Eureka Editor
archim-eureka@srcf.net
The Archimedeans
Centre for Mathematical Sciences
Wilberforce Road
Cambridge CB3 0WA
United Kingdom

Published by The Archimedeans, the mathematics student society of the University of Cambridge

Thanks to the Betty & Gordon Moore Library, Cambridge

54

Eureka

Eureka is the journal of the Archimedeans, which is the Cambridge University Mathematical Society. It is published approximately annually, but since it, like the Society, is run entirely by student volunteers, it is impossible to guarantee precise publication dates. The Society also publishes QARCH, a problems journal.

Subscriptions

A subscription account may be opened by sending at least ten pounds to the Business Manager at the address below; issues will be sent as they are published and you will be informed when your account is running low. It is also possible to subscribe to *Eureka* by standing order. Back numbers of some of the more recent issues are available, and photocopies of any issue may be made available by special arrangement.

Electronic mail

The Society can receive electronic mail via the Internet, and can read most electronic media. Our electronic mail address is given below. There may be long delays in dealing with messages of any kind in the University vacations (especially June to October), and you are asked not to send large amounts of text electronically without prior warning.

Addresses

Correspondence, clearly marked "Editor" or "Business Manager", should be sent to:

Eureka, The Arts School, Bene't Street, Cambridge CB2 3PY, England

Our electronic mail address is:

`soc-archim@lists.cam.ac.uk`

EUREKA

Editor: Alan Bain — Number 54, March 1996

Table of Contents

ISSN 0071-2248

Editorial

Here is *Eureka* **54**, albeit considerably later than was originally intended. This delay was not caused by a lack of articles submitted—in fact some have had to be held over until next issue.

When producing *Eureka* some of the articles have caused me to wonder whether the contibutor's secondary aim was to produce the most fiendish diagrams possible to draw.

Acknowledgements

I should like to thank Dave Harris and Jon Peatfield for arranging computing and printing facilities in DAMTP which have been indispensable for the completition of *Eureka*. I was grateful for the accounts on Austin Donnelly's and Ian Jackson's computers. Colin Bell deserves special thanks for his helpful advice and experience gained from producing *Eureka* **53**, which he willingly shared. Michael Fryers did considerable amounts of work on the diagrams for his joint article with Michael Greene on Sokoban, which helped enormously. Of course I must also thank the authors of the articles without whose efforts there would be no *Eureka*. Peter Benie, the Business Manager has helped in many ways, without which producing *Eureka* would have been impossible. I would like to thank Richard Tucker for drawing the cover. I must also thank everyone who helped proofread countless versions of the articles. All remaining typographical mistakes are of course mine.

The Society

1993-1994

The Archimedeans' year began with the AGM in March. A number of events took place in the Easter term. Some members of the Oxford University Invariant Society were invited to Cambridge for a croquet match, which was won by the Archimedeans. Prof. Frank Kelly won the Lecturer of the Year for the second year running. In the last week of term, we had the garden party in Clare Memorial Court. The weather was very fine and hot, and about fifty people attended. The Barbershop Subgroup entertained those present with a number of songs. Later in term, there was a very enjoyable nocturnal punt trip to Grantchester.

An Extraordinary General Meeting was held in the Easter term, and a new agent, the Minorities Officer, was unofficially introduced to the Society. Two *Eurekas* were published during the year; *Eureka 52* in the Easter term, and *Eureka 53* in the Lent term.

In the Michaelmas term, there was the annual Puzzle Hunt, which was a very enjoyable event. Later in the term, various members of the committee went to Oxford for a treasure hunt organised by the Invariants. This was won by a team consisting of the President, the Chronicler and Juliette White.

We had six speaker meetings during the Michaelmas and Lent terms. Prof. R. Brown of Bangor, Prof. P. I. P. Kalmus of London, Dr F. H. Berkshire of Imperial College, Prof. P. M. Stocker of East Anglia, Prof. W. Ledermann of Sussex and Dr R. Webster of Sheffield all gave interesting talks.

During the Lent term, the Problems Drive was attended by a large number of people from Warwick and Oxford as well as Cambridge. It was a very pleasant event. The main event of the Lent term was the Triennial Dinner. Sir Michael Atiyah and Ian Stewart were invited, and both of them gave very interesting and amusing after-dinner speeches, as did the President.

The college mathematical societies had a large number of meetings. The most exciting of these was the New Pythagoreans' "Just A Minute" competition during the Lent term. Four senior Archimedeans had to give one-minute speeches on vaguely mathematical topics without repetition, hesitation or deviation, and then members of the audience were called upon to have a go.

The Puzzles and Games Ring and the Music Appreciation Subgroup met regularly during the year, and the Barbershop Subgroup met during the Easter term. The year ended in March 1994 with a very exciting AGM at which four people stood for the post of President.

Eva Myers

1994-1995

It seemed as though the year would get off to a shaky start when in April, Prof. Roger Penrose began his speaker meeting by announcing that he had never spoken on the subject before and so would "make up the theory as he went along," but initial fears were allayed by an excellent talk. In June, with revision and Tripos papers a distant memory, Peterhouse played host to the exquisite annual garden party, whose many participants included the lecturer of the year, Dr Ray Lickorish. The next event in our packed social calendar was, naturally, the nocturnal punt trip to Grantchester. This veritable extravaganza of riotous revelry, frenzied feasting, and punting prowess was enlivened further when the Junior Treasurer plunged head first into the murky Cam. To conclude May Week, the Astronomical Society was foolhardy enough to challenge our august society to a game of cricket. The mathematicians emerged victorious after a quite devastating display of erratic bowling from the stargazers had enabled 'Extras' to become our top scorer with 27 not out.

October, and the start of a new academic year brought a host of fresh faces to the society. Membership was down, but with such a scintillating programme of events for the year, headed by an astutely chosen series of speaker meetings, hopes were high. Thus, despite being hit by no fewer than four cancellations, eight of these meetings were given during the year by speakers from establishments such as Oxford, London and the Open University. Their content was varied to say the least, ranging from the continuum hypothesis to quantum computation, and from algebraic equations to hurricane dynamics. Secondly, no Archimedean's year is complete without their having attended at least one of those perennial faves, the Puzzle Hunt, the Problems Drive, or the croquet match against the Invariants. Finally, the Archimedeans were ably represented on the national stage, when two of our committee members formed part of the newly resurrected series of University Challenge. It just goes to show how regular attendance at speaker meetings can increase one's general knowledge!

None may say for sure what the future has set aside for us, but I see more talks, partying and general fun. Don't miss out!

Sean Blanchflower

One-Dimensional Tilings

Michael Greene and Robin Michaels

1. Introduction

The idea of tiling the plane $\mathbb{R}^2$ is familiar; there are many ways this can be done by just repeating one tile. It may seem that the scope for tilings of the line is somewhat limited but in fact, with suitable definitions, this is not the case. Initially, we look at the simplest case of one dimensional tilings.

2. Finite tilings of an interval

DEFINITION. A *one-dimensional tile* T is a finite non-empty subset of $\mathbb{Z}$. Without loss of generality we shall take the smallest integer in T to be 0. A *partial tiling* is a non-empty set of disjoint translates of a tile. The region tiled is the union of these translations. A *tiling of an interval* is a partial tiling where the region tiled is an interval of $\mathbb{Z}$. We use the notation $[n]$ to denote the interval of $\mathbb{Z}$ of length n, starting at 0. For example:

The tile $\{0, 3, 6\}$ is ●○○●○○● and $\{5, 6, 8, 9, 11, 12, 101, 104, 107\}$ is a region tiled by a partial tiling with this tile, as shown:

●○○●○○●·····●○○●○○●
 ●○○●○○●

Also $\{0, 3, 6\}$ tiles the interval $\{0, 1, 2, \ldots, 8\} = [9]$:

●○○●○○●	$0, 3, 6$
●○○●○○●	$1, 4, 7$
●○○●○○●	$2, 5, 8$

The solid interval being tiled here is of length 9, so we speak of the tiling having length 9 (or even, being a tiling of $[9]$).

We can consider a tile as a binary number. If T is a tile then $\tau = \sum_{a \in T} 2^a$ is the *number* of the tile. This can be thought of as writing a 1 at each $a \in T$ and 0s in between, and then reading the number from right to left. Thus the above example has number $1001001_2 = 1 + 8 + 64 = 73$. We will abuse notation a little and not worry too much about distinguishing between a tile T and the binary number τ representing T.

Note that here we are only considering finite tiles, tiling finite intervals of $\mathbb{Z}$, purely by translation. Infinite tiles, tiling of all of $\mathbb{Z}$, and tilings where reflections are permitted, are all worthy of interest and will be considered later.

The interpretation of a tile as a binary number is not merely a representational device—it is also useful in proving results about tiles, as the following proposition shows.

PROPOSITION 1. *For any integer n greater than 1, there are an even number of tilings of $[n]$.*

PROOF. Suppose T is a tile, tiling $[n]$. To translate T by m places, we multiply it by 2^m. Thus, if the tiling of $[n]$ by T involves translating T by $m_1, m_2, m_3, \ldots, m_r$, then we can write

$$T(2^{m_1} + 2^{m_2} + \cdots + 2^{m_r}) = 2^n - 1,$$

since $[n]$ has binary representation $2^n - 1$. Let $S = 2^{m_1} + 2^{m_2} + \cdots + 2^{m_r}$ so that $TS = 2^n - 1$.

Then S is also a tile for $[n]$, since it can fill the interval with translations given by the positions of 1s in T. We call S a co-tile of T (and T a co-tile of S). In the above example, $T = 1001001_2$ and has co-tile $S = 111_2$ with $ST = 2^9 - 1$.

There can be no 'carries' in the multiplication $TS = 2^n - 1$ since that would correspond to tiles overlapping in the tiling. If $T = S$ then $T^2 = 2^n - 1 \equiv 3 \pmod 4$ for $n \geq 2$, which is impossible.

Thus, we have a bijection from the set of tilings of $[n]$ to itself, given by tile $\to$ co-tile, and this has no fixed point, so we have an even number of tiles. This implies that a tile has at most one co-tile with respect to $[n]$. □

In the course of proving this result, the concept of co-tile has been introduced. We write $\mathrm{co}_n(T)$ for the co-tile of T with respect to a tiling of $[n]$. Where no ambiguity exists we write $\mathrm{co}(T)$. We can rapidly deduce another result.

DEFINITIONS. $\#T =$ the number of 1s in the binary representation of $T = |T|$

THEOREM 2. $x = \mathrm{co}_n(y)$ *if and only if* $xy = 2^n - 1$ *and* $\#x \cdot \#y = n$

PROOF. ($\Rightarrow$) We have already shown that $xy = 2^n - 1$. Since $[n]$ contains n 1s, and the tile x contains $\#x$ ones, exactly $n/\#x$ copies of the tile must be used. Hence $\#y = n/\#x$.

($\Leftarrow$) Provided there are no 'carries' in the multiplication $xy = 2^n - 1$, then $x = \mathrm{co}_n(y)$. Now, if there are no 'carries', then $\#x \cdot \#y = \#xy$. Further, it is easy to check that any 'carry' occurring in the multiplication must decrease $\#xy$ from this value. Hence $\#xy = \#x \cdot \#y$ if and only if no 'carries' occur, so if $\#x \cdot \#y = n$, our conditions are satisfied. □

If the reader has attempted to find some tilings, they will invariably find that all the tiles they encounter which lead to tilings are symmetric (i.e. palindromic), for example:

$$\begin{array}{lcll} 11000011_2 & \times & 10101_2 & = & 111111111111_2 \\ 111_2 & \times & 10001_2 & = & 111111_2 \\ 1_2 & \times & 1111_2 & = & 1111_2 \end{array}$$

THEOREM 3. *All tiles leading to tilings of intervals are symmetric.*

PROOF. If T has binary representation $a_r a_{r-1} \cdots a_1 a_0$ then we define the polynomial of T,

$$P_T(x) = a_0 + a_1 x + \cdots a_{r-1} x^{r-1} + a_r x^r.$$

Then by the above results, if $S = \mathrm{co}_n(T)$ we have:

$$P_T(x) P_S(x) = 1 + x + \cdots + x^{n-1} = (x^n - 1)/(x - 1)$$

Now, the roots of $x^n - 1$ in $\mathbb{C}$ are all of modulus 1, and so are of the form $e^{i\theta}$, so all the roots of $P_T(x)$ are also of this form. Since $P_T(x)$ has real coefficients its roots occur in complex conjugate pairs. Thus if $P_T(e^{i\theta}) = 0$, then $P_T(e^{-i\theta}) = 0$.

Since all roots of $P_T(x)$ occur in complex conjugate pairs, we have that

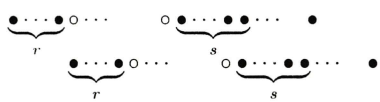
$$P_T(\alpha) = 0 \Leftrightarrow P_T(1/\alpha) = 0$$

But this implies that $P_T(x)$ is a palindromic polynomial; $a_0 = a_r$, $a_{r-1} = a_1$, etc. so T is a symmetric tile. $\square$

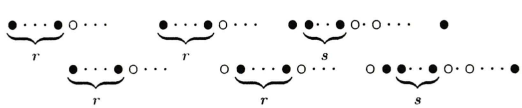
Still more can be said about finite tilings. We ask how many distinct tilings of $[n]$ there are. We denote this by $f(n)$. Then Proposition 1 gives that $2|f(n)$ for $n > 1$. In order to solve this problem, we need to look a little more closely at the structure of acceptable tiles.

Lemma 4 (The Blocks Lemma). *Suppose T tiles an interval. Then*
i) *all blocks of 1s in T are the same length, and*
ii) *all blocks of 0s are a multiple of this length.*

Proof. We think of placing tiles from the left, i.e. starting with the interval empty and putting down one tile each time to cover the leftmost remaining zero. Say T starts with r 1s. If T has just one block of 1s, we are done. Otherwise, the first two tiles go down r apart, so no block of 1s is longer than r. If $s > r$ the tiles overlap:



If any block of 1s has length less than r, take the leftmost such; then nothing can fill the gap between the copies of this block from the first two tiles. If $s < r$ the marked position can never be covered:



This proves part (i). Hence all blocks of 0s in T are completely filled by blocks of length r, so we are done. $\square$
We are now in a position to prove:

Theorem 5. *$f(1) = 1$ and $f(n) = g(n)$ for $n > 1$ where*

$$g(m) = \begin{cases} 2, & m = 1 \\ \sum_{\substack{k|m \\ k \neq m}} g(k), & m > 1. \end{cases}$$

PROOF. We start with yet another definition:

An *isolate tile* is one in which all the blocks of ones are of length 1. It can be seen that the co-tile of an isolate tile is non-isolate (providing the tiling is of length > 1), for otherwise the product of co-tile and tile would start $10\ldots$ and this is clearly not an interval.

Also, the co-tile of a non-isolate tile must be isolate, since otherwise overlaps would result. So, if $g_*(n)$ is the number of isolate tiles, which tile $[n]$ and $g^*(n)$ the number of non-isolate tiles which tile $[n]$, then we have $g(n) = g^*(n) + g_*(n)$, where $f(n) = g(n)$ for $n > 1$, but $g(1) = 2$. (We define $g^*(1) = g_*(1) = 1$.)

We need some more definitions here: *Expanding* a tile by a factor r involves replacing each 1 in the tile by r 1s and each 0 by r 0s. *Diluting* a tile by a factor r involves replacing each 1 by a 1 and $r-1$ 0s, and each 0 by r 0s. For example

$$\begin{array}{llll} 10101_2 & = T & 110011_2 & = S \\ \big\downarrow \text{Expansion by factor 3} & & \big\downarrow \text{Dilution by factor 2} & \\ 111000111000111_2 & = Exp_3(T) & 101000001010_2 & = Dil_2(S) \end{array}$$

By the Blocks Lemma, every non-isolate tiling of n with blocks of length r, is the r-expansion of an isolate tiling of $\frac{n}{r}$; conversely, every r-expansion of an isolate tiling of $\frac{n}{r}$ is a tiling of n.

Thus, since $r|n$, $r \neq 1$ we have

$$g^*(n) = \sum_{\substack{r|n \\ r>1}} g_*(n/r) = \sum_{\substack{k|n \\ k\neq n}} g_*(k).$$

But, for $n > 1$ isolate and non-isolate tiles are in bijection, so that

$$g(n) = 2g^*(n) = 2g_*(n).$$

Thus

$$g(n) = \sum_{\substack{k|n \\ k\neq n}} g(k) \text{ as required, for } n > 1.$$

□

This theorem is not very useful for calculating values of $f(n)$, but a slightly better formula for $g(n)$ can rapidly be derived.

