Latin Transversals in Long Rectangular Arrays *†

S. Akbari, O. Etesami, H. Mahini, M. Mahmoody, A. Sharifi

Abstract

In this paper it is shown that every $m \times n$ array in which each symbol appears at most $(mn-1)/(m-1)$ times has a latin transversal, when n is large enough in comparison to m.

Introduction

An $m \times n$ *array* is a table of m rows and n columns and therefore mn cells, where each cell contains exactly one symbol. A *partial transversal* in an array is a set of cells in which no two cells are in the same row or column. A *transversal* in an $m \times n$ array is a partial transversal of size $\min(m, n)$. A *partial latin transversal* is a partial transversal in which no two cells contain the same symbol. A *latin transversal* is a partial latin transversal of size $\min(m, n)$. We define $L(m, n)$ as the largest integer z such that every $m \times n$ array in which no symbol appears more than z times has at least one latin transversal. A *row latin rectangle* is an array in which no symbol occurs more than once in any row. A *column latin rectangle* is defined in a similar way. A *latin square* is an $n \times n$ row-column latin rectangle that contains exactly n different symbols. Ryser [7] conjectured that every latin square of odd order has a latin transversal. Moreover Brualdi [5] conjectured that every latin square of order n has a partial latin transversal of size at least $n-1$. These conjectures have remained unsettled.

The problems on the existence of large partial latin transversals in latin squares and rectangular arrays are among the most beautiful problems in com-

* *Key Words*: Latin transversal, rectangular array.

† 2000 *Mathematics Subject Classification*: 05B15, 05D15.

binatorics. There are many theorems and conjectures in this area. The lower bounds of $n-\sqrt{n}$ [1, p.256] and $n-5.53(\log n)^2$ [8] for the size of the largest partial latin transversal in latin squares of order n are well-known. Drisko [6] proved that if $n \geq 2m-1$, then any $m \times n$ column latin rectangle has a latin transversal. Stein [10] showed that $L(m,n) \leq n-1$ for $m \leq n \leq 2m-2$ by a simple construction and conjectured that $L(n-1,n) = n-1$. Clearly, if this conjecture is true, then Brualdi's Conjecture is also true. A result due to Hall [3] supports this conjecture: Any $n-1 \times n$ row latin array constructed from $n-1$ (not necessarily distinct) rows of the group table of an abelian group of order n, has a latin transversal. Stein [9] showed that in an $n \times n$ array where each symbol appears exactly n times there is a partial latin transversal with length at least approximately $(0.63)n$. Also Erdős and Spencer [2] proved that $L(n,n) \leq (n-1)/16$.

Result

In [10] it is shown that $L(m,n) \leq (mn-1)/(m-1)$. To see this by contrary suppose that $L(m,n) > (mn-1)/(m-1)$. It implies that $(m-1)L(m,n) \geq mn$, and we could assign $m-1$ different symbols to the cells of A such that each symbol appears at most $L(m,n)$ times. Obviously, $m-1$ symbols is not sufficient for a latin transversal, showing that $L(m,n) \leq (mn-1)/(m-1)$. It is shown that $L(3,n) = \lfloor (3n-1)/2 \rfloor$, for $n \geq 5$, see [10].

In the following we will give a theorem which shows when n is large enough in comparison to m, then the above upper bound on $L(m,n)$ is tight.

Theorem. *If $m \geq 2$ and $n \geq 2m^3-6m^2+6m-1$, then $L(m,n) = \lfloor \frac{mn-1}{m-1} \rfloor$.*

Proof. Define $f(m) = 2m^3-6m^2+6m-1$ and $g(m,n) = \lfloor (mn-1)/(m-1) \rfloor$. By the above discussion, we have $L(m,n) \leq g(m,n)$. Hence we just need to show that for $n \geq f(m)$, if each symbol appears at most $g(m,n)$ times in an $m \times n$ array A, then A has a latin transversal. We will prove this by applying induction on m. For $m=2$, it is clear that $L(2,n) = 2n-1$ when $n \geq 3$. We may thus assume $m > 2$.

