at a few chess puzzles. The most famous of these is the 8-queens puzzle. The problem is stated as follows: can you place eight queens on a chessboard so that no queen is threatening any other queen? If so, how many ways can this be done?

Most people might raise the question of this problem's relevance. After all, it is highly unlikely that eight queens would exist in a single chess game. Each chess game begins with two queens. For a game to contain eight queens, six pawns must have reached the opposite edge of the board, with no queens being captured. It is true that the event of having eight queens is highly unlikely, but the purpose of the puzzle is to think critically and solve a problem.

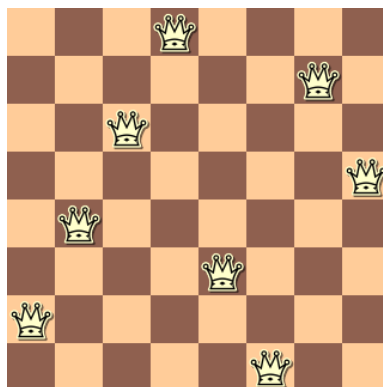
First we shall note that the maximum number of non-attacking queens that could exist on a chessboard is eight. That is, nine or more queens could not be placed on a chessboard so that none would threaten any other. Proving this is simple. When placing a queen on the board, the queen will be placed on the intersection of a row and a column. The next queen cannot be placed on the same row or the same column as the first queen. The third queen cannot be placed on a row or a column occupied by the first two, and so on. Once eight queens are placed, each of the eight columns will contain a queen. Since there are no remaining unoccupied columns, a ninth queen cannot be placed.

There are 64 individual squares on a chessboard. How many different ways can eight queens be placed on the board? If we place our queens one at a time, the first queen

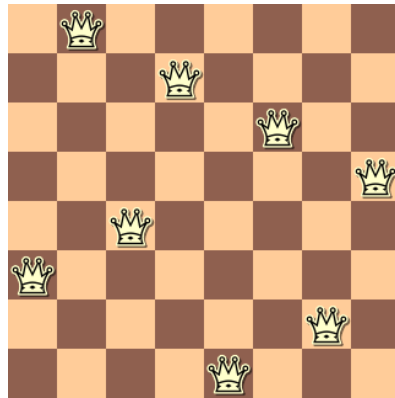
will have 64 options. The second queen will then have 63 options and so on. So the number of ways eight queens can be placed on the board would be  $64 \times 63 \times 62 \times 61 \times 60 \times 59 \times 58 \times 57$ . We note that swapping two queens will produce the same solution. Assume the first queen, say  $Q_1$ , is on square b2 and the second queen, say  $Q_2$ , is on square c5. If we move  $Q_1$  to c5 and move  $Q_2$  to b2, the overall configuration would still be the same. So to eliminate all the redundant combinations, we divide by  $8!$ . Therefore, the number of unique ways that eight queens can be placed on the board is  $(64 \times 63 \times 62 \times 61 \times 60 \times 59 \times 58 \times 57) / (8!)$ , which equals 4,426,165,368.

We can solve this problem alternatively through combinatorics. There are  $64!$  ways to arrange 64 objects (squares on the chess board). Of these objects, 56 are empty squares and can be swapped without changing the formation. Therefore we divide by  $56!$ . We also divide by  $8!$  to account for the eight interchangeable queens. Therefore the number of possible solutions is  $(64!) / (56!8!)$ , which equals 4,426,165,368.

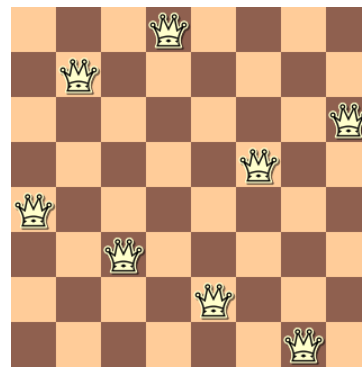
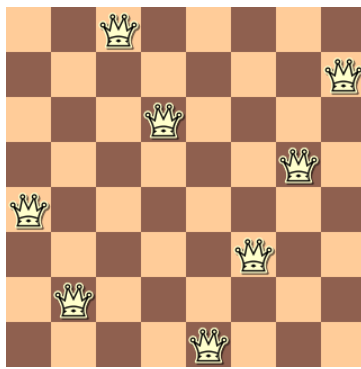
The following board is a solution to the 8-queens puzzle. Each queen has its own row and column. Also, no two queens exist in the same diagonal. We now see that there is at least one solution for the 8-queens puzzle:



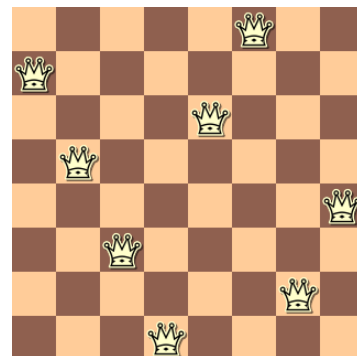
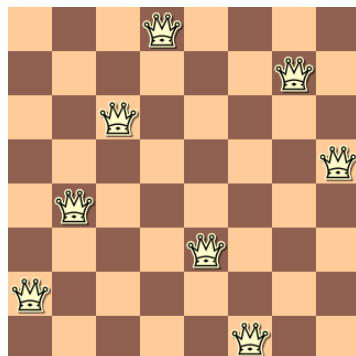
But is this the only solution? On examining this placement, we see that rotating the queens 90 degrees about the center of the board will produce a new solution.



