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

eureka 51

Eureka

Eureka is the journal of the Archimedeans, which is the Cambridge University Mathematical Society, and a Junior Branch of the Mathematical Association. *Eureka* 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 five 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. TEX source for issues since issue 49 may also be available, subject to copyright; contact the Business Manager for details.

Electronic Mail

The Society can receive electronic mail via JANET, ARPAnet and so on, and can read most electronic media. Our electronic mail addresses are 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

Electronic mail from the UK should be sent to:

`archim@uk.ac.cambridge.phoenix`

From the rest of the World, the following should work:

`archim@phoenix.cambridge.ac.uk`

But if it doesn't, try whichever of these seems the more appropriate:

`archim%phx.cam.ac.uk@nss.cs.ucl.ac.uk`
`archim%uk.ac.cam.phx@ukacrl`

EUREKA

Number 51 Editor: Mark Wainwright March 1992

Table of Contents

 ISSN 0071-2248

Archimedeans' Committee 1991-1992

ACKNOWLEDGEMENTS

This is the present Editor's second (and last) issue of *Eureka*, and thus for the second time he has sorely tried the patience of numerous friends and acquaintances and indeed almost anyone else within reach. Again the first debt is to the contributors: this time, as well as for providing articles, for their patience in waiting while the journal was almost a year overdue. The journal was typeset in TEX; anyone who has used TEX will know that the program has a mind of its own, and Robert Hunt was helpful in keeping it, if not under control, at least within manageable limits. A number of people have lent assistance with typing in text; David Jones has been invaluable in this and other respects. Richard Tucker, Aldabra Stoddart and Dilip Sequeira, among others, have also been especially helpful. Colin Bell again came to the rescue by executing the unpleasant diagrams in the Problems Drive in CAMPLOT. Finally the Department of Applied Mathematics and Theoretical Physics, this year through the office of Dr David Harris, again gave kind permission to use the departmental laser-printing facilities for preparation of hard copy to be sent to the printers.

Editorial

The present issue of *Eureka* is somewhat later than it would, in the ordinary course of things, have been. Readers who subscribe will have noticed that it is also more expensive than previous issues. Both of these facts call for some comment.

The first (let it be said at once) is in part due to sheer delinquency on the part of the Editor, in not having prepared the journal for publication before now. However, the fact that it was not ready by last Easter is also due, in large part, to the fact that there was then no money with which to publish it. The last four issues of *Eureka* have been considerably bulkier than any of their predecessors, and while this is a very encouraging thing for an editor, it is the scourge of a distractedly impecunious committee. The result of this runaway success has been that the Archimedeans have been losing money hand over fist, and this could not continue indefinitely.

The price of *Eureka* has not recently increased so dramatically. Indeed, it is a part of the problem that it has consistently failed to keep pace with inflation for several years. (Yet it has been getting bigger, not smaller.) Furthermore, in any event, the major part of the money used to print it comes, not directly in the form of subscriptions, but from membership fees collected by the Archimedeans—as is appropriate since it is distributed free to members. And there is no very coherent policy on how much the Society should pay (for example, whether it should be a fixed proportion of the cover charge, or of the membership fee, or determined in some other way). So in practice the Society has paid the amount needed after subscriptions have been accounted for to print the journal the editor has put together. The current difficulty arises because this has of late always been greater than the revenue raised in membership charges after other Society expenses have been deducted. A few years ago there was a surplus which could offset this. There is a surplus no longer.

It was therefore decided to delay the printing of this *Eureka*, to increase the subscription charge, and to put some limit on the size of the issue. This appears to be (for the moment at least) a satisfactory solution. Yet *Eureka* remains the major expenditure of the Archimedeans year. No doubt this is an excellent thing, and the journal has an international circulation—it is bought by departments and libraries from Tashkent to Salt Lake City. Yet in the darker moments of an Editor's life, it is sometimes tempting—considering the minimal feedback that *Eureka* generates—to wonder whether anyone actually reads the thing. (Getting articles is a matter of jumping up and down at people until, in a moment of weakness, they agree to write something, and then jumping up and down some more until they do write it. It is rare for people to submit articles if the Editor is not in a position to jump at them quite a lot.) No doubt this is undue pessimism and *Eureka* is really avidly devoured in six hundred households, but (gentle reader)—if you have the best interests of the journal at heart, try to remember, before laying yourself down to sleep at night, to murmur, "I *do* believe in *Eureka* Editors!" and to send in your articles, frivolous or serious but at any rate unsolicited and in plenty of time, to one of my successors.

