Circular designs balanced for neighbours at distances one and two

R. A. Bailey
University of St Andrews / QMUL (emerita)

Scottish Combinatorics Meeting,
University of St Andrews, 24 April 2017

Joint work with Tank Aldred (University of Otago, New Zealand), Brendan McKay (ANU, Australia) and Ian Wanless (Monash University, Australia)

David Finney's 100th birthday cake, January 2017

Take your knife and cut this into ten rows.

(i) Each row has each of ten numbers (0–9) once.
Lay the rows out one after the other to give a sequence of 100 numbers.

(ii) Each ordered pair of numbers (0–9) occurs precisely once as ordered neighbours
(if we imagine that the last entry is repeated before the first entry).

Serially balanced sequences

Finney and Outhwaite (1956) called these serially balanced sequences of type 1 (allow self-neighbours) and index 1 (each pair once). They

- gave such sequences for 2, 6
- showed that there are none for 3, 4 or 5.

Sampford (1957)

- found some for 2, 6, 7, 8, 9, 10, 11, 14, 18, 22
- with no general construction apart from getting a sequence for $2n$ from one for n when n is odd.

Nonyane and Theobald (2007)

- described a computer algorithm which had succeeded in finding such a sequence for all values of n which had been tried, viz. $8, 9, \ldots, 34$.

Our designs: one condition the same, one different

We consider experiments where the experimental units are arranged in a circle or in a single line in space or time.
It is desirable that each ordered pair of treatments occurs just once as neighbours and just once with a single unit in between.
I shall show that a circular design with these properties is equivalent to a special type of quasigroup.

In one variant of this, self-neighbours are forbidden.
In a further variant, it is assumed that the left-neighbour effect is the same as the right-neighbour effect, so all that is needed is that each unordered pair of treatments occurs just once as neighbours and just once with a single unit in between.

I shall report progress on finding methods of constructing the three types of design.

An experiment in marine biology

Richard Cormack (St Andrews) posed me this question in 1993.

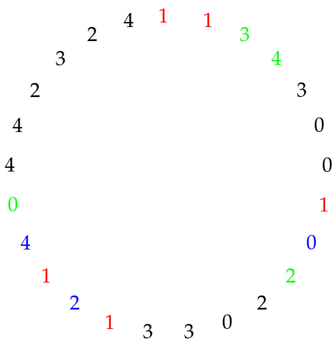
A marine biologist (M. Bayer at St Andrews) wanted to compare 5 genotypes of bryozoan by suspending them in sea water around the circumference of a cylindrical tank. Each genotype was replicated 5 times, so that altogether 25 items were suspended in the tank.

The marine biologist required that

(i) each ordered pair of items should occur just once as ordered neighbours around the circumference of the tank;

(ii) each ordered pair of items should occur just once with a single item in between them, in order.

A circular design for 5 treatments with neighbour balance at distances one and two



Convention

2 4 1 1 3
3 4
2 3
4 0
4 0
0 1
4 0
1 2
2 2
1 3 3 0

Look at the design while standing in the centre of the circle.
Then 'right neighbour' = 'clockwise neighbour'
and 'left neighbour' = 'anti-clockwise neighbour'.

The lazy way to write the design

(1 1 3 4 3 0 0 1 0 2 2 0 3 3 1 2 1 4 0 4 4 2 3 2 4)

2 4 1 1 3
3 4
2 3
4 0
4 0
0 1
4 0
1 2
2 2
1 3 3 0

Statistical model

Denote by $\tau(i)$ the treatment on plot i.

Denote by Y_i the response on plot i.

$$Y_i = \lambda_{\tau(i-1)} + \delta_{\tau(i)} + \rho_{\tau(i+1)} + \varepsilon_i$$

where the ε_i are independent random variables with mean 0 and common variance σ^2.

The direct treatment effects δ,
the left neighbour effects λ
and the right neighbour effects ρ
can be estimated orthogonally of each other
in a experiment of this size
if and only if each pair (λ_j, δ_k) occurs equally often
and each pair (δ_j, ρ_k) occurs equally often
and each pair (λ_j, ρ_k) occurs equally often;
in other words, the design has neighbour balance at distances one and two.

Generalize the original problem

I wanted to prepare myself for future design requests like this.

Can we construct such a neighbour-balanced design
for n treatments each replicated n times
around a circle with space for n^2 items?

Those conditions again

Among the triples of the form

$$(\tau(i-1), \tau(i), \tau(i+1)),$$

each ordered pair of treatments occurs once in positions 1 and 2, once in positions 1 and 3, and once in positions 2 and 3.

Among the triples of the form

$$(\text{row}, \text{column}, \text{letter}),$$

each ordered pair of symbols occurs once in positions 1 and 2, once in positions 1 and 3, and once in positions 2 and 3.