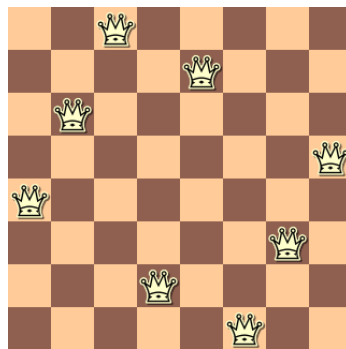
We can then rotate it 90 degrees two more times to get the following two solutions:



In addition to rotating, we can also reflect each solution across the center line (the imaginary horizontal line between rows 4 and 5). The following solutions are reflections of each other. Reflecting the solution on the left across the center line yields the solution on the right:



As we have seen, each solution can generate other related solutions through rotation and reflection. We will call each group of these related configurations solution families. So how many solution families exist? According to Guti'erez-Naranjo et al., there are twelve families of solutions (199). Every family except for one consists of eight related configurations (four rotations and four reflections). The following board belongs to the solution family that only contains four configurations. This is because the board has rotational symmetry. Rotating the board 180 degrees yields the same configuration.



This leaves us with a total of 92 solutions for the 8-queens puzzle out of the 4,426,165,368 possibilities. This means that 0.000002% of the possible combinations are solutions to the puzzle.

Let us now look at another puzzle of non-attacking pieces. This time we will work out the solution for ourselves. How many ways can you arrange eight rooks on a chessboard so that no rook threatens another? First we note that we cannot fit more than eight non-attacking rooks on a chessboard. After eight are placed, each row will contain a rook and any future rooks that are placed would be put in an already occupied row. Similar to the queens, there are 4,426,165,368 possible combinations of eight rooks on a chessboard.



We can turn this puzzle into a fairly elementary combinatorics problem. First, we assign each rook its own column. The rook we place in the first column will have eight options (each of the eight rows). The rook we place in the second column will have seven options for the seven untaken rows. This pattern continues until the eighth rook has only one option. This is essentially a problem of arranging a permutation of 8 objects. Therefore there are  $8!$ , or 40,320, solutions. This formula may be used for  $n$  rooks on an  $n \times n$  board. There will be  $n!$  ways to arrange  $n$  rooks on an  $n \times n$  chess board so that they are non-attacking.

### **Methodology**

I hypothesized that being the first player to lose the queen would dramatically reduce the chances of winning. In order to test this hypothesis, my professor and I wrote a program that analyzes a large data set of chess games. We found a data set consisting of 20,058 games at kaggle.com. These games were compiled from lichess.org, a free chess website used by players around the world.

The data set came in the form of a Microsoft Excel spreadsheet. The data included the players' ratings, the amount of turns, and even the type of opening deployed in each game. For the purposes of testing my hypothesis, I stripped away everything except the moves and the winner of each game. Each game is represented by the moves made and the piece that made each move. The notation "Nf6" means that a knight was moved to square f6. When a coordinate appears with no capitalized letter in front of it, it is implied that a pawn was moved to that coordinate. The "+" sign is used to represent a move that

puts the opposing king in check. The notations “O-O” and “O-O-O” are used to denote castling. The following example shows what a typical game looks like.

```
d4 d5 Nf3 Bf5 Nc3 Nf6 Bf4 Ng4 e3 Nc6 Be2 Qd7 O-O O-O-O Nb5 Nb4 Rc1
Nxa2 Ra1 Nb4 Nxa7+ Kb8 Nb5 Bxc2 Bxc7+ Kc8 Qd2 Qc6 Na7+ Kd7 Nxc6
bxc6 Bxd8 Kxd8 Qxb4 e5 Qb8+ Ke7 dxe5 Be4 Ra7+ Ke6 Qe8+ Kf5 Qxf7+ Nf6
Nh4+ Kg5 g3 Ng4 Qf4+ Kh5 Qxg4+ Kh6 Qf4+ g5 Qf6+ Bg6 Nxc6 Bg7 Qxg7#
```

The program analyzes the data set one game at a time; in fact one move at a time.

We first store each queen’s starting position. If a queen moves, her new position will be stored. If there is a move in which a piece is captured, an “x” will appear, followed by the square where the capture took place. If this square is the position of one of the queens, we know that queen was taken. We then store the winner of that particular game (in this study we simply ignore games where neither queen was captured).

It is important to note that these games were played online by everyday people. They were not grandmasters. It might be interesting in the future to evaluate the play of the world’s greatest players. Perhaps they are savvier about losing their queen.

