Comment.Math.Univ.Carolin. 51,2 (2010) 175–184

Overlapping latin subsquares and full products

Joshua M. Browning, Petr Vojtěchovský, Ian M. Wanless

Abstract. We derive necessary and sufficient conditions for there to exist a latin square of order n containing two subsquares of order a and b that intersect in a subsquare of order c. We also solve the case of two disjoint subsquares. We use these results to show that:
(a) A latin square of order n cannot have more than $\frac{n}{m}\binom{n}{h}/\binom{m}{h}$ subsquares of order m, where $h = \lceil (m+1)/2 \rceil$. Indeed, the number of subsquares of order m is bounded by a polynomial of degree at most $\sqrt{2m}+2$ in n.
(b) For all $n \geq 5$ there exists a loop of order n in which every element can be obtained as a product of all n elements in some order and with some bracketing.

Keywords: latin square, latin subsquare, overlapping latin subsquares, full product in loops

Classification: 05B15, 20N05

1. Overlapping latin subsquares

A $k \times n$ *latin rectangle* is a $k \times n$ matrix containing n different symbols, with each symbol occurring exactly once in each row and at most once in each column. If $k = n$ the latin rectangle is a *latin square*. A *subsquare* in a latin square L is a submatrix of L that is a latin square in its own right. The cells in a subsquare are not required to be contiguous.

It is well-known that if two subsquares of a latin square intersect then their intersection is itself a subsquare. Also, a subsquare of a latin square is either the whole square or it has at most half the order of the whole square. Another fact that we will frequently use is the following result due to Ryser [5]:

Theorem 1. *Suppose that R is an $r \times s$ matrix with symbols from $\{1,\dots,n\}$ such that no symbol occurs more than once in any row or column. For $1 \leq i \leq n$, let $\Gamma(i)$ be the number of occurrences of i in R. Then R can be embedded into a latin square of order n if and only if $\Gamma(i) \geq r+s-n$ for every $i \in \{1,\dots,n\}$.*

The goal of this first section is to find conditions under which a latin square may have subsquares of two specified orders. We begin by treating the case when the two subsquares overlap. The simplest way for this to happen is for one subsquare to contain the other (for this to be possible it is necessary and sufficient that the

P. Vojtěchovský supported by Enhanced Sabbatical grant of the University of Denver. I. Wanless supported by ARC grants DP0662946 and DP1093320.

larger subsquare is at least twice the order of the smaller one). A more interesting case is when the subsquares intersect, but neither is inside the other:

Theorem 2. *Suppose $0 < c < a \leq b < n$ are integers. In order for there to exist a latin square of order n containing two subsquares of order a and b whose intersection is a subsquare of order c, it is necessary and sufficient that*

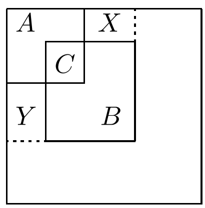
$$n - 2b \geq a - 2c \geq 0 \tag{1}$$

and

$$(n-2a)(n-2b) \geq c^2 - (n - 2a - 2b + 3c)^2. \tag{2}$$

Proof: To prove necessity, assume that L is a latin square of the desired type. By permuting rows and columns if necessary, we may assume that L has the form

(3)



where A is a subsquare of order a, B is a subsquare of order b and they intersect in C, a subsquare of order c.

Let S be the set of symbols of L that do not occur in A or B. Note that $|S| = n-(a+b-c)$. The regions X and Y are $(a-c)\times(b-c)$ and $(b-c)\times(a-c)$ submatrices, respectively, and they contain only symbols from S. To have enough symbols to fill the first row of X we must have $|S| \geq b - c$, which gives the first inequality in (1). The inequality $a \geq 2c$ is immediate, given that C is a subsquare of A and $c < a$.

To prove (2) we apply Ryser's condition to $R = A \cup B \cup X \cup Y$ and conclude that each symbol in S must occur at least $2(a+b-c)-n$ times in $X \cup Y$. To fit this many symbols of S into $X \cup Y$, we require

$$2(a-c)(b-c) = |X \cup Y| \geq \big(2(a+b-c)-n\big)|S|. \tag{4}$$

Upon multiplying this inequality by 2 and rearranging, it becomes (2).