The Archimedeans, 1990–1991

Michael Aird (Chronicler 1990–91)

The Archimedeans' year began in June. (I suppose I could start with April or May but nothing ever happens apart from the subgroups and the odd TMS meeting. Everyone is hard at work studying to get those grades, you see, and no-one wants to fiddle about with Puzzle Hunts and stuff.) Three punts full of Archimedeans made a long, arduous, nocturnal journey to Grantchester where the rain was minimal, and the first true sunshine of June greeted the garden party held on Pembroke library lawn.

October brought a new year and the Societies' Fair and squash. The intake was slightly down on last year, but the Puzzle Hunt was (almost) as successful as ever.

Disaster struck, however, in the middle of Michaelmas term when five members of the committee found themselves trapped in an old abandoned mine shaft. Luckily a nearby quick-thinking communist saved the day by concocting an ingenious plan to rescue them involving a rope, a phone and an oak tree.

A fine selection of talks were delivered this term, given by Dr R. Wilson of the Open University, Dr P. Neumann of Oxford, Prof Churchouse of Cardiff and Mr J. Beasley, author of the book *The Mathematics of Games*. Also lunchtime speaker meetings were revived: Mark Wainwright, editor of *Eureka*, gave a talk on "Maths, the Universe and Everything" after a light lunch had been provided by the catering manager.

The Musical Appreciation subgroup was also revived. Unfortunately, the music to be appreciated just wasn't appreciated enough and the subgroup returned to its restless sleep. Other subgroups, however, had a very nice time of it. The Puzzles and Games Ring gave out lots of coffee and lots of people came along every week to play on David Moore's computer. The Othello subgroup played Othello and did numerous other excruciatingly interesting things.

In Lent Term, talks were given by Dr N. Hitchin of Warwick, Prof Piper of Royal Holloway & Bedford New College, Prof Maynard-Smith of Sussex and Dr Edward de Bono. Lunchtime talks were given by Mr Paul Balister of Cambridge, Miss Emma Body of Cambridge, and Mr Graham Nelson of Oxford.

But the highlight of the Lent term (and the whole year) must have been the Triennial Dinner. Guests of honour were Prof P. M. Cohn and the Revd Dr J. C. Polkinghorne who entertained us with excellent after dinner speeches. (I must add that the President gave a fine speech also, which almost made up for the lack of the Christmas party he had promised at the AGM.)

And so finally the first Wednesday in March and the AGM. Thanks to some rather accurate minuting on the part of last year's Chronicler, the reading of the last AGM's minutes took about three-quarters of an hour, but we soon got round to the real business of trying to re-arrange the constitution and expelling members. The old Committee resigned their posts, weary, battle-hardened and bitter, to be replaced by fresh young faces, full of life, joy and ambition, eager to please the membership with new ideas and fresh orange juice.

At this point I'm supposed to sum it all up by saying it was another successful year, so I will. All in all, it was another successful year for the Archimedeans. Let's hope that next year will be just as successful.

Near Metric Theory

Paul Nicholls

Often mathematicians use the notion of a 'distance' between members of a set. The distance has to obey certain conditions: a point is zero distance from itself, but a positive distance from other points, and the distance between any two points via a third is at least that of going directly. Formally we define a *metric* on a set X to be a function $d: X \times X \to \mathbb{R}$ satisfying

$(M1)$ $\quad d(x,y) = d(y,x)$ for all $x, y \in X$

$(M2)$ $\quad d(x,y) + d(y,z) \geq d(x,z)$ for all $x, y, z \in X$.

$(M3)$ $\quad d(x,y) = 0$ if and only if $x = y$

Notice that we can deduce $d(x,y) \geq 0$ for all x and y: $2d(x,y) = d(x,y) + d(y,x) \geq d(x,x) = 0$. A pleasing feature of this proof is that it uses each condition exactly once.

Often proving a function is a metric requires a certain amount of tedious case-checking. Sometimes this may be simplified by use of the following method.

