



Some of the oldest combinatorial objects, whose study apparently goes back to ancient times, are the *Latin squares*. To obtain a Latin square, one has to fill the n^2 cells of an $n \times n$ square array with the numbers $1, 2, \dots, n$ so that every number appears exactly once in every row and in every column. In other words, the rows and columns each represent permutations of the set $\{1, \dots, n\}$. Let us call n the *order* of the Latin square.

Here is the problem we want to discuss. Suppose someone started filling the cells with the numbers $\{1, 2, \dots, n\}$. At some point he stops and asks us to fill in the remaining cells so that we get a Latin square. When is this possible? In order to have a chance at all we must, of course, assume that at the start of our task any element appears at most once in every row and in every column. Let us give this situation a name. We speak of a *partial Latin square* of order n if some cells of an $n \times n$ array are filled with numbers from the set $\{1, \dots, n\}$ such that every number appears at most once in every row and column. So the problem is:

When can a partial Latin square be completed to a Latin square of the same order?

Let us look at a few examples. Suppose the first $n - 1$ rows are filled and the last row is empty. Then we can easily fill in the last row. Just note that every element appears $n - 1$ times in the partial Latin square and hence is missing from exactly one column. Hence by writing each element below the column where it is missing we have completed the square correctly.

Going to the other end, suppose only the first row is filled. Then it is again easy to complete the square by cyclically rotating the elements one step in each of the following rows.

So, while in our first example the completion is forced, we have lots of possibilities in the second example. In general, the fewer cells are pre-filled, the more freedom we should have in completing the square.

However, the margin displays an example of a partial square with only n cells filled which clearly cannot be completed, since there is no way to fill the upper right-hand corner without violating the row or column condition.

1	2	3	4
2	1	4	3
4	3	1	2
3	4	2	1

A Latin square of order 4

1	4	2	5	3
4	2	5	3	1
2	5	3	1	4
5	3	1	4	2
3	1	4	2	5

A cyclic Latin square

1	2	...	$n-1$	
				n

A partial Latin square that cannot be completed

If fewer than n cells are filled in an $n \times n$ array, can one then always complete it to obtain a Latin square?

1	3	2
2	1	3
3	2	1

R : 1 1 1 2 2 2 3 3 3

C : 1 2 3 1 2 3 1 2 3

E : 1 3 2 2 1 3 3 2 1

If we permute the lines of the above example cyclically,

$R \rightarrow C \rightarrow E \rightarrow R$, then we obtain the following line array and Latin square:

1	2	3
3	1	2
2	3	1

R : 1 3 2 2 1 3 3 2 1

C : 1 1 1 2 2 2 3 3 3

E : 1 2 3 1 2 3 1 2 3

This question was raised by Trevor Evans in 1960, and the assertion that a completion is always possible quickly became known as the Evans conjecture. Of course, one would try induction, and this is what finally led to success. But Bohdan Smetaniuk's proof from 1981, which answered the question, is a beautiful example of just how subtle an induction proof may be needed in order to do such a job. And, what's more, the proof is constructive, it allows us to complete the Latin square explicitly from any initial partial configuration.

Before proceeding to the proof let us take a closer look at Latin squares in general. We can alternatively view a Latin square as a $3 \times n^2$ array, called the *line array* of the Latin square. The figure to the left shows a Latin square of order 3 and its associated line array, where R , C and E stand for rows, columns and elements.

The condition on the Latin square is equivalent to saying that in any two lines of the line array all n^2 ordered pairs appear (and therefore each pair appears exactly once). Clearly, we may permute the symbols in each line arbitrarily (corresponding to permutations of rows, columns or elements) and still obtain a Latin square. But the condition on the $3 \times n^2$ array tells us more: There is no special role for the elements. We may also permute the lines of the array (as a whole) and still preserve the conditions on the line array and hence obtain a Latin square.

Latin squares that are connected by any such permutation are called *conjugates*. Here is the observation which will make the proof transparent: A partial Latin square obviously corresponds to a partial line array (every pair appears at most once in any two lines), and any conjugate of a partial Latin square is again a partial Latin square. In particular, a partial Latin square can be completed if and only if any conjugate can be completed (just complete the conjugate and then reverse the permutation of the three lines).

We will need two results, due to Herbert J. Ryser and to Charles C. Lindner, that were known prior to Smetaniuk's theorem. If a partial Latin square is of the form that the first r rows are completely filled and the remaining cells are empty, then we speak of an $r \times n$ *Latin rectangle*.