It remains to show sufficiency. Assuming (1), we have $b \geq a \geq 2c$ and so it is possible to construct A, B and C as in (3). We aim to fill in $X \cup Y$ in such a way that the symbols in S each occur at least $\lfloor 2(a-c)(b-c)/|S| \rfloor$ times in $X \cup Y$. By (4) this would mean that each symbol in S occurs at least $\lfloor 2(a+b-c)-n \rfloor = 2(a+b-c)-n$ times in R. Furthermore, every symbol of $A \cup B$ occurs at least a times in R, and $a \geq 2(a+b-c)-n$ by (1). Therefore, if $X \cup Y$ can be filled as desired, Ryser's condition holds for R, and L exists.

To fill $X \cup Y$, we order the cells of $X \cup Y$ in a sequence such that any subsequence of $2(a-c)$ consecutive terms contains cells from distinct rows and columns. This is easily achieved by alternating "diagonals" of X and "diagonals" of Y. We then fill the cells in the order determined by the sequence, using all occurrences of one symbol before starting with the next symbol. Some symbols (it does not matter which) should be designated to occur $\lfloor 2(a-c)(b-c)/|S| \rfloor$ times, while the others occur $\lceil 2(a-c)(b-c)/|S| \rceil$ times. Note that $|S| \geq b-c$ and hence $\lceil 2(a-c)(b-c)/|S| \rceil \leq 2(a-c)$, which means that our construction will not violate the latin property. $\square$

The following corollary gives a sufficient (though not in general necessary) condition, and it is easier to apply than Theorem 2.

Corollary 3. *Suppose $0 < c < a \leq b < n$ are integers. If $n - 2b \geq a - 2c$ and $a \geq \frac{5}{2}c$ then there is a latin square of order n containing subsquares of order a and b whose intersection is a subsquare of order c.*

Proof: The two conditions $n - 2b \geq a - 2c$ and $a \geq \frac{5}{2}c$ imply (1), so it remains to show (2).

If $n \geq 2b + c$ then $(n-2a)(n-2b) \geq c^2$ and (2) follows. If $n \leq 2a + 2b - 4c$ then $(n - 2a - 2b + 3c)^2 \geq c^2$, so the right hand side of (2) is negative, and (2) follows once again.

Otherwise we have $n < 2b + c$ and $n > 2a + 2b - 4c$, which gives $5c > 2a$, a contradiction. $\square$

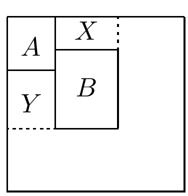
Next, we examine the case of a latin square with two subsquares that do not overlap.

We say that two subsquares *share a row* if there is some row that intersects both subsquares. Sharing a column or symbol is defined analogously. By the operation known as *conjugacy* or *parastrophy* (a permutation of the roles of rows, columns and symbols in a latin square) we may assume that if the subsquares share anything, they share a row. If two disjoint subsquares share a row, it immediately follows that they do not share a column or symbol.

Lemma 4. *Suppose $0 \leq c < a \leq b < n$ are integers. In order for there to exist a latin square of order n containing two subsquares of order a and b that share exactly c rows and do not share any columns or symbols, it is necessary and sufficient that $n \geq a + 2b$.*

Proof: Up to permutation of the rows and columns, our latin square, if it exists, looks like

(5)



where X and Y are (non-empty) submatrices of dimensions $(a-c)\times b$ and $(b-c)\times a$ respectively. The symbols in $X\cup Y$ must be distinct from those in $A\cup B$, and by assumption, the symbols in subsquare A are distinct from the symbols in subsquare B. Each row of X must contain b symbols that are different from the $a+b$ symbols in $A\cup B$, which demonstrates the necessity of the condition $n\geq 2b+a$.

To prove the sufficiency of the condition $n\geq 2b+a$, we apply Theorem 1 to $R=A\cup B\cup X\cup Y$. Symbols in A each occur a times in R, so we need $a\geq a+b-c+a+b-n$, that is, $n\geq a+2b-c$. Similarly, considering the symbols in B leads to $n\geq 2a+b-c$. Both of these inequalities hold if $n\geq 2b+a$, given that $a\leq b$.