In their 1973 paper, “Skill in Chess,” Herbert Simon and William Chase estimate roughly that the typical grandmaster has spent 10,000 to 50,000 hours staring at chess positions (Simon and Chase 402). A vital component in gaining expertise in chess is deliberate practice. Grandmasters spend hundreds of hours honing a particular skill (Charness et al. 152). Perhaps grandmasters spend a lot of time learning to play without a queen and are therefore better equipped to win without one. This would be worth looking into in future studies.

## Conclusion

Chess is a beautiful game. It is not overly complex and yet it takes years to master. It is played by people of all ages from countless different backgrounds. A myriad of different strategies may be employed in the game. In short, chess is one of the greatest board games ever invented. Chess is a game of perfect information. At any point in a game, each player knows every move that has taken place up to that point. Each player also has knowledge of every move their opponent can possibly make.

Through probability, statistics, and combinatorics, we have observed possibilities, probabilities, and solutions to chess puzzles. Using combinatorics, we demonstrated the vast amount of possible chess games. This amount is so massive that the most powerful computers in the world have still not been able to provably solve the game. We used probabilities and statistics to establish the importance of the queen. We also used math to analyze and solve some chess puzzles. All these things can be used to better understand chess and even influence playing strategy.

At this moment, there is no such thing as a perfect chess player, computer or human. There are just too many possible moves and strategies in chess for the existence of an all-knowing, all-powerful player. Even the world's strongest computers don't have the capacity to evaluate every possible move of every possible game. With more analysis and more computing power, however, we may one day be able to solve the game.

Unfortunately, until chess is provably solved, a player's skills cannot be perfected. While this may be true, chess skills can still be greatly improved upon. The best way to improve chess skill is to take the time to analyze and understand the game. As we have shown, math exists inherently in the game. A better understanding of the

mathematics of chess will lead to a better understanding of the game itself. This will ultimately lead a player to improve their chess playing abilities.

## Works Cited

- Adams, Jenny. *Power Play: The Literature and Politics of Chess in the Late Middle Ages*. University of Pennsylvania Press, 2006. *ProQuest Ebook Central*, <http://ebookcentral.proquest.com/lib/seu/detail.action?docID=3442075>.
- Charness, Neil, et al. "The Role of Deliberate Practice in Chess Expertise." *Applied Cognitive Psychology*, vol. 19, no. 2, Mar. 2005, pp. 151–65. *onlinelibrary.wiley.com*, doi:10.1002/acp.1106.
- Chess Diagram Setup*. <http://www.jinchess.com/chessboard/composer/>. Accessed 25 Mar. 2018.
- Chess Game Dataset (Lichess)*. <https://www.kaggle.com/datasnaek/chess>. Accessed 25 Mar. 2018.
- Chuan-Chong, Chen, and Koh Khee-Meng. *Principles and Techniques in Combinatorics*. World Scientific, 1992.
- Gutiérrez-Naranjo, Miguel Ángel, et al. "Solving the N-Queens Puzzle with P Systems." *Proceedings of the Seventh Brainstorming Week on Membrane Computing*, Vol. I, 199-210. Sevilla, ETS de Ingeniería Informática, 2 Feb. 2009.
- Kasparov, Garry. *The Chess Master and the Computer*. <http://www.nybooks.com/articles/2010/02/11/the-chess-master-and-the-computer/>. Accessed 25 Mar. 2018.
- Kaufman, Larry. "Evaluation of Material Imbalances." *Dan Heisman*, <http://www.danheisman.com/evaluation-of-material-imbances.html>. Accessed 25 Mar. 2018.
- "Let's Play Chess: Summary of the Moves of Chess." *Uschess.org*, United States Chess Federation, [www.uschess.org/docs/forms/LetsPlay.pdf](http://www.uschess.org/docs/forms/LetsPlay.pdf).

- Mazur, David R. *Combinatorics: A Guided Tour*. Mathematical Association of America, 2009. *ProQuest Ebook Central*, <http://ebookcentral.proquest.com/lib/seu/detail.action?docID=3330436>.
- O'Connor, J., and E. Robertson. *Zermelo Biography*. <http://www-groups.dcs.st-and.ac.uk/history/Biographies/Zermelo.html>. Accessed 25 Mar. 2018.
- Schaeffer, Jonathan, and Robert Lake. "Solving the Game of Checkers." *Games of No Chance*, vol. 29, 1996, pp. 119–133.
- Schaeffer, J., et al. "Checkers Is Solved." *Science*, vol. 317, no. 5844, Sept. 2007, pp. 1518–22. *CrossRef*, doi:10.1126/science.1144079.
- Schwalbe, Ulrich, and Paul Walker. "Zermelo and the Early History of Game Theory." *Games and Economic Behavior*, vol. 34, no. 1, 2001, pp. 123–137.
- Simon, Herbert, and William Chase. "Skill in Chess." *American Scientist*, vol. 61, no. 4, 1973, pp. 394–403.
- Shannon, Claude. "Programming a Computer for Playing Chess." *Philosophical Magazine*, vol. 41, no. 314, ser. 7, Mar. 1950. 7.
- "The Best Free, Adless Chess Server." *Lichess.org*, <https://lichess.org>. Accessed 25 Mar. 2018.