We first define a *near metric*. This is defined like a metric, but without the condition that $d(x,x) = 0$ for all x. Precisely, a function $d': X \times X \to \mathbb{R}$ is a near metric if it satisfies $(M1)$, $(M2)$ and

$(M3')$ $\quad d'(x,y) \geq 0$ for all $x, y \in X$

$(M4')$ $\quad d'(x,y) = 0$ implies $x = y$

Now we define the *derived function* d of a near metric d' by

$$d(x,y) = \begin{cases} 0 & \text{if } x = y \\ e(x,y) & \text{otherwise.} \end{cases}$$

We can now state the Near Metric Theorem:

THEOREM. *The derived function of a near metric is a metric.*

PROOF. As usual this is a question of checking cases. $(M1)$ is still satisfied, since the distances have not changed unless $x = y$, when they are both zero. Similarly $(M2)$ will certainly still be satisfied unless two of the points are now equal; then either $x = z$, when the RHS is zero, or $x = y$ or $y = z$, when one of the LHS terms is the same as the RHS term. Finally we must satisfy $(M3)$. By definition $d(x,x) = 0$ for all x. Also $d(x,y) = 0$ implies either that $x = y$, or that $d'(x,y) = 0$, when again $x = y$. $\square$

Applications of the Near Metric Theorem

The *discrete metric* on a set X is defined by $d(x,y) = 0$ if $x = y$ and $d(x,y) = 1$ otherwise. It is now easy to show that this is indeed a metric. We simply note that the function $d'(x,y) = 1$ is a near metric, and that d is its derived metric.

Let S be any set of non-negative reals, with $0 \in S$. We wish to construct a metric space X whose distance set (the set of distances between points of X) is S. An easy solution is to take S itself to be the set X, and define $d'(x,y) = \max(x,y)$. Note that d' is a near metric; now we define d to be its derived metric. All the distances are certainly elements of S; and every element of S is a distance, as $d(x,0) = \max(x,0) = x$.

Galois' Group

Amites Sarkar

Not long before Evariste Galois was shot in 1832, he had spent six months in jail, been thought of as one of the most dangerous men in France, and created group theory. He had also failed to enter the Ecole Polytechnique, publicly threatened to kill the king, attempted suicide, and, during his last weeks alive, fallen in love. These facts are known. Later, whilst biographers tried to extract more details from his turbulent twenty year life, mathematicians drew out the mathematics from the few memoirs he had left after his death.

Soon after some of his writings were published in 1846, Galois' genius was universally recognised. In modern terminology, Galois' contributions to mathematics include the theory of finite fields, group theory, and Galois Theory, each of which happens to depend on foundations laid generations after his death. In particular, the modern version of Galois Theory took a hundred years to evolve, and so gives little idea of its creator's views on his work. All this is fine, but since Galois anticipated so much of modern algebra these views are worth seeking and tracing to their birth.

Galois Theory did not spring into his head in a day or a dream, and was not dragged into the world on the night before the duel that killed him. It was the result of prolonged incisive reasoning on a specific problem in the theory of equations, described and analysed in detail in several memoirs written in the early nineteenth century. This was to find a formula involving only addition, subtraction, multiplication, division and extraction of roots, that would express a root of the equation

$$x^n + a_1x^{n-1} + \ldots + a_{n-1}x + a_n = 0$$

in terms of its coefficients $a_1, \ldots, a_{n-1}, a_n$. For $n = 1$ we have $x = -a_1$, while the next case is solved using the quadratic formula. When $n = 3$ we first write $3x = 3y - a_1$, to get $y^3 = b_1y + b_2$, and then introduce t, u such that $3tu = b_1$, $t^3 + u^3 = b_2$ (these yield a quadratic with roots t^3, u^3) giving the three roots as $x_1 = t+u$, $x_2 = wt + w^2u$, $x_3 = w^2t + wu$, where $w = e^{2\pi i/3}$. The case $n = 4$ is dealt with by removing the cubic term and comparing coefficients in the identity

$$x^4 + px^2 + qx + r = (x^2 + kx + l)(x^2 - kx + m)$$

to derive a cubic in k^2, hence finding k, l, m and four values of x. All four types had been solved by 1545 when Girolanio Cardano published his *Ars Magna* with detailed discussions of such equations. Cardano put little emphasis on quartic equations, since they had no obvious geometric interpretation, and thought it pointless to pursue things any further. It seemed fitting, therefore, to end his book with the words "Written in five years, may it last as many thousands."