We fill in the symbols in $X\cup Y$ in the same way that we did in Theorem 2. Each symbol in $X\cup Y$ will occur at most

$$\begin{aligned}\left\lceil\frac{(a-c)b+(b-c)a}{n-a-b}\right\rceil &\leq \left\lceil\frac{(a-c)b+(b-c)a}{b}\right\rceil\\ &= a-c+\left\lceil\frac{(b-c)a}{b}\right\rceil\\ &\leq a-c+\min\{b-c,a\}\end{aligned}$$

times, which means that we will not allocate the same symbol to two different cells in the same row or column.

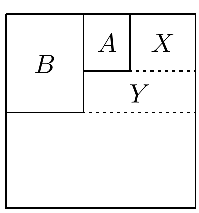
Moreover, the symbols in $X\cup Y$ will satisfy Ryser's condition, given that $(a-c)b+(b-c)a\geq(n-a-b)(2a+2b-c-n)$. This last condition is algebraically equivalent to $(n-2a-b)(n-a-2b+c)+(a-c)b\geq 0$, which is obviously true. $\square$

Lemma 4 did not cover the case when $a=c$. It needs to be treated separately:

Lemma 5. *Suppose $0<a\leq b<n$ are integers. In order for there to exist a latin square of order n containing two subsquares of order a and b that share exactly a rows and do not share any columns or symbols, it is necessary and sufficient that either $a=b=n/2$ or $n\geq\max\{2a+b,2b\}$.*

Proof: First suppose that $n>a+b$. Then, up to permutation of the rows and columns, our latin square looks like this:

(6)



where as usual, A and B are the subsquares of order a and b.

Once we are given A, Theorem 1 shows that to be able to fill in $A \cup X \cup Y$ it is necessary and sufficient that $0 \geq 2a - (n-b)$ and $b \leq n-b$. The remainder of the latin square can then always be completed. Hence, for the latin square to exist when $n > a+b$, it is necessary and sufficient that $n \geq \max\{2a+b, 2b\}$.

That leaves the case $n = a+b$. In that case, the submatrix X in (6) has no columns. Also, there are no symbols available to fill the submatrix Y, so we are forced to make $a = b$. With that condition, it is trivial to complete a latin square of order $n = 2a = 2b$. □

We have now considered all possible ways that a latin square might contain subsquares of order a and b. Theorem 2 (and its preamble) handles the case when the subsquares intersect. Disjoint subsquares that share any row, column or symbol are covered (up to conjugacy/parastrophy) by Lemma 4 and Lemma 5. Subsquares that do not share any row, column or symbol are covered explicitly by Lemma 4 (the same result would be obtained by taking $c = 0$ in Theorem 2, although for simplicity we did not allow that case in our phrasing of the theorem). Putting these results together, we get:

Theorem 6. *Suppose $1 < a \leq b < n$ are integers. In order for there to exist a latin square of order n containing two distinct subsquares of order a and b, the following set of conditions is necessary and sufficient:*

(a) $n \geq 2b$,
(b) *if $n = 2b+1$ and a is even then $a \leq \frac{2}{3}(b+1)$ and*
(c) *if $n = 2b$ and a is odd then $a = b$ or $a \leq b/2$.*

Proof: As usual, A and B denote subsquares of order a and b in a latin square L of order n.

Condition (a) is clearly necessary. If $n \geq 2b+a$ then we can use Lemma 4, so we may assume that $2b \leq n < 2b+a$.

Suppose first that $n \equiv a \mod 2$. Take $c = (a+2b-n)/2$ and note that (1) holds. Moreover,

$$(n-2a)(n-2b) \geq 0 \geq a(2b-n) = c^2 - (n-2a-2b+3c)^2$$

so (2) holds as well and we are done. Henceforth we may assume that $n \not\equiv a \mod 2$. At this point we split into three cases.

Case: $n \geq 2b+2$

Take $c = (a+2b-n+1)/2$ and note that (1) holds. Moreover,

$$(n-2a)(n-2b) \geq 0 \geq (a-1)(2b-n+2) = c^2 - (n-2a-2b+3c)^2$$

so (2) holds as well and we are done.

