Substitutes for the non-existent square lattice designs for 36 treatments

R. A. Bailey
University of St Andrews

British Combinatorial Conference
University of Birminham, 30 July 2019
Mini-Symposium on Designs and Latin Squares

Joint work with Peter Cameron (University of St Andrews), Leonard Soicher (Queen Mary University of London) and Emlyn Williams (Australian National University)

Resolvable block designs

Trials of new crop varieties typically have a large number of varieties.
Even at a well-run testing centre,
inhomogeneity among the plots (experimental units) makes it desirable to group the plots into homogeneous blocks,
usually too small to contain all the varieties.
For management reasons, it is often convenient if the blocks can themselves be grouped into replicates, in such a way that each variety occurs exactly once in each replicate. Such a block design is called resolvable.

Square lattice designs

Yates (1936, 1937) introduced square lattice designs for this purpose. The number of varieties has the form n^2 for some integer n, and each replicate consists of n blocks of n plots.
Imagine the varieties listed in an abstract $n \times n$ square array.
The rows of this array form the blocks of the first replicate, and the columns of this array form the blocks of the second replicate.

Let r be the number of replicates. If $r > 2$ then $r - 2$ mutually orthogonal Latin squares of order n are needed. For each of these Latin squares, each letter determines a block of size n.

Mutually orthogonal Latin squares

Definition
A pair of Latin squares of order n are orthogonal to each other if, when they are superposed, each letter of one occurs exactly once with each letter of the other.

Here are a pair of orthogonal Latin squares of order 4.

A	B	C	D
B	A	D	C
C	D	A	B
D	C	B	A

α	β	γ	δ
γ	δ	α	β
δ	γ	β	α
β	α	δ	γ

Definition
A collection of Latin squares of the same order is mutually orthogonal if every pair is orthogonal.

Square lattice designs for 16 varieties in 2–4 replicates

1	2	3	4
5	6	7	8
9	10	11	12
13	14	15	16

A	B	C	D
B	A	D	C
C	D	A	B
D	C	B	A

α	β	γ	δ
γ	δ	α	β
δ	γ	β	α
β	α	δ	γ

Replicate 1

1	5	9	13
2	6	10	14
3	7	11	15
4	8	12	16

Replicate 2

1	2	3	4
5	6	7	8
9	10	11	12
13	14	15	16

Replicate 3

1	2	3	4
6	5	8	7
11	12	9	10
16	15	14	13

Replicate 4

1	2	3	4
7	8	5	6
12	11	10	9
14	13	16	15

Using a third Latin square orthogonal to the previous two Latin squares gives a fifth replicate, if required.

All pairwise variety concurrences are in $\{0, 1\}$.

Efficiency factors and optimality

Given an incomplete-block design for a set $\mathcal{T}$ of varieties in which all blocks have size k and all treatments occur r times, the $\mathcal{T} \times \mathcal{T}$ concurrence matrix Λ has (i, j)-entry equal to the number of blocks in which treatments i and j both occur, and the scaled information matrix is $I - (rk)^{-1}\Lambda$.
The constant vectors are in the null space of the scaled information matrix.

The eigenvalues for the other eigenvectors are called canonical efficiency factors: the larger the better.

Let μ_A be the harmonic mean of the canonical efficiency factors.
The average variance of the estimate of a difference between two varieties in this design is

$$\frac{1}{\mu_A} \times \begin{array}{l}\text{the average variance in an experiment}\\ \text{with the same resources but no blocks}\end{array}$$

So $\mu_A \leq 1$, and a design maximizing μ_A, for given values of r and k and number of varieties, is A-optimal.

Square lattice designs are optimal

Cheng and Bailey (1991) showed that, if $r \leq n + 1$, square lattice designs are optimal among block designs of this size, even over non-resolvable designs.

We have a problem when $n = 6$

If $n \in \{2, 3, 4, 5, 7, 8, 9\}$ then there is a complete set of $n - 1$ mutually orthogonal Latin squares of order n.

