A few remarks on avoiding partial Latin squares

Jaromy Scott Kuhl* Tristan Denley†

November 20, 2007

Abstract

Let P be an $n \times n$ array of symbols. P is called avoidable if for every set of n symbols, there is an $n \times n$ Latin square L on these symbols so that corresponding cells in L and P differ. Due to recent work of Cavenagh and Öhman, we now know that all $n \times n$ partial Latin squares are avoidable for $n \geq 4$. Cavenagh and Öhman have shown that partial Latin squares of order $4m + 1$ for $m \geq 1$ [1] and $4m - 1$ for $m \geq 2$ [2] are avoidable. We give a short argument that includes all partial Latin squares of these orders of at least 9. We then ask the following question: given an $n \times n$ partial Latin square P with some specified structure, is there an $n \times n$ Latin square L of the same structure for which L avoids P? We answer this question in the context of generalized sudoku squares.

1 Introduction

A Latin square of order n is an $n \times n$ array of n symbols so that each symbol appears exactly once in each row and column. For a partial Latin square of order n we require that each symbol appear at most once in each row and column. We will always assume that the symbol set of a Latin square and a partial Latin square of order n is $[n] = \{1, 2, \ldots, n\}$, unless otherwise stated.

Let P be a partial Latin square of order n. For $1 \leq j, k, l \leq n$, symbol j appearing in cell (k, l) of P is denoted by $(j, k, l) \in P$. We say that P is avoidable if there is a Latin square L of order n so that corresponding cells in L and P differ. More formally, P is avoidable if there is a Latin square L of order n so that for each $(j, k, l) \in P$ we have $(i, k, l) \in L$ for $j \neq i$.

We begin our discussion by mentioning two trivial though useful properties. The first concerns partial Latin squares of order n that can be completed to Latin squares of order n and the second concerns isotopic partial Latin squares. An isotopism of P (or a Latin square) is any reordering of the rows or columns or any relabeling of the symbol set, or any combination of these. First we mention that if a partial Latin square P can be completed, then P can be avoided. That is, if we complete P and then permute the symbols so that no symbol remains

*University of West Florida
†University of Mississippi

fixed, then we have found a Latin square of order n avoiding P. Secondly, P is avoidable if and only if any isotopism of P is avoidable. Both properties will be used in the next section.

The question of which $n \times n$ arrays are avoidable was first asked by Häggkvist [4]. Chetwynd and Rhodes [3] conjectured that all partial Latin squares of order $n \geq 4$ are avoidable and proved the following.

Theorem 1 *For $m \geq 2$, all partial Latin squares of order $2m$ and $3m$ are avoidable.*

Additionally, Cavenagh [1] shows that all partial Latin squares of order $4m+1$ for $m \geq 1$ are avoidable. He does so separately for $m = 1$ and for $m \geq 3$ while leaving $m = 2$ to Theorem 1. Cavenagh and Öhman [2] confirm Chetwynd and Rhodes' conjecture by avoiding all partial Latin squares of order $4m - 1$ for $m \geq 2$ and they do so separately for $m = 2$, $m = 3$, and $m \geq 4$. Thus we have the following result.

Theorem 2 *Any partial Latin square P of order $n \geq 4$ is avoidable.*

The condition $n \geq 4$ is necessary since one can find unavoidable partial Latin squares of orders two and three. However, as Chetwynd and Rhodes have shown [3], there is only one unavoidable partial Latin square for each order up to isotopisms. The following partial Latin squares are unavoidable.

1	
	2

1	2	3
3		1
2	1	

The purpose of this paper is to give a more concise proof for avoiding odd ordered partial Latin squares of at least nine. This proof, which mimics Chetwynd and Rhodes's proof, does fail for orders of five and seven but avoiding these small orders are shown independently in [1] and [2]. Thus we have reduced Theorem 2 to two cases, namely the even and odd ordered partial Latin squares. In section 3 we ask the following less general question. Given a partial Latin square P with some kind of specified structure, can we avoid P with a Latin square containing the same kind of structure? We consider this question in the context of generalized sudoku squares.

2 Avoiding odd partial Latin squares

We will employ the following results for the proof of the main theorem. A proof of the first lemma can be found in [3], and a constructive proof can be found in [2].