These are conditions for a Latin square
whose rows and columns have the same labels as the letters
—a quasigroup.

Building the design from a quasigroup (Latin square)

The quasigroup operation $\circ$ is defined by

$a \circ b =$ letter in row a and column b of the Latin square.

In the circular design, each triple should have the form

$$(a, b, a \circ b).$$

We can start with any ordered pair (x, y) and successively build the circular design from the quasigroup as

$$x \quad y \quad x \circ y \quad y \circ (x \circ y) \quad (x \circ y) \circ (y \circ (x \circ y)) \quad \cdots$$

Latin square to circle

$\circ$	A	B	C	D
A	B	A	D	C
B	C	D	A	B
C	D	C	B	A
D	A	B	C	D

$(A \quad A \quad B \quad A \quad C \quad D \quad A \quad A$ oops!

This quasigroup gives a design with four separate circles, not one.

$$(A \quad A \quad B \quad A \quad C \quad D)$$
$$(A \quad D \quad C \quad C \quad B \quad C)$$
$$(B \quad B \quad D)$$
$$(D)$$

Eulerian quasigroups

Let's call a quasigroup Eulerian if it gives a single large circle: that is, a sequence with maximal period.

	0	1	2	3	4
0	1	0	2	3	4
1	2	3	1	4	0
2	3	4	0	2	1
3	0	2	4	1	3
4	4	1	3	0	2

(1 1 3 4 3 0 0 1 0 2 2 0 3 3 1 2 1 4 0 4 4 2 3 2 4)

Do Eulerian quasigroups of order n exist?

If $n \leq 4$, a manual check shows that there are none.

For $n = 5$, we have shown an example.

For every other value of n that we have tried,
we have found an Eulerian quasigroup by computer search;
and we can prove that existence for coprime n and m implies existence for mn;
BUT we have been unable to prove that they always exist.

It is quite easy to show that, if $Q = \mathbb{Z}_{p^s}$ or $Q = \mathrm{GF}(p^s)$, then no binary operation of the form

$$x \circ y = ax + by + c$$

makes Q into an Eulerian quasigroup.

Some history

After Richard Cormack posed me the question, Nick Cavenagh (then a PhD student at QMUL, now head of the Department of Mathematics at the University of Waikato) and I got this far, and then got stuck.

In July 1999 I posed the question in the Problem Session at the British Combinatorial Conference in the University of Kent.

Brendan McKay (Computer Science, Australian National University) became interested, and worked on the question with Ian Wanless (then his PhD student, now in the School of Mathematical Sciences at Monash University) and Tank Aldred (Department of Mathematics and Statistics, University of Otago). They invented two variants of the question.

In September 2004 I spent two weeks at ANU working with BDM and IMW (and remotely with RELA). We solved the two variants completely.

Variant I: no self-neighbours

Sometimes it is undesirable to have the same treatment on neighbouring plots.

We need a circular design with $n(n-1)$ plots in which each

- each ordered pair of distinct treatments occurs just once as ordered neighbours;
- each left-neighbour treatment occurs just once with all but one of the right-neighbour treatments.

The incidence of
direct treatments with left-neighbour treatments is a symmetric balanced incomplete-block design (BIBD, aka 2-design);
direct treatments with right-neighbour treatments is a symmetric BIBD;
left-neighbour treatments with right-neighbour treatments is a symmetric BIBD.

Preece (1975 ACC, Adelaide) showed that, for overall balance, the missing pairs at distance two must also be the self-pairs.

Idempotent Eulerian circular sequences

We need a circular design with $n(n-1)$ plots in which each

- each ordered pair of distinct treatments occurs just once as ordered neighbours;
- each left-neighbour treatment occurs just once with every right-neighbour treatment except itself.

The results of Druilhet (1999) show that such designs are optimal for the estimation of direct effects and neighbour effects, in the sense of minimizing average variance of these estimators.

A quasigroup is idempotent if $x \circ x = x$ for all x.

Our circular design is equivalent to an idempotent quasigroup in which the $n(n-1)$ off-diagonal cells give a single circle.