Case: $n = 2b+1$

Then a is even since $n \not\equiv a \mod 2$. If A and B intersect in a smaller subsquare then (1) implies that we must have $a = 2c$. In turn, (2) tells us that $2b+1 \geq$

$3a-1$. Hence if $a \le \frac{2}{3}(b+1)$ then Theorem 2 shows that L exists. Otherwise $a > \frac{2}{3}(b+1) > \frac{1}{2}b$, so the only hope of constructing L is that A and B do not overlap. However, we have $n = 2b+1 < b+\frac{3}{2}a < b+2a \le 2b+a$, which breaches the necessary conditions in both Lemma 4 and Lemma 5. So A and B cannot be disjoint.

Case: $n = 2b$

A latin square of order n containing a subsquare of order $n/2$ must be composed of four disjoint subsquares of order $n/2$. Thus by changing our choice of B if necessary, we can assume that $a = b$ or that A and B intersect. Since $n \not\equiv a \mod 2$ we find that a is odd, but this makes (1) impossible to satisfy. So A and B cannot intersect unless A is wholly inside B, in which case $a = b$ or $a \le b/2$ (and both options are achievable). □

2. A bound on the number of subsquares

The following simple bound on the number of subsquares in a latin square generalises results by Heinrich and Wallis [2] and van Rees [6] who proved the $m = 2$ and $m = 3$ cases, respectively.

Theorem 7. *Let m, n be positive integers, and define $h = \lceil (m+1)/2 \rceil$. No latin square of order n may have more than*

$$\frac{n\binom{n}{h}}{m\binom{m}{h}}$$

subsquares of order m.

Proof: Suppose L is a latin square of order n, and consider a set H of h entries from within one row of L. Since $h > m/2$, we know from (1) that at most one $m \times m$ subsquare of L can contain H. There are $n\binom{n}{h}$ possible choices for H, and each $m \times m$ subsquare contains $m\binom{m}{h}$ of them. □

By refining the idea of Theorem 7 we can prove a better bound when m is large. All asymptotics in the remainder of this section are for $n \to \infty$ with other quantities fixed.

Theorem 8. *Let m, n, t be positive integers with $n \ge m \ge t$. Define $\psi(m,t)$ by*

$$\psi(m,t) = \begin{cases} \lfloor m/(2t) \rfloor + 1 & \text{if } m \text{ is even,} \\ \lceil \frac{1}{2} \lfloor m/t \rfloor \rceil & \text{if } m \text{ is odd.} \end{cases}$$

No latin square of order n can have more than $O(n^{\psi(m,t)+t})$ subsquares of order m.

Proof: Let L be a latin square of order n. Define a *block* to be a latin subrectangle of L with exactly t rows, that is minimal in the sense that it does not contain any smaller such latin rectangle.

Suppose that we choose up to $\psi(m,t)$ blocks that lie in the same t rows of L. We claim that for every subsquare S of order m in L there is (at least) one such choice of blocks that lies in S and not in any other subsquare of order m. It will then follow that the number of $m \times m$ subsquares is not more than $O(n^{\psi(m,t)+t})$ since there are $O(n^t)$ ways to choose t rows and $O(n^{\psi(m,t)})$ ways to choose up to $\psi(m,t)$ blocks in those rows.

To prove our claim, we choose blocks from the first t rows of S until the chosen blocks cover strictly more than $m/2$ columns. By (1) this will guarantee that no other subsquare of order m contains these blocks. At each step we select a block that is at least as large as any of the remaining blocks. This ensures that we need no more than $\psi(m,t)$ blocks, as the following argument shows.

There are at most $\lfloor m/t \rfloor$ blocks within the first t rows of S, since each block uses at least t columns. If m is odd, we need only choose at least half of them and we are done. So suppose m is even. If $m/2$ is divisible by t then $t\psi(m,t) = \frac{1}{2}m + t > \frac{1}{2}m$ and otherwise $t\psi(m,t) = t\lceil m/(2t) \rceil > \frac{1}{2}m$. Thus, as claimed, there is no case where we need more than $\psi(m,t)$ blocks to determine S. □

Corollary 9. *No latin square of order n can have more than $O(n^{\sqrt{2m}+2})$ subsquares of order m.*

