PACIFIC JOURNAL OF MATHEMATICS
Vol. 59, No. 2, 1975

TRANSVERSALS OF LATIN SQUARES AND THEIR GENERALIZATIONS

S. K. STEIN

The main theme in this paper is the existence of a transversal with many distinct elements in an array more general than a latin square.

A transversal of a latin square of order n is any set of n cells such that no two come from the same column or same row. There has been a good deal of effort spent on establishing the existence of a transversal that has many distinct elements, e.g. [4, 5]. A close inspection of the argument in [5] reveals that the results there apply in a context far more general than that explicitly considered. Indeed, the assumptions that there are no duplications in a row or column can in some cases be dropped.

A variety of conjectures conclude the paper.

1. Definitions. An *n-square* is an n by n array of n^2 cells in each of which one of the symbols $1, 2, 3, \cdots$ appears. An n-square in which each symbol from 1 to n appears n times is called an equi-n-square.

If $m < n$, an (m, n)-rectangle is an m by n array of mn cells in each of which one of the symbols $1, 2, 3, \cdots$ appears. There are m rows and n columns.

A *transversal* of an n-square or an (m, n)-rectangle is a set of cells, one from each row and no two from the same column. A *partial transversal* is a subset of a transversal. A transversal is *latin* if no two cells have the same symbol. Since a latin transversal need not contain all the symbols in the array, we do not use the traditional term, "complete". A row (or column) of an n-square is *latin* if no two of its cells contain the same symbol.

Note that a usual latin square can be described as an equi-n-square for which each row and each column is latin. Observe that a latin square is an equi-n-square.

2. Survey of results. Ryser in [10] conjectured that a latin square of odd order n has a latin transversal. Koksma [4] proved that a latin square of order n has a transversal with at least $(2n + 1)/3$ distinct symbols. Lindner and Perry, in a mimeographed publication [5], proved that the average number of distinct symbols in transversals of a latin n-square (taken over all transversals) is precisely

$$n\left(1-\frac{1}{2!}+\frac{1}{3!}-\cdots\pm\frac{1}{n!}\right).$$

From this it follows that there is a transversal with at least $[(1-1/e)n] = .63n$ elements. Because Koksma's result is stronger, [5] was not formally published.

This paper utilizes the technique of Lindner and Perry, which might be called "existence by averaging", to establish, for instance, that an equi-n-square has a transversal with a least $[(1-1/e)n]$ distinct symbols (Corollary 3.3). Koksma's technique, on the other hand, using all his assumptions, does not seem to be easily generalized.

Bruck proved that the Cayley table of a group of odd order has a transversal (namely the main diagonal). This follows from the fact that such a group is the union of cyclic groups of odd order and hence every element is the square of some element. Paige [7] proved that any finite abelian group that is not of the form $C(2^n) \times H$, where $C(2^n)$ is the cyclic group of order 2^n, $n \geqq 1$, and H has odd order, possesses a latin transversal. In [3] Hall generalized this result.

3 Transversals of n-squares.

For a subset X of the n^2 cells of an n-square, let $t(X)$ denote the number of transversals that meet X. This number, examined in the context of determinants of matrices with 0-entries in X, has been the subject of some study (see Netto ([6], p. 73)). In the case where X is itself a transversal or a subset of a transversal a formula for $t(X)$ is known (see [6], [9]). It is given in the following lemma, which is another version of the "hatcheck problem".

LEMMA 3.1. *Let X be a set of q cells in an n-square such that no two lie in the same column or in the same row. Then*

$$t(X) = n!\left(\frac{q}{n} - \binom{q}{2}\frac{1}{n(n-1)} + \binom{q}{3}\frac{1}{n(n-1)(n-2)} - \cdots \pm \binom{q}{q}\frac{1}{n(n-1)\cdots(n-q+1)}\right).$$

The next lemma implies that a set that is not a partial transversal meets at least as many transversals as does a partial transversal of the same cardinality.

LEMMA 3.2. *Let X be a set consisting of q cells, $q \leqq n$, in an n-square. Then $t(X) \geqq t(Z)$, where Z is a set of q cells in an n-*

square that form a partial transversal.