COROLLARY 6. *For $n > 1$, $g(n) = 2\sum_{k|n} \mu(k)g(n/k)$ where $\mu(1) = 1$, $\mu(m) = 0$ if m is not square free, otherwise $\mu(m) = (-1)^s$ where m has s prime factors.*

PROOF. Applying the Möbius Inversion formula, which states that

$$H(x) = \sum_{y|x} h(y)$$

$$\Rightarrow h(y) = \sum_{x|y} \mu(x)H(y/x)$$

to

$$2g(n) = \sum_{k|n} g(k)$$

gives the result. □

In the process of proving Theorem 3, we introduced the concepts of dilution and expansion. Actually, these are dual concepts, since diluting a tile corresponds to expanding its co-tile, and *vice versa*. Thus, the following diagram commutes:

$$\begin{tikzcd} \text{tiles of } [\mathrm{n}] \arrow[r, "\mathrm{co}_n"] \arrow[d, "\mathrm{dil}_r"] & \text{tiles of } [n] \arrow[d, "\exp_r"] \\ \text{tiles of } [\mathrm{rn}] \arrow[r, "\mathrm{co}_{nr}"] & \text{tiles of } [rn] \end{tikzcd}$$

$$\exp_r(\mathrm{co}_n(T)) = \mathrm{co}_{nr}(\mathrm{dil}_r(T))$$

For example:

$$1\quad 0\quad 0\quad 1\quad 0\quad 0\quad 1 \quad\times\quad 1\quad 1\quad 1 \quad=\quad 111111111$$

$$\overbrace{111}\ \overbrace{000}\ \overbrace{000}\ \overbrace{111}\ \overbrace{000}\ \overbrace{000}\ \overbrace{111} \quad\times\quad \overbrace{001}\ \overbrace{001}\ \overbrace{001} \quad=\quad 111111111111111111111111111$$

Notice that $\#\mathrm{dil}_r(T) = \#T$ and $\#\exp_r(T) = r\#T$.

Now, we know that every non-isolate tile is the expansion of some smaller isolate tile, from the Blocks Lemma. Since an isolate tile has a non-isolate co-tile which is the expansion of a smaller tile, we know from the duality of expansion and dilution, an isolate tile is therefore the dilution of a non-isolate tile.

So, by repeated inverse dilution (*concentration*) and inverse expansion (*contraction*), every tile can ultimately be reduced to the unit tile. For example :

●●○○○○●●○○○○●●○○○○
↓ contract 2
●○○●○○●○○
↓ concentrate 3
●●●
↓ contract 3
●

●○●○●○○○○○○○●○●○●○○○○○○○●○●○●○○○○○○○
↓ concentrate 2
●●●○○○●●●○○○●●●○○○
↓ contract 3
●○●○●○
↓ concentrate 2
●●●
↓ contract 3
●

These numbers, the factors of contraction and concentration occuring in the reduction of a tile, uniquely define the tile. The sequence, written in reverse order, is called the *signature* of the tile, so the examples above have signatures [3, 3, 2] and [2, 2, 3, 3] respectively.†

By reversing this process, any sequence of integers (all greater than 1) is the signature of a unique tile. For example, the sequence [4, 2, 3, 2] generates the following tile:

●○●○●○○○○○○○○○○●○●○●○○○○○○○○○○●○●○●○○○○○○○○○○●○●○●○○○○○○○○○○

If there are an even number of numbers in the sequence, then the final operation is a dilution so the tile is isolate, while if not, an expansion is final, so the tile is non-isolate.

Also $\#T$ is equal to the product of the odd position integers in the signature, since it is the expansions that increase $\#T$. Further, the length of the shortest inverval tiled by T is the product of all the integers in the signature. We leave this to the reader to prove. As a summarising example, $f(12) = 16$, and the tiles are:

1	11	10101010101	[2]	[6, 2]
2	111	1001001001	[3]	[4, 3]
3	1111	100010001	[4]	[3, 4]
4	111111	1000001	[6]	[2, 6]
5	111111111111	1	[12]	[1]
6	1100110011	101	[3, 2]	[2, 2]
7	111000111	1001	[2, 2, 3]	[2, 3]
8	11000011	10101	[2, 3, 2]	[3, 2]

The reader may care to consider the effects of dilution and expansion on the polynomial representations of a tile.

3. Higher dimensional tilings

In $\mathbb{Z}^n$, the argument of the Blocks Lemma extends to give

THEOREM 7. *If T tiles a cuboid in $\mathbb{Z}^n$ by translation then T is the cartesian product of one-dimensional tiles.*

PROOF. Write $[k]$ for $\{0, 1, \ldots, k-1\}$. Suppose T tiles $A = [a_1] \times [a_2] \times \ldots \times [a_n]$ and no smaller cuboid. Write T_0 for the tile 'in the corner', that is, the copy of T which contains $(0, 0, \ldots, 0)$, and define T_0^* similarly, where T^* is the co-tile of T.

If $A = \{(0, 0, \ldots, 0)\}$, we are done; otherwise, we may assume that $a_1 > 1$. Look at the edge

$$E_1 := [a_1] \times \{0\} \times \ldots \times \{0\}.$$

The only part of T which meets E_1 is the corresponding edge of the tile, that is, the edge $T_0 \cap E_1$, so this edge must be a one-dimensional tile, say with block length r. Suppose E is any edge parallel to E_1, and T' is any tile in the tiling. We claim that any block of 1s in $T' \cap E$ has length r, and any block of 0s has length a multiple of r (including those at the ends)—if not, set $T' = T_0$ and choose an E 'nearest' E_1, say with minimum nth coordinate, then minimum $(n-1)$th coordinate, etc. down to

† Note that this doesn't lead to an ambiguity of notation!

minimum 2nd coordinate, such that the claim fails for $T' \cap E$, and argue as in the Blocks Lemma. Therefore, if $r > 1$ we may contract T in the 1-direction and get a tile which tiles $[a_1/r] \times [a_2] \times \ldots \times [a_n]$, so we are done by induction on the volume of A (since the expansion of a cartesian product is a cartesian product).

If $r = 1$, look at $T_0^* \cap E_1$. This must be non-isolate, so by applying the above reasoning to T^* we can contract T^* in the 1-direction. T^* tiles A, so the volume of the smallest cuboid tiled by T^* is at most the volume of A, and the co-tile of a cartesian product is also a cartesian product (by induction on dimension), so we are done. □

4. Tilings of $\mathbb{Z}$

So far, we have considered only finite tilings of finite intervals. Now tilings of all of $\mathbb{Z}$ are considered. Clearly, if a finite tile is to tile $\mathbb{Z}$, then it cannot have a finite co-tile. We can slightly modify our definition of the co-tile in a tiling to deal with this (the binary representation method is obviously flawed here). The co-tile is a sequence of 0s and 1s arranged so that there is a 1 precisely over the left end of each tile in the tiling. Note that this corresponds to the odd definition in the finite case. For example

```
●○○●
 ●○○●
  ●○○●
```

This tiling of $\mathbb{Z}$ is generated by repeating a finite tiling. However, not all tilings of $\mathbb{Z}$ are of this form. For example:

```
      ●○○●
   ●○○●●○○●
●○○●    ●○○●· · ·
```

Note that the way a particular tile can tile $\mathbb{Z}$ is not unique, unlike in the finite case. ●○○● can also tile $\mathbb{Z}$ as shown, with a different co-tile:

```
 ●○○●
  ●○○●
●○○●
```

All the tilings of $\mathbb{Z}$ we have seen so far are periodic. We will now prove that this is always the case.

THEOREM 8. *All finite tilings of $\mathbb{Z}$ are periodic.*

PROOF. The partial tiling *induced* by a region of $\mathbb{Z}$ and a tiling is the smallest partial tiling covering the region. We consider the sequence of partial tilings induced by increasing intervals of $\mathbb{Z}$, so that tiles are adjoined to the right end of the block one by one. In such a tiling, the regions at the end of a solid block which are not solidly tiled are called the *interfaces.*

For example

```
Interface
●○○○○○●
       ●○○○○●
    Interface
```

Now, given a tile and interface, with a partial tiling induced by an interval of length greater than that of the tile, there is only one place the next tile can go, since the left-most 0 of a right interface can only be plugged by a tile starting there. Similarly for the left interface. It is clear from this that both interfaces are no longer than the tile.

Another example, using the tile ●○●○○○○○○○●:

```
Left Interface
           ●○●○○○○○○○●
●○●○○○○○○○● ●○●○○○○○○○●
             ●○●○○○○○○○●
                Right Interface
```

From such a sequence of partial tilings, we get a sequence of right interfaces, However, if the tile is of length n, there are at most 2^n possible interfaces, so the pattern must eventually repeat. This however, doesn't preclude the possiblity of an initial non-periodic starting sequence. Starting well into the section of the tiling now known to be periodic we can now work left by the same method, using the left interface, and hence the same argument must apply. Thus the whole tiling must be periodic.

This shows that the task of tiling $\mathbb{Z}$ can be reduced to that of tiling C_n where this is the circle of circumference n.

Now, every tiling of $[n]$ gives clearly also a tiling of $\mathbb{Z}$, and a method of constructing tilings of $\mathbb{Z}$ from such tilings is shown. An example will demonstrate the general principle: ●○●○● tiles [6]
Now, an integer is assigned to each • of the tile:

```
a b c
0 -1 2
●○●○●
```

and then that • is shifted by that number multiplied by the tiling length:

```
b    a              c
-1   0              2
●○○○○●○○○○○○○○○○○○●
```

The reader may care to check that this tiles $\mathbb{Z}$. □

DEFINITIONS. A tile is *prime* if it isn't the dilution of some tile. A tiling is *irreducible* if there is no partial tiling in it with an empty interface. It is conjectured that all prime tilings of $\mathbb{Z}$ are of this form, but this has not been proved. The canonical production of a tiling of $\mathbb{Z}$ from a tiling of $[n]$ leads to a reducible tiling, but the above methods allows irreducible tilings to be produced.

5. Tilings with reflections

Now the extra structure given by allowing reflections of the tile to be used is considered. In this situation, the results proved earlier are all false. For example, new structure *is* added; asymmetric tilings exist. Here is the smallest known such tiling of [9].

```
 ●●○○○●
   ●○○○●●
●○○○●●
```

In this case the ratio of left-handed to right-handed tiles is 2 : 1. It is weakly conjectured that no more extreme ratio is possible for asymmetric tiling of a finite interval. This is known as the $2:1$ conjecture, and is weakly believed by the authors, but little evidence has been found either way.

The tile-co-tile duality found in the symmetric case no longer applies, and asymmetric tilings are far harder to work with. It is strongly suspected that a tiling of a symmetric region with a ratio of $3:1$ is possible, but the example has been lost.

The *discrepancy* of an asymmetric tiling is the difference between number of left-handed and right-handed tiles. We now show that prime tilings exist with arbitrarily large discrepancy.

PROPOSITION 9. *There exist prime tilings with arbitrarily large discrepancy.*

PROOF. We use the prime tile $T =$●●○○○○○○● This can be used to tile the regions $A =$●●●○○○●●●●●●●●●○○○●●● and $B =$●●●○○○●●●●●●●●●●●●●●●●●●●●●●●●●●●●●● ($\bar{B}$ is the reverse of B.) as follows:

```
●●●○○○●●●●●●●●●○○○●●●
●○○○○○○●●
 ●●○○○○○○●
      ●○○○○○○●●
          ●●○○○○○○●
            ●○○○○○○●●
```

```
●●●○○○●●●●●●●●●●●●●●●●●●●●●●●●●●●●●●
●○○○○○○●●
 ●●○○○○○○●
      ●○○○○○○●●
          ●○○○○○○●●
           ●●○○○○○○●
               ●●○○○○○○●
                    ●●○○○○○○●
                      ●○○○○○○●●
                        ●○○○○○○●●
                         ●●○○○○○○●
                           ●○○○○○○●●
```

An interval of length $15n + 66$ can be tiled by T with discrepancy n by use of the units A, B, $\bar{B}$ which fit together, as shown:

```
●●●●●●· · ·●●●○○○●●●
           ●●●○○○●●●●●●●●●○○○●●●
                         ●●●○○○●●●· · ·●●●●●●
```

Since A has discrepancy 1, this has discrepancy n, and the result follows. It is worth noting that this tile also leads to a tiling of $\mathbb{N}$, by using the same units:

$$\bar{B}AAA\cdots$$

This tiling can be made to have any non-negative integral discrepancy,

Infinite tiles with infinite co-tiles

EXAMPLE. Write each integer n in 'base -2 with digit set $\{0,1\}$, i.e.,

$$n = \sum_{i=0}^{\infty} \epsilon_i(-2)^i, \quad \epsilon_i \in \{0,1\}\ \forall\, i.$$

It is easy to show that such an expression is unique. Set

$$T = \{n : \epsilon_i(n) = 0 \ \ \forall\, i \text{ even}\},$$
$$T^* = \{n : \epsilon_i(n) = 0 \ \ \forall\, i \text{ odd}\}.$$

Then any n is expressible uniquely as $t + t^*$ where $t \in T$ and $t^* \in T^*$, so T tiles $\mathbb{Z}$ with co-tile T^*.

REMARK. This example has half-infinite tiles with half-infinite co-tiles.

The proof that all tilings of $\mathbb{Z}$ are periodic can be modified to show that any tile which tiles $\mathbb{Z}$ with reflection can be made to do it in a periodic way. We leave this to the reader.

6. Acknowledgements

Since starting this investigation of tiles we have had valuable assistance from many people. We are indebted to Imre Leader for Theorem 3, Michael Fryers for the final example, and to Kevin Buzzard, Adam Chalcraft, Richard Tucker and others for much valuable assistance.

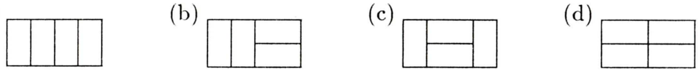
Tatami Mats

Adam Chalcraft

In Japan, a *Tatami* room is a rectangular room, $m \times n$, whose floor is covered by non overlapping *tatami* mats, which are 2×1 rectangles.

There seems to be an unwritten rule that it is not allowed for four tatami mats to meet at a point. Thus of the following possible ways of covering the floor of a 4×2 room, (a), (b) and (c) are allowed, while (d) is not.

(a) (b) (c) (d)



Can all $m \times n$ rooms (with mn even) be covered with tatami mats without violating thls rule? If not, what is the smallest room which cannot be so covered?

The Principle of Least Action

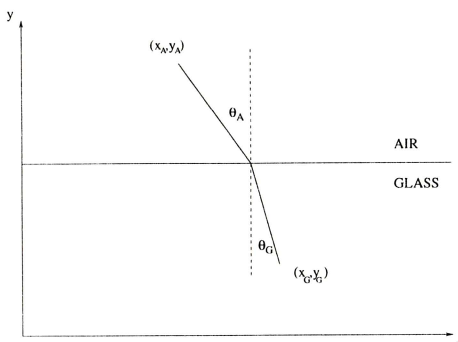
Mark Hayes

Fermat's Principle

Around the year 1650, Fermat made a great discovery – that light rays will always travel from point to point along the path which takes the minimum time. This simple statement explained all the laws of diffraction, reflection and refraction in geometrical optics.

Example – Snell's Law
The time taken to travel an infinitesimal distance dx is proportional to $n(x)\,dx$, where $n(x)$ is the refractive index at the point x. The total time is thus the minimum of:

$$T = \int_{(x_A,y_A)}^{(x_G,y_G)} n\,dx$$



The light rays are observed to be straight lines, and we assume that n_A and n_G are constants, making this integral particularly easy:

$$T = n_A[(x_A - x_0)^2 + {y_A}^2]^{1/2} + n_G[(x_0 - x_G)^2 + {y_G}^2]^{1/2}.$$

Differentiating with respect to x_0 and using the angles θ_A, θ_G we obtain Snell's Law,

$$\frac{n_A}{n_G} = \frac{\sin\theta_G}{\sin\theta_A},$$

where n_A and n_G are the refractive indices of light in air and glass respectively. This means that the apparent speed of light in glass is $\frac{c}{n_G}$ (and similarly in air).

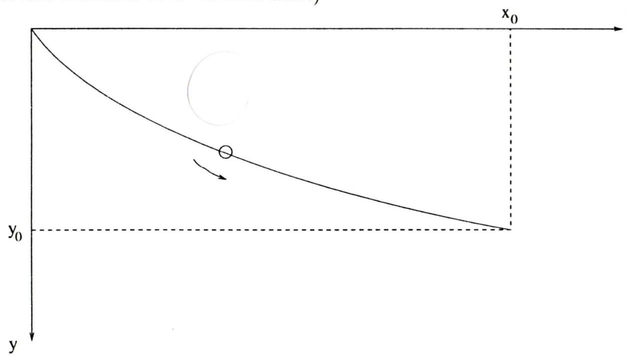
The brachistochrone problem

Fermat's work was followed in 1696 by the formulation of the Brachistochrone problem by Johann Bernoulli.

It is:

If we have a bead on a piece of (frictionless) curved wire, what shaped curve allows the bead to fall in the minimum time?

(Apart from the trivial solution of a vertical line.)