Using these gives a square lattice design
for n^2 treatments in $n(n + 1)$ blocks of size n,
which is a balanced incomplete-block design.

There is not even a pair of mutually orthogonal Latin squares of order 6, so square lattice designs for 36 treatments are available for 2 or 3 replicates only.

Patterson and Williams (1976) used computer search to find a design for 36 treatments in 4 replicates of blocks of size 6.
All pairwise treatment concurrences are in $\{0, 1, 2\}$.
The value of its A-criterion μ_A is 0.836, which compares well with the unachievable upper bound of 0.840.

A new design problem: sesqui-arrays

A sesqui-array of order n is an allocation of $n(n + 1)$ letters to the cells of rectangle with $n + 1$ rows and n^2 columns, satisfying conditions (i) and (ii) below.

Example with $n = 3$

D	*H*	*F*	*L*	*E*	*K*	*I*	*G*	*J*
A	*K*	*I*	*B*	*J*	*G*	*C*	*L*	*H*
J	*A*	*L*	*D*	*B*	*F*	*K*	*E*	*C*
G	*E*	*A*	*H*	*I*	*B*	*D*	*C*	*F*

Condition (i) Each letter occurs in all rows except one.

Condition (ii) Each row has n letters in common with each column.

Constructing sesqui-arrays

Tomas Nilson (University of mid-Sweden) and Peter Cameron hoped to give a general construction of sesqui-arrays for all $n \geq 3$.

TN found a general construction, using a pair of mutually orthogonal Latin squares of order n. So this works for all positive integers n except for $n \in \{1, 2, 6\}$.

This motivated PJC to find a sesqui-array for $n = 6$.

Later, RAB found a simpler version of TN's construction, that needs a Latin square of order n but not orthogonal Latin squares. So $n = 6$ is covered. If this had been known earlier, PJC would not have found the nice design for $n = 6$.

Naughty but nice

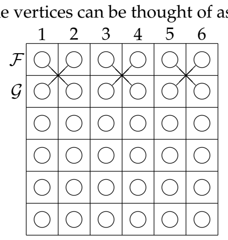
6 is uniquely **BAD** amongst positive integers in that
it is big enough to have a pair of orthogonal Latin squares
but there are no such squares.

6 is uniquely **GOOD** amongst positive integers in that
the symmetric group S_6 of all permutations of $\{1, 2, 3, 4, 5, 6\}$
has an automorphism σ which is not of the form $\sigma(g) = h^{-1}gh$.

This can be used to construct the Sylvester graph,
which has 36 vertices, all with valency 5.

The Sylvester graph

The vertices can be thought of as the cells of a 6×6 grid.



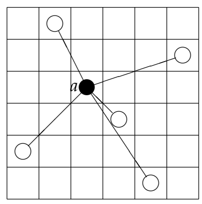
Rows are labelled by the one-factorizations (edge-colourings) of K_6.

$$\mathcal{F} = ||12|34|56||13|25|46||14|26|35||15|24|36||16|23|45||$$
$$\mathcal{G} = ||12|34|56||23|15|46||24|16|35||25|14|36||26|13|45|| = \mathcal{F}^{(12)}$$

Automorphisms: S_6 on rows and on columns at the same time;
the outer automorphism of S_6 swaps rows with columns.

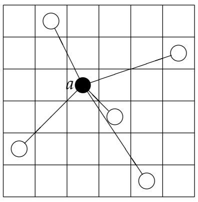
The Sylvester graph and its starfish

The Sylvester graph Σ has a transitive group of automorphisms (permutations of the vertices which take edges to edges), so it looks the same from each vertex.



At each vertex a, the *starfish* $S(a)$ defined by the 5 edges at a has 6 vertices, one in each row and one in each column.

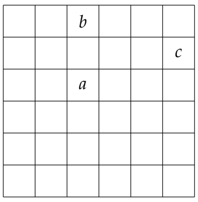
Pedantic naming



When I started to explain these ideas,
I called this set of six vertices the spider centred at a.
Peter Cameron pointed out that spiders usually have more than five legs, whereas some starfish have five.