Lemma 1 *For any $n \neq 3$ there exists a Latin square of order n which has symbol n on the leading diagonal and has the symbols of the last row appearing in the same order as the symbols in the last column.*

For the following lemma, we will assume that A is a partial Latin square of order 3 on the symbol set $\{a, b, c\}$ with cell $(3,3)$ empty. We will also assume that a does not appear in both the last row and last column of A and similarly b does not appear in both the last row and last column of A.

Lemma 2 *If c does not appear in the upper left 2×2 subsquare of A, then there is a Latin square of order 3 avoiding A so that $(c, 3, 3) \in L$ unless A contains the following partial Latin square.*

a	b	
b	a	

PROOF: Suppose the last column of A contains a. Relabeling the first two rows if need be, we consider $(a, 2, 3) \in A$. Then for a Latin square L of order 3 avoiding A with $(c, 3, 3) \in L$, we have $(b, 2, 3) \in L$ and $(a, 1, 3) \in L$. Clearly we are permitted either $(a, 2, 1) \in L$ or $(a, 2, 2) \in L$. If $(b, 3, i) \in A$ for $i \in [2]$, then we take $(a, 3, i)$, $(a, 2, j)$, $(b, 1, i)$, $(b, 3, j) \in L$ for $j \in [2]$ and $j \neq i$ and certainly L, filling in the empty cells with c, avoids A. Otherwise the last row of A does not contain a or b. We assume there is a diagonal of the upper left 2×2 subsquare of A that contains b, otherwise A does not contain b, except possibly in the last column, and one can easily find L. In this case, we consider c on the corresponding diagonal in L. There is only one Latin square with this arrangement of symbols and certainly L avoids A.

If the last column of A does not contain a but b, then we relabel the symbols and repeat the argument above. If the last column of A is empty but the last row nonempty, then we interchange columns with rows and repeat the argument above. Therefore we suppose that both the last row and the last column of A are empty. If this is the case, then finding L is trivial unless A contains the partial Latin square given in the statement of the Lemma for which L does not exist. □

Let P be a partial Latin square of order $2k+1$ for $k \geq 2$. In addition to this, we will view P as a $(k+1) \times (k+1)$ array denoted P' with cells denoted $\langle i, j \rangle$ where

$$\langle i, j \rangle = \begin{cases} 2 \times 2 \text{ subsquare} & \text{if } i < k+1 \text{ and } j < k+1 \\ 2 \times 1 \text{ sub rectangle} & \text{if } i < k+1 \text{ and } j = k+1 \\ 1 \times 2 \text{ sub rectangle} & \text{if } i = k+1 \text{ and } j < k+1 \\ 1 \times 1 \text{ subsquare} & \text{if } i = k+1 \text{ and } j = k+1 \end{cases}$$

We use the following terminology for the proof of our main theorem. If a partial Latin square A of order 2 contains a diagonal with distinct symbols, then we say that these symbols form a bad diagonal. If a partial Latin square A of order 3 contains the partial Latin square given in Lemma 2, then we say that the pair of symbols $\{a, b\}$ form a bad subsquare in A. The reason for denoting these as 'bad' is because if A contains a diagonal with symbols a and b, then there is no Latin square of order 2 on the symbol set $\{a, b\}$ avoiding A. Similarly, if A contains the partial Latin square given in Lemma 2 then there is no Latin square L of order 3 on the symbol set $\{a, b, c\}$ with $(c, 3, 3) \in L$ avoiding A.

Theorem 3 *Let P be a partial Latin square of order $2k+1$ for $k \geq 4$. Then P is avoidable.*

PROOF: We begin by reordering the rows of P so that in the last column symbol i appears in row i, if at all, and cell $(2k+1, 2k+1)$ is empty. Since there must be a column containing an empty cell, we take this column to be the last. If symbol i does not appear in the last column of P, then either we take cell $(i, 2k+1)$ to be empty or to contain symbol $2k+1$. Similarly, we reorder the first $2k$ columns of P so that symbol i appears in column $2k+1-i$ of the last row using empty cells and the symbol $2k+1$ as before. If possible, we allow symbol $2k+1$ to appear both in row i of the last column and column $2k+1-i$ of the last row for some i. Since $k \geq 4$, this arrangement guarantees that cells $\langle i, k+1\rangle$ and $\langle k+1, i\rangle$ do not share symbols provided k is even. If k is odd, it could be that cells $\langle \frac{k+1}{2}, k+1\rangle$ and $\langle k+1, \frac{k+1}{2}\rangle$ share symbols. Therefore we interchange columns $\frac{k+1}{2}$ and k of P' guaranteeing that cells $\langle i, k+1\rangle$ and $\langle k+1, i\rangle$ do not share symbols for all i.

