COMPLETING PARTIAL LATIN SQUARES WITH PRESCRIBED DIAGONALS

MARTIN GRÜTTMÜLLER

ABSTRACT. This paper deals with completion of partial latin squares $L = (l_{ij})$ of order n with k cyclically generated diagonals ($l_{i+t,j+t} = l_{ij} + t$ if l_{ij} is not empty; with calculations modulo n). There is special emphasis on cyclic completion. Here, we present results for $k = 2, \dots, 7$ and odd $n \leq 21$, and we describe the computational method used (hill-climbing). Noncyclic completion is investigated in the cases $k = 2, 3$ or 4 and $n \leq 21$.

1. INTRODUCTION

A *partial latin square* L of order n is an $n \times n$ array in which each cell is either empty or contains a single element from an n-set S of symbols, such that each element occurs at most once in each row and at most once in each column. If every cell is filled, then L is a *latin square*. If not explicitly stated differently, we assume the elements of S to be the integers $0, 1, \dots, n-1$ and also that the rows and columns are indexed by $0, 1, \dots, n-1$. All calculations are performed modulo n. A *partial transversal* of a partial latin square of order n is a set of filled cells, at most one in each row, at most one in each column, and such that no two of the cells contain the same symbol. A partial transversal with n cells is called a *transversal*. We refer the reader to [4, 5] for undefined terms as well as a general overview of latin squares.

Completion of partial latin squares has been investigated in a number of papers. Best known is Evans' conjecture [6] that an $n \times n$ partial latin square which has $n - 1$ cells occupied can always be completed to a latin square of order n. Based on work by Marica and Schönheim [11] and Lindner [10] this conjecture was proved to be true by Häggkvist [9] for $n \geq 1111$ and independently by Smetaniuk [13] and by Andersen and Hilton [2] for all n. We also like to mention a still unsolved conjecture stated by Daykin and Häggkvist [3] that says if L is a partial $n \times n$ latin square where each row, column and symbol is used at most un times (where u is some constant, e.g $u = \frac{1}{4}$), then L can be completed. Daykin and Häggkvist proved this for $n = 16k$ and $un = \sqrt{k}/32$ where $k \in \mathbb{N}$.

In connection with questions from design theory the following problem was posed by Alspach and Heinrich in 1990 [1]: Does there exist an $N(k)$ such that if k transversals of a partial latin square of order $n \geq N(k)$ are prescribed, the square can always be completed? For $k = 1$ one has $N(1) = 3$ since there exists an idempotent latin square for every order $n \neq 2$. Giving an uncompletable example Alspach and Heinrich also showed that $N(4) \geq 10$. A more specific

$$\begin{array}{ccccccc} 0 & 2 & * & * & 6 & 4 & * \\ * & 1 & 3 & * & * & 0 & 5 \\ 6 & * & 2 & 4 & * & * & 1 \\ 2 & 0 & * & 3 & 5 & * & * \\ * & 3 & 1 & * & 4 & 6 & * \\ * & * & 4 & 2 & * & 5 & 0 \\ 1 & * & * & 5 & 3 & * & 6 \end{array} \qquad \begin{array}{ccccccc} 0 & 2 & 5 & 1 & 6 & 4 & 3 \\ 4 & 1 & 3 & 6 & 2 & 0 & 5 \\ 6 & 5 & 2 & 4 & 0 & 3 & 1 \\ 2 & 0 & 6 & 3 & 5 & 1 & 4 \\ 5 & 3 & 1 & 0 & 4 & 6 & 2 \\ 3 & 6 & 4 & 2 & 1 & 5 & 0 \\ 1 & 4 & 0 & 5 & 3 & 2 & 6 \end{array}$$

FIGURE 1. A partial latin square of order 7 with 4 prescribed diagonals and its unique completion

version of their question was posed by Rees [12]: Does there exist an N such that if four cyclically generated transversals $l_{i+t,j+t} = l_{ij} + t$ of a partial latin square of order $n \geq N$ are prescribed, the square can always be completed to one which contains a further five transversals?