PROOF: Taking $t = \lceil \sqrt{m/2}\,\rceil$, we will show that $\psi(m,t) + t \le \sqrt{2m} + 2$. If m is even then

$$\begin{aligned}\psi(m,t) + t &= \left\lfloor \frac{m}{2\lceil \sqrt{m/2}\,\rceil} \right\rfloor + 1 + \lceil \sqrt{m/2}\,\rceil \\ &\le \lfloor \sqrt{m/2} \rfloor + 1 + \sqrt{m/2} + 1.\end{aligned}$$

Similarly, if m is odd,

$$\psi(m,t) + t = \left\lceil \frac{1}{2} \left\lfloor \frac{m}{\lceil \sqrt{m/2}\,\rceil} \right\rfloor \right\rceil + \lceil \sqrt{m/2}\,\rceil \le 2\lceil \sqrt{m/2}\,\rceil.$$

In either case, $\psi(m,t) + t \le 2(\sqrt{m/2} + 1)$, as required. □

The bound in Theorem 7 is obviously achieved when $m = n$ and is known [2], [6] to be achieved for infinitely many n when $m \in \{2,3\}$. It is also not hard to show that the elementary abelian 2-groups achieve the bound in Theorem 7 for $m = 4$. On the other hand, Corollary 9 shows that the bound in Theorem 7 can only be achieved for finitely many n when $m > 9$. In fact, a more careful analysis using Theorem 8 reveals that, for $m > 4$, the number of subsquares of order m is $o(n^{1+h})$. Hence Theorem 7 is not best possible except for $m \le 4$.

See [3] for some results on how many subsquares are 'typical' in a random latin square. In that paper it is conjectured that the proportion of latin squares of order n with a subsquare of order greater than 3 tends to zero as $n \to \infty$.

3. Full products

A *quasigroup* is a groupoid $(Q, \cdot)$ such that the equations $ax = b$ and $ya = b$ have unique solutions x, $y \in Q$ for every a, $b \in Q$. A *loop* is a quasigroup with a neutral element. Latin squares are precisely multiplication tables of finite quasigroups. Multiplication tables of finite loops correspond to normalized latin squares.

For a subset S of a loop Q, denote by $P(S)$ the set of all elements of Q that are obtained as products of all elements of S with each element of S being used precisely once. We refer to elements of $P(Q)$ as *full products* of Q. Full products play an important role in a recent non-associative interpretation for the Hall-Paige conjecture [4].

For a loop Q, let Q' be the *derived subloop*, that is, the least normal subloop H of Q such that Q/H is an abelian group.

It is not difficult to show that $P(Q)$ is contained in a coset of Q'. The Dénes-Hermann Theorem [1] states that if Q is a group then $P(Q)$ is equal to a coset of Q'; more precisely, either $P(Q) = Q'$ or $P(Q) = xQ'$ where $x^2 \in Q'$.

It is not true for a general loop Q that $P(Q)$ is a coset of Q', but the only known counterexamples are of order 5. For instance, the loop

Q_1	1	2	3	4	5
1	1	2	3	4	5
2	2	1	5	3	4
3	3	4	1	5	2
4	4	5	2	1	3
5	5	3	4	2	1

satisfies $P(Q_1) = \{2, 3, 4, 5\}$.

Note that $|P(Q)|$ is not an isotopy invariant, since the loop

(7)

Q_2	1	2	3	4	5
1	1	2	3	4	5
2	2	1	4	5	3
3	3	4	5	2	1
4	4	5	1	3	2
5	5	3	2	1	4

is isotopic to Q_1 but satisfies $P(Q_2) = Q_2$.

While we can certainly have $P(Q) < Q$ in an arbitrarily large loop (an abelian group Q will do), it is to be expected that a sufficiently large loop Q chosen at random will satisfy $Q' = Q$ and, in fact, $P(Q) = Q$. However, we are not aware of any argument that would show $P(Q) = Q$ for a 'typical' sufficiently large loop Q. As an application of our previous results we show here that for every $n \geq 5$ there is a loop Q of order n satisfying $P(Q) = Q$.