If the y-axis points downwards then the velocity of the bead at any point on the wire will be $\sqrt{2gy}$. Hence the time required to reach the point (x_0, y_0) will be:

$$T = \int_0^{x_0} \left(\frac{1+y'^2}{2gy}\right)^{1/2} dx.$$

This is a slightly harder problem than the path of a light ray since we have both y and y' depending on x, which means that we cannot simply differentiate to find the minimum time. Essentially, the solution to this problem is to deform the wire slightly, so that the bead requires the same amount of time to fall as it did on the undeformed wire. What this means mathematically is that we replace y by $y + \delta y$ and look at the Taylor series to first-order in δy. In general, we have

$$S = \int_{x_0}^{x_1} f[x, y(x), y'(x)]\, dx$$

letting y become $y + \delta y$, so that S becomes $S + \delta S$. We find on subtracting:

$$\begin{aligned} \delta S &= \int_{x_0}^{x_1} \left(\frac{\partial f}{\partial y}\delta y + \frac{\partial f}{\partial y'}\delta y'\right) dx \\ &= \int_{x_0}^{x_1} \left[\frac{\partial f}{\partial y} - \frac{d}{dx}\left(\frac{\partial f}{\partial y'}\right)\right] \delta y\, dx + \left[\frac{\partial f}{\partial y'}\delta y\right]_{x_0}^{x_1}, \end{aligned}$$

where the second term has been integrated by parts. Fixing the wire at the endpoints x_0, x_1 while deforming it, i.e. $\delta y(x_0) = \delta y(x_1) = 0$, gives:

$$\delta S = \int_{x_0}^{x_1} \left[\frac{\partial f}{\partial y} - \frac{d}{dx}\left(\frac{\partial f}{\partial y'}\right)\right] dx.$$

For S to be a minimum we must have $\delta S = 0$. For an arbitrary δy this means

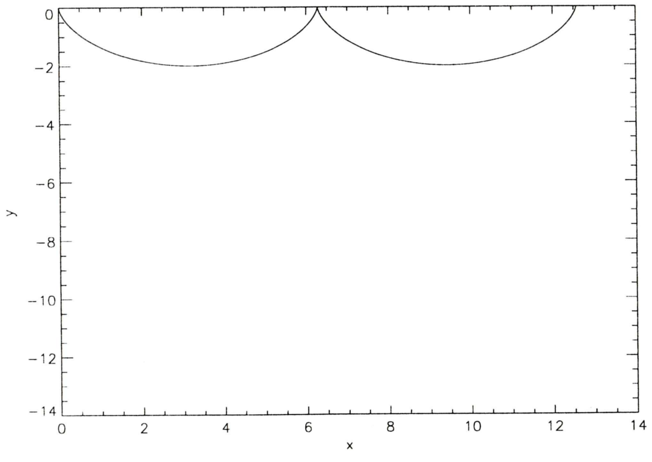
$$\frac{\partial f}{\partial y} - \frac{d}{dx}\left(\frac{\partial f}{\partial y'}\right) = 0,$$

which is known as the Euler-Lagrange equation. Applying it to our falling bead, with

$$f = \left(\frac{1 + y'^2}{2gy}\right)^{1/2},$$

we get a differential equation for the shape of the wire which gives the minimum time of descent:

$$yy'' + \frac{1}{2}(1 + y'^2) = 0,$$



which has the parametric solution:

$$x = a(\phi - \sin\phi),$$
$$y = a(1 - \cos\phi).$$

The solution to this "wonderful and unheard of problem", as Leibniz called it, is the cycloid and was first given by Leibniz himself, along with Newton, Bernoulli and the Marquis de L'hôpital. Actually the cycloid is the curve traced out by a circle of radius a rolling along the x-axis. A particularly neat solution!

This problem made many people wonder how similar ideas could be applied to more general problems in mechanics. In 1669 Leibniz defined the mechanical 'action' of a body of mass m which travels a distance x in time t with velocity v to be $mvx = mv^2t$. The Principle of Least Action was then first formulated by Maupertuis who, in 1746, declared it to be "the most general principle of Nature". Unaware of the work of Leibniz, he considered the problems of elastic and inelastic collisions as well as the refraction of light. However he also appeared to be looking for some theological foundation for mechanics. In his opinion, "perfection of the Supreme Being in His divine wisdom would be incompatible with anything other than utter simplicity and minimum expenditure of action." This interpretation was the subject of much argument and controversy, with Euler, d'Arcy, König, d'Alembert, Voltaire and even the Prussian King Frederick II all taking part. Euler was the first mathematician to consider the Principle of Least Action as an exact analytical statement. For a particle moving along a path between points A and B, he said that the integral

$$S = \int_A^B mv\, ds$$

must be a minimum for the actual path the particle takes, assuming only that the energy is conserved:

$$\frac{1}{2}mv^2 + U(x) = T + U = E = \text{constant}$$

Lagrangians

The mathematician Lagrange took as his starting point a general mechanical system for which he introduced "generalised co-ordinates" $q_1, q_2 \ldots$ which may be angles, distances, etc. With enough generalised co-ordinates to completely specify the system at any given time, he then defined generalised velocities:

$$\dot{q}_i = \frac{dq_i}{dt}.$$

The system can then be thought of as evolving over time in $(q_1, q_2 \ldots)$ "configuration space". Hamilton developed Lagrange's ideas further, producing the modern form of the Principle of Least Action, without having to assume the conservation of energy. He derived the Euler-Lagrange equations from the Principle of Least Action in the form

$$S = \int_{t_1}^{t_2} L(q, \dot{q}, t)\, dt = \min,$$

where L is called the Lagrangian. For example, a particle moving under gravity has $L = \frac{1}{2}m\dot{x}^2 - mgx$. The trajectory of this particle between times t_0 and t_1 will be

$$S = \int_{t_1}^{t_2} L(q, \dot{q}, t)\, dt = \int_{t_1}^{t_2} \left(\frac{1}{2}m\dot{x}^2 - mgx\right) dt = \min.$$

Applying the Euler-Lagrange equation, we recover Newton's equation of motion:

$$\frac{d}{dt}(m\dot{x}) = -mg.$$

Note that in an elementary context, we can only consider conservative forces, like gravity. That is, forces that can be derived from a potential, in this case $-gx$. The final form of the Principle of Least Action is extremely general and finds a place in virtually all areas of physics.

Example – Electrons in an electromagnetic field

If we have an electron with charge e moving with velocity $\dot{\mathbf{x}} = \mathbf{v}$ in an electromagnetic field (where $\mathbf{E}(\mathbf{x}, t) = -\nabla\phi - \frac{\partial \mathbf{A}}{\partial t}$ and $\mathbf{B} = \nabla \times \mathbf{A}$) then the correct Lagrangian to use is

$$L = -m_0\sqrt{1 - v^2} - e(\phi - \mathbf{v} \cdot \mathbf{A}). \qquad (\dagger)$$

The algebra is fairly involved but the Euler-Lagrange equation is equivalent to the Lorentz equation of motion

$$\mathbf{F} = e(\mathbf{E} + \mathbf{v} \times \mathbf{B})$$

We see that the Principle of Least Action and the Lagrangian (†) contain the complete theory of an electron moving in an electromagnetic field.

The Principle of Least Action is perhaps the most basic principle of classical physics in the sense that given the Lagrangian for any particular situation, the laws of motion and everything else follow naturally and easily from it. So the real question behind every modern theory in physics is: "What is the Lagrangian?". A deeper question might be, "where does the Principle of Least Action come from?"

The quantum realm

In classical physics, the equations of motion and the Principle of Least Action are completely equivalent. However this equivalence breaks down in quantum mechanics. Here the equations of motion (e.g. Schrödinger's equation) are only the end-product of a deeper concept, that of a *Feynman path integral*, which in turn will lead to the classical principle of least action.

$$\begin{tikzcd}
\text{Path Integrals} \arrow[r] \arrow[d] & \text{Quantum Mechanics} \arrow[d] \\
\text{Least Action} \arrow[r] & \text{Classical Mechanics}
\end{tikzcd}$$

We can summarise everything we need to know about quantum mechanics in two principles:

(1) The probability $P(a, b)$ of a particle moving from a point a to point b is the square of the absolute value of a complex number, the transition amplitude $\phi(a, b)$:

$$P(a, b) = |\phi(a, b)|^2.$$

(2) The transition amplitude is given by a sum of phase factors, one for each possible path the particle might take. In the simplest case, of non-relativistic QM, this means:

$$\phi(a, b) = \sum_{\text{paths}} c e^{i2\pi S/\hbar},$$

where S is the classical action for each path and c is a normalisation constant which ensures that the probabilty of the particle doing **something** is unity (we shall omit it from here on).

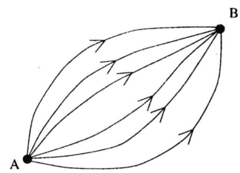
The second principle essentially says that the particle "sniffs out" all the possible paths it might take from a to b, no matter how complicated they may be. We calculate the phase factor for every one of these infinite number paths and then add them up to find the transition amplitude for the particle to go from a to b.

Of course, this sum is really an integral which we could write as:

$$\int_A^B e^{i2\pi S/\hbar}\, Dx,$$

where the Dx indicates that this is actually a product of an infinite number of integrals over all the intermediate points between A and B:

$$\int Dx = \lim_{N\to\infty} \prod_{n=1}^{N} \int d^3x_n$$



Remarkably, the essence of quantum mechanics is completely contained in these two principles. In particular, the strange results of the double slit experiment are easily understood (see the Feynman Lectures on Physics, Vol. III Chapter 1).

The Principle of Least Action says that the actual path the particle will take is the one for which the action is a minimum, that is, to first order, $\delta S = 0$. If we consider the long, complicated paths for which $\delta S \gg \hbar$ then we see that the phase factor $e^{i2\pi S/\hbar}$ for these paths will oscillate rapidly, cancelling out their contribution to the total sum over all the paths:

$$\sum e^{i2\pi S/\hbar} \to 0 \quad \textit{for paths with } \delta S \gg \hbar.$$

The only contributions to the path integral that survive are those for which $\delta S \sim \hbar$. In the classical limit $\hbar \to 0$ this is equivalent to the principle of least action.

So the Principle of Least Action has its origin in the laws of quantum mechanics, where we have to consider all the possible paths a particle might take, or all the possible states a system may be found in. It is difficult to imagine a more fundamental approach and so in the search for the ultimate laws of nature, the principle of least action and the path integral approach to quantum mechanics remain fundamentally important.

Problems Drive 1994

Colin Bell and Michael Fryers

Chapter 1 *In which Kanga and Baby Roo come to the Forest, and Piglet has a bath.*

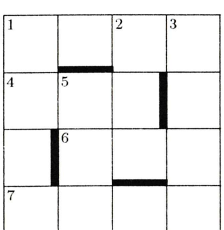
What are the next two terms of the following sequences?
(a) 10, 11, 12, 13, 14, 20, 22, ..., ...
(b) 1, 2, 6, 24, 20, 20, 15, ..., ...
(c) 1, 2, 4, 8, 16, 17, 23, 24, 31, ..., ...
(d) 8, 18, 80, 88, 85, 84, 89, 81, 87, 86, 83, 82, ..., ...
(e) 2, 5, 5, 3, 4, 3, 4, 2, 3, 6, 3, 6, 3, ..., ...

Chapter 2 *Ni cwihh rea ultress eb dunfo ot nyma fuilubate.*

Each of the following is the name of a famous result, conjecture, etc., with the letters of each word jumbled, and the word order changed. Find the original text of each. No credit was given for working out what the title was.
(a) letilt stafmer hotreem
(b) merohet saryphagost
(c) het posheysith innemar
(d) olrocu urof eermoth eth
(e) hoemret netnuflamad eth slauccul fo
(f) openheglio niclipper het
(g) eth hometre nidermare ensheic

Chapter 3 *In which numbers are entirely supplanted by letters.*

In the clues to the following cross-number, all digits have been replaced by letters; each of EILNOPSTUW stands for a different decimal digit. The clues are not necessarily given in numerical order. NB: Juxtaposition of letters in the clues does not represent multiplication—'ELS' is a three digit number, not E × L × S. The answer should be given in the same cipher; in this form, each answer is an English word. No number in the clues or answers is written with any unnecessary leading zeroes.



Across:
E: $E! + ELS$
S: O^{SL}
T: $E \times NN$
W: $O^{T} \times SS$

Down:
N: $O^{E} + N^{T} + T^{P}$
O: $O^{U} + S$
P: $E! - EEI$
S: $SOW + SOWI$

Chapter 4 *In which Timothy invents a new game and the PGR joins in.*

The PGR has Timothy Luffingham to thank for introducing it to the game of Number Boggle; however, as is its wont, it has changed the rules slightly and the game is now known as Prime Number Boggle. Five members of the PGR are having a Prime Number Boggle tournament, in which each of them will play each of the others once in a two-player match. For the sake of anonymity, we shall refer to them as Colin, David, Michael, Paul and Richard.

A Prime Number Boggle *match* consists of five *rounds*, each of which is either won by one of the two players or drawn. If you win the match (by winning more rounds than your opponent) you get two points: if you draw the match you get one point. The tournament will be won by the player with most points; ties are resolved by considering the difference between the total number of rounds won and the total lost (and if this is the same for two players, they share that position).

You have the following information:

A prime number of matches have been played, and a prime number of matches are yet to be played. All five players have a prime number of points, and have won and lost a prime number of rounds. Drawn rounds are common, in fact more than half the individual rounds have been drawn, but there have been at least two non-drawn rounds in each match. As you might expect, the total number of non-drawn rounds is also prime, as is the winning margin in all the non-drawn matches.

Richard currently heads the table, despite losing to Paul. Michael is second and Colin third. Paul and David, who have the same number of points, play each other next week.

The following facts are also of interest:

One match has been a 5-0 whitewash. One player has yet to either win or lose a match. There is a group of three players each of whom has beaten one of the other two by 2-0, and lost to the other by 2-0.

Reconstruct the scores in the matches played.

Chapter 5 *In which it is demonstrated that snakes do not climb ladders.*

Julian is playing Snakes and Ladders on a board of 25 squares, labelled S, $1, \ldots, 23$, F, with 3 snakes and 3 ladders. He starts on S, and finishes on square F with an exact die roll, having in the process landed on each numbered square precisely once. (When he goes up a ladder or down a snake he is considered to have landed on both the start square and the end square.) His rolls are as follows:

1 L 2 6 1 S 5 1 L 5 1 S 1 4 1 5 3 S 2 3 L 3 1 5

where L indicates 'went up a ladder' and S indicates 'went down a snake'.

On which squares are the two ends of each snake and each ladder? (They are all disjoint.)

Chapter 6 *In which someone from Nether Wallop does not have a birthday but receives four presents.*

In Ashby de la Zouch there thrives a bizarre sect which worships Scrawl Eorill, god of logical paradox. Every year at the vernal equinox they perform the following rite: on each of two plates of copper are inscribed four sentences, chosen such that a paradox occurs (i.e. so that it is impossible to assign truth or falsity to all of the sentences simultaneously and consistently). Each plate is then broken in two, and the four pieces, each bearing two sentences, are posted to randomly chosen addresses from the Nether Wallop telephone directory.

By a remote chance, the four half-plates from the 1993 ceremony were all sent to the same address. Do not calculate the probability of this happening; instead work out how the half-plates should be stuck together to reconstruct the original plates (and thereby according to the sect's beliefs, bring about the end of the world).

Here are the four half-plates, in no particular order:

Half-plate A:

Either this is the third sentence on this plate, or the second sentence on this plate is true.
Either this is the second sentence on this plate, or the second sentence on this plate is false.

Half-plate B:

The number of true sentences on this plate is odd.
This is the last sentence on this plate.

Half-plate C:

The second sentence on this plate is true.
The first sentence on this plate is true.

Half-plate D:

The second sentence on this half of the plate is true.
The last sentence on this plate is true.

(References to 'this plate' refer to the complete four-sentence plate.)

Chapter 7 *In which the Sesquipedes learn to count.*

The Sesquipedes of Elmyrk are strange creatures with $1\frac{1}{2}$ legs each. As they also have $1\frac{1}{2}$ fingers on each hand, it is unsurprising that they have developed a notation for numbers based on powers of $1\frac{1}{2}$. Their notation is derived from the simple Roman system—the value represented by a string is the sum of the values of the letters in the string. A=1, B=$1\frac{1}{2}$, C=$(1\frac{1}{2})^2$, D=$(1\frac{1}{2})^3$ etc. (and the Sesquipedes use an infinite alphabet).

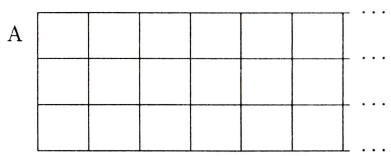
Naturally, Sesquipede mathematicians are interested in the minimum representing string of any number. For any positive integer n, the *length* of n is the length of the shortest string whose value is n. Find:

(a) the least positive integer of length 4;
(b) the greatest positive integer of length 4;
(c) the least positive integer of length 5;
(d) the greatest positive integer of length 5.