Let Q be a Latin square of order $k+1$ described by Lemma 1 with symbols $X_1, X_2, \ldots, X_{k+1}$. For $i \in [k+1]$, we identify X_i with the set of 2×2 subsquares appearing in the corresponding cells of P'. For X_i in cells $(i, k+1)$ and $(k+1, i)$ of Q, we identify it with the 3×3 subsquare consisting of cells $\langle i, i\rangle$, $\langle i, k+1\rangle$, $\langle k+1, i\rangle$, and $\langle k+1, k+1\rangle$ of P'. Since there are $4k+1$ cells in the subsquares identified by X_{k+1} and since cell $\langle k+1, k+1\rangle$ is empty, there is a symbol appearing at most once in these cells which, without loss of generality, we will assume to be $2k+1$. We wish to arrange the first $2k$ columns of P so that symbol $2k+1$ does not appear in these cells while still guaranteeing that $\langle i, k+1\rangle$ and $\langle k+1, i\rangle$ do not share symbols. Let column j of P be the column intersecting the cells identified by X_{k+1} containing symbol $2k+1$. Due to the arrangement of P, there are at most seven columns of P for which we can not interchange with column j in order that $\langle i, k+1\rangle$ and $\langle k+1, i\rangle$ share no symbols and symbol $2k+1$ does not appear in the cells identified by X_{k+1}. Note that these seven columns include the final column of P and the column in the same pair as j. We will always have a column to choose outside these seven provided $k \geq 4$.

Let S be the set of symbols $[2k+1]$. Furthermore, let $S_1, S_2, \ldots, S_{k+1}$ be an ordered partition of S so that S_i consists of a pair of symbols for $i \in [k]$, and S_{k+1} consists of symbol $2k+1$. There are

$$\frac{(2k)!}{(2!)^k}$$

such ordered partitions of S.

For each appearance of symbol X_i not in the last row and column of Q, S_i is suitable for X_i if we can form a Latin square of order 2 on S_i avoiding the

2×2 subsquares identified by X_i. For X_i appearing in the last row and column of Q, we say that S_i is suitable for X_i if we can form a Latin square of order 3 on $S_i \cup \{2k+1\}$ avoiding the 3×3 subsquare identified by X_i so that $2k+1$ appears in cell $(3,3)$. If S_i is suitable for X_i for all i, then we can construct a partial $(2k+1) \times (2k+1)$ Latin square from Q avoiding P. Our goal then is to find an ordered partition of $S \setminus \{2k+1\}$ into k pairs of symbols such that S_i is suitable for X_i for all i.

Consider a 2×2 subsquare X of P' identified by X_i for $i \in [k]$ and X_i appearing outside the last row or column. If X contains a bad diagonal, then the two symbols giving the bad diagonal is an unsuitable pair of symbols for X_i. For X_i appearing in the last row and column of Q, if the 3×3 subsquare identified by X_i contains a bad subsquare, then by Lemma 2 this too gives an unsuitable pair of symbols for X_i. Let B denote the set that consists of all the unsuitable pairs of symbols for X_i for each i. There are

$$\frac{(2k-2)!}{(2!)^{k-1}}$$

partitions for which S_i has been fixed for some i. Therefore there are at most

$$|B|\frac{(2k-2)!}{(2!)^{k-1}}$$

partitions for which we must exclude. Then

$$\frac{(2k)!}{(2!)^k} - |B|\frac{(2k-2)!}{(2!)^{k-1}} > 0$$

if $|B| < 2k^2 - k$. Excluding the cells of the 3×3 subsquares identified by X_i in the last row and column of Q, $|B| \leq \frac{4k^2-4k}{2}$ since two cells make up a bad diagonal. Furthermore, of these excluded cells, we have at most k bad sub squares by Lemma 2. Therefore $|B| \leq 2k^2 - k$ and finally $|B| < 2k^2 - k$ since we can choose the empty cell $\langle k+1, k+1 \rangle$ so that there is another empty cell outside the last row and column of P, otherwise P can be completed. □

Note that for this argument to hold we must have $k \geq 4$. For $k = 1$, we have already mentioned that there are unavoidable partial Latin squares of order 3. For $k = 2$, we can not use Lemma 1 in the proof above. And for $k = 3$, we can not exclude symbol 7 from the cells of P identified by X_4 while guaranteeing that $\langle i, 4 \rangle$ and $\langle 4, i \rangle$ do not share symbols for $1 \leq i \leq 3$. It would seem then, as for Cavenagh and Öhman, we would need a separate argument for these two small orders. However, there is potentially a brief argument for these.