Proof. Assume that X has at least two cells in the same row. (A similar argument applies if some column contains at least two cells of X.) Let c be a cell of X in the row mentioned. Let Y be the set of q cells obtained from X by deleting cell c and adjoining a cell c' in a row not meeting X, but in the same column as c. Let X' be the set of cells in X that are not in the row containing c. Let $X'' = X - \{c\}$. Thus $X'' \supset X'$.

Now, $t(X)$ equals:

the number of transversals that meet X'', but not c or c'
$+\ t(\{c\})$
$+$ the number of transversals that meet X'' and also c'.

On the other hand, $t(Y)$ equals:

the number of transversals that meet X'', but not c or c'
$+\ t(\{c'\})$
$+$ the number of transversals that meet both X'' and c.

To compare these two sums, observe first that the first terms of each are the same and that $t(\{c\}) = t(\{c'\})$. Also,

the number of transversals that meet both X'' and c
equals
the number of transversals that meet both X' and c,
which equals
the number of transversals that meet both X' and c'.

Since $X'' \supset X'$, it follows by comparsion of the third terms of the sums for $t(X)$ and $t(Y)$ that $t(X) \geqq t(Y)$. Repeated application of this argument, at most $q - 1$ times, establishes the lemma.

The following theorem and its corollary generalizes the result of Lindner and Perry from latin n-squares to n-squares.

THEOREM 3.2. *In an n-square in which each symbol* $1, 2, \cdots, s$ *appears at least q times, $q \leqq n$, there is a transversal that contains at least*

$$s\left[\frac{q}{n} - \binom{q}{2}\frac{1}{n(n-1)} + \binom{q}{3}\frac{1}{n(n-1)(n-2)} - \cdots \pm \binom{q}{q}\frac{1}{n(n-1)\cdots(n-(q-1))}\right]$$

distinct symbols.

Proof. Let U be the set of ordered pairs (t, i) where symbol i

is contained in transversal t. The cardinality of U is equal to

$n! \cdot$(the average number of distinct symbols in all transversals of the n-square) .

On the other hand, since there are s symbols, U has cardinality

$s \cdot$(the average number of transversals that contain a given symbol).

Let X_i be the set of cells occupied by the symbol i. Since $|X_i| \geqq q$, $t(X_i)$ is greater than or equal to the number of transversals that meet q diagonal elements, by Lemma 3.2. Comparison of these two expressions for the cardinality of U together with Lemma 3.1 establishes the theorem.

The case $q = n$ is singled out in the following corollary.

COROLLARY 3.3. *In an equi-n-square there is a transversal that contains at least*

$$n\left(1 - \frac{1}{2!} + \frac{1}{3!} - \cdots \pm \frac{1}{n!}\right)$$

distinct symbols.

It is not clear how much Corollary 3.3 can be strengthened. Koksma's argument for $(2n + 1)/3$ does not apply to equi-n-squares, since it makes use in several places of the assumption that each row and each column is latin. Moreover, Ryser's conjecture is not valid for equi-n-squares, where n is odd and at least 3. To see this, consider the equi-n-square whose first $n - 1$ rows each consist of the symbols $1, 2, \cdots, n$ in order, and whose n^{th} row is the same set of symbols, in the order $2, 3, \cdots, n, 1$. It is a simple matter to show that it does not have a latin transversal. Note, incidentally, that each row of this equi-n-square is latin.

The proofs of the next two theorems, being similar to that of Theorem 3.2, are only sketched.

THEOREM 3.4. *Let n be even and at least 4. Let each of $n^2/2$ symbols appear twice in an n-square. Then there is a transversal that contains n distinct symbols.*

Proof.

$n! \cdot$(the average number of distinct symbols in a transversal)

$$= \frac{n^2}{2} \cdot \text{(the average number of transversals that contain a given symbol)} .$$

Thus

$$n! \text{ (the average number of distinct symbols in a transversal)} \geqq \frac{n^2}{2} \cdot n! \left(\frac{2}{n} - \frac{1}{n(n-1)} \right) .$$

Hence the average number of distinct symbols

$$\geqq n - \frac{1}{2} \cdot \frac{n}{n-1} .$$

If $n \geqq 4$, there is consequently a transversal with n distinct symbols.