Answers should be written in decimal notation.

Chapter 8 *In which Arthur comes to the hallway and has breakfast.*

My accomplished pet mouse Arthur has now been retrained to run through a 3 by infinite rectangular grid of squares I have set up in my (infinitely long) hallway:



If he runs over a square with cheese in, he will eat the cheese, turn left or right at the centre of the square with equal probability, and then continue to the next square. If there is no cheese, he will continue straight on. Arthur enters the grid at A. All of the squares initially contain cheese. What is the probability he leaves it at A?

Chapter 9 *In which a well-known position is turned on its head.*

Given a standard chess starting position but with the pawns removed what is the minimum number of chess moves required to get to the same position but with all the pieces the opposite colour. There is no requirement to move white and black pieces alternately, nor any requirement to keep kings out of check. Castling is not allowed. Partial marks were given for answers which are close but too high.

Chapter 10 *In which certain simplifying assumptions are made about a great city.*

For the purposes of this question, London should be considered to be a cylindrical hole whose base is horizontal, whose axis is vertical, whose radius is 18 miles and whose depth is sufficiently large, say 2000 miles.

A_0, A_1, A_2, etc. are an infinite sequence of solid spheres of identical uniform density, A_n having radius 2^{4-n} miles for all n. They are placed in London so that their overall centre of gravity is as low as possible.

For which n is A_n in contact with the base of London?

Chapter 11 *In which you lead an expedition to Sowmidlandia.*

The Sowmidlandians have kidnapped the noted Cambridge polymath Dr A. N. Other and are holding him prisoner in the Citadel of Notrem. Since no-one can be found to take over his extensive lecturing duties, the University has reluctantly stumped up the ransom of 100 packets of Jaffa Cakes (henceforward abbreviated to pJC) and you have been sent to Sowmidlandia with 102pJC (including taxi fares to and from the station) to bring him back. However, on the long and arduous journey to Droxfo, the capital of Sowmidlandia, you feel somewhat peckish and eat a number of the Jaffa Cakes. Fortunately your guide points out that the Sowmidlandians have yet to discover economics and hence by careful trading you can increase your amount of currency back to the ransom amount. The following locations are of interest:

1. Drofxo Station, where you must bring Dr Other.
2. The Citadel, where you may exchange 100pJC for Dr Other.
3. Pizza Hut, where you may purchase a pizza for 18pJC.
4. Bleckwalls, where you may purchase or sell a copy of *Eureka* for 3pJC or a copy of QARCH for 2pJC.
5. The Ertnyvoc Exchange Emporium, where you may exchange a pizza for seven copies of *Eureka* (either way around) or three copies of *Eureka* for four copies of QARCH (again either way around).

Sowmidlandia is a somewhat dangerous country, particularly when carrying large amounts of currency, and hence the only safe way to travel is by taxi, for which a flat fare of 1pJC is charged regardless of distance.

(a) How many of the original 102pJC can you eat on the journey and still get Dr Other to the station?

(b) Most of the time taken involves travel between Ertnyvoc and Droxfo, so assuming you have eaten as many Jaffa Cakes as possible, find the minimum number of trips to Ertnyvoc required.

Partial marks were given for feasible but non-optimal answers.

Chapter 12 *In which a hypergoat and a hypersheep are taken to a hyperfield and we leave them there.*

An (n^2+1)-dimensional hyperfarmer has two (n^2+1)-dimensional hyperanimals grazing in an (n^2)-dimensional hyperfield: one hypersheep and one hypergoat. The hyperfield is in the shape of an (n^2)-dimensional hypercube with an edge length of 1 furlong, and is currently entirely convered in hypergrass. The hypersheep eats hypergrass at the rate of n^{n^2} furlong$^{n^2}$/fortnight, and the hypergoat at $(n+1)^{n^2}$ furlong$^{n^2}$/fortnight. Hypergrass grows at a negligible rate. The hyperanimals must be tethered, each by a single (1-dimensional) rope to a point in the hyperfield, so that they cannot get at one another to fight, and so that they cannot damage the hyperhawthorn hedgehyperrow, which borders the hyperfield. They must be tethered so aS to maximise the length of time before the hyperfarmer has to move them to new hyperpasture. Find the length of the hypergoat's tether.

Sokoban

Michael Fryers and Michael Greene

In 1983, or thereabouts, a Japanese company called Thinking Rabbit brought out a new kind of sliding block puzzle: the excellent (and now well-known) *Sokoban*†. Guiding the eponymous character around a square grid, you aim to move all the 'barrels' onto marks on the floor denoting 'storage sites'. The only way to move the barrels is to push them, one at a time. That's the whole game. Yet despite its I-could-have-thought-of-that simplicity, Sokoban's levels are varied, addictive, and, in some cases, extremely hard. Here is a simple example level:

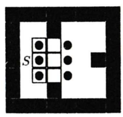


Figure 1.

Walls are immovable and drawn as solid blocks, barrels are round and solid, and storage sites are shown as squares—barrels can be pushed over them in exactly the same way as over unmarked floor. The position of the sokoban will be marked 'S'.

The abstract feel of the kinds of planning needed to solve a level suggests that the game is, in some technical sense, hard. We'll think of a 'level' as a finite connected playing area surrounded by a solid wall with empty space outside it, and the 'size' of the level as its area. To *solve* a level, or part of a level, is to move all barrels there onto storage sites. (We will always have the same number of each.)

Could we write a program which solved, or found impossible, prospective levels? Well, yes: there are only finitely many possible arrangements of the sokoban, S, and the barrels in the playing area, and from each position there are at most four legal moves, so we could turn the whole thing into an enormous directed graph and see if it's possible to get from the start to the end. Hmm. This procedure is finite, but in general the time taken to carry it out is at least an exponential function of the area of the level (see below—even the number of positions with each barrel either where it starts or just below is an exponential function of the area).

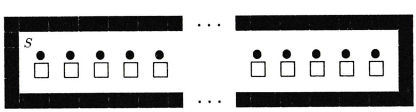


Figure 2.

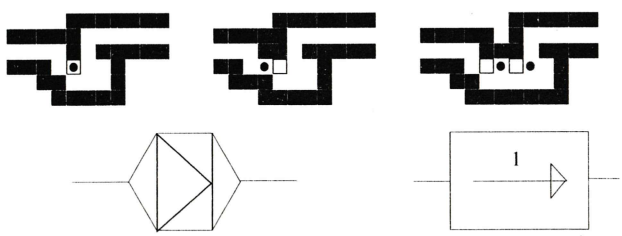
We start with a simplistic summary of some Complexity Theory. It is natural to think of a 'hard' problem as one which takes lots of time to solve. A problem is said to be P (polynomial) if this time is at most a polynomial function of the size of the

† or 'warehouse-keeper'

input. We call a problem *NP* (non-deterministic polynomial) if we can check a suggested solution in polynomial time, or, equivalently, if a 'non-deterministic' computer (which we may assume to be one which makes a choice leading directly to a solution whenever its program is asked to check different cases) could solve it in polynomial time. For example, determining whether a number is composite is NP, because we can check a 'solution', two integers which allegedly factorise our input, simply by multiplying them together and comparing (which takes polynomial time); or, given a program which happened to try dividing the product by one of the factors as its first step, we could get a solution in polynomial time. Any P problem is NP. It is unknown whether *all* NP problems are in fact P; this is an important and emotive question for complexity theorists.

Some problems are so rich and general that they are as hard as any NP problem; that is, any NP problem can be encoded in them in polynomial time. These special, rich problems are said to be *NP-hard*. An NP-hard problem which is also NP is called *NP-complete*. If anyone found a polynomial-time solution to an NP-hard problem, we would know that $NP = P$. Perhaps the simplest NP-complete problem is that of Boolean satisfiability: given a Boolean expression $B(x_1, x_2, \ldots, x_n)$, a string of ands, ors, and nots like '$((x_1 \text{ and } \neg x_2) \text{ or } x_3)$', does there exist a valuation on $x_1, x_2, \ldots, x_n$ (that is, a choice of true or false for each) which makes B true?

Our first aim in this article is to build a given Boolean expression into a Sokoban level in such a way that the level has a solution if and only if the expression can be true, and such that we can construct (in polynomial time) a valuation from a solution, and vice versa. We are thus attemping to show that Sokoban is NP-hard. We shall build levels by plugging together small regions whose behaviour we understand. We can always do this 'cleanly' (i.e. in such a way that barrels can never leave one region and interfere with another)—for example, we can put sharp bends or kinks on the connecting paths so that trying to push a barrel along them gets it stuck. Figure 3 shows some basic units we need, with their symbols.



(a) (b) (c)

Figure 3.

On the left is a *turnstile*, which S can travel through in only one direction. Its barrel must end up on its own storage site (since there is no way to push the barrel out or other barrels in), and there are only two places the barrel can sit without being forever stuck against a wall (recall that S can only push barrels); in either case S cannot travel from right to left. The other direction is possible, however, and in this case, the turnstile is left solved.

Purists who object to barrels starting on storage sites may prefer 3(b), where the barrel can be in a different position. Notice that the input side of this must be visited at least once for it to end up solved (though it need not be traversed). We will use the first turnstile by default, pointing out where it is possible to replace it with the other.

More usefully, Figure 3(c) shows a *1-gate*, which must be traversed exactly once (left to right). It is easy to make n-gates: connect n 1-gates in parallel. Readers may wish to try using these to construct ($\geqslant r$)*-gates*, for any $r > 0$, and ($\geqslant r$)*-and-*($\leqslant s$)*-gates* for $0 < r < s$; each of these should only be traversable left to right, and must be traversed a number of times in the specified range.

Suppose we have a Boolean expression $B(x_1, x_2, \ldots, x_n)$. For definiteness, we construct a level for the string

$$B(x,y,z) = (z \text{ and } ((\neg y \text{ and } \neg z) \text{ or } (\neg x \text{ or } y)))$$
$$\text{where } x \equiv x_1,\ y \equiv x_2, \text{ and } z \equiv x_3.$$

We represent this schematically, as shown in Figure 4.

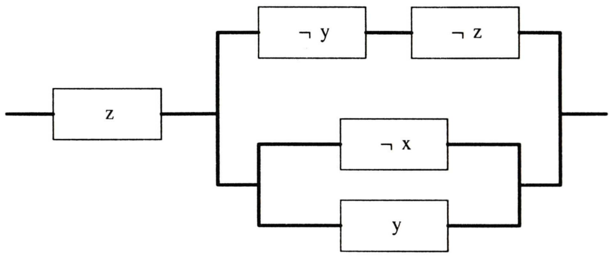


Figure 4.

The idea is as follows. Let's assume S is somehow constrained never to go leftwards. Before S is allowed to attempt to traverse this diagram, he will be forced to block *either* all the x nodes *or* all the $\neg x$ nodes, then *either* all the ys *or* all the $\neg y$s, and so on. If after this S is still able to get right through Figure 4, there is a choice of true or false for each x_i such that B is true, as wanted—if S went through (and therefore blocked) the x_is, set x_i false; if not, set x_i true. If, whatever the choices, S can never get through Figure 4, B is always false. The parts of levels known to the authors, however, are too limited to do this directly; we therefore add more pieces to the level so that S can go round after all this has happened and 'tidy up'.

Choose a minimal set of left-right paths across Figure 4 such that every node is visited—say there are r of them. (In the example above there is no choice, but in

$$((a \text{ or } b) \text{ and } (c \text{ or } (d \text{ or } e))),$$

for example, there is.) To the jth node, associate a number $r_j \geqslant 1$, the number of paths which pass through it. We use the following piece of level for this node:

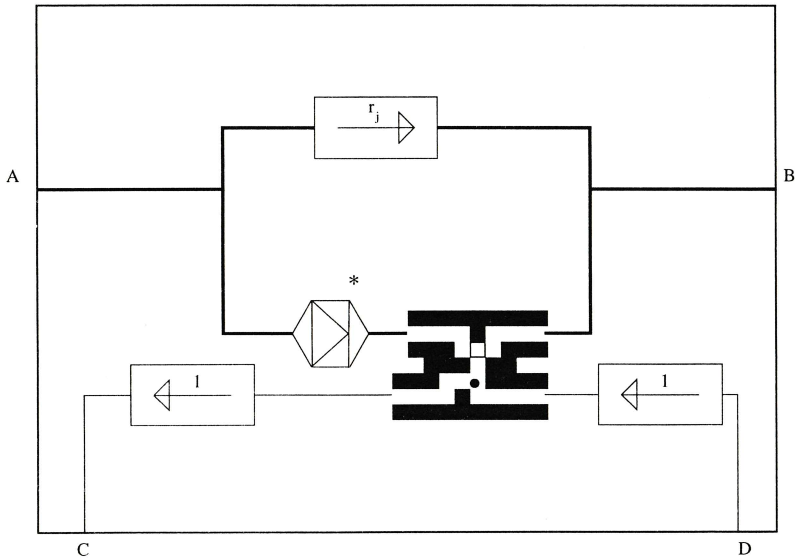


Figure 5.

This block is well-behaved in various ways. In the above state, S can only move the barrel (without it becoming forever stuck in a corner) if he enters by D, so he gets nowhere entering by B or C, and from A must leave by A (pointless) or B. If S enters D and passes through the 1-gate, he must then move the barrel. He is going to have to leave by C, and once he has done so he will never be below the barrel again, so if he only moves it one square and leaves, this node will never be solved; therefore he moves it two, onto the storage site. In this state, entering C or D is useless, and from the top S cannot move the barrel again without getting it stuck—thus the lower bold arm is now impassable. In summary, then, bold and faint tracks are separate, and once S has travelled through from D to C, he can travel through from A to B at most r_j times.

Connect these nodes as shown in Figure 6.

In a certain kind of solution (described below), each crossover unit will be used vertically exactly once, and then horizontally at least once (both in the directions indicated). We use the block shown in Figure 7, as it is possible to leave it solved when traversed in this way. Again, checking the various possible states shows that faint and bold tracks are separate. Now we have specified the whole level, we notice that its total area can be made a polynomial function of the length of the Boolean input string.

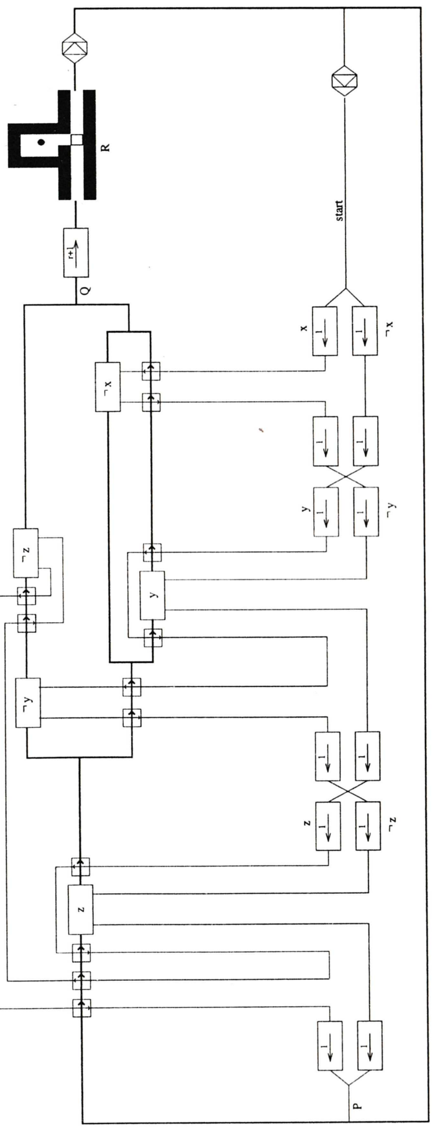
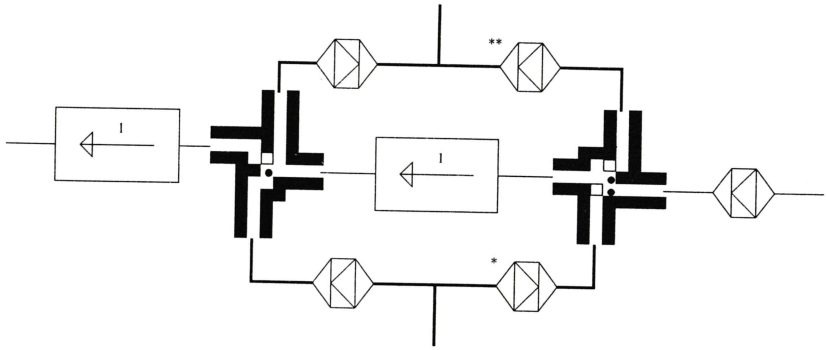


Figure 6



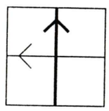


Figure 7.