Lemma 10. *Let Q be a loop of order $m \geq 5$ such that $P(Q) = Q$. Then for every $3m - 2 \leq n \leq 4m - 3$ there is a loop H of order n satisfying $P(H) = H$.*

Proof: Let $\overline{Q}$ be an isomorphic copy of Q. Set $a = b = m$ and $c = 1$ in Corollary 3 to see that for every $n \geq 3m - 2$ there is a loop H of order n that contains Q and $\overline{Q}$ as subloops and such that $Q \cap \overline{Q} = \{1\}$, where 1 is the neutral element of H.

Assume further that $n \leq 4m-3$ and let $x \in H \setminus (Q \cup \overline{Q})$. By Ryser's condition, x appears in $(Q \cup \overline{Q}) \times (Q \cup \overline{Q})$ at least $2(2m-1) - n \geq 1$ times, and hence either $x \in Q \times \overline{Q}$ or $x \in \overline{Q} \times Q$.

We are going to show that $P(Q \cup \overline{Q}) = H$. This will imply that $P(H) = H$ since the cardinality of $P(Q \cup \overline{Q})$ cannot decrease upon multiplying if by the elements of $H \setminus (Q \cup \overline{Q})$ in any way.

Let $x \in H$. If $x \in Q$, we see that $x \in P(Q \cup \overline{Q})$ since $P(Q) = Q$ and $1 \in P(\overline{Q})$. Similarly, $x \in P(Q \cup \overline{Q})$ for every $x \in \overline{Q}$. Suppose that $x \in H \setminus (Q \cup \overline{Q})$. By the argument above, we know that either $x = y\overline{y}$ or $x = \overline{y}y$ for some $y \in Q$, $\overline{y} \in \overline{Q}$. This means that $x \in P(Q \cup \overline{Q})$, as $P(Q) = Q$ and $P(\overline{Q}) = \overline{Q}$. □

Theorem 11. *There is a loop Q of order n satisfying $P(Q) = Q$ if and only if $n = 1$ or $n \geq 5$.*

Proof: The statement is true for $n = 1$. Assume that $1 < n < 5$. Then Q is an abelian group, $Q' < Q$, and thus $P(Q)$ is a proper subset of Q'. In (7) we gave a loop Q_2 of order 5 where $Q_2 = P(Q_2)$. It is easy to check by computer that for every $5 < n \leq 12$ there is a loop Q of order n satisfying $P(Q) = Q$. (In fact we did not find any example where $P(Q) < Q$, except those with $Q' < Q$.) By Lemma 10, the theorem is also true for all n in the intervals $[3 \cdot 5 - 2, 4 \cdot 5 - 3] = [13, 17]$, $[3 \cdot 6 - 2, 4 \cdot 6 - 3] = [16, 21]$, $[3 \cdot 7 - 2, 4 \cdot 7 - 3] = [19, 25]$, and so on, obviously accounting for every $n \geq 5$. □

Acknowledgement. The authors are grateful to Doug Stones for useful feedback on a draft of this paper.

References

[1] Dénes J., Hermann P., *On the product of all elements in a finite group*, Ann. Discrete Math. **15** (1982), 105–109.

[2] Heinrich K., Wallis W.D., *The maximum number of intercalates in a latin square*, Lecture Notes in Math. **884** (1981), 221–233.

[3] McKay B.D., Wanless I.M., *Most latin squares have many subsquares*, J. Combin. Theory Ser. A **86** (1999), 323–347.

[4] Pula K., *Products of all elements in a loop and a framework for non-associative analogues of the Hall-Paige conjecture*, Electron. J. Combin. **16** (2009), R57.

[5] Ryser H.J., *A combinatorial theorem with an application to latin rectangles*, Proc. Amer. Math. Soc. **2** (1951), 550–552.

[6] van Rees G.H.J., *Subsquares and transversals in latin squares*, Ars Combin. **29**B (1990), 193–204.

School of Mathematical Sciences, Monash University, Vic 3800, Australia

E-mail: joshua.browning@sci.monash.edu.au

Department of Mathematics, University of Denver, Denver, Colorado 80208, U.S.A.

E-mail: petr@math.du.edu

School of Mathematical Sciences, Monash University, Vic 3800, Australia

E-mail: ian.wanless@sci.monash.edu.au

(*Received* October 22, 2009, *revised* January 25, 2010)