It lasted less than two hundred. Mathematicians took Cardano's statement as a challenge and immediately tried to settle the problem when $n = 5$. It was hard work. Until Lagrange, few attempts were made to systematize the problem, the general idea being to try as many tricks as possible in the hope that one of them would kill the quintic. The area was a testing ground for mathematicians' ingenuity, in the sense that the above solution of the cubic makes ingenious use of the identity

$$(a+b)^3 = 3ab(a+b) + a^3 + b^3.$$

Sadly this was just the problem. Galois did not approve: he wanted everything in its logical place. In his *Discussions sur les Progrès de l'Analyse Pure*, he wrote in disgust:

> Take an Algebra book, be it a textbook or some original work, and you will see only a confused mass of propositions whose structures contrast strangely with the chaos of the whole. It seems that the ideas have already cost the author too much for him to take the trouble to tie them together and that the conception of the ideas underlying his work has so exhausted him that he cannot create a unifying idea. If you do find a method, a connection, a coordination, these are artificial and false.

In any case, attempts to solve the quintic were doomed: it cannot be solved. This is not because there is no formula "wide" enough to cover all cases, but because the roots of x^5+4x+2, for example, cannot be expressed using the four arithmetic operations and extractions of roots at all. No "radical expression" satisfies $x^5+4x+2=0$. Abel gave an incomplete proof in 1824, but still no "connection" was in sight. Galois found one: namely the fact that A_5 is a simple group.

Clues as to the state of Galois' knowledge in 1829 are easy to find. The 1896 biography by Paul Dupuy mentions that the young Evariste read Lagrange: references to Abel's memoirs appear frequently in Galois' manuscripts. The relevant treatise of Lagrange is his 1771 *Réflections sur la Résolution Algèbrique des Equations*, a lengthy, relaxed work with a double purpose best explained by its author:

> I propose in this memoir to examine the various methods found so far for the algebraic solution of equations, to reduce them to general principles, and to show *a priori* why these methods succeed for the third and fourth degrees, and fail for higher degrees.

Lagrange's extensive analysis of his predecessors' ways with cubics and quartics carried out in the first two sections of his memoir, led him to announce in the third:

> It follows ... that it is doubtful whether the methods we have just spoken of can give the complete solution of equations of the fifth degree.

Lagrange's results in the fourth section of his memoir are inspired by those of the first three, but can be motivated as follows:

Suppose we have a cubic equation $x^3+a_1x^2+a_2x+a_3=0$ with rational coefficients. Whether or not we can find its roots a, b, c we can certainly find some special functions of them: simply compare coefficients in $x^3+a_1x^2+a_2x+a_3=(x-a)(x-b)(x-c)$ to get $a+b+c=-a_1$, $ab+bc+ca=a_2$ and $abc=-a_3$. We can also find the value of $a^2+b^2+c^2$ since it is equal to $(a+b+c)^2-2(ab+bc+ca)=a_1^2-2a_2$. As it happens, any polynomial $S(a,b,c)$ unchanged under all six permutations of a, b, c can be written as a polynomial in a_1, a_2 and a_3. But this will not help us with our equation, as we have to find a, b and c themselves. In a precise sense, you can measure the progress you have made in solving an equation by the lack of symmetry in the functions of its roots that you have calculated. The less symmetry you've got, the better you're doing. The "symmetric polynomials" $S(a,b,c)$ referred to above have the most symmetry, being unchanged under six permutations of a, b, c: they have rational values and are easy to calculate in terms of the rational numbers a_1, a_2, a_3. The polynomial $\Delta = a^2b+b^2c+c^2a$ is unchanged under only three permutations, the other three taking it to $\Delta' = ab^2 + bc^2 + ca^2$. If we can find Δ in terms of a_1, a_2 and a_3 we will have taken a step forward. To do this, consider $(\Delta - \Delta')^2$. The three cyclic permutations of a, b, c leave it untouched, while the rest interchange Δ and Δ', negating $\Delta - \Delta'$ but leaving $(\Delta - \Delta')^2$ as it was before. So $(\Delta - \Delta')^2$ is unchanged under all six permutations, and can be expressed as a polynomial in a_1, a_2 and a_3. Explicitly, we have the identity

$$\begin{aligned}(\Delta - \Delta')^2 &= (a^2b + b^2c + c^2a - ab^2 - bc^2 - ca^2)^2 \\ &= a_1^2a_2^2 - 4a_1^3a_3 - 4a_2^3 - 27a_3^2 + 18a_1a_2a_3.\end{aligned}$$