Figure 1 shows as an example a partial latin square with 4 cyclically generated diagonals together with its unique completion. Throughout this paper an asterisk indicates an empty cell. Notice that the remaining 3 diagonals in the completed latin square are also cyclically generated. Therefore, it seems natural to try a completion to a cyclically generated latin square. We show in Section 2 that such a cyclic completion is impossible if n is even. This suggests the following question. Does there exist a constant $C(k)$ such that if k cyclically generated diagonals $l_{i+t,j+t} = l_{ij} + t$ of a partial latin square of odd order $n \geq C(k)$ are prescribed, the square can always be cyclically completed? For example, an idempotent latin square $L = (l_{ij})$ can be constructed for all odd n by defining $l_{ij} = n - j + 2i$. Note that L is cyclically generated. This implies that $C(1) = 1$. To avoid trivial cases, we assume $C(k) \geq k$ if k is odd and $C(k) \geq k + 1$ if k is even.

In Sections 2 and 3, we prove lower bounds for $C(k)$ and $N(k)$. Moreover, we conjecture the bound for $C(k)$ to be sharp and provide strong evidence for this claim by some computer constructions. Using hill-climbing (Section 4) we show that every partial latin square L of order n with k cyclically generated diagonals is cyclically completable for all k in the range $2 \leq k \leq 7$ if n is odd and $3k - 1 \leq n \leq 21$. Furthermore, we show that L is (noncyclically) completable for $k = 2, 3$ or 4 and arbitrary n with $4k - 1 \leq n \leq 21$.

2. A Lower Bound for $C(k)$

In this section, we discuss some obvious necessary conditions and prove a lower bound for $C(k)$.

Clearly, a cyclically generated square of order n is completely described by its first row $(l_{0,0}, l_{0,1}, \ldots, l_{0,n-1})$ (if there is no chance of confusion we write $(l_0, l_1, \ldots, l_{n-1})$) and it is a latin square if and only if all elements l_i and all differences $l_i - i$ (modulo n) are mutually distinct. The later condition ensures that the elements in every column are pairwise different. It is easily

$$\begin{array}{ccccccc}
0 & 1 & 2 & \boxed{3} & 4 & 5 & 6 \\
1 & 2 & 3 & 4 & \boxed{5} & 6 & \\
\boxed{2} & 3 & 4 & 5 & 6 & & \\
3 & \boxed{4} & 5 & 6 & & & \\
4 & 5 & \boxed{6} & & & & \\
5 & 6 & & & & & \\
6 & & & & & &
\end{array}$$

FIGURE 2. A partial transversal in an LCUTLS of order 7

checked that there is no cyclically generated latin square of even order n since $\sum_{i=0}^{n-1} i \equiv \frac{n}{2} \mod n$ but $\sum_{i=0}^{n-1} l_i - i \equiv 0 \mod n$. A *proper partial row* is a row $(l_0, l_1, \ldots, l_{n-1})$ where some of the l_i are empty and all nonempty l_i and the corresponding differences are mutually distinct. Of course, a proper partial row with exactly k nonempty l_i corresponds to a partial latin square with k cyclically prescribed diagonals.

A partial latin square $L = (l_{ij})$ of order n with

$$l_{ij} = \begin{cases} i+j & \text{if } i+j < n, \text{ or} \\ \text{empty} & \text{otherwise} \end{cases}$$

is called a *left cyclic upper triangle latin square* (LCUTLS). We now give the following preliminary result.

Lemma 2.1. *Let L be an LCUTLS of order n. Then the number of cells in a partial transversal of L of maximum size is*