A real starfish

Starfish whose centres are in the same column



If there is an edge from a to c and an edge from b to c
then the starfish $S(c)$ has two vertices in the third column.
This cannot happen,
so the starfish $S(a)$ and $S(b)$ have no vertices in common.
So, for any one column,
the 6 starfish centred on vertices in that column do not overlap,
and so they give a single replicate of 6 blocks of size 6.

The galaxy of starfish centered on column 3

D	A	B^*	C	E	F
F	E	C^*	B	D	A
E	B	A^*	D	F	C
B	F	D^*	A	C	E
A	C	E^*	F	B	D
C	D	F^*	E	A	B

This is a Latin square.

Constructing resolved designs with r replicates

For $r = 2$ or $r = 3$:

Replicate 1	the blocks are the rows of the grid
Replicate 2	the blocks are the columns of the grid
Replicate 3	the blocks are the starfish of one particular column

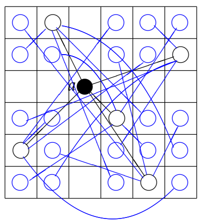
These are square lattice designs.

For $r = 4$, $r = 5$, $r = 6$, $r = 7$ or $r = 8$ we can construct very efficient resolved designs using some of

- all rows of the grid
- all columns of the grid
- all starfish of some columns.

Note that, if there is an edge from a to c in the graph, then varieties a and c both occur in both starfish $S(a)$ and $S(c)$.
So if we use the galaxies of starfish of two or more columns then some treatment concurrences will be bigger than 1.

More properties of the Sylvester graph



Vertices at distance 2 from a are all in rows and columns different from a.
The Sylvester graph has no triangles or quadrilaterals.

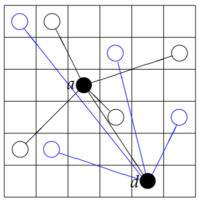
This implies that, if a is any vertex, the vertices at distance 2 from vertex a are precisely those vertices which are not in the starfish $S(a)$ or the row containing a or the column containing a.

Consquence: concurrences

The Sylvester graph has no triangles or quadrilaterals.

Consequence

If we make each starfish into a block, then the only way that distinct treatments a and d can occur together in more than one block is for vertices a and d to be joined by an edge so that they both occur in the starfish $S(a)$ and $S(d)$.



Our designs

- $*^m$ galaxies of starfish from m columns, where $1 \le m \le 6$
- R, $*^m$ all rows; galaxies of starfish from m columns
- C, $*^m$ all columns; galaxies of starfish from m columns
- R, C, $*^m$ all rows; all columns; galaxies of starfish from m columns,

If $m = 6$ then the design is partially balanced with respect to the association scheme whose classes are edges; pairs at distance 2; pairs in the same row; pairs in the same column: so we can easily calculate the canonical efficiency factors. Otherwise, we use computational algebra (GAP) to calculate them exactly.

The large group of automorphisms tell us that

- the design R, $*^m$ has the same canonical efficiency factors as the design C, $*^m$;
- if we use the galaxies of starfish from m columns it does not matter which subset of m columns we use.

Values of μ_A for our designs

r	R, C, $*^{r-2}$	C, $*^{r-1}$	$*^r$	HDP/ERW 1976	square lattice
3	0.8235				0.8235
4	0.8380	0.8341	0.8285	0.836	0.8400
5	0.8453	0.8422	0.8383		0.8485
6	0.8498	0.8473	0.8442		0.8537
7	0.8528	0.8507			0.8571
8	0.8549				0.8547

Highlighted entries correspond to partially balanced designs.
Blue entries correspond to designs which do not exist.

Personal communication from Emlyn Williams

I gave a talk about these designs in August 2017 at the meeting on *Latest advances in the theory and applications of design and analysis of experiments* in the Banff International Research Station in Canada.