Let's suppose B can be true. Take a valuation of the variables which makes it so. Then the following solves the level:

(i) from 'start', travel along the path x_1 if x_1 is true and $\neg x_1$ if x_1 is false, and so on for $x_2, \ldots, x_n$;
(ii) travel along bold paths to Q along a route using only the lower parts of nodes, i.e., leaving all the r_i-gates untouched (this is possible since B is true);
(iii) traverse the remaining faint paths, returning to P;
(iv) travel from P to Q along each of the r chosen paths through the bold part of the diagram, returning each time round the edge, and finally complete block R.

When the last of these paths is traversed, each unit will be solved, except perhaps (if we are using the purists' turnstile) for the asterisked turnstiles in nodes and crossovers. However, two of these kinds are visited on the side from which it is possible to solve them, and the third, the double-asterisked turnstiles in the crossovers, can be removed altogether, so the whole level can be completed.

Suppose, conversely, that there is a solution to the level. S passes through the $(r+1)$-gate $(r+1)$ times, and so there must be a set of nodes which were not blocked when S first reached P (though, notice that they need not be traversed immediately)—these connect one side of the bold subdiagram to the other (easy to show by induction). Thus the solution gives a valuation of $\{x_1, \ldots, x_n\}$ such that B is true.

Hence Sokoban is NP-hard. $\square$

We have so far been considering the problem of working out whether a level has a solution. We now turn to the question of the number of moves necessary to solve a level. In particular, if $M(n)$ is the maximum number of moves needed to complete any solvable level of size (area) n, how does $M(n)$ grow as $n \to \infty$?

An upper bound on $M(n)$ can be found by considering the number of states of the level: there are no more than $n\binom{n}{b}$ positions of b barrels and one warehouseman, so if the level can be solved at all, it can be solved in at most this many moves. We shall prove that $M(n)$ is at least exponential in n. Before reading the following construction, the reader is encouraged to try to solve this level, L_1:

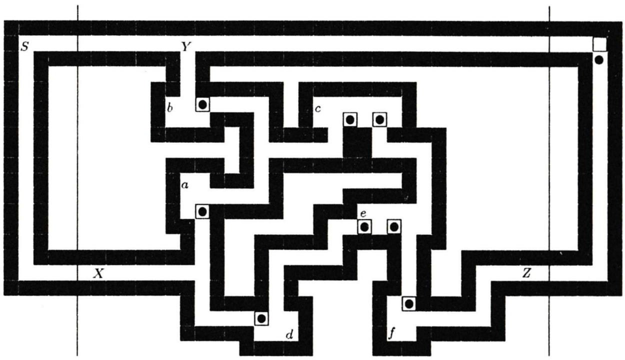


Figure 8.

Call the section of L_1 between the vertical lines A. A contains six 'rooms' containing barrels; four of these (a, b, d, and f) are of a type we've seen before—They are turnstiles. c and e are new, however.

c is a *ferry*. Its two barrels always block at least one of its entrances, so whenever S leaves c and later re-enters it, he must do so by the same entrance. Initially, as shown in Figure 8, its left-hand entrance is open. Its final state must be the same as its initial state, so the last time S leaves c he must do so by the left door.

e is a *1-gate with reset*. It has two states, blocked and unblocked, and it starts out blocked. S can enter from the left only when it is unblocked, but he can enter from the top whichever state it is in. When S leaves, if he does so through the bottom opening, he must leave the gate blocked, while if he leaves through the top he may leave it in either state.

Now suppose S is at X and wishes to get to Z, there being no way to Z except through A. He has the following options:

(i) Go through the gate d. But then he is locked between d and e, unable to pass through either.

(ii) Go through a and then through b to Y. Then even if he can go round the rest of the level back to X, nothing has been gained.

(iii) Go through a and c to e, leaving e by the bottom right and thus reaching Z. But now he will be unable to enter e from the left, f from the right, or c from the left, so he will never be able to restore c to its solved state.

(iv) Go through a and c to e, unblock e and return through c, then go out through b to Y. Then if he or she can get to X again, there is a new option: pass through d, e, and f to Z, leaving the entirety of A in the same state as it started.

This shows that *in order to reach* Z *once, S must reach* X *twice*, and since the states of the rooms in A are reset by this process, we see that *in order to reach* Z k *times, S must reach* X $2k$ *times.*

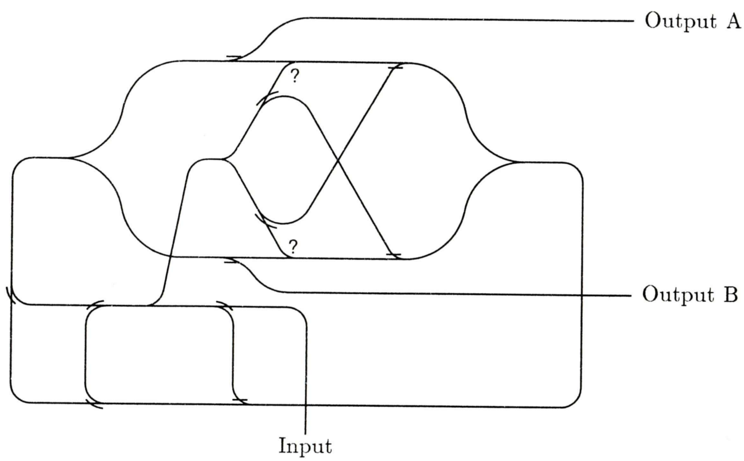
Now make level L_m by plugging together m copies of A, and then putting on the ends just what L_1 has outside the vertical lines. The area of L_m is linear in m, but since to reach the last exit Z once S must pass through the first entrance X 2^m times, the time taken to complete L_m is $O(2^m)$. Thus $M(n)$ is at least exponential in n, as claimed. $\square$

Acknowledgements

Similar results about Sokoban and other sliding-block puzzles were obtained by Dorit Dor and Uri Zwick at about the same time as this article was written; they appeared in SODA 1996, the Seventh Annual ACM-SIAM Symposium on Discrete Algorithms. The second author would like to thank Uri Zwick for a fruitful correspondence and a minor correction.

Train Sets Conjecture Disproved

Clive Monk



In *Eureka* **53**, Adam Chalcraft and Michael Greene conjectured that it is impossible to make a distributor out of finitely many lazy, sprung and random points [1]. This is false as the above counterexample shows. In the diagram the random points are indicated by the ? adjacent to them.

Reference

[1] A. Chalcraft, M. Greene. "Train Sets". *Eureka* **53** (1994) 5–12

Lectures You May Have Missed

Nonlinear Dynamical Systems
'The fixed points will be unique, and there'll only be one of them.'

Methods
'We could take $A = 1$, but then we'd have to multiply by the multiplier of the thing that was multiplying the incident wave.'
'Instead of 5 words, you could write 16 symbols which say the same thing.'

Quadratic Maths
'I've got more energy to clean the board today because I spent my weekend eating black puddings in Leeds.'
'I'm somewhat shell-shocked today, as I was in Oxford yesterday, trying terribly hard to be bright.'

Differentiable Manifolds
'The subject (Differential Topology) is alive and well, and living in Oxford.'
'We can always assume it's finite because:
(i) The manifold is compact
(ii) We're told it is.'

Foundations of Quantum Mechanics
'To become intelligent sometimes requires a bit of thought.'

Logic and Set Theory
'We have a stock of variables x, y, z and so on off the end of the alphabet.'

Methods of Mathematical Physics
'You would have thought that the 1 should be there. But I've put it here. Perhaps it's a misprint.'
'Divergent Series are politically incorrect.'
'And this will be, at the same time, more general and less general!'

Partial Differential Equations
'This lecture is, by definition, almost the most boring lecture that we do.'

Perturbation Methods
'I expect to be heckled – otherwise you all go to sleep.'
'Waves tend to be linear things, except when they're awkward.'

Geological Fluid Mechanics
'We'll start at the beginning. We may not end at the end, but we'll definitely start at the beginning.'
'Roughly half means somewhere between 10 and 90 percent.'
'You're taking too much notice of me.'

Dynamics of Rotating Stratified Flows
'π is one, 2π is ten.'

Waves and Stability Theory
'The solution can't tell whether it's had a mathematical boundary condition imposed, or whether there's a real wall.'

Juggling Generators

Adam Chalcraft

Site-swap notation

There is a mathematical notation describing some juggling patterns which has been known for some time. Independently discovered in at least Cambridge and America at roughly the same time, it has come to be known by the impersonal name of *site-swap* notation.

The idea is that a juggler has h hands (usually $h = 2$) and juggles to a metronome (usually going at about 3 beats per second). The beats are indexed by $t \in \mathbb{Z}$ and the hands cycle, so that hand i is active on beat t when $i \equiv t \pmod{h}$. A hand being *active* means that one ball is thrown just before the beat and another is caught just after. This exchange is known as a *swap*.

The height to which the ball is thrown and the hand to which it is thrown are both determined by the time at which it is supposed to land. This brings us to site-swap notation. A site-swap notation for a juggle is a doubly infinite sequence

$$(\ldots, x_{-1}, x_0, x_1, x_2, \ldots)$$

such that

(i) x_t is a non-negative integer, and

(ii) The map $t \mapsto t + x_t$ is a bijection on $\mathbb{Z}$.

The idea behind the notation is that the ball thrown on beat t lands at the right time for it to be thrown again on beat $t + x_t$. The map in condition (ii) must be an injection so that no two balls land at the same time. There is less reason to force the map to be surjection, but it always is in practice and it makes the maths look nicer.

If we look at the constant sequence $x_t = 3$, for example, a little experimentation will convince the reader that (with $h = 2$) this defines the standard 3-ball juggle, with the ball thrown on beat t caught on beat $t + 1$, to be thrown again on beat $t + 3$.

Indeed, the ball caught on beat t will always be thrown again on beat $t + h$, and so in general the ball thrown on beat t must be caught at time $t + x_t - h$ if it is to be thrown again at the right time. This looks as if it might cause a problem when $x_t \leqslant h$, but we simply define our way out of the difficulties.

Let us assume that $h = 2$. If $x_t = 2$, then on beat t the ball is not thrown at all, it is simply held for that beat. If $x_t = 1$, then the ball is rapidly passed to the other hand (this is used in the pattern usually called a *shower*). Finally, if $x_t = 0$, then there will be no ball in the hand to throw, so no ball is thrown. This generalises neatly to $h = 1$ and slightly less neatly to $h \geqslant 3$. All this is easier to demonstrate than to explain. It is very common for a juggle to be cyclic, in the sense that the notation is a cyclic sequence. In this case, an obvious abbreviation is used, so (3) means $(\ldots, 3, 3, 3, \ldots)$, $(5, 1)$ means $(\ldots, 5, 1, 5, 1, \ldots)$ and so on.

The pattern $(5, 1)$ is a 3-ball juggle, which illustrates a theorem that the average of the numbers in the notation for a cyclic juggle is the number of balls in the pattern.

In particular, it is an integer, which is not immediately obvious from the definition of site-swap notation.

For a cyclic juggle, the word "bijection" in condition (ii) can be replaced by either surjection or injection – they are all equivalent.

All this was very interesting to juggling mathematicians. Almost as soon as the notation had been invented, tables of juggling patterns were written out, and the elegant $(5,0,4)$ and the frenetic $(5,3,1)$ were discovered. Juggles like $(6,7,2)$ and $(6,6,6,6,1)$ are now part of the standard repertoire of some jugglers, though not that of the author, sadly.

Some questions

At this point, the practical issue arises of how to join two cyclic juggles together. To get from (3) to $(5,1)$ and back again, for example, one of the "best" transation sequences seems to be

$$\ldots,3,3,3,3,4,5,1,5,1,5,1,5,1,4,1,3,3,3,3,\ldots \qquad (\dagger)$$

but it is unclear what "best" means in this context.

On an apparently completely different track, it was recently discovered [1] that the number of cyclic juggles using strictly fewer than b balls and of period dividing p is b^p. For example, if $b=2$ and $p=3$ they are the $2^3=8$ juggles

$$(0,0,0)\ (1,1,1)$$
$$(0,1,2)\ (1,2,0)\ (2,0,1)$$
$$(0,0,3)\ (0,3,0)\ (3,0,0)$$

and if $b=3$ and $p=2$, they are the $3^2=9$ juggles

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

The question asked in [1] is *why* this happens. It seems that there ought to be some natural reason for such a neat phenomenon. It turns out that there is a natural reason, and that it also answers the question of how to join two cyclic juggles together.

Juggle generators

A *juggle generator* is a doubly infinite sequence $(\ldots, y_{-1}, y_0, y_1, y_2, \ldots)$ such that

(i) y_t is a non negative integer, and

(ii) $\forall t, \exists u > t : y_u \geqslant y_t$.

The second condition is purely technical, and is automatically satisfied in the interesting cases.

Now there is a transformation from a generator $Y=(y_t)$ to a site-swap notation $X=(x_t)$ which can be described as follows. Write out Y along $\mathbb{Z}$, and at each point in $\mathbb{Z}$ write a copy of $\mathbb{N}$. Now between t and $t+1$ write the permutation

$$0 \mapsto y_t \mapsto y_t - 1 \mapsto \cdots \mapsto 2 \mapsto 1 \mapsto 0.$$

Suppose, for example, that Y is the cyclic generator $(3, 3, 0, 1)$. Then the diagram looks like Figure 1. From this diagram, we can read off x_t. Starting on beat t, follow the path of 0 (the first place it goes to is y_t). Condition (ii) forces it to return to 0 at some point, so suppose it takes n beats to return to 0, then $x_t = n - 1$.

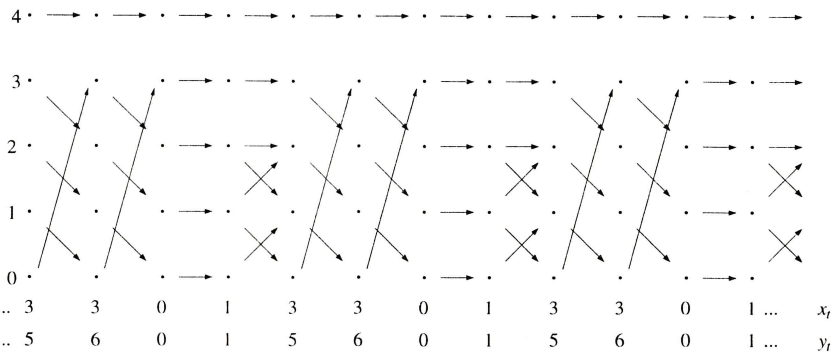


Figure 1

The inverse of this operation is easily read off the diagram:

$$\begin{aligned} y_t &= |\{u : u < t < u + x_u + 1 < t + x_t + 1\}| \\ &= |\{u : t < u < t + x_t + 1 < u + x_u + 1\}| . \end{aligned}$$

It is annoying that it seems to be $x_t + 1$ which is naturally related to y_t, but it does seem to be true. The notation as described has many nice properties.

The site-swap notation (3) corresponds to the juggle generator (3), and so on. Remember that the *average* of the numbers in a cyclic site-swap notation is the number of balls. In a juggle generator, it is the *maximum* of the numbers which gives the number of balls.

Some answers

It is this last fact that explains the b^p observation earlier. It is trivial that the number of ordered p tuples of non-negative integers with entries less than b is b^p and these are exactly the juggle generators for the juggles we are trying to count.

We can also use juggle generators to join two juggles together. Suppose we wish to juggle the site-swap notation (5), switch into the pattern $(6, 7, 2)$, and then drop back down to (5) again.‡ Converting the site-swap (5) into its juggle generator gives (5) and converting $(6, 7, 2)$ gives $(5, 5, 2)$. These can be put together in only one way, giving the juggle generator

$$\ldots, 5, 5, 5, 5, 5, 5, 5, 2, 5, 5, 2, 5, 5, 2, 5, 5, 2, 5, 5, 5, \ldots$$

‡ This may seem an unlikely feat, but it is a question the author was genuinely asked at a juggling convention.

The advantage of using juggle generators here is that there is no need to check the validity of the result. If the two patterns have the same number of balls, then the result will be valid. We can now convert this back into site-swap notation, getting

$$\ldots, 5, 5, 5, 6, 6, 7, 2, 6, 7, 2, 6, 7, 2, 6, 6, 2, 5, 5, 5, \ldots$$

This operation does indeed generate the transitions between (3) and (5, 1) in (1), which shows that it might not be too far from the 'right' answer, whatever that might mean.

Final thoughts

One unforseen use for juggle generators is for generating random juggling patterns. Take an infinite stream of integers, each chosen randomly from $\{0, 1, \ldots, n\}$. Write these down, and prefix them by a left-infinite string of ns. Convert this from a juggle generator to site-swap notation, and the result is an infinite random juggle. A computer program which could accept an infinite site-swap notation would just juggle n balls randomly for ever.

There are various attempts to generalise site-swap notation to cope with two different hands throwing (possibly to different heights) at the same time. This is useful for representing juggling patterns involving more than one juggler. There are also attempts to represent one hand throwing or catching two balls at the same time. This is called *multiplex* juggling.