$$t(n) = \left\lfloor \frac{2n+1}{3} \right\rfloor .$$

Proof. We first show that for every n there is a partial transversal in L with $\lfloor \frac{2n+1}{3} \rfloor$ cells. Let $n \equiv 1 \mod 3$ and $T = \{(i + \frac{n-1}{3}, i) : i = 0, 1, \ldots, \frac{n-1}{3}\} \cup \{(i - \frac{n+2}{3}, i) : i = \frac{n+2}{3}, \ldots, \frac{2n-2}{3}\}$ be a set of $\frac{2n+1}{3}$ cells containing the elements $\frac{n-1}{3}, \frac{n-1}{3} + 2, \ldots, n-1, \frac{n+2}{3}, \frac{n+2}{3} + 2, \ldots, n-2$. Clearly all cells are from different rows and columns and all elements are pairwise distinct. Therefore, T is a partial transversal. See Figure 2 for an example with $n = 7$. Now, if $n \equiv 0$ or $2 \mod 3$ construct a partial transversal T' in an LCUTLS L' of order $n+1$ or $n+2$ with $\frac{2n+3}{3}$ or $\frac{2n+5}{3}$ cells. Then removing the back diagonal containing $n-1$ or the back diagonals containing $n-1$ and $n-2$ yields an LCUTLS L of order n. Removing cells in T' which are in the deleted back diagonals gives then a partial transversal in L with at least $\lfloor \frac{2n+1}{3} \rfloor$ cells.

It remains to show that there is no partial transversal T in an LCUTLS L of order n with more than $\lfloor \frac{2n+1}{3} \rfloor$ cells. We do so by counting edges in a bipartite graph $G = (A \cup (B \cup C), E)$ in two ways. Let t denote the number of cells in an arbitrary partial transversal T and assume $|A| = \frac{n(n+1)}{2}$, $|B| = t$ and $|C| = 3(n-t)$ where the vertices of A are labeled by the filled cells of L, the vertices of B are labeled by the cells of T, and the vertices in C are labeled by those rows, columns and back diagonals which do not contain a cell from T (note, two cells from the same back diagonal contain the same symbol implying that T contains at most one cell from each back diagonal). Two vertices from A and B are connected by an edge if and only if the corresponding cells occur together in a common row, column or back diagonal. Furthermore, two vertices from A and C are connected if and only if the corresponding cell occurs in the corresponding row, column or back diagonal. Obviously, a vertex in A has either degree 1 if it is labeled by a cell from T or degree 3 otherwise. Hence, $|E| = 3\frac{n(n+1)}{2} - 2t$. Every vertex in B has degree $2n-1$ and every vertex in C has degree at least 1. Moreover, vertices in C which correspond to distinct rows have different degrees. The same is true for distinct columns and back diagonals. Therefore, $|E| \geq t(2n-1) + 3\frac{(n-t)(n-t+1)}{2}$. This in turn implies $0 \geq t(t - \frac{2n+1}{3})$. Hence, $t \leq \frac{2n+1}{3}$. □

Theorem 2.2. *The following inequality holds for every $k > 2$:*

$$C(k) \geq 3k - 1.$$

Proof. Define k entries of a proper partial row R as follows: $l_{2i} = i$ for $i = 0, \ldots, \lfloor \frac{k-1}{2} \rfloor$, $l_{2i+1} = n - \lfloor \frac{k}{2} \rfloor + i$ for $i = 0, \ldots, \lfloor \frac{k-2}{2} \rfloor$ and l_i is empty for $i = k, \ldots, n-1$. (For example, when $k = 5$ and $n = 13$ we obtain $R = (0, 11, 1, 12, 2, *, *, *, *, *, *, *, *)$.) We prove that this partial row R cannot be cyclically completed if $n \leq 3k-2$ and k is odd, or $n \leq 3k-3$ and k is even.

Let k be odd. In this case we have not used indices $k, \ldots, n-1$, elements $\frac{k+1}{2}, \ldots, n-1-\frac{k-1}{2}$ and differences $1, \ldots, n-k$. We may represent possible relations between indices, differences and elements by an $(n-k) \times (n-k)$ array A whose rows are indexed by unused differences and whose columns are indexed by unused indices. An entry in this array contains the element corresponding to the row and column index if that element is unused or is empty otherwise, i.e. $a_{rc} = r + c$ if there is no l_i in R with $l_i = r + c$, or $a_{rc} = *$ otherwise. In Figure 3 the general case is exhibited. It is easy to see that R is cyclically completable if and only if the partial latin square A (on the set of symbols $S = \{\frac{k+1}{2}, \ldots, n-1-\frac{k-1}{2}\}$) has a transversal. Note that both the upper left triangle and the lower right triangle in A are equivalent to an LCUTLS of order $n - \frac{3k-1}{2} - 1$. If there is a transversal in A, then one of these triangles contains at least $\frac{n-k}{2}$ cells. Hence, Lemma 2.1 implies $\frac{2(n-\frac{3k-1}{2}-1)+1}{3} \geq \frac{n-k}{2}$. Thus, $n \geq 3k$.