Construction when $n = 6$ (in general, when n is even)

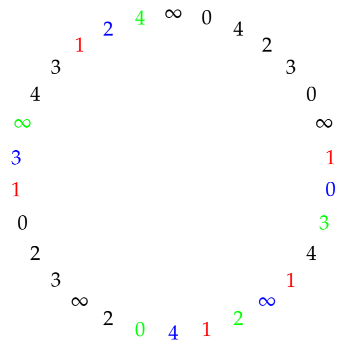
The treatments are the integers modulo 5, together with ∞.

linear sequence	$[4, 3, 1, 2]$	all different, non-zero
neighbour sums	$[2, 4, 3]$	all different, non-zero, non-1
sum of ends	1	must be 1 (consequence)
cumulative sums	$[0, 4, 2, 3, 0]$	last must be 0 (consequence)

$(\infty\ 0\ 4\ 2\ 3\ 0\ \infty\ 1\ 0\ 3\ 4\ 1\ \infty\ 2\ 1\ 4\ 0\ 2\ \infty\ 3\ 2\ 0\ 1\ 3\ \infty\ 4\ 3\ 1\ 2\ 4)$

Neighbours of ∞ at distances one and two are OK, by cyclic construction.
Differences at distance one come from the original sequence; most differences at distance two are the neighbour sums.
$1 -$ last cumulative sum $= 1 - 0 = 1 =$ missing neighbour-sum so differences at distance two either side of ∞ give this.

A circular design for 6 treatments with no self-neighbours at distance one or two



$(\infty\ 0\ 4\ 2\ 3\ 0\ \infty\ 1\ 0\ 3\ 4\ 1\ \infty\ 2\ 1\ 4\ 0\ 2\ \infty\ 3\ 2\ 0\ 1\ 3\ \infty\ 4\ 3\ 1\ 2\ 4)$

Construction when $n = 7$ (in general, when n is odd)

The treatments are the integers modulo 6, together with ∞.

linear sequence	$[4, 1, 2, 5, 3]$	all different, non-zero
neighbour sums	$[5, 3, 1, 2]$	all different, non-zero, non-4
sum of ends	1	must be 1 (consequence)
cumulative sums	$[0, 4, 5, 1, 0, 3]$	last must be 3 (consequence)

$(\infty\ \ 0\ \ 4\ 5\ 1\ 0\ 3\ \infty\ \ 1\ \ 5\ 0\ 2\ 1\ 4\ \infty\ \ 2\ \ 0\ 1\ 3\ 2\ 5$
$\ldots\ \infty\ 3\ 1\ 2\ 4\ 3\ \ 0\ \ \infty\ 4\ 2\ 3\ 5\ 4\ \ 1\ \ \infty\ 5\ 3\ 4\ 0\ 5\ 2)$

Neighbours of ∞ at distances one and two are OK, by cyclic construction.
Differences at distance one come from the original sequence; most differences at distance two are the neighbour sums.
$1 -$ last cumulative sum $= 1 - 3 = 4 =$ missing neighbour-sum so differences at distance two either side of ∞ give this.

Solution for variant I

Theorem
Given an initial sequence of the non-zero integers modulo $n - 1$ satisfying those conditions, that construction always produces an idempotent Eulerian circular sequence.

Theorem
Such an initial sequence can be constructed whenever $n \geq 6$.

Paradigm

$$\begin{array}{c} \text{Small thing} \\ \text{beautiful} \end{array} \xrightarrow{\text{algorithm}} \begin{array}{c} \text{Large thing} \\ \text{good} \end{array}$$

Theorem
If small is beautiful then large is good.

- Work out the algorithm.
- Find the appropriate definition of 'beautiful'.
- Prove the theorem.

Theorem
I can construct a small beautiful thing for almost all values of n.

- Find a construction (which may differ for different residues modulo something).
- Prove that it works.

Variant II: undirectional neighbour effects

Suppose that the effect of the neighbouring treatment is the same whether it is from the left or the right.

$$Y_i = \lambda_{\tau(i-1)} + \delta_{\tau(i)} + \lambda_{\tau(i+1)} + \varepsilon_i,$$

where the ε_i are independent random variables with mean 0 and common variance σ^2.

Can we arrange that every treatment have every treatment as neighbour just once, on one side or the other?

A self-pair gives a self-neighbour on both sides, so we must ban self-pairs. So we need a circle of $n(n-1)/2$ plots.

Each plot has two neighbours, so each treatment has an even number of neighbours, so $n - 1$ must be even.

Any triple (a, b, a) gives b as a neighbour of a on both sides, so there can be no such triples.

Construction when $n = 9$

The treatments are the integers modulo 9.

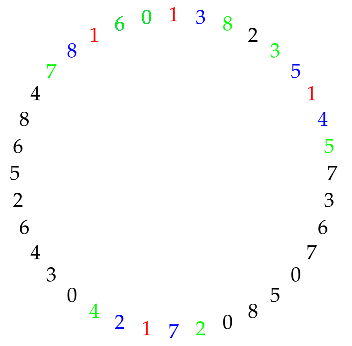
circular sequence	$(1, 2, 5, 3)$	$\pm$ entries are all different
circular neighbour sums	$(3, 7, 8, 4)$	$\pm$ entries are all different
cumulative sums	$[1, 3, 8, 2]$	last one is coprime to 9

(1 3 8 2 3 5 1 4 5 7 3 6 7 0 5 8 0 2 7 1 2 4 0 3 4 6 2 5 6 8 4 7 8 1 6 0)

We keep adding 2 to the original (cumulative) sequence of length 4.
Because 2 is coprime to 9, every pair in the original sequence gets all its shifts modulo 9.