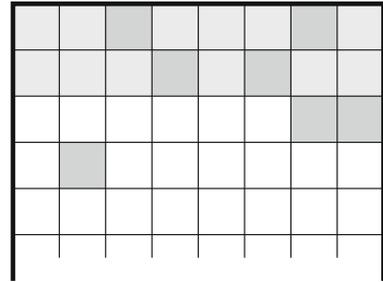
Lemma 1. *Any $r \times n$ Latin rectangle, $r < n$, can be extended to an $(r + 1) \times n$ Latin rectangle and hence can be completed to a Latin square.*

■ **Proof.** We apply Hall's theorem (see Chapter 30). Let A_j be the set of numbers that do *not* appear in column j . An admissible $(r + 1)$ -st row corresponds then precisely to a system of distinct representatives for the collection A_1, \dots, A_n . To prove the lemma we therefore have to verify Hall's condition (H). Every set A_j has size $n - r$, and every element is in precisely $n - r$ sets A_j (since it appears r times in the rectangle). Any m of the sets A_j contain together $m(n - r)$ elements and therefore at least m different ones, which is just condition (H). □

Lemma 2. *Let P be a partial Latin square of order n with at most $n - 1$ cells filled and at most $\frac{n}{2}$ distinct elements, then P can be completed to a Latin square of order n .*

■ **Proof.** We first transform the problem into a more convenient form. By the conjugacy principle discussed above, we may replace the condition “at most $\frac{n}{2}$ distinct elements” by the condition that the entries appear in at most $\frac{n}{2}$ rows, and we may further assume that these rows are the top rows. So let the rows with filled cells be the rows $1, 2, \dots, r$, with f_i filled cells in row i , where $r \leq \frac{n}{2}$ and $\sum_{i=1}^r f_i \leq n - 1$. By permuting the rows, we may assume that $f_1 \geq f_2 \geq \dots \geq f_r$. Now we complete the rows $1, \dots, r$ step by step until we reach an $r \times n$ rectangle which can then be extended to a Latin square by Lemma 1.

Suppose we have already filled rows $1, 2, \dots, \ell - 1$. In row ℓ there are f_ℓ filled cells which we may assume to be at the end. The current situation is depicted in the figure, where the shaded part indicates the filled cells.



A situation for $n = 8$, with $\ell = 3$, $f_1 = f_2 = f_3 = 2$, $f_4 = 1$. The dark squares represent the pre-filled cells, the lighter ones show the cells that have been filled in the completion process.

The completion of row ℓ is performed by another application of Hall’s theorem, but this time it is quite subtle. Let X be the set of elements that do *not* appear in row ℓ , thus $|X| = n - f_\ell$, and for $j = 1, \dots, n - f_\ell$ let A_j denote the set of those elements in X which do *not* appear in column j (neither above nor below row ℓ). Hence in order to complete row ℓ we must verify condition (H) for the collection A_1, \dots, A_{n-f_ℓ} .

First we claim

$$n - f_\ell - \ell + 1 > \ell - 1 + f_{\ell+1} + \dots + f_r. \tag{1}$$

The case $\ell = 1$ is clear. Otherwise $\sum_{i=1}^r f_i < n$, $f_1 \geq \dots \geq f_r$ and $1 < \ell \leq r$ together imply

$$n > \sum_{i=1}^r f_i \geq (\ell - 1)f_{\ell-1} + f_\ell + \dots + f_r.$$

Now either $f_{\ell-1} \geq 2$ (in which case (1) holds) or $f_{\ell-1} = 1$. In the latter case, (1) reduces to $n > 2(\ell - 1) + r - \ell + 1 = r + \ell - 1$, which is true because of $\ell \leq r \leq \frac{n}{2}$.

Let us now take m sets A_j , $1 \leq m \leq n - f_\ell$, and let B be their union. We must show $|B| \geq m$. Consider the number c of cells in the m columns corresponding to the A_j ’s which contain elements of X . There are at most $(\ell - 1)m$ such cells above row ℓ and at most $f_{\ell+1} + \dots + f_r$ below row ℓ , and thus

$$c \leq (\ell - 1)m + f_{\ell+1} + \dots + f_r.$$

On the other hand, each element $x \in X \setminus B$ appears in each of the m columns, hence $c \geq m(|X| - |B|)$, and therefore (with $|X| = n - f_\ell$)

$$|B| \geq |X| - \frac{1}{m}c \geq n - f_\ell - (\ell - 1) - \frac{1}{m}(f_{\ell+1} + \dots + f_r).$$

It follows that $|B| \geq m$ if

$$n - f_\ell - (\ell - 1) - \frac{1}{m}(f_{\ell+1} + \dots + f_r) > m - 1,$$

that is, if

$$m(n - f_\ell - \ell + 2 - m) > f_{\ell+1} + \dots + f_r. \tag{2}$$

Inequality (2) is true for $m = 1$ and for $m = n - f_\ell - \ell + 1$ by (1), and hence for all values m between 1 and $n - f_\ell - \ell + 1$, since the left-hand side is a quadratic function in m with leading coefficient -1 . The remaining case is $m > n - f_\ell - \ell + 1$. Since any element x of X is contained in at most $\ell - 1 + f_{\ell+1} + \dots + f_r$ rows, it can also appear in at most that many columns. Invoking (1) once more, we find that x is in one of the sets A_j , so in this case $B = X$, $|B| = n - f_\ell \geq m$, and the proof is complete. \square

Let us finally prove Smetaniuk’s theorem.