If k is even a similar argument shows that one needs a partial transversal in an LCUTLS of order $n - \frac{3}{2}k - 1$ having $\frac{n-k-1}{2}$ cells. Consequently, $n \geq 3k-1$. □

We remark that in the case $k = 2$, not covered in the above discussion, we have only the trivial bound $C(2) \geq 3$ since the two possible proper partial rows of length 3 $(0, 2, *)$ and $(0, *, 1)$ are completable to $(0, 2, 1)$.

	k	$k+1$	$k+2$	$\dots$	$\frac{3k+1}{2}$	$\dots$	$n-\frac{k-1}{2}-2$	$\dots$	$n-1$
1	$k+1$	$k+2$	$k+3$	$\dots$		$\dots$	$n-\frac{k-1}{2}-1$	$*$	$*$
2	$k+2$	$k+3$				$\iddots$	$*$		$\vdots$
3	$k+3$								
$\vdots$	$\vdots$								$*$
$\frac{k+3}{2}$								$*$	$\frac{k+1}{2}$
$\vdots$	$\vdots$	$\iddots$			$*$			$\iddots$	$\vdots$
$n-\frac{3k-1}{2}-1$	$n-\frac{k-1}{2}-1$	$*$							
$\vdots$	$*$				$*$	$\iddots$			$\vdots$
$n-k$	$*$	$\dots$		$*$	$\frac{k+1}{2}$	$\dots$		$\dots$	$n-k-1$

FIGURE 3. Partial latin square A of order $(n-k)$ containing unused elements; indexed by unused differences (rows) and unused indices (columns)

3. A LOWER BOUND FOR $N(k)$

Similarly as in the previous section we obtain a lower bound for $N(k)$ by showing that a special type of partial latin squares with k prescribed transversals (we use again k cyclically generated diagonals for that purpose) is not completable.

Theorem 3.1. *The following inequality holds for every $k \geq 1$:*

$$N(k) \geq 4k-1.$$

Proof. Let L be a partial latin square of even order n with k cyclically generated diagonals where all elements in the first row (row index 0) are even and contained in a cell with even column index, take $l_{2i} = n-2i$ for $i = 0, 1, \dots, k-1$ as an example for $n = 4k-2$. Let L' be a completion of L. Consider the set A of cells in L' with even row index and odd column index. An arbitrary even element x occurs in k prescribed cells (with even row and column index). Thus, x occurs in at most $\frac{n}{2} - k$ cells of A. Therefore, there are at most $\frac{n}{2}(\frac{n}{2} - k)$ cells in A containing even elements. On the other hand consider the cells from A which are in a fixed column. At least k of these cells contain an even element since the column contains k odd elements in cells with odd row index. Hence, there are at least $\frac{n}{2}k$ cells in A containing even elements. Consequently, $\frac{n}{2}k \leq \frac{n}{2}(\frac{n}{2} - k)$. This implies $n \geq 4k$, but then we cannot complete the example from the beginning of the proof with $n = 4k-2$. $\square$

4. Computational Construction Methods

In this section, we describe the way in which a nonexhaustive search technique called hill-climbing was applied to construct cyclically generated latin squares with prescribed diagonals. Hill-climbing has been successfully applied to a variety of combinatorial problems, for background information see e.g. [7, 8]. A hill-climbing problem can be specified as a set Σ of feasible solutions, together with a cost $c(R)$ associated with each feasible solution $R \in \Sigma$. Here, let Σ be the set of proper partial rows and define $c(R)$ of a proper partial row R to be the number of empty cells in R. A cyclically generated latin square has no empty cells and, therefore, corresponds to a feasible solution R with minimum cost $c(R) = 0$. Starting with an initial solution R (the first row of the prescribed partial latin square) our hill-climbing algorithm works by transforming $R = (l_0, \dots, l_{n-1})$ into another feasible solution in which the cost either remains the same or decreases by one. We use two different transformations:

Transformation T_1

---

1. Choose an unused index i at random;
2. Choose an unused difference d at random;
3. Let $a := i + d \bmod n$;
4. **If** there is an l_j with $l_j = a$ **and** a is not prescribed **then**
   put $l_j := *$;
5. **If** a is not prescribed **then**
   put $l_i := a$;

Transformation T_2

---

1. Choose an unused index i at random;
2. Choose an unused element a at random;
3. Let $d := a - i \bmod n$;
4. **If** there is an l_j with $l_j - j \equiv d \mod n$ **and** l_j is not prescribed **then**
   put $l_j := *$;
5. **If** l_j is not prescribed **then**
   put $l_i := a$;

The hill-climbing algorithm is now given below. There are situations in which the construction gets stuck. Therefore, we use a threshold value $t_{\max}$ to specify the maximum number of consecutive cost-preserving transformations allowed before we abandon the algorithm and restart with a new random seed.

Hill-climbing algorithm to complete a proper partial row R

---

1. Mark all elements in the given proper partial row R as prescribed;
2. Let $t := 0$;
3. **While** $c(R) > 0$ **and** $t < t_{\max}$ **do**

3.1 Increase t by one;

3.2 choose $r = 1$ or 2 at random with equal probability;

3.3 perform T_r;

3.4 **If** $c(R)$ has been decreased by one in T_r **then**
put $t := 0$;

In order to complete all possible partial latin squares of order n with k cyclically prescribed diagonals one has to compute all proper partial rows of length n with exactly k nonempty cells. Two proper partial rows $R = (l_0, \ldots, l_{n-1})$ and $R' = (l'_0, \ldots, l'_{n-1})$ are called *isomorphic* if there are integers r, s such that $l_i = l'_{i+r} + s$ for $i = 0, \ldots, n-1$. It is easy to verify that if R and R' are isomorphic and R is cyclically completable, then also is R'. We avoid unnecessary work by considering only one *canonical* representative from each isomorphism class, for example the lexicographically smallest row (assume $* = \infty$).

We also used hill-climbing for the completion of partial latin squares of even order as very briefly explained in the sequel. Define Σ to be the set of all partial latin squares of order n. For $L \in \Sigma$ define the cost $c(L)$ to be the number of empty cells in L. Clearly, L is a latin square if and only if $c(L) = 0$. An admissible transformation consists of choosing an empty cell l_{ij} and filling it with an element x not used in row i (column j), i.e. $l_{ij} := x$. If there is a cell $l_{i'j}$ in column j (a cell $l_{ij'}$ in row i) containing x and x is not prescribed in cell $l_{i'j}$ ($l_{ij'}$), then put $l_{i'j} := *$ ($l_{ij'} := *$). If x is prescribed in cell $l_{i'j}$ ($l_{ij'}$), then put $l_{ij} := *$. Clearly, the cost of a partial latin squares during such a transformation remains unchanged or decreases by one. Again, it might be possible that at some point the algorithm makes no further progress (in particular with partial latin squares described in the proof of Theorem 3.1 and $n = 4k$), so that one needs to restart the construction.

5. Results and Conjectures

We constructed all canonical proper partial rows for given k and odd n with a backtracking algorithm (see Table 1 for the number of rows considered) and tried to complete these rows with the hill-climbing algorithm for proper partial rows. Using $t_{\max} = 1000$ and at most 5 restarts we made the following observation.