There is currently no similar generalisation of juggle generators.

Acknowledgements

In Cambridge, site-swap notation was invented by the Mike Day, Colin Wright and the author. Colin Wright now frequently gives very popular talks on the mathematics of juggling.

Reference

[1] Joe Buhler, David Eisenbud, Ron Graham and Colin Wright, "Juggling Drops and Descents". *The American Mathematical Monthly* **101** **[6]** (1994) 507—519.

1995 Problems Drive

Luke Pebody and Richard Tucker

1. to slay the Nemean lion

In Oneida, integers are represented by expressions involving only the number 1 and the operations of addition, multiplication and exponentiation (with brackets to indicate the order in which they should be applied). The Oneidans can represent all the natural numbers this way, except for zero, which strangely does not seem to bother them. They define the wonderful function, $wf(n)$, to be the minimal number of 1s needed to represent n in this way. A natural number n is said to be wondrous if $n = wf(n)$, and wanton otherwise. What is the difference between the largest number n such that $wf(wf(n))$ is wondrous and the smallest n such that $wf(wf(n))$ is wanton?

2. to kill the Lernean Hydra

In the following, different capital letters stand for different digits 0–9; if

$$EUREKA + EUREKA + EUREKA + EUREKA + EUREKA = ARCHIM,$$

what is $QARCH$?

3. to catch and retain the Arcadian stag

Arrange the ten jumbles of letters below in five pairs, then use the letters in each pair to make the surnames (ignoring accents) of two famous mathematicians.

1.	ACHLOY	6.	AEGNOPRST
2.	ACMRT	7.	AHNOTWY
3.	AEFNORT	8.	AIMNOR
4.	AEGILNS	9.	BISTZ
5.	AEGILNSU	10.	CDEGU

4. to destroy the Erymanthian boar

Given λ, let $f : \mathbb{Z} \to \mathbb{R}$ be such that $f(x+1) + f(x-1) = \lambda f(x)$.
For which $\lambda \in \mathbb{R}$ must f be periodic?

5. to cleanse the Augean stables

This year we subcontracted the task of setting a number crossword to a group of friendly aliens. However, all they sent back was the grid below, together with the information that all the answers are different, and that amongst them there are two pairs which multiply to give the number one. The answers (8 in all) read left to right and top to bottom as usual, but unfortunately they must all be written in the aliens' own number system, which uses two symbols to represent the binary digits 0 and 1, and a third to stand for the binary point. Needless to say, no number is ever written using more symbols than are necessary (for example, using our symbols, a half would be .1, not 0.1 or .10). Thankfully the aliens have made a start on filling in the grid themselves.

6. to destroy the cannibal birds of the lake Stymphalis

Abel and Bela have a regular polyhedron with perfectly thin and rigid faces, but completely flexible edges. They play a game in which they take turns (Abel starting) to cut one of the remaining edges of the solid. The result is determined in one of the four following ways:

a. you win if, just after your move, some part of the solid will flex;
b. you win if, just before your move, some part of the solid will flex;
c. you win if, just after your move, the solid will fold flat;
d. you win if, just before your move, the solid will fold flat.

Assuming best play, who wins under each of these conditions, when the solid is: (1) a tetrahedron; (2) an octahedron?

7. to take captive the Cretan bull

Regular solvers will remember my pet mouse, Arthur, who has been taught to run through grids of coloured squares. When he runs over a red square he turns right at the centre of it (I retrained him); when he runs over a blue square he turns left at the centre of it; and when he runs over a white square he continues straight on. It takes him exactly a second to run through a square, irrespective of its colour. Of the 16 squares in the grid below, I painted 4 white, and each of the others either red or blue. I let Arthur run into the grid at various points, next to each of which I have written how many seconds elapsed between him entering and leaving the grid. Fill in the grid

8. to catch the horses of the Thracian Diomedes

In the year 2043 tastes have changed, and the once repugnant colours of the Archimedeans scarf have become quite fashionable, if not actually trendy. Accordingly, it falls to the vice-president to redesign the scarf so as to return it to its previous level of vileness. The local Scarves'R'us, has a Wool-O-Mat, a robotic knitting machine which will manufacture scarves to order. The machine has a large button marked GO!, a coin slot, and colour keys marked R (red), B (blue), P (pink), G (green) and O (orange). Operating it is simply a question of inserting a few ECUs, pressing a sequence of colour keys, and pressing the GO! button. However, not all colour sequences are acceptable—the instructions explain that:

the sequence R is acceptable;
any acceptable sequence with an occurrence of R replaced by BPB is acceptable;
any acceptable sequence with an occurrence of OB replaced by G is acceptable;
any acceptable sequence with an occurrence of B replaced by OR is acceptable;
any acceptable sequence with an occurrence of BP replaced by PO is acceptable;
EXCEPT that no sequence of more than 8 keypresses can ever be acceptable.

If the sequence is acceptable, the machine will produce a scarf with stripes in colours corresponding to the letters in the sequence (two identical letters in a row will produce a double-width stripe). The vice-president wants a scarf that has at least five stripes, she wants it to be symmetrical (so the sequence should be palindromic), and since she thinks the colour blue is rather pleasant, she will not allow any blue stripes in the pattern. What sequence of colour keys can she press to be sure of getting a scarf which meets her requirements?

9. to get possesion of the girdle of Hippolyta

Many centuries ago the queen of the Amazons placed four caskets, one each of lead, bronze, silver and gold, in a secret chamber in her palace. In each casket she deposited coins, again of lead, bronze, silver or gold, so that each casket contained a different type of coin from the others. Three of the caskets were booby-trapped so that anyone opening them would inhale toxic gases and die, but the fourth, called the prize casket was safe. She positioned the caskets in a row, and on each one were inscribed seven statements ,of which three were true and the other four false. On each casket, the first four statements (starred) said what type of coins were in the four caskets, from left to right. Among these sixteen starred statements, only five were true. On the prize casket none of the starred statements was true.

Years later, you stumble across the secret chamber, where the caskets have now been moved to the four corners of the room. The inscriptions read as follows:

Lead casket
* The leftmost casket contains silver.
* The second casket from the left contains bronze.
* The third casket from the left contains gold.
* The rightmost casket contains lead.
This casket is adjacent to the bronze casket.
The prize casket is third from the left.
The lead casket contains bronze coins.

Silver casket
* The leftmost casket contains bronze.
* The second casket from the left contains silver.
* The third casket from the left contains gold.
* The rightmost casket contains lead.
This casket is adjacent to the gold casket.
The prize casket is on the left.
The bronze casket contains gold coins.

Gold casket
* The leftmost casket contains gold.
* The second casket from the left contains silver.
* The third casket from the left contains bronze.
* The rightmost casket contains lead
This casket is adjacent to the lead casket.
The prize casket is on the right.
The gold casket contains lead coins.

Bronze casket
* The leftmost casket contains silver.
* The second casket from the left contains lead.
* The third casket from the left contains bronze.
* The rightmost casket contains gold.
This casket is adjacent to the silver casket.
The prize casket is second from the left.
The silver casket contains silver coins.

What were the original positions of the caskets, which is the prize casket, and what type of coins does it contain?

10. to capture the oxen of the monster Geryon

The game of Orengy requires three teams of three players. It is played in three rounds of three minutes each, at the end of which one team has won (and scores 3 points), another come second (scoring 1 point), and the third lost (0 points). In a recent tournament, Alphonse, Beatrix, Celia, Dempster and Eugene entered teams, and each threesome played, making 10 matches in all. A local newspaper carried the scores and a report on the tournament:

	played	won	second	lost	points
Alphonse	6				13
Beatrix	6				10
Celia	6				7
Dempster	6				6
Eugene	6				4
		10	10	10	

Sadly, the printers obliterated most of the table, but from the report we know:
Celia beat Alphonse twice, but never beat Dempster;
Two zeroes have been covered up in the above table;
Eugene's team once beat Alphonse's.

Who won, came second and lost in each of the ten matches?

11. to obtain the apples of the Hesperides

What are the next two elements in each of the following sequences?

a. 1, 2, 6, 10, 11, 12, 17, 20, 21, 22, 26, 30, 31, 32, 36, ?, ?.
b. 0, 0, 2, 0, 3, 0, 6, 6, 5, 0, 9, 0, ?, ?.
c. 1, 5, 13, 41, 85, 257, ?, ?.
d. 34, 20, 22, 14, 10, 8, ?, ?.
e. 36, 27, 43, 65, ?, ?.

12. to bring Cerberus from the infernal regions

A unit cube is placed on a grid of unit squares, and rolled around on the grid from square to adjacent square. A corner of the cube traces out a path in three-dimensional space which finishes where it began but is otherwise not self-intersecting, and which is knotted. What is the minimum number of grid squares on which the cube has rested?

The Bluffer's Guide to Poker

Mark Wainwright

Introduction

Poker is a betting game. There are many forms, but the idea at the heart of all of them is this: each player puts a fixed amount of money (the *ante*) into a central pool (the *pot*), and receives a hand of five cards. Each player in turn then has the option of *betting* on his hand. If one of your opponents bets, you may *call*, in which case you accept the challenge by matching his bet; or *fold*, that is, decline the challenge. Players who have folded are out of the hand and lose any money they have already put into the pot. Yet another option is to *raise*: in this case you increase the size of the bet, and the player who made the original bet must either fold or pay the difference.

At the end of the hand there is a *showdown*, where all the players still in the hand show their cards and the one with the best hand wins the pot. Hands are ranked according to certain consistency criteria. For example, a hand with three cards of the same value (*trips*) is better than one with only a *pair* (two cards of the same value). Five cards all of the same suit make a *flush*, quite a strong hand.

Almost all forms of poker have more than one betting round. Between rounds, the players' hands change in some way. For instance, in *draw*, each player may discard some cards, and is dealt replacements from the pack. In *stud* games, each player receives an extra face-up or face-down card between rounds. (Players may end up with more than five cards and have to choose the best five cards to make their hands.) The showdown happens only after the last round of betting. Often, all players except one have folded before this time, and there is no showdown: the player left in scoops the pot regardless of his cards.

Betting structure

You have a good hand, and you want to bet. What size of bet may you make? There are three possible rules, and most poker games use one of them in some form or another:

(i) Limit: all bets must be the same pre-determined size—say £1. In this case, as the hand progresses, the size of the pot increases in proportion to the size of a bet—the *pot odds* improve.

(ii) Pot limit: the maximum allowable bet is the current size of the pot. In this case the pot odds are not directly affected as the hand is played.

(iii) No limit: there is no upper limit on the amount you may bet. In this case you can only make intelligent guesses about what will happen to the pot odds.

Incidentally, an opponent cannot force you to fold by betting more money than you have. At the start of the game, each player puts on the table all the money he is prepared to use in the game (his *stack*). If at any point the bet is more than the current size of your stack, you may stay in the hand by putting your entire stack into the pot. Bets beyond this amount between your opponents are made in a *side pot*, which you are not eligible to win, but you may still win the main pot if you have the best hand at the end.

The problem

You have four cards of the same suit, but your hand is otherwise uninteresting; you are certain to be losing at the moment. However, you are about to receive an extra card, after which the last round of betting will take place. You might make a flush! We will suppose that, for some reason, you are sure none of your opponents can beat a flush. In the jargon, you are on a *draw*, where you need an extra card or cards to make your hand. Should you bet at the end?

If you do make the flush, of course you will bet. If you do not make it, you may still decide to bet. You hope your opponents will think you have made the draw, and will fold, allowing you to win the pot. In summary, you are *bluffing*. Bluffing too often is an expensive mistake, but if you never bluff, your play becomes too predictable, and good players will easily take money off you. To keep them guessing, you must bluff sometimes at random when you fail to make your hand.

David Sklansky, in his book *The Theory of Poker*, points out that this position can be analysed using *game theory*. This article is an extended look at that idea, with a few diagrams which hopefully make the mathematics clearer. We confine our attention to the final round of betting, and assume the game is *heads up*, that is, there are only two players. (It is very common for all but two of the players to have folded by this point.) We assume there is one card to come, and it is face-down; if we are on a draw, our opponent will not know whether we make it or not. This condition applies, for instance, to five-card draw before the draw, or seven-card stud on *sixth street* (the point after the sixth, i.e. penultimate, card is dealt). We shall analyse the case where one player is drawing to (i.e. hoping to make) a *lock*: a hand he knows cannot be beaten.

Game theory

The principal concept we need from game theory is that of a *mixed strategy*. The idea is that a player has several options, and consistently following any one of them leads to a loss, as it can be exploited by the opponent. (For a simple example, consider the game of scissors–paper–stone.) Instead, the player must adopt a mixed strategy where he chooses randomly from among the options, assigning each a certain probability in advance.

A happy feature of this idea is that, for any given (possibly mixed) strategy of your opponent, the payoff (expected gain) from a mixed strategy of your own is just a linear combination of the payoffs from each of the pure strategies that compose it. This makes many situations pleasantly tractable, and allows us to draw various graphs with straight lines.

The analysis

In a game with a hidden card to come (e.g. 7-card stud after 6th street or 5-card draw before the draw), Player 1 ($P1$) is on a draw to a lock, which he makes with probability m. The pot is being contested heads up, and its current size is 1. If $P1$ makes his draw, he will bet n. However, Player 2 ($P2$) will not know when $P1$ has made his hand, so $P1$ should also sometimes bet on the end with a bust, as a bluff.

Generally speaking, if $P1$ always bluffs when he fails to make his hand, $P2$ will obviously make a profit by calling, whereas if $P1$ never bluffs, $P2$ has an easy fold. To make life more difficult for $P2$, $P1$ should bluff at random a certain proportion of the time when he fails to make his hand.

The following diagram shows how the payoffs for each player change as $P1$ bluffs more. The x-axis is $P1$'s probability of betting; it starts at m, not 0, as he always bets when he makes his hand. The y-axes are the payoffs for $P1$ and $P2$. Graphs are shown for when $P2$ always folds, and when he always calls. Intermediate, probabilistic strategies for $P2$ will be interpolations between these two lines.

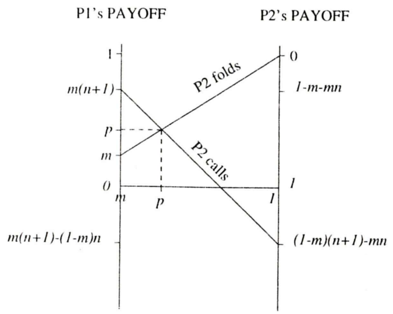


Figure 1

How much of the detail of this graph can we guarantee? It is easy to verify that the two lines always slope in the given directions. For the analysis to work, we also need the crossing point of the graphs to exist. For this, notice that the points on the LH axis are always the given way round, and the points on the RH axis are so, as long as

$$0 < (1-m)(n+1) - mn = n + 1 - 2mn - m;$$
$$m(2n+1) < n+1$$

If n is small (limit poker), this is a negligible restriction; if $n = 1$ (pot limit), it tells us $m < 2/3$; if n is large (no limit), then $m < 1/2$. A draw which succeeds 2/3 of the time is a pretty odd creature and can hardly be called a draw at all, so in practice only the last of these restrictions is ever important; and even a requirement that the draw succeed less than 1/2 of the time is scarcely stringent. (The rare case in which it is violated is covered near the end of this article.)

Note that the scales on $P1$'s axis and on $P2$'s axis are the same, though they increase in opposite directions. In fact, the two payoffs in a given hand must add up to 1 (the amount in the pot).

Given all this, the graph makes clear a fundamental principle of game theory: $P1$ should fix his strategy so that $P2$'s payoff (and, for that matter, his own) is the same whether $P2$ calls or folds. Thus his overall betting probability should be the value marked p on the diagram. If $P1$ bets less often (bluffs too little), $P2$ increases his payoff by folding; if $P1$ bets more often (bluffs too much), $P2$ increases his payoff by calling. This is just the observation that the two lines slope in opposite directions.

How should $P1$ locate this optimal value of p? To see this most perspicuously, we consider a different diagram. This one is rather more like a bar chart; the horizontal axis is the probability line, and the vertical axis is $P2$'s payoff. If $P1$'s overall betting probability is p, then we know he makes his hand m of the time, bluffs $(p-m)$ of the time, and folds the rest of the time, and we draw bars whose height is the payoff in each case.

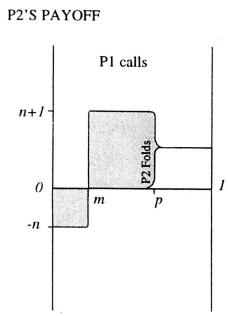


Figure 2

If $P1$ has found the optimal value of p, $P2$'s payoff is the same whether he calls or folds. This tells us that the area of the two shaded regions on the chart must be equal. The ratio of their heights is $(n+1)/n$—precisely the pot odds $P2$ is getting when $P1$ bets. The ratio of their widths is $(p-m)/m$, the relative frequency with which $P1$ bluffs and bets for value (i.e. bets with a made hand). This gives us our first theorem:

THEOREM 1. P1 *should fix his bluffing strategy so that the relative frequency with which his bets are bluffs is equal to the pot odds he is offering* P2. □

It is hard to make any decision with a given probability, but another remark of Sklansky's is that $P1$ can use the cards to randomise when he bluffs: if his number of outs (cards that will make his hand) is k, he should choose an extra $kn/(n+1)$ cards on which he will bluff.