$\Delta + \Delta'$ is also a symmetric polynomial: it is $-a_1a_2 + 3a_3$, so that Δ and Δ' are given by

$$\tfrac{1}{2}\left(-a_1a_2 + 3a_3 \pm \sqrt{a_1^2a_2^2 - 4a_1^3a_3 - 4a_2^3 - 27a_3^2 + 18a_1a_2a_3}\right).$$

Not only have we had to resort to a trick to determine Δ; it is not generally a rational number. We have got closer to the solution, but we have paid the price. To find a, b and c, and hence functions which remain unchanged under only the identity permutation, we must extract a further root, a cube root. It turns out the the extraction of roots is the only tool we have for breaking symmetry, and this also has its limitations. For quintic equations, no roots of any degree can destoy enough symmetry to yield a solution.

Lagrange puts things slightly more precisely in Article 104:

> If t and y are any two functions of the roots x', x'', x''', ... of the equation
>
> $$x^\mu + mx^{\mu-1} + nx^{\mu-2} + \ldots = 0,$$
>
> and if these functions are such that all the permutations of the roots x', x'', ... which change the function y also change the function t, one can, generally speaking, calculate the value of y in terms of t and $m, n, p, \ldots$ using a rational expression, so that on knowing a value of t one will also know immediately the corresponding value of y ...

Lagrange writes "generally speaking" because his theorem fails if the equation has two or more equal roots. The functions t and y are, loosely speaking, related to each other as follows: y is more symmetrical than t. Any two functions with the same amount of symmetry are rationally expressible in terms of one another. On finding the function Δ above, we have not only found one function unchanged under only three permutations; we have got all of them. An example is the function $a^3b^5 + b^3c^5 + c^3a^5$, which can be expressed as a polynomial in the symmetric functions, Δ and Δ'. No roots need be taken.

Sixty years later, all of Galois' expositions of Galois theory were rejected by the establishment. There are many reasons why the mathematical community did not idolise Galois during his lifetime: one may be that his ideas were new, while a more certain one is that they were badly explained. Galois invented groups: these make their debut in the statement of Proposition 2 in his *Mémoire sur les Conditions de Résolubilité des Equations par Radicaux*. No definition appears, only some clarification inserted a few days before his death. It seems that his own conception of groups was as follows:

Write down some letters: $a\ b\ c\ d\ e$. Rearrange them: $d\ c\ a\ b\ e$. Rearrange them again: $b\ c\ e\ d\ a$. Then write the arrangements in a column:

$$\begin{array}{ccccc} a & b & c & d & e \\ d & c & a & b & e \\ b & c & e & d & a \end{array}$$

Now we define three substitutions: the identity substitution which leaves the letters as they are, the substitution $a \to d,\ b \to c,\ c \to a,\ d \to b,\ e \to e$, which takes the first row to the second, and the substitution $a \to b,\ b \to c,\ c \to e,\ d \to d,\ e \to a$, which takes the first row to the third. (We only consider substitutions operating on the first row and ignore the four other possibilities.) Galois sees only the substitutions as significant, attaching no importance to the static arrangements. They are there only because "one can hardly form the idea of a substitution without that of a permutation". Next rearrange the rearrangements:

$$\begin{array}{ccccc} b & c & e & d & a \\ d & c & a & b & e \\ a & b & c & d & e \end{array}$$

This time the substitutions are different. They are the identity, $b \to d,\ c \to c,\ e \to a,\ d \to b,\ a \to e$ and $b \to a,\ c \to b,\ e \to c,\ d \to d,\ a \to e$. This means that there is no group in sight. For the arrangements to define substitutions that form a group, these substitutions should not depend on the order in which the arrangements are written. Such arrangements are, for example:

$$\begin{array}{ccccc} a & b & c & d & e \\ a & c & b & d & e \\ a & b & c & e & d \\ a & c & b & e & d \end{array}$$

All four of the usual group axioms follow easily: Galois' substitutions form groups. Galois verifies one axiom on the spot in his memoir:

> As one is always concerned with questions where the original position of the letters is irrelevant, in the groups which we consider, one must have the same substitutions whatever the permutation which one leaves. Thus if in such a group one has the substitutions S and T, one is sure of having the substitution ST.