For $1 \leq i \leq m$ and $1 \leq j \leq n$, we use (i,j) to denote the cell at the intersection of row i and column j, and we refer to the symbol contained in that cell by $A(i, j)$. Notice the distinction between cells (which are positions) and symbols (which are values assigned to the positions). Without loss of generality, we assume that symbols are positive integers. Consider the following three *primary operations* that can be applied to array A, i) Interchanging two rows; ii) Interchanging two columns; iii) Permutation on symbols. Each of the above operations preserves the existence of latin transversals. We may hence apply each of them wherever needed in the proof, and pretend that A has not been changed during the proof.

Let B be the $(m-1) \times n$ array consisting of the first $m-1$ rows of A. We have $n \geq f(m) \geq f(m-1)$ for $m \geq 2$, and each symbol appears at most $g(m, n) \leq g(m-1, n)$ times in B. Thus, by the induction hypothesis, B has a latin transversal T of size $m-1$. Without loss of generality and by using the primary operations, we may assume that T is on the main diagonal of B, and $A(i, i) = i$ for $1 \leq i \leq m-1$. Let $S = \{(m, j) \,|\, m \leq j \leq n\}$. If a cell in S has a symbol greater than $m-1$, then adding that cell to T creates a latin transversal, and we are done. Assume that k is the number of distinct symbols in S. Using the primary operations, without loss of generality we can assume that all symbols of S are less than or equal to k, $k \leq m-1$, and $A(m, m+i-1) = i$ for $1 \leq i \leq k$.

We will now show that by a suitable sequence of primary operations, we can find sets $T_1, T_2, \ldots, T_{m-1}$, each consisting of $m-1$ cells, such that for every $1 \leq i \leq m-1$, the following conditions hold:

1. $A(i, i) = i$.
2. T_i represents a partial latin transversal.
3. Each row, except row i, has a cell in T_i.
4. For every $i < j < m$, cell (j, j) is included in T_i.
5. All of the symbols $1, 2, \ldots, m-1$ appear in T_i.

We start by defining $T_i = \{(j, j) \,|\, 1 \leq j \leq m-1, j \neq i\} \cup \{(m, m+i-1)\}$ for $1 \leq i \leq k$. It is easy to verify that $T_1, T_2, \ldots, T_k$ satisfy Conditions 1–5. Now suppose that we have constructed $T_1, T_2, \ldots, T_p$ for $p \geq k$. We will show how to construct T_{p+1}, when $p < m-1$.

Let $X = \{(i, j) \mid 1 \leq i \leq p \text{ and } T_i \text{ has no cell in column } j\}$. We claim that at least one symbol of X is greater than p. If this does not hold, then all symbols in X and S are less than or equal to p. We know that $|X| = p(n - m + 1)$ and $|S| = n - m + 1$. Hence, one symbol appears at least $(n - m + 1)(p + 1)/p \geq (n - m + 1)(m - 1)/(m - 2)$ times in A. An elementary calculation shows that $f(m) > m^3 - 3m^2 + 2m + 1$, which implies $(n - m + 1)(m - 1)/(m - 2) > g(m, n)$. But this requires that a symbol appears more than $g(m, n)$ times in A, which is a contradiction. Consequently, there is $(r, s) \in X$ such that $A(r, s) = t > p$.

None of the cells of T_r is in row r. Besides, T_r has no cell from column s since $(r, s) \in X$. It follows that if $t \geq m$, then $T_r \cup \{(r, s)\}$ represents a latin transversal, and we are done. Thus one may assume that $t < m$. We will show how to make $A(r, s) = p + 1$ in case $t \neq p + 1$: Cells (t, t) and $(p + 1, p + 1)$ have rows and columns different from (r, s). Therefore, one way to ensure that $A(r, s) = p + 1$ is by first interchanging symbols t and $p + 1$, next interchanging rows t and $p + 1$, and finally interchanging columns t and $p + 1$. To see why this works, we note that the last two primary operations (i.e. interchanges of rows and columns t and $p + 1$) swap the values of $A(t, t)$ and $A(p + 1, p + 1)$. Furthermore the above primary operations preserve Conditions 1–5. This is because $T_1, \ldots, T_p$ all have the entries $A(i, i) = i$ for $p < i < m$, and the above operations only act on rows, columns, and symbols $p + 1, \ldots, m - 1$. Now that we have $A(r, s) = p + 1$, it is not hard to check that $T_{p+1} = T_r \cup \{(r, s)\} \backslash \{(p + 1, p + 1)\}$ does not violate any of the Conditions 1–5. By this method, all $T_1, \ldots, T_{m-1}$ are constructed.