The next theorem is a companion of Theorem 3.4.

THEOREM 3.5. *Let q be greater than 2 and let n be a positive multiple of q. Let each of n^2/q symbols appear q times in an n-square. Then some transversal contains more than n-$q/2$ distinct symbols.*

Proof. There is a transversal for which the number of distinct symbols is at least

$$\frac{n^2}{q}\left(\frac{q}{n}\frac{1}{n} - \frac{q \cdot q-1}{1 \cdot 2} \cdot \frac{1}{n(n-1)} + \frac{q \cdot q-1 \cdot q-2}{1 \cdot 2 \cdot 3}\frac{1}{n} \cdot \frac{1}{n-1} \cdot \frac{1}{n-2} + \cdots \frac{q(q-1)\cdots 1}{1 \cdot 2 \cdots q}\frac{1}{n} \cdot \frac{1}{n-1} \cdots \frac{1}{n-q+1} \right) .$$

Hence, there is one with more than

$$\frac{n^2}{q}\left(\frac{q}{n} - \frac{q(q-1)}{2}\frac{1}{n(n-1)} \right)$$

distinct symbols. Since $n \geqq q$ the theorem follows.

4. **Transversals of (m, n)-rectangles.** The method used in Section 3 also applies to (m, n)-rectangles. However, this section will illustrate a different averaging process, much simpler, and only slightly weaker. It employs the notion of a "singular" pair of cells. Two cells in different rows and different columns form a *singular* pair if they contain the same symbol. The method is based on a count of incidences of transversals and singular pairs.

THEOREM 4.1. *Let q divide mn and let each symbol in an (m, n)-rectangle appear q times. Then there is a transversal with at most*

$$\frac{m(q-1)}{2(n-1)}$$

singular pairs.

Proof. Count the set of ordered pairs (t, p), where t is a transversal and p is a singular pair in t. Counting in both orders yields

$$\begin{aligned} n(n-1)\cdots(n-m+1)\ &(\text{average number of singular} \\ &\text{pairs on a transversal}) \\ &\leqq \frac{mn}{q}\cdot\frac{q(q-1)}{2}\cdot(\text{average number of transversals on} \\ &\text{a singular pair}) \\ &\leqq \frac{mn}{q}\cdot\frac{q(q-1)}{2}(n-2)(n-3)\cdots(n-m+1)\,. \end{aligned}$$

The theorem follows immediately.

The following corollaries are immediate consequences.

COROLLARY 4.2. *If each symbol in an (m, n)-rectangle appears n times, then there is a transversal with at least $m/2$ distinct symbols.*

COROLLARY 4.3. *If each symbol in an (m, n)-rectangle appears q times and if*

$$\frac{m(q-1)}{2(n-1)} < 1\,,$$

then there is a latin transversal.

A special case of Corollary 4.3 is given by the following.

COROLLARY 4.4. *If each symbol in an (m, n)-rectangle appears m times, and if*

$$n > \frac{m^2 - m + 2}{2}\,,$$

then there is a latin transversal.

The method of Section 3 yields a slightly stronger result, which implies that "$>$" can be replaced by "$\geqq$" in Corollary 4.4.

5. Rows or columns with many distinct symbols. The "existence by averaging" technique may also be applied to establish the existence of a row or column in an n-square with "many" distinct symbols.

THEOREM 5.1. *Let the cells of an n-square be occupied by the symbols $1, 2, \cdots, k$, with i appearing n_i times, $1 \leqq i \leqq k$. Then some row or column contains at least*

$$\frac{1}{n}(\sqrt{n_1} + \sqrt{n_2} + \cdots + \sqrt{n_k})$$

different symbols.