Incidentally, it is worth remarking that we assume that when $P2$ calls he has enough money to call the full size of $P1$'s bet. If this is not the case, it is as if $P1$ had bet only the size of $P2$'s stack; he should remember to decrease his bluffing probability accordingly.

If $P1$ deviates from the optimal strategy we have calculated, $P2$ can improve his odds by always calling or always folding. If $P1$ bluffs with the correct frequency, it doesn't matter what $P2$ does, so we can assume he folds to calculate payoffs. That is, when $P1$ bets, he effectively claims the pot; put yet another way, the case where $P2$ folds is the line $y = x$ on Figure 1. So $P1$'s payoff averaged over all hands is p.

This point that his payoff for betting is just the size of the pot is very telling—it graphically illustrates the fact that betting is all about the battle for the antes or, more generally, the money already in the pot. This is why good poker books recommend that your strategy should depend a good deal on the size of the antes in relation to the size of a bet. If there were no antes, it would never be correct to bet except with a lock.

Of course, if $P2$ always calls (or always folds), $P1$ can do better than this probabilistic strategy. Let us assume $P2$ knows $P1$ was on the draw, and wants to keep $P1$ to his minimum gain; then clearly $P2$ should call some of the time (with some probability), to 'keep $P1$ honest'. The associated graph on Figure 1 must be a horizontal line through the critical point; that is, $P2$'s strategy will be such that $P1$ cannot alter his payoff by whether he bluffs or not. If we suppose that $P2$ calls $P1$'s bets with probability q, we get the following chart in the spirit of Figure 2:

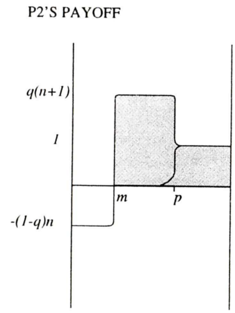


Figure 3

$P1$ changing his strategy corresponds to moving the point p left or right. If this is to make no difference to the payoffs, the two shaded regions in fact have the same height (unlike on the diagram)—that is, $q(n+1) = 1$, our second theorem. (Game theorists call this sort of position, where either player independently can force a fair result, a *Nash equilibrium*.)

THEOREM 2. *Suppose* P1 *bets. Then* P2*'s optimal strategy, assuming he knew* P1 *was on the draw, is to call with probability* $1/(n+1)$*, the reciprocal of the size of the pot after* P1*'s bet.* □

This appears cock-eyed, as it is the relative frequency with which $P1$ is betting for value—we have shown that (if $P1$ plays correctly), the more often $P1$ bluffs, the less often $P2$ should call! In fact this is correct; it is by forcing $P2$ to call less often (by increasing n) that $P1$ enables himself to bluff more. I am grateful to Paul Pudaite who, on seeing this, recalled an anecdote about Johnny Moss. Moss was a legendary no-limit poker player and three times world champion, who died last December. He used to say something like, 'We both know I'm bluffing, but I'm going to put so much money in this pot that you can't call me!'

Notice the highly satisfactory point that $P2$'s correct calling frequency depends only on n, the amount of the raise. He does not need to know m exactly (except that, as we shall see later, if the bet is large he is relying on the fact that $m < 1/2$). Sometimes he might actually have more information about m than $P1$ himself; for instance, if in 7-card stud he can see that $P1$ is on a flush draw, then his own hole cards (face-down cards) give him information denied to his opponent. This is likely to be only a small advantage. There is no such luxury in draw, where it is much harder to guess which of many draws one's opponent is on.

What should the value of n be, that is, how much should $P1$ bet if he has a choice? As his payoff is p, he wants to maximise p, his betting frequency. Equivalently he wants to maximise $p-m$, his bluffing frequency. By the above, $p-m = mn/(n+1)$. The more he bets, the closer $n/(n+1)$ to unity and the more often he can bluff. So the answer here is that the bigger the bet, the better. In no-limit, $P1$ should bet his entire stack when he bets on the end in this position.

Remember the assumptions at this point. We assume $P1$ knows he will be winning if he makes his draw. In practice, this is usually at best an approximation, though it might be true in 7-stud if the right cards are out. We assume $P2$ does not know whether $P1$ made his draw, so the analysis can hardly be applied as it stands to a game like *hold 'em*, where there are community cards shared by all players. We are also assuming at the moment (since we are talking about no-limit) that $m < 1/2$. All these conditions being met, it is true that the more $P1$ bets, the greater his expectation.

As we saw, by making n very large, $P1$ makes $p - m$ very close to m: that is, he can profitably bluff as often as he makes a legitimate hand. If $m \geqslant 1/2$, namely $P1$ makes his draw more than half the time, this means $P1$ always bets on the end, irrespective of his last card. Now the crossing point p does not exist, and we see that even calling $1/(n+1)$ of the time is wrong for $P2$. Any call has negative expectation, as $P1$ is more likely than not to have made his hand.

We have proved that if $P1$ has a draw that comes off more than half the time, and if he can bet a large amount compared to the pot, then he can pick up the pot every time, whether he makes the draw or not. In this case, $P1$ need no longer bet arbitrarily large amounts. As long as he bets enough to make it correct for $P2$ to fold, he does not have to bet his entire stack.

The pot-limit case is also interesting. $P1$ can bet at most 1, so now $n/(n+1) = 1/2$; so $P1$ can bluff half as often has he makes his hand. On the other hand, in limit, if the pot is k times the size of the bet, $P1$ can bluff only $1/(k+1)$ times as often as he makes his hand—confirming our intuition that in limit poker, one must bluff less and stick closer to the odds currently offered by the pot.

Some examples

To finish, consider a specific case where in 7-card stud, $P1$ has a draw which he will make with probability $1/5$. (This is plausible for a flush draw or straight draw.) In limit poker, common sense tells us that when $P2$ bets on 6th street, the pot (including $P2$'s bet) needs to be 4 times the size of the bet to give $P1$ the 4-to-1 (1 in 5) odds he needs to make it correct for him to call. The extra bluffing possibilities improve the situation only slightly: it turns out pot odds of 3.5-to-1 will suffice. If $P1$'s draw succeeds one time in M, as M grows large, the odds $P1$ requires from the pot on 6th street are only very slightly better than $(M-2)$-to-1, rather than the expected $(M-1)$-to-1. That is, the pot can contain almost a whole bet less than current pot odds would suggest. (The calculation is omitted, but it is an easy exercise.)

If the game is no-limit, $P1$ will bluff as often as he bets on the end, so he will bet a total of $2/5$ of the time; his profit from betting is the size of the pot going into 7th street, so his average profit is $2/5$ of that pot. If the pot going into 6th street has size 1, how much must $P2$ bet to raise $P1$ out? If he raises 2 and $P1$ calls, the pot is now 5, and $P1$ will make $(2/5) \times 5 = 2$ on the end, just recouping his investment on 6th street. That is, $P2$ must raise at least *twice the pot* on 6th street to make it correct for $P1$ to fold his draw.

If the game is pot-limit, $P1$ will bluff half of $1/5$ of the time, namely $1/10$ of the time, so will bet a total of $3/10$ of the time. He figures to gain 30% of the pot on the end. By putting in a pot-sized raise on 6th street, $P2$ can make $P1$ pay 33% of that final pot, and so just swing the odds against $P1$, who must now fold. However, it was a close call. If the cards that are out give $P1$ odds only slightly better than $1/5$, it may become correct to call on 6th street with his draw.

A Voting Problem

Eva Myers and Joseph Myers

1. Introduction

A voting problem due to Merrick Furst, Jim Aspnes, Richard Beigel and Steven Rudich appeared in [1]. It is:

"n people are seated around a table and each has a 0 or 1 painted on his forehead randomly, and each of the 2^n possibilities has equal probability. As is usual in this situation, each person can see everyone else's forehead but not his own. A ballot is held in an attempt to determine the overall parity, each person voting either '0' (meaning an even number of 1s) or '1'. Do there exist voting strategies such that the probability that a (strict) majority vote correctly is greater than $\frac{1}{2}$ and if so, what is the greatest probability that can be attained?

Does the answer change if:

(a) weighted votes are allowed, i.e. each person may submit a number of votes of his choice?

(b) random strategies are allowed?"

At first glance it might seem that the probability of success (a strict majority's voting correctly) cannot be raised above $\frac{1}{2}$, as each person has a probability of only $\frac{1}{2}$ of voting for the correct parity. However, a simple strategy which does better than this in the case of unweighted, non-random votes is the "vote-with-majority" strategy; i.e. if a person sees more 1s than 0s, he should pretend he has a 1 and vote accordingly, and if he sees more 0s than 1s he should pretend he has a 0. (For odd n, if he sees the same number of 0s and 1s he has two possible strategies. We shall assume for definiteness that, in this case, he pretends to have a 1.)

For example, if there are 70 0s and 30 1s, each person sees more 0s than 1s and votes as though he has a 0. For the 70 0s, this is a correct vote, and for the 30 1s, this is a wrong vote, but as there are more 0s than 1s, the majority of votes will be correct. This strategy fails if and only if there are exactly as many 0s as 1s for even n, or one more 0 than 1 for odd n. Thus its probability of failure can be shown to be on the order of $\frac{1}{\sqrt{n}}$. However, this is not the best possible strategy.

This article will prove that with weighted votes, the greatest possible probability of success is $1 - (\frac{1}{2})^n$, that this can be obtained for all n, and that allowing random strategies does not change the answer.

We shall also find bounds on the maximum possible probability of success where neither weighted votes nor random strategies are allowed.

It will be convenient from now on to talk in terms of failure rather than success. We will write "vote Ax" to mean "vote as if you are x" and also "vote y" to mean "vote that the parity is y". For convenience the people will be called $P_1, P_2, ..., P_n$.

2. Weighted votes

THEOREM 1 *With weighted votes, a probability of failure of* $(\frac{1}{2})^n$ *can be achieved.*

PROOF. The strategy is:
(i) P_1 votes A1 with weight 1.
(ii) P_2 abstains (votes with weight 0) if P_1 has number 1, and otherwise votes A1 with weight 2 (to outvote P_1's error).
(iii) P_k $(2 \leqslant k \leqslant n)$ abstains if *any* P_j $(j < k)$ has number 1, and otherwise votes A1 with weight 2^{k-1} (to outvote $P_1, \ldots, P_{k-1}$). This clearly works unless everyone has number 0. □

THEOREM 2 *The minimum possible probability of failure (if random strategies are not allowed) is* $(\frac{1}{2})^n$*; i.e. the strategy in the proof of* Theorem 1 *is a best possible strategy; there may be others.*

PROOF. Such a strategy has already been exhibited (Theorem 1), so it will suffice to prove that no strategy can achieve a lower probability of failure. In any position (choice of numbers on foreheads), we define the *success margin* as the total of the correct votes minus the total of the incorrect votes. We must now prove that some position has success margin less than or equal to zero.

A strategy consists of the weights with which each person votes in favour of an even parity, for each choice of numbers on other people's foreheads. The success margin of a position will be the sum of these weights for each person if the parity is even, and minus this sum if the parity is odd. Each weight will occur in just one position with even parity and just one with odd parity. Now we consider the sum of the success margins of all positions. From the form of a success margin, this will be the sum of some weights and minus some weights. However, each weight occurs once in a position with even parity and once in a position with odd parity. Thus it appears once positively and once negatively. Therefore all weights cancel from the sum of success margins, which is thus 0. Thus some success margin is less than or equal to zero. □

THEOREM 3 *Random strategies cannot reduce the probability of failure to below* $(\frac{1}{2})^n$.

PROOF. We consider E_i, for $1 \leqslant i \leqslant 2^n$, the 2^n propositions of the form: "In position X_i, the ballot gives the correct result." By Theorem 2, it is not possible for all these propositions to be true for the same choice of weights, so we need to prove that

$$\sum_{i=1}^{2^n} \mathbb{P}(E_i) \leqslant 2^n - 1,$$

where the left hand side is the expected number of true propositions. Let E_i^c be the event "not E_i". So $\mathbb{P}(E_1^c \cup \cdots \cup E_{2^n}^c) = 1$. But then we must prove

$$\sum_{i=1}^{2^n} \mathbb{P}(E_i^c) \geqslant 1.$$

This is trivial, so the theorem is true. □

We shall now move on to the main and much more difficult problem of where weighted votes are not allowed.

3. Unweighted votes

THEOREM 4 *Random strategies are no use.*

PROOF. Consider any person, any assignment of strategies to the other people and any arrangement of numbers on the other people's foreheads. If that person votes 0, there is a probability p_0 of success, and if he votes 1, there is a probability p_1 of success. If he has a random strategy, the probability of success will be a weighted average of p_0 and p_1, which cannot be greater than $\max(p_0, p_1)$. Thus he may adopt a non-random strategy in this case without reducing the probability of success. As the same reasoning applies to all the people and all arrangements of visible numbers, for any random strategy there is a non-random strategy which is at least as good. □

The remaining problem of finding the best strategy for each n is only partly solved; however we have found a lower bound on the minimum probability of failure and an algorithm that has at most twice this probability of failure.

THEOREM 5 *For odd n the minimum probability of failure is at least $\frac{1}{n+1}$; for even n the minimum probability of failure is at least $\frac{2}{n+2}$.*

PROOF. We have already shown in the proof of Theorem 2 that the sum of success margins of all positions is zero. Each success is by a margin of at least 1 (or 2 if n is even). Each failure is by a margin of at most n. Therefore for every n successes ($\frac{n}{2}$ if n is even) there must be at least one failure. Thus the minimum probability of failure is at least $\frac{1}{n+1}$ (or $\frac{1}{\frac{n}{2}+1} = \frac{2}{n+2}$ if n is even). □

We shall now exhibit the positive results (algorithms) that we have. In some of these, it will be convenient to divide the people into groups. A group can abstain (only if it has even size) by alternate members voting 1 and 0, or can all vote the same way, assuming that the total parity of the group is either odd ("vote G1") or even ("vote G0"). Clearly the group's members must decide whether or not to abstain purely on the basis of numbers on foreheads of people outside the group.

4. The grouping strategy

(The following strategy is due to Imre Leader.)

First suppose n is odd and $2^a - 1 \leqslant n \leqslant 2^{a+1} - 3$. Then the probability of failure by the following algorithm will be $(\frac{1}{2})^a$:

divide the people into disjoint groups: Group 1 of 1 person, Group 2 of 2 people, up to Group $a-1$ of 2^{a-2} people, and Group a containing all remaining people (at least 2^{a-1}).

These groups then vote like the people in the weighted votes strategy (Theorem 1): Group 1 votes G1, Group 2 abstains if the parity of Group 1 is 1 and otherwise votes G1, Group i abstains if the parity of any Group j ($1 \leqslant j < i$) is 1 (odd) and otherwise votes G1, and so on. This clearly fails if and only if all groups have even parity. □

Now suppose n is even and $2^{a+1} - 2 \leqslant n \leqslant 2^{a+2} - 4$. The probability of failure will be $(\frac{1}{2})^a$ by the following algorithm: Divide the people into disjoint groups: Group 1 of 2 people, Group 2 of 4 people, ... , Group $a-1$ of 2^{a-1} people and Group a containing all remaining people (at least 2^a). These groups then vote as above. This clearly fails if and only if all groups have even parity. □

Some refinements

We know that for certain values of n it is possible to improve on the grouping strategy. For example, consider $n = 4$. In this case, the grouping strategy reduces to "vote 1" with probability $\frac{1}{2}$ of failure. "Vote-with-majority" is better in this case (and is also better for the cases $n = 10$ and $n = 12$ but for no other n): here it fails if and only if there are two 0s and two 1s, i.e., $\frac{3}{8}$ of the time, which by Theorem 5 is the best possible (as the exact minimum $5\frac{1}{3}/16$ clearly cannot be achieved, so it is not possible to do better than $\frac{6}{16} = \frac{3}{8}$). Unfortunately, for large n the probability of failure of "vote-with-majority" is on the order of $\frac{1}{\sqrt{n}}$, whereas that of the grouping strategy is of order $\frac{1}{n}$.

THEOREM 6 *If there is an algorithm for* $2a$ *people with probability of failure equal to* f, *there is an algorithm for* $2a + 1$ *people with probability of failure equal to* $\frac{f}{2}$.

PROOF. Divide the $2a + 1$ people into two groups: one containing one person and the other with the other $2a$ people. The one-person group votes G1. The other group abstains if the one person has a 1 on his forehead, and otherwise follows the $n = 2a$ strategy; a success in this will have a success margin of at least 2 and will therefore outvote the one person. □

Using Theorem 6, we derive a strategy for $n = 5$ with probability of failure $\frac{3}{16}$, the best possible.