Later on, in some rough working in the margin, Galois denotes the combination of the permutations S and T by $S + T$. In a note on the same page, he writes "The substitutions are independent even of the number of roots." What Galois did not write, although he probably knew it, was that the substitutions are independent of just about everything. It doesn't even matter that the word substitution means what it does: it can just refer to a "thing" as long as "things" in a group satisfy some rules of combination. The combination of S and T can be written as ST, $S + T$, or $S\,!\,T$, according to taste. Once Galois had founded groups, his position in the history books was reserved, even if certain facts about equations couldn't have been deduced from them. As it happens, they can.

As an illustration, we solve the cubic with the help of groups, using the method Galois outlined in his memoir. Consider the general cubic $x^3 + a_1x^2 + a_2x + a_3 = 0$, with solutions α, β, γ. We must express α, β, γ in terms of $\alpha+\beta+\gamma=-a_1$, $\alpha\beta+\beta\gamma+\gamma\alpha=a_2$ and $\alpha\beta\gamma=-a_3$. Assume, then, that we know the numerical values of a_1, a_2 and a_3. Write out the six arrangements of α, β, γ (these define permutations which form a group). Here they are:

$$(\alpha\beta\gamma, \beta\gamma\alpha, \gamma\alpha\beta, \alpha\gamma\beta, \beta\alpha\gamma, \gamma\beta\alpha). \qquad \text{(list 1)}$$

Now I'll gratuitously invent some terminology: we call a set of arrangements which defines a group a *list*. There are two stages in the solution, each one involving (a) a division of a list into two smaller lists, (b) three functions p, q and r, and (c) the extraction of a root (e.g. a cube root).

Stage 1

Divide list 1 above into the two lists

$$(\alpha\beta\gamma, \beta\gamma\alpha, \gamma\alpha\beta) \quad \text{(list 2)} \qquad \text{and} \qquad (\alpha\gamma\beta, \beta\alpha\gamma, \gamma\beta\alpha) \quad \text{(list 3)}.$$

Both lists 2 and 3 define the same substitutions $((\alpha,\beta,\gamma)\to(\alpha,\beta,\gamma)$, $(\alpha,\beta,\gamma)\to(\beta,\gamma,\alpha)$, $(\alpha,\beta,\gamma)\to(\gamma,\alpha,\beta))$: this is crucial. I'll also want to refer to the lists 1a, 2a and 3a which are obtained by replacing α, β and γ by x, y and z in lists 1, 2 and 3 respectively.

Now choose a function $p_3(x,y,z)$ such that $p_3(x,y,z)$ is formally unchanged under the substitutions of list 2a (and hence unchanged under those of list 3a), whilst $p_3(x,y,z)\neq p_3(x,z,y)$. Our aim is to find the numerical values of $p_3(\alpha,\beta,\gamma)$ and $p_3(\alpha,\gamma,\beta)$. Next define

$$\begin{aligned} q_3(x,y,z) &= p_3(x,y,z) - p_3(x,z,y), \\ r_3(x,y,z) &= p_3(x,y,z) + p_3(x,z,y). \end{aligned}$$

Now $q_3(x,y,z)^2$ and $r_3(x,y,z)$ are functions unchanged under all the substitutions of list 1a. A theorem similar to Lagrange's says that each of $q_3(x,y,z)^2$ and $r_3(x,y,z)$ can be expressed as ratios of polynomials with rational coefficients in $(x+y+z)$, $(xy+yz+zx)$ and xyz. But this means that we can find the numerical values of $q_3(\alpha,\beta,\gamma)^2$ and $r_3(\alpha,\beta,\gamma)$.

Finally, observe that

$$\left.\begin{array}{l} p_3(\alpha,\beta,\gamma) \\ p_3(\alpha,\gamma,\beta) \end{array}\right\} = \tfrac{1}{2}\{r_3(\alpha,\beta,\gamma) \pm \sqrt{q_3(\alpha,\beta,\gamma)^2}\}$$

Note that $p_3(\alpha,\beta,\gamma)$ is indistinguishable from $p_3(\alpha,\gamma,\beta)$; which is which depends on the numbering of the roots. As an example, we could take $p_3(x,y,z)=x^2y+y^2x+z^2x$.