Also note that if P had order $2k$, then X_i would identify a 2×2 subsquare for all i and so we would need only to guarantee that P does not have too many bad diagonals. This can easily be accomplished by examining the number of empty cells and by arranging P so that the diagonals of some of the 2×2 subsquares contain the same symbol.

3 Avoiding partial sudoku squares with sudoku squares

A Latin square of order n^2 is a generalized sudoku square if the $n \times n$ square given by the rows $\{in+1, \ldots (i+1)n\}$ and the columns $\{jn+1, \ldots (j+1)n\}$ for $i, j \in [n-1]$ contains $[n^2]$. Theorem 2 clearly asserts that all partial generalized sudoku squares are avoidable. However, we wish to consider the following problem: if given a partial generalized sudoku square P of order n^2, is there a generalized sudoku square of order n^2 avoiding P? If P can be completed as a generalized sudoku square, then the answer to this question is clearly yes. Therefore we will assume that P can not be completed.

In answering this, we consider the following object which generalizes the Latin square. Define $A(ns, nr)$ to be an $ns \times nr$ array of n symbols so that each symbol appears r times in each row and each symbol appears s times in each column. Immediately one sees that for $s = r = 1$, $A(ns, nr)$ is a Latin square of order n.

Theorem 4 *Let P be a partial generalized sudoku square of order n^2. Then P can be avoided by a generalized sudoku square of order n^2.*

PROOF: Let T be an $A(n^2, n)$ array on symbols $\{t_1, \ldots, t_n\}$. We may assume that T is filled so that for each column of T, $\{t_1, \ldots, t_n\}$ appears in rows $\{in + 1, \ldots (i + 1)n\}$ of T for $0 \le i \le n$. Partition $[n^2]$ so that t_i is assigned to n symbols for each i and without loss of generality we may suppose that t_i is assigned to symbols $\{(i-1)n+1, \ldots, in\}$.

In column j of T, suppose that t_i appears in rows $\{\rho_{i_1}, \ldots, \rho_{i_n}\}$. In columns $\{jn+1, \ldots (j+1)n\}$ of P, consider the rows $\{\rho_{i_1}, \ldots, \rho_{i_n}\}$ of P and let D_i be the symbols appearing in these rows. These rows, in columns $\{jn+1, \ldots (j+1)n\}$, joined together form an $n \times n$ array of symbols. We form a partial Latin square P_i of order n by removing all symbols $s \in D_i$ for which $s \notin \{(i-1)n+1, \ldots, in\}$. Then, by Theorem 2, there is a Latin square of order n on the symbols $\{(i-1)n+1, \ldots, in\}$ avoiding P_i when $n \ge 4$. We then replace the kth appearance of t_i in column j of T with the kth row of this Latin square for $1 \le k \le n$. We do this for each i and then for each j, yielding a sudoku square which avoids P.

For $n = 2, 3$ we use the same argument, however, we need to be more careful on which symbols t_i receives since there are unavoidable partial Latin squares of order 2 and 3. There are

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

partitions in which $t_1, \ldots, t_n$ receive their symbols. Also, there can be at most n unavoidable partial Latin squares for which the symbols given to t_i must avoid. Since there can be at most

$$\frac{n(n^2-n)!}{n!^{n-1}}$$

unsuitable partitions with regard to t_i and since

$$\frac{n^2(n^2-n)!}{n!^{n-1}} < \frac{n^2!}{n!^n}$$

for $n = 2, 3$, there is a partition of $[n^2]$ for which the symbols given to t_i can avoid the corresponding partial Latin square in P. $\square$

References

[1] N.J. Cavenagh, Avoidable partial Latin squares of order $4m + 1$, *Ars Combinatoria*, to appear.

[2] N. J. Cavenagh and L-D Öhman, *Partial Latin squares are Avoidable*, Research Report in Mathematics **2** (2006) Dept. of mathematics and mathematical statistics, Umeå University, Umeå.

[3] A. G. Chetwynd and S. J. Rhodes, Avoiding partial Latin squares and intricacy, *Discrete Mathematics*, **177** (1997), 17-32.

[4] R. Häggkvist, A note on Latin squares with restricted support, *Discrete Mathematics*, **75** (1989), 253-254.