Assume that $T_1, \ldots, T_k$ are as we introduced before. Let $C_i = \{j \mid 1 \leq j < m \text{ or column } j \text{ intersects } T_1 \cup \cdots \cup T_i\}$. By the above construction for $T_1, \ldots, T_{m-1}$, we have $|C_k| = m - 1 + k$. Furthermore we have $|C_{i+1}| \leq |C_i| + 1$ for $k \leq i < m - 1$, since T_{i+1} does not introduce more than one new cell (see definition of T_{p+1}). Hence we have $|C_{m-1}| \leq (m - 1 + k) + (m - 1 - k) = 2m - 2$. Let Q be the set of all cells residing in the columns of C_{m-1}, and let Q' be the complement of Q relative to A. Since each symbol appears at most $g(m, n)$ times, one of the following two cases always happens:

Case 1. There exists $(x, y) \in Q'$ such that $A(x, y) \geq m$. It is clear that $x \neq m$ and since T_x does not intersect column y, $T_x \cup \{(x, y)\}$ is a latin transversal.

Case 2. All symbols appearing in Q' are less than m, and there exists $(x, y) \in Q$ such that $A(x, y) \geq m$. Let Z be an $(m-1) \times (n-2m+2)$ array that is obtained from Q' by removing row x and $2m - 2 - |C_{m-1}|$ arbitrary columns. To show that Z has a latin transversal by induction on m, we check that:

1. There are $n - 2m + 2 \geq f(m) - 2m + 2 \geq f(m - 1)$ columns in Z.

2. Every symbol appears at most $g(m - 1, n - 2m + 2)$ times in Z, because $n \geq f(m)$ implies $g(m, n) \leq g(m - 1, n - 2m + 2)$.

Now since all symbols of Z are less than m, the latin transversal of Z can be extended to a latin transversal for A by adding (x, y) and the proof is complete. □

The above theorem shows that $L(3, n) = \lfloor (3n-1)/2 \rfloor$, for $n \geq 17$ and $L(4, n) = \lfloor (4n - 1)/3 \rfloor$, for $n \geq 55$.

Acknowledgement. The authors are indebted to the research council of Sharif University of Technology for support.

References

[1] R. A. Brualdi and H. J. Ryser, " Combinatorial Matrix Theory ", Cambridge University Press, 1992.

[2] P. Erdős and J. Spencer, Lopsided Lovasz Local Lemma and latin transversals, *Discrete Applied Math.* **30** (1991) 151-154.

[3] M. Hall, Jr., A Combinatorial problem on groups, *Proc. Amer. Math. Soc.* **3** (1952) 584-587.

[4] J. Dénes and A. D. Keedwell, " Latin Squares and Their Applications ", Academic Press, New York, 1974.

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

[6] A. A. Drisko, Transversals in Row-Latin Rectangles, *J. Combin, Theory Ser. A,* **84** (1998), No. 2, 181–195.

[7] H. J. Ryser, Neuere Probleme der Kombinatorik, *in* " Vortrage uber Kombinatorik Oberwolfach ", Mathematiches Forschungsinstitut Oberwolfach. pp, 24-29, Juli 1967.

[8] P. W. Shor, A lower bound for the length of a partial transversal in a latin square. *J. Combin. Theory Ser. A* **33** (1982), 1-8.

[9] S. K. Stein, Transversals of latin squares and their generalizations, *Pac. J. Math.* **59** (1975), 567-575.

[10] S. K. Stein, Latin transversals of rectangular arrays, arXiv:math. CO/0107066 v3, (2001).

SAIEED AKBARI `s_akbari@sina.sharif.ac.ir`
Department of Mathematical Sciences
Sharif University of Technology
P. O. Box 11365-9415, Tehran, Iran.
OMID ETESAMI `etesami@ce.sharif.edu`
HAMID MAHINI `mahini@ce.sharif.edu`
MOHAMMAD MAHMOODY `mahmoody@ce.sharif.edu`
ALI SHARIFI `asharifi@ce.sharif.edu`
Department of Computer Engineering
Sharif University of Technology
Tehran, Iran.