Proof. Let U be the set of ordered pairs (L, i), where L is a line (either a row or a column) that contains the symbol i. Since there are $2n$ such lines, U has

$$2n \cdot (\text{average number of distinct symbols in a line}) .$$

On the other hand, U has

$$k \cdot (\text{average number of lines that contain a given symbol}).$$

To evaluate the second average, let $L(i)$ be the number of lines that contain the symbol i. Let $R(i)$ be the number of rows and $C(i)$ be the number of columns that contain i. Thus $L(i) = R(i) + C(i)$.

Now, the set of cells occupied by i is contained in the intersection of $R(i)$ rows and $C(i)$ columns. Consequently

$$R(i) \times C(i) \geqq n_i .$$

It follows that

$$R(i) + C(i) \geqq 2\sqrt{n_i} ,$$

hence that

$$L(i) \geqq 2\sqrt{n_i} .$$

Thus

$$\sum_{i=1}^{k} L(i) \geqq \sum_{i=1}^{k} 2\sqrt{n_i} ,$$

from which the theorem follows.

The specialization of Theorem 5.1 to an equi-n-square is described in the next corollary.

COROLLARY 5.2. *In an equi-n-square there is a row or a column that contains at least* $\sqrt{n}$ *distinct symbols.*

G. D. Chakerian and D. Hickerson have independently shown that Corollary 5.2 is best possible if it is not required the set of cells occupied by a given symbol be topologically connected.

6. Conjectures. The following conjectures, some of which are logically related, may suggest directions for further study.

(1) An equi-n-square has a transversal with at least $n - 1$ distinct symbols.

(2) An n-square in which each symbol appears at most $n - 1$ times has a latin transversal. (It is easy to show by induction, or by either averaging method that if each symbol in an n-square, $n \geqq 3$, appears at most two times, the n-square has a latin transversal.)

(3) An $(n - 1, n)$-rectangle in which each symbol appears at most n times has a latin transversal.

(4) A row-latin $(n - 1, n)$-square has a latin transversal.

(5) An (m, n)-rectangle in which each symbol appears at most n times has a latin transversal.

Note that Conjectures (3) and (5) are equivalent. Moreover, for $m = 1$, Conjecture (5) is immediate. For $m = 2$, Conjecture (5) is valid with the weaker assumption that each symbol appears at most $2n - 1$ times.

(6) An $(n - 1, n)$-rectangle in which each symbol appears exactly n times has a latin transversal.

(7) An (m, n)-rectangle in which each symbol appears at most $m + 1$ times has a latin transversal.

REFERENCES

1. L. D. Baumert, et al, *A combinatorial packing problem*, Computers in Algebra and Number Theory, vol. 4, SIAM/AMS Proceedings, 1971.

2. G. D. Chakerian and S. K. Stein, *On measures of symmetry of convex bodies*, Canad. J. Math., **17** (1965), 497-504.

3. M. Hall, *A combinatorial problem on abelian groups*, Proc. Amer. Math. Soc., **3** (1952), 584-587.

4. K. K. Koksma, *A lower bound for the order of a partial transversal of a Latin square*, J. Combinatorial Theory, **7** (1969), 94-95.

5. C. C. Lindner and N. C. Perry, *Transversals of Latin sequares*, 1970, Auburn University, mimeographed.

6. E. Netto, *Lehrbuch der Combinatorik*, Chelsea, N.Y.

7. L. J. Paige, *A note on finite abelian groups*, Bull. Amer. Math. Soc., **53** (1947), 590-593.

8. ———, *Complete mappings of finite groups*, Pacific J. Math., **1** (1951), 111-116.
9. H. J. Ryser, *Combinatorial Mathematics*, Wiley, 1963.
10. ———, *Neure Probleme der Kombinatorik in Vorträge über Kombinatorik*, Oberwolfach 24-29 Juli, 1967. Mathematischen Forschgsinstitut, Oberwolfach, 1968.
11. S. K. Stein, *The symmetry function in a convex body*, Pacific J. Math., **6** (1956), 145-148.

Received November 27, 1974. For a general survey of transversals in latin n-sequares see J. Dénes and A. D. Keedwell, Latin squares and their applications, Academic Press, 1974. Incidentally, it is mentioned there that the analog of Conjecture (1) has been proposed for latin n-squares.

UNIVERSITY OF CALIFORNIA, DAVIS