Stage 2

Divide the list $(\alpha\beta\gamma, \beta\gamma\alpha, \gamma\alpha\beta)$ into the three lists

$$(\alpha\beta\gamma) \quad \text{(list 4)}, \qquad (\beta\gamma\alpha) \quad \text{(list 5)} \qquad \text{and} \qquad (\gamma\alpha\beta) \quad \text{(list 6)}.$$

Let $p_1(x,y,z)$ be such that $p_1(x,y,z)\not\equiv p_1(y,z,x)$ so that $p_1(y,z,x)\not\equiv p_1(z,x,y)$. We must find $p_1(\alpha,\beta,\gamma)$, $p_1(\beta,\gamma,\alpha)$ and $p_1(\gamma,\alpha,\beta)$. Define

$$\begin{aligned} q_1(x,y,z) &= p_1(x,y,z) + \omega p_1(y,z,x) + \omega^2 p_1(z,x,y), \\ r_1(x,y,z) &= p_1(x,y,z) + p_1(y,z,x) + p_1(z,x,y), \end{aligned}$$

where $\omega = e^{2\pi i/3}$.

Then $q_1(x,y,z)^3$ is unchanged under the substitutions of list 2a, since if we define $p_1(x,y,z) = A$, $p_1(y,z,x) = B$, $p_1(z,x,y) = C$, we obtain

$$\begin{aligned}
q_1(y,z,x)^3 &= (B + \omega C + \omega^2 A)^3 = (\omega^2(A + \omega B + \omega^2 C))^3 \\
&= (A + \omega B + \omega^2 C)^3 = q_1(x,y,z)^3, \\
q_1(z,x,y)^3 &= (C + \omega A + \omega^2 B)^3 = (\omega(A + \omega B + \omega^2 C))^3 \\
&= (A + \omega B + \omega^2 C)^3 = q_1(x,y,z)^3.
\end{aligned}$$

So $q_1(x,y,z)^3$ and $r_1(x,y,z)$ are unchanged under the substitutions in question. But by an extension of the theorem quoted above, these functions can be written as ratios of polynomials in $(x+y+z)$, $(xy+yz+zx)$, xyz, $p_3(x,y,z)$ and $p_3(y,z,x)$. This means that we know the numerical values of $q_1(\alpha,\beta,\gamma)^3$, $q_1(\alpha,\gamma,\beta)^3$ and $r_1(\alpha,\beta,\gamma)$.

Lastly for some choice of cube roots,

$$\left.\begin{array}{c} p_1(\alpha,\beta,\gamma) \\ p_1(\beta,\gamma,\alpha) \\ p_1(\gamma,\alpha,\beta) \end{array}\right\} = \tfrac{1}{3}\left\{ r_1(\alpha,\beta,\gamma) + \sqrt[3]{q_1(\alpha,\beta,\gamma)^3} + \sqrt[3]{q_1(\alpha,\gamma,\beta)^3} \right\}$$

We can choose $p_1(x,y,z) = x$ to complete the solution at this stage.

Having failed the entrance exam for the Ecole Polytechnique, Galois was officially enrolled as a student at the Ecole Normale (a far less prestigious place than the Polytechnique) in the early months of 1830. It was here that he suffered a series of rejections of manuscripts, involving some of the most famous French mathematicians of the day. As already mentioned, this was mainly his fault each time, but it is worth noting that Cauchy fell ill so that he couldn't present Galois' work to the Academy of Sciences, Fourier died so that he couldn't judge it, and Poisson declared it "incomprehensible" having had several months to read it.

When Galois was thirteen, Charles X had succeeded Louis XVIII. Six years later, however, the liberal opposition had made a majority in the elections: Charles attempted a coup d'état, resulting in a revolution. It seems likely that at this stage in the proceedings the young Evariste would have loved to be fighting on the streets. The director of the Ecole Normale, M Guiguiault, had other ideas: he locked all his students in. An incensed Galois later wrote a letter to the Gazette des Ecoles, attempting to expose this man who was by then "engaging in polemics against students in the pages of several newspapers". The letter was published anonymously while Galois was expelled. He made an attempt at organizing a private class in mathematics, but people were aware of his political activities and, after the first week, few turned up.