We now have the best possible strategy for $n \leqslant 7$, using the above and the grouping strategy (which is the best possible for $2^a - 1$, $2^a - 2$: 1, 2, 3, 6, 7). We conjecture that the minimum probability of failure is $\frac{k}{2^n}$ where k is the integer such that

$$\frac{2^n}{n+1} \leqslant k < 1 + \frac{2^n}{n+1},$$

where n is odd, and

$$\frac{2^{n+1}}{n+2} \leqslant k < 1 + \frac{2^{n+1}}{n+2},$$

where n is even. By Theorem 5 this would be the best possible.

For $8 \leqslant n \leqslant 13$ we have results which, although improvements on the grouping strategy, are suboptimal if the conjecture is true. For odd n, the strategy is derived by Theorem 6. For even n, the people are divided into a group of 2 people and a group of $n - 2$ people. The group of 2 people votes G1. If the total parity of the group of 2 people is odd the group of $n - 2$ people abstains. Otherwise they follow the following strategy (to increase the probability of succeeding by a margin of 4 or more above $\frac{1}{2}$). Each person, on seeing in that group $\frac{n}{2}$ or more of a number x, votes Ax. On seeing $(\frac{n}{2}) - 2$ exactly of a number x he votes Ax. The probabilities achieved by our various strategies for $1 \leqslant n \leqslant 15$ are shown in the table below.

n	$\mathbb{P}$(Failure with grouping strategy)	$\mathbb{P}$(Failure with best known strategy)	$\mathbb{P}$(Failure with conjectured theoretical best strategy)
1	$\frac{1}{2} = 0.500$	$\frac{1}{2^1} = 0.500$	$\frac{1}{2^1} = 0.500$
2	$\frac{1}{2} = 0.500$	$\frac{2}{2^2} = 0.500$	$\frac{2}{2^2} = 0.500$
3	$\frac{1}{4} = 0.250$	$\frac{2}{2^3} = 0.250$	$\frac{2}{2^3} = 0.250$
4	$\frac{1}{2} = 0.500$	$\frac{6}{2^4} = 0.375$	$\frac{6}{2^4} = 0.375$
5	$\frac{1}{4} = 0.250$	$\frac{6}{2^5} \simeq 0.188$	$\frac{6}{2^5} \simeq 0.188$
6	$\frac{1}{4} = 0.250$	$\frac{16}{2^6} = 0.250$	$\frac{16}{2^6} = 0.250$
7	$\frac{1}{8} = 0.125$	$\frac{16}{2^7} = 0.125$	$\frac{16}{2^7} = 0.125$
8	$\frac{1}{4} = 0.250$	$\frac{60}{2^8} \simeq 0.234$	$\frac{52}{2^8} \simeq 0.203$
9	$\frac{1}{8} = 0.125$	$\frac{60}{2^9} \simeq 0.117$	$\frac{52}{2^9} \simeq 0.102$
10	$\frac{1}{4} = 0.250$	$\frac{224}{2^{10}} \simeq 0.219$	$\frac{171}{2^{10}} \simeq 0.167$
11	$\frac{1}{8} = 0.125$	$\frac{224}{2^{11}} \simeq 0.109$	$\frac{171}{2^{11}} \simeq 0.083$
12	$\frac{1}{4} = 0.250$	$\frac{840}{2^{12}} \simeq 0.205$	$\frac{586}{2^{12}} \simeq 0.143$
13	$\frac{1}{8} = 0.125$	$\frac{840}{2^{13}} \simeq 0.103$	$\frac{586}{2^{13}} \simeq 0.072$
14	$\frac{1}{8} = 0.125$	$\frac{2048}{2^{14}} = 0.125$	$\frac{2048}{2^{14}} = 0.125$
15	$\frac{1}{16} \simeq 0.063$	$\frac{2048}{2^{15}} \simeq 0.063$	$\frac{2048}{2^{15}} \simeq 0.063$

We would welcome improvements to these results or any further results on the problem. They may be sent to Eva at Newnham College, Cambridge, or e-mailed (*erm1001@cam.ac.uk*), or sent to Joseph at Trinity College, Cambridge, or e-mailed (*jsm28@cam.ac.uk*).

Acknowledgements

We thank Dr Leader for his helpful comments on an earlier draft of this article.

Bibliography

[1] $\mathbb{QARCH}$ **12** (The Archimedeans 1991), problem 57.

Solutions to the Problems Drives

1994 Colin Bell and Michael Fryers

1. (a) 101,1010. (10 in bases 10,9,...)

(b) 20, 5. (Factorials modulo 25)

(c) 33,36. (The difference between the nth and $n+1$th numbers is the nth digit in the sequence)

(d) 11,50. (Numbers, less than 100, in alphabetical order in English).

(e) 1,4. (Number of letters in each word of the title)

2. (a) Fermat's Little Theorem.

(b) Pythagoras's Theorem.

(c) The Riemann Hypothesis.

(d) The Four Colour Theorem.

(e) The Fundamental Theorem of Calculus.

(f) The Pigeonhole Principle.

(g) The Chinese Remainder Theorem.

3.

1 S	L	2 O	3 T
4 P	5 U	N	O
I	6 S	E	W
7 N	E	T	S

(S across can only be 2^{10} or 2^{13}; there is no L clue so it's 2^{10}. This gives us S, L, and O. P down must have three or four digits, so E can only be 7, and P must be 3. E across is 5741, so we get I=9. U is the remaining letter with no clue, so is 8, and N the remaining down clue, 5. T across ends in a 5, so it must be 4 across, (6 across ends in a 6), leaving W as 6.)

4. R 2-0 M, R 2-0 C, R 1-1 D, R 0-2 P, M 2-0 P, M 1-1 D, C 5-0 P; C-D, P-D and M-C yet to be played.

(Number of matches played is 3, 5 or 7. We are given the results of four, all non-draws. If 5 have been played, then all players have two points, so all have won one match. Contradiction. So seven matches have been played.

Distributing 14 points between 5 primes gives us 5-3-2-2-2 or 3-3-3-3-2. We're told the bottom two have the same number of points, so it's the former.

R has 5 points including a defeat which must be WWDL. M has 3, giving us DDD or WD (plus possibly some defeats). If DDD then he's drawn against all but P—all P's points come from the win over R, but then neither C nor D can have won a game, which contradicts our 2-0 triad. So M is WD. R must have beaten M since M's draw needs to be against the all-draw player.

One of C and D is the all-draw player: call him X, and the other Y. So R has beaten Y and drawn with X. M has drawn with D and beaten someone. It must be P since this is the only way to get a 2-0 triad. There is one match left, and Y needs a win which must be against M or P.

The 5-0 game can either be R-Y, Y-M or Y-P. R has 2-0, 0-2 and a draw (1-1 or 2-2) already: 5-0 takes him to 8-3 or 9-4 in total: no good. Y-M can be eliminated for similar reasons. So Y has beaten P, which means Y is C and X is D.

We've now allowed for 11 won rounds in the four games we know about. The total number of won rounds is 11,13 or 17, but we need at least two won rounds per game, so it's 17. So R-C is 2-0 and the two draws are 1-1.

Unsuprisingly, nobody got this even nearly correct...)

5. Snakes 21-5, 23-2 and 16-4. Ladders 1-12, 9-15, 11-17. (We know that a ladder starts at square 1, a ladder ends at 15, and have used squares 18 and 19. A snake must end at 2. It can't be the first one, since then another one ends at 3, and clashes with it. If it's the third one, then again another one ends at 3, and it must be the first one, so the second snake ends at 4, and we then go to 5,6,10,11,16,19, which doesn't work. So the snake ending at 2 is the second.

Similarly, a snake must start at 23. It must the second one, by what precedes it. The remaining snake starts at 21, and then the rest is easy.)

6.

D	B
A	C

(Try cases.)

7. (a) 5; (b) 15; (c) 8; (d) 81.

(We give the working for (d); (b) is similar and the other parts trivial. We use place-value notation with base $1\frac{1}{2}$: BBBBAA$=42_S$, EEDB$=21010_S$. The length is then the sum of the digits. In a shortest representation, no digit is > 2.

Consider the maximal integer N with length 5, and suppose its leftmost nonzero digit is in the $(1\frac{1}{2})^n$ place for $n > 0$.

$N \in \mathbb{Z}$, so $N \equiv 0 \pmod{2^{-(n-1)}}$, and so the nth digit of N is 2. $N - 20...0_S = N - 3(1\frac{1}{2})^{n-1} \not\equiv 0 \pmod{2^{-(n-2)}}$, so the $(n-1)$th digit is 1. $N - 210...0_S = N - 6(1\frac{1}{2})^{n-2} \equiv 0 \pmod{2^{-(n-3)}}$, so the $(n-2)$th digit is 0 or 2.

If $N = 212\{\text{something}\}_S$, it is length 5 only if {something} is all zeroes. $212_S = 8$, so we can add 3 zeroes to give 27 (and no more, since this is odd).

If $N = 210\{\text{something}\}_S$, $N - 2100...0_S = N - 9(1\frac{1}{2})^{n-2} \not\equiv 0 \pmod{2^{-(n-4)}}$, so $(n-3)$th digit is 1. $N - 21010...0_S = N - 15(1\frac{1}{2})^{n-3} \not\equiv 0 \pmod{2^{-(n-5)}}$, so $(n-4)$th digit is 1. Again, we've got to length 5 and may only add zeroes. $21011_S = 16$, so we can add 4 zeroes to give 81. This is therefore the answer.)

8. 1/15. (The only sequence of turns which sends him back to A is $(RLLR)^n RLLL$. which sends him straight back to A. The chance of this is $2^{-4} + 2^{-8} + \ldots$. Once he's facing leftwards for the first time, he'll either hit a piece of cheese and go off the side, since there are two pieces of cheese already eaten in that column, or go off the left hand side of the grid if not.)

9. 50. (Kings take 7 each, bishops 2, knights 4, and the two queens 4 between them and these can be done in order. A pair of rooks can be swapped in four if both knights are out of the way, which can be easily arranged. Try solving the problem when one or both of the restrictions is added in.)

10. All but A_1. (If we put A_0 in first, A_1 sits with its centre $16 + \sqrt{24^2 - (36-24)^2} = 16 + 12\sqrt{3}$ miles up (draw a picture); all the other balls can then be placed on the base.

If we put A_1 in first, A_0 sits with its centre $8 + 12\sqrt{3}$ miles up, and again all other balls can easily go on the base. So we need to compare $16 \cdot 16^3 + (16 + 12\sqrt{3}) \cdot 8^3$ and $(8 + 12\sqrt{3}) \cdot 16^3 + 8 \cdot 8^3$. A bit of manipulation makes it obvious that the latter is much bigger.)

11. (a) 75. (There are two ways of getting money: buying pizzas, swapping for Eurekas and selling the Eurekas, and buying $\mathbb{Q}ARCH$es, swapping for Eurekas and selling the Eurekas. Taking into account the taxi fares, the former requires you to buy two pizzas to make money: so 39pJC is required (including three taxi fares). The latter requires you to buy 12 $\mathbb{Q}ARCH$es to make a profit: again you need three taxi fares to get your money back, so you can start with 27pJC.)

(b) 23. (Consider what your best option is when you get to Bleckwalls with a given number of pJC. 11 $\mathbb{Q}ARCH$ loops, then 3 pizza loops, 3 $\mathbb{Q}ARCH$ and then 6 pizza loops is optimal.)

12. $n/(2n + 1)$. (We obviously need to have two hyperspheres in opposite corners of the hypercube: calling the radii G and S, we have the conditions $G/S = (n + 1)/n$ and $(1 + n)(S + G) = n$ (looking at the diagonal).)

1995 Richard Tucker and Luke Pebody

1. $2^{512} - 11$. (Note that if $wf(i) < j$ then $wf(i + k) < j + k$ for all $k > 0$. $wf(5) = 5$, $wf(6) = 5$, so 5 is the largest wondrous number and 6 the smallest wanton number. 11 is the smallest n with $wf(wf(n)) = 6$, i.e. wanton. The biggest number obtainable with 5 ones is $3^2 = 9$, and the biggest number obtainable with 9 ones is: $2^{2^{2^3}} = 2^5 12$.)

2. $QARCH = 38956$. (By word lengths, $E = 0$ or 1. 5 divides both M and C, so $E = 1$, and $A > 5$, $H > 5$. From its position in both words, $R = 0$ or 9, so $R = 9$, whence $C = 5$, $M = 0$, U is odd and A is even. If $A = 6$ then $U = 3$, $I = 8$, $K = 7$, $H = 8$: contradiction. Therefore $A = 8$, so $U = 7$, $I = 4$, $H = 6$, $K = 2$. By elimination $Q = 3$.)

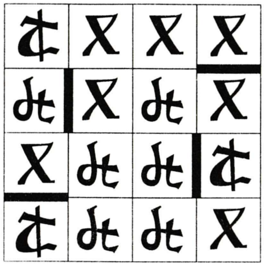
3. Cauchy–Gödel, Galois–Riemann, Gauss–Leibnitz, Pythagoras–Newton and Fermat–Cantor (or Carnot).

4. $\lambda \in \{2\cos q\pi, q \in \mathbb{Q}\} - \{2, -2\}$. (Clearly, if $|\lambda| \geqslant 2$ then $|f(x)|$ is unbounded, hence non-periodic. For $|\lambda| < 2$, write $\lambda = 2\cos\theta$. Then $(f(x + 1), f(x))$ is obtained from $(f(x), f(x - 1))$ by the matrix

$$\begin{pmatrix} 0 & 1 \\ -1 & \lambda \end{pmatrix} = \frac{1}{4x}\begin{pmatrix} x+1 & x-1 \\ x-1 & x+1 \end{pmatrix}^{-1}\begin{pmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix}\begin{pmatrix} x+1 & x-1 \\ x-1 & x+1 \end{pmatrix}$$

where $x^2 = \frac{2+\lambda}{2-\lambda}$. Thus it is conjugate to A rotation through θ and so f will be periodic if and only if θ is a rational multiple of π.)

5. The two inverse pairs must be 1000, ·001 and 100, ·01. (No answer begins with 0 or ends with a point, so ·001 must go in the third column or the bottom row. If in the third column, there is nowhere for 1000 to go. Therefore ·001 goes in the bottom row, and 1000 in the third column. 100 must then go in the third row, and now we know which symbol is which. The rest is trivial.)



6. Bela always wins except in case 2c. Note that cutting two edges of an octahedron can enable it to fold flat.
7.

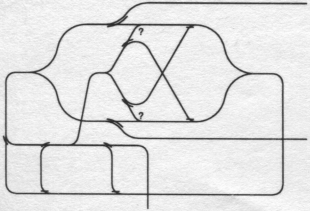
B	B	W	R
B	R	B	B
B	R	W	W
B	W	R	R

(Begin by deducing that the second square of the bottom row can't be red.)
8. GPOPOPG. (From the rules and the minimum length of five, the first and last letter in the sequence must be G, hence it must derive from OB...OB, by B...R giving GPB...PG. To get B...R, we must have had B...B giving B...OR. Continuing this analysis gives the sequence: R $\rightarrow$ BPB $\rightarrow$ ORPB $\rightarrow$ OBPBPB $\rightarrow$ GPBPB $\rightarrow$ GPORPB $\rightarrow$ GPOBPBPB $\rightarrow$ GPOPOPOB $\rightarrow$ GPOPOPG.)
9. Caskets, left to right: gold, silver, lead, bronze. Prize is bronze coins in lead casket.
10. Matches (win-2nd-lose): C-A-B, A-D-B, A-B-E, A-D-C, C-E-A, A-D-E, B-D-C, B-C-E, B-D-E, E-D-C.
11. (a) 44, 45 (when written out, numbers' lengths are divisible by three)
(b) 7, 10 (largest number less than n and not coprime with n)
(c) 517, 1553 (take previous number and alternately multiply by 3 and add 2, or multiply by 2 and add 3)
(d) 7, 1 (add proper factors of previous number including 1)
(e) 109, 277 (sum of sequences a-d)
12. 22 squares. (Start with the marked corner on the southwest corner of the cube's top face, and roll it as follows:

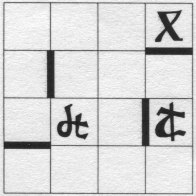
$$NWNNWSWWWSESSSENEENWNNWWWSESSEEENE).$$

Eureka 54 Corrections

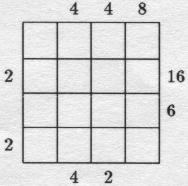
Unfortunately several errors crept into *Eureka* **54**. These are corrected below:
Page 32 The train sets diagram contains three points with errors and should looks as below;
Page 38 The diagram below of the partially completed crossword was omitted from question **5**;
Page 39 The diagram for question **7** should have been as below.



32



38



39