They video all lectures, and make them available on the web.

Emlyn Williams learnt about this, and watched the video of my lecture.

This motivated him to re-run that computer search from the 1970s with a more up-to-date version of his search program on a more up-to date computer.

Thus he found resolvable designs for 36 varieties in up to eight replicates of blocks of size six.

All concurrences are in $\{0, 1, 2\}$.

He emailed me these results in September 2017.

Another connection

I gave another talk about these designs in February 2018 in a seminar in St Andrews.

As I was preparing the talk (the day before), I realised a connection with some other designs that I have studied, called semi-Latin squares.

What is a semi-Latin square?

Definition

A $(n \times n)/s$ semi-Latin square is an arrangement of ns letters in n^2 blocks of size s which are laid out in a $n \times n$ square in such a way that each letter occurs once in each row and once in each column.

A $(6 \times 6)/2$ semi-Latin square

A L	F K	C H	B G	D I	E J
C I	B J	E F	H L	G K	A D
E K	H I	D G	A F	J L	B C
D J	A E	I L	C K	B F	G H
F G	C D	A B	I J	E H	K L
B H	G L	J K	D E	A C	F I

This one is not made from two Latin squares.

From semi-Latin square to block design

Suppose that we have a $(n \times n)/s$ semi-Latin square.

Construction

1. Write the n^2 varieties in an $n \times n$ square array.
2. Each of the ns letters gives a block of n varieties.

If the semi-Latin square is made by superposing s Latin squares then the block design is resolvable.

Good leads to good

Theorem

If the block design has A-criterion μ_A and the semi-Latin square has A-criterion λ_A then

$$\frac{35}{\mu_A} = 6(6-s) + \frac{6s-1}{\lambda_A}.$$

So maximizing μ_A is the same as maximizing λ_A (among semi-Latin squares which are superpositions of Latin squares, if we insist on resolvable designs).

What is known about good semi-Latin squares with $n = 6$?

Good designs have been found by RAB, Gordon Royle and Leonard Soicher, partly by computer search. Independently, Brickell (1984) found some in communications theory.
In 2013, LHS gave a $(6 \times 6)/6$ semi-Latin square made superposing Latin squares,
so it gives $(6 \times 6)/s$ semi-Latin squares for $2 \leq s \leq 6$.
The table shows values of λ_A.

		not superposed Latin squares				
s	$*^s$	Brickell RAB 1990	RAB/GR 1997	Brickell LHS web	LHS 2013	MOLS SLS
2	0.4889	0.5127	0.5133		0.5116	0.5238
3	0.6730			0.6922	0.6745	0.6939
4	0.7604				0.7614	0.7753
5	0.8111				0.8111	0.8227
6	0.8442				0.8442	0.8537

partially balanced — do not exist

Semi-Latin square to block design: again

Just as with the designs made from the Sylvester graph, if we make a block design from a semi-Latin square then we have the option of including another replicate whose blocks are the rows and another replicate whose blocks are the columns.

As before, these two special replicates give us better designs than just using a semi-Latin square with 12 more letters.

Are any of these designs the same?

r	RAB/PJC R, C, $*^{r-2}$	LHS +R, C	ERW	square lattice
4	0.8380	0.8393	0.8393	0.8400
5	0.8453	0.8456	0.8464	0.8485
6	0.8498	0.8501	0.8510	0.8537
7	0.8528	0.8528	0.8542	0.8571
8	0.8549	0.8549	0.8549	0.8547

It is possible that the LHS and ERW designs for $r = 4$ are isomorphic, and that the RAB/PJC and LHS designs for $r = 7$ are isomorphic. Otherwise, for $4 \leq r \leq 7$, the efficiency factors of the three new designs differ slightly, so no pair of the new designs are isomorphic.

For $r = 8$, all three new designs have the same efficiency factor. Their concurrence matrices are the same up to permutation of the treatments. Their automorphism groups have order 1440, 144 and 1 respectively, so no pair are isomorphic.