Theorem. Any partial Latin square of order n with at most $n - 1$ filled cells can be completed to a Latin square of the same order.

s_1		2				7		
s_2			5		4			
s_3				5				
s_4		4						



•								
2	7							
		•						
			•					
	4	5		•				
			5		•			
4								•

2	3	4	1	6	5		
5	6	1	4	2	3		
1	2	3	6	5	4		
6	4	5	2	3	1		
3	1	6	5	4	2		
4	5	2	3	1	6		

■ **Proof.** We use induction on n , the cases $n \leq 2$ being trivial. Thus we now study a partial Latin square of order $n \geq 3$ with at most $n - 1$ filled cells. With the notation used above these cells lie in $r \leq n - 1$ different rows numbered s_1, \dots, s_r , which contain $f_1, \dots, f_r > 0$ filled cells, with $\sum_{i=1}^r f_i \leq n - 1$. By Lemma 2 we may assume that there are more than $\frac{n}{2}$ different elements; thus there is an element that appears only once: after renumbering and permutation of rows (if necessary) we may assume that the element n occurs only once, and this is in row s_1 .

In the next step we want to permute the rows and columns of the partial Latin square such that after the permutations all the filled cells lie below the diagonal — except for the cell filled with n , which will end up on the diagonal. (The diagonal consists of the cells (k, k) with $1 \leq k \leq n$.) We achieve this as follows: First we permute row s_1 into the position f_1 . By permutation of columns we move all the filled cells to the left, so that n occurs as the last element in its row, on the diagonal. Next we move row s_2 into position $1 + f_1 + f_2$, and again the filled cells as far to the left as possible. In general, for $1 < i \leq r$ we move the row s_i into position $1 + f_1 + f_2 + \dots + f_i$ and the filled cells as far left as possible. This clearly gives the desired set-up. The drawing to the left shows an example, with $n = 7$: the rows $s_1 = 2, s_2 = 3, s_3 = 5$ and $s_4 = 7$ with $f_1 = f_2 = 2$ and $f_3 = f_4 = 1$ are moved into the rows numbered 2, 5, 6 and 7, and the columns are permuted “to the left” so that in the end all entries except for the single 7 come to lie below the diagonal, which is marked by •s.

In order to be able to apply induction we now remove the entry n from the diagonal and ignore the first row and the last column (which do not contain any filled cells): thus we are looking at a partial Latin square of order $n - 1$ with at most $n - 2$ filled cells, which by induction can be completed to a Latin square of order $n - 1$. The margin shows one (of many) completions of the partial Latin square that arises in our example. In the figure, the original entries are printed bold. They are already final, as are all the elements in shaded cells; some of the other entries will be changed in the following, in order to complete the Latin square of order n .

In the next step we want to move the diagonal elements of the square to the last column and put entries n onto the diagonal in their place. However, in general we cannot do this, since the diagonal elements need not

In our example the “exchange case” happens for $k = 5$: the element $x_5 = 3$ does already occur in the last column, so that entry has to be moved back to column $k = 5$. But the exchange element $x'_5 = 6$ is not new either, it is exchanged by $x''_5 = 5$, and this one is new.

2	7	4	1	6	5	3
5	6	7	4	2	3	1
1	2	3	7	5	4	6
6	4	5	2	3	1	7
3	1	6	5	4	2	
4	5	2	3	1	6	

2	7	4	1	3	5	6
5	6	7	4	2	3	1
1	2	3	7	6	4	5
6	4	5	2	7	1	3
3	1	6	5	4	2	
4	5	2	3	1	6	

Finally, the exchange for $k = 6 = n - 1$ poses no problem, and after that the completion of the Latin square is unique:

2	7	4	1	3	5	6
5	6	7	4	2	3	1
1	2	3	7	6	4	5
6	4	5	2	7	1	3
3	1	6	5	4	2	7
4	5	2	3	1	6	

2	7	4	1	3	5	6
5	6	7	4	2	3	1
1	2	3	7	6	4	5
6	4	5	2	7	1	3
3	1	6	5	4	7	2
4	5	2	3	1	6	

7	3	1	6	4	2	4
2	7	4	1	3	5	6
5	6	7	4	2	3	1
1	2	3	7	6	4	5
6	4	5	2	7	1	3
3	1	6	5	4	7	2
4	5	2	3	1	6	7

... and the same occurs in general: We put an element n into the cell (n, n) , and after that the first row can be completed by the missing elements of the respective columns (see Lemma 1), and this completes the proof. In order to get explicitly a completion of the original partial Latin square of order n , we only have to reverse the element, row and column permutations of the first two steps of the proof. \square

References

- [1] T. EVANS: *Embedding incomplete Latin squares*, Amer. Math. Monthly **67** (1960), 958-961.
- [2] C. C. LINDNER: *On completing Latin rectangles*, Canadian Math. Bulletin **13** (1970), 65-68.
- [3] H. J. RYSER: *A combinatorial theorem with an application to Latin rectangles*, Proc. Amer. Math. Soc. **2** (1951), 550-552.
- [4] B. SMETANIUK: *A new construction on Latin squares I: A proof of the Evans conjecture*, Ars Combinatoria **11** (1981), 155-172.