Proposition 5.1. *Let k be an integer in the range $2 \le k \le 7$. Every partial latin square of odd order n with $3k - 1 \le n \le 21$ and k cyclically generated diagonals can be completed to a cyclically generated latin square.*

(k,n)	(2,5)	(2,7)	(3,9)	(4,11)	(4,13)
NCPPR(k,n)	6	15	346	11,030	41,885

(k,n)	(5,15)	(6,17)	(6,19)	(7,21)
NCPPR(k,n)	2,172,003	133,073,720	582,669,528	47,765,113,158

TABLE 1. NCPPR(k,n) – Number of canonical proper partial rows of length n with k nonempty cells

n	4	6	8	10	12	14	16	18	20
NCPPR(2)	3	10	21	36	55	78	105	136	171
NUCPPR(2)	3	7	0	0	0	0	0	0	0
NCPPR(3)	-	34	182	600	1504	3172	5950	10246	16530
NUCPPR(3)	-	34	30	10	0	0	0	0	0
NCPPR(4)	-	34	674	4972	22300	74110	201614	475384	1006872
NUCPPR(4)	-	34	590	1396	291	181	0	0	0

TABLE 2. NCPPR(k) – Number of canonical proper partial rows of length n with k nonempty cells; NUCPPR(k) – Number of uncompletable canonical proper partial rows of length n with k nonempty cells

We believe that this result provides significant evidence in support of the following conjecture.

Conjecture 5.2. *Every partial latin square of odd order n with k cyclically generated diagonals can be cyclically completed if $n \geq 3k-1$. This means $C(k) = 3k-1$.*

In addition, we constructed all canonical proper partial rows for even n (see Table 2) and the corresponding partial latin squares. These partial latin squares were tried to complete with the second hill-climbing approach described. Choosing $t_{\max} = 5000$ and at most 100000 restarts we proved the following result.

Proposition 5.3. *Let $k = 2, 3$ or 4. Every partial latin square of order n with $4k-1 \leq n \leq 21$ and k cyclically generated diagonals can be completed to a latin square.*

In view of the foregoing, we are prompted to pose the following.

Conjecture 5.4. *Every partial latin square of order n with k cyclically generated diagonals can be completed if $n \geq 4k-1$.*

References

[1] B. Alspach and K. Heinrich. Matching designs. *Australasian J. of Combin.*, **2**:39–55, 1990.

[2] L.D. Anderson and A.J.W. Hilton. Thank Evans! *Proc. London Math. Soc.*, **47**:507–522, 1983.

[3] D.E. Daykin and R. Häggkvist. Completion of sparse partial latin squares. In *Graph Theory and Combinatorics, Proc. Conf. Hon. P. Erdős, Cambridge 1983*, pages 127–132. Academic Press, London, 1984.

[4] J. Dénes and A.D. Keedwell. *Latin Squares and Their Applications.* English University Press, London, 1974.

[5] J. Dénes and A.D. Keedwell, editors. *Latin Squares: New Developments in the Theory and Applications.* Annals of Discrete Math. **46**. North-Holland, Amsterdam, 1991.

[6] T. Evans. Embedding incomplete latin squares. *Amer. Math. Monthly*, **67**:959–961, 1960.

[7] P.B. Gibbons. Computational methods in design theory. In C.J. Colbourn and J.H. Dinitz, editors, *The CRC Handbook of Combinatorial Designs*, pages 718–740. CRC Press, 1996.

[8] P.B. Gibbons and R. Mathon. The use of hill-climbing to construct orthogonal Steiner triple systems. *J. of Combin. Designs*, **1**:27–50, 1993.

[9] R. Häggkvist. A solution of the Evans conjecture for latin squares of large size. In A. Hajnal and V.T. Sos, editors, *Combinatorics, Coll. Math. Soc. Janos Bölyai*, volume **18**, pages 495–514. North Holland, Amsterdam, 1978.

[10] C.C. Lindner. On completing latin rectangles. *Canad. Math. Bull.*, **13**:65–68, 1970.

[11] J. Marica and J. Schönheim. Incomplete diagonals of latin squares. *Canad. Math. Bull.*, **12**:235, 1969.

[12] R.S. Rees. Mandatory representation designs MR$(3, k; v)$ with k even and v odd. *J. Stat. Plann. Inference*, **86**:567–594, 2000.

[13] B. Smetaniuk. A new construction on latin squares I. A proof of the Evans conjecture. *Ars Combinatoria*, **11**:155–172, 1981.

Department of Mathematics, University of Rostock, 18051 Rostock, Germany

E-mail address: `m.gruettmueller@mathematik.uni-rostock.de`