Differences at distance one come from the original sequence; difference at distance two are the neighbour sums.

A circular design for 9 treatments with undirectional neighbour balance at distances one and two



(1 3 8 2 3 5 1 4 5 7 3 6 7 0 5 8 0 2 7 1 2 4 0 3 4 6 2 5 6 8 4 7 8 1 6 0)

Solution for variant II

Theorem
Given an initial circular sequence of $(n-1)/2$ of the integers modulo n satisfying those conditions, that construction always produces a circular sequence balanced for undirected neighbours at distances one and two.

Theorem
Such an initial sequence can be constructed whenever n is odd and $n \geq 9$. There is also such a circular sequence when $n = 7$.

Back to the original question

A quasigroup of order n with operation $\circ$ is Eulerian if the sequence

$$x \quad y \quad x \circ y \quad y \circ (x \circ y) \quad (x \circ y) \circ (y \circ (x \circ y)) \quad \cdots$$

does not repeat before n^2 steps.

Conjecture
If $n \geq 5$ then there exists an Eulerian quasigroup of order n.

Coprime sizes

Theorem
If $(Q_1, \bullet)$ and $(Q_2, \circ)$ are Eulerian quasigroups of orders n and m, where n and m are coprime, then $Q_1 \otimes Q_2$ is an Eulerian quasigroup of order nm.

Proof.
In the sequence

$$(a, x) \quad (b, y) \quad (a \bullet b, x \circ y) \quad (b \bullet (a \bullet b), y \circ (x \circ y)) \quad \cdots$$

the first coordinates repeat every n^2 steps, but not earlier, and the second coordinates repeat every m^2 steps, but not earlier. □

Some more history

Email from Ian Wanless to RAB in March–April 2010: we have to finish that paper, so I am coming to visit you in June–July.

RAB thinks: gulp! but Chris Brien is coming to work with me in July–August on two-phase experiments, and I also have to talk at LinStat 2010 in July at Tomar, Portugal.

RAB gives a talk about the problem in Tomar (and starts new work there with Pierre Druilhet on a related problem).

IMW comes to London, and fruitful work gets done.

Email from Ian Wanless on 11 July 2010:

> *Back in Australia now and awake in the middle of the night... but wanted to let you know that in my sleeplessness I've solved that parity question.*

We still have no general construction,
but a paper eventually got written and submitted.

Strategy

Because of the 'coprime' theorem, and because there is no solution for 2, 3 or 4, all we have to do is to find an Eulerian quasigroup for all of the following orders:

- q where q is an odd prime power and $q \geq 5$
- $3q$ where q is an odd prime power
- $2q$ where q is an odd prime power
- $4q$ where q is an odd prime power
- powers of 2 bigger than 4

(and the paper had been accepted before we realised that we also need)

- $3 \times$ all non-trivial powers of 2.

Reminder: the obvious way is no good

If p is prime and $Q = \mathbb{Z}_p$, then no binary operation of the form

$$x \circ y = ax + by + c$$

makes Q into an Eulerian quasigroup.

If $a + b - 1 \neq 0$ and $x = -(a + b - 1)^{-1}c$ then $x \circ x = x$.

If $a + b - 1 = 0$ and $b \neq 2$ and $t = -(b-2)^{-1}c$
then $mt \circ (m+1)t = (m+2)t$ for all integers m,
so we get a circle of size p.

If $a + b - 1 = 0$ and $b = 2$
then ${}^{m}\mathrm{C}_2 c \circ {}^{m+1}\mathrm{C}_2 c = {}^{m+2}\mathrm{C}_2 c$ for all positive integers m,
so we get a circle of size p.

Technique to avoid brute search

If q is odd, try taking $Q = \mathbb{Z}_q$ and putting

$$x \circ y = \pi(x + y)$$

where π is a relatively simple permutation.

For example, when $q = 7$ put $\pi = (0\,1\,2)(3\,4)$ so that

$$4 \circ 5 = \pi(4 + 5) = \pi(2) = 0.$$

This
(the permutation $(0\,1\,2)$ with some adjacent transpositions)
works for all odd numbers that we have tried.

That parity obstacle

Theorem
If n is even then no Eulerian quasigroup can be obtained from a group of order n by permutions of rows, columns or symbols.

… so IMW found another technique to cut down the computer search when n is even.

… for all practical purposes

Theorem
If $n \geq 5$ and there is no Eulerian quasigroup of order n then n is divisible by a prime power exceeding 1000.

But, just as for the problem with serially balanced sequences,
we do not have a general construction and
we do not have a proof that they exist for all large enough n.