There are two schools of thought concerning Galois' entry into republican revolutionary politics. One is that his republican involvement was directly due to the repeated rejection of his papers: he had come to hate the whole French "establishment", not just the mathematical one, and so he turned against it. The other, perhaps more plausible, one is that Galois had been a revolutionary all along: one with a fierce temper and little patience.

Of course, neither of these claims can be proved. But if Galois' manuscripts are anything to go by, we can give a good argument for the second. In the memoir referred to above, the phrase "It is clear that ..." appears six times, usually in place of what should be a long proof. Throughout the work, the "given equation" is referred to: not once does

Galois write this equation out. On the first page, he erases the telling statement: "I leave out proofs that are too easy." Many proofs are completed in the margin, and some finish simply with "Therefore, etc." It seems likely that Galois' mathematical impatience and anger is on a par with, and closely related to, his impatience in general. In his essay *Sur l'Enseignment des Sciences* he writes words to the effect that any hierarchy condemns genius in favour of mediocrity.

In a similar way (i.e. much of this is speculation too) Galois' attitude to the future standing of his discoveries can be glimpsed through a few of his offhand comments. At the end of the string of definitions which begins his memoir, he writes "These are the definitions I believed necessary to recall", followed by this, crossed out: "Perhaps they appear superfluous. But we prefer verbosity to obscurity." His reluctance to include them indicates his eagerness to get to the heart of things, in the second part of his work. Lemmas I and II get a one-sentence proof each. Poisson described the proof of Lemma III as "insufficient". Not that the proofs get any better. Anyone reading this memoir would notice something strange happen after Lemma IV. The tone becomes less terse and even slightly persuasive. In the next theorem (Proposition I) he not only gives two examples, but attaches two notes as well. Also, the word "group" starts to appear with increasing frequency. By the end of the manuscript, it seems that Galois has almost completely forgotten about the "given equation". The proofs continue to get worse, the emphasis being placed not on the logic, nor even on the results, but on the form of the results. Galois knew that he had found the governing principle in the theory of equations and that the acknowledged experts of the day were not only proving the wrong theorems but the wrong kind of theorems. What was more frustrating for Galois was that groups are actually mathematically simpler objects than transformations of polynomial equations.

Suppose someone approaches you in the park and tells you that all the world's problems could be solved if only everyone would worship the sun. You tell him he's talking rubbish because, you say, "it's far more complicated than that, and what's sun-worshipping got to do with it anyway?" An early nineteenth century mathematician might have had a similar attitude to Galois' work, believing that "A load of permutations is not only far simpler than, but also completely irrelevant to, the theory of equations." It's quite easy to sympathize with Poisson. The situation is worsened by Galois sacrificing local explanations in favour of a global understanding of his theory. Maybe he was making a bet with the establishment, with the future as referee. (Something like "When all this has finally been sorted out, groups will be of central importance, you'll see!")

Here is what happened next. Galois was arrested twice, imprisoned for six months, and put on parole. Then he fell in love with a Stéphanie-Félicie Poterin du Hutel. This is a translation of part of a letter she wrote to him (taken from a copy Galois made; it seems he omitted some phrases that were incriminating or distasteful):

> Please let us break up this affair. ... I do not have the wit to follow ... a correspondence of this nature ... but I will try to have enough to converse with you as I did before anything happened. [...] and do not think about those things which did not exist and which never would have existed.

Soon afterwards, Evariste was challenged to a duel, probably over Stéphanie, which he lost. He died a day later, perhaps believing that what he had discovered would remain hidden from the world forever. I doubt it.

Letter Addition and Doubly Self-Referential Haiku

Anton Cox

In English (as in other languages) it is possible to construct self-referential sentences, for example, "This sentence refers to itself".

As mathematicians appear to be the group who most enjoy constructing such sentences, it seems a little ironic that they must "abandon" their calling (at least for the duration of the exercise) to indulge in this pastime. Indeed, it is possible to view the introduction of Gödel numbering earlier this century as an (albeit unconscious) attempt to rid mathematicians of their collective guilt.

It is however noticeable that very few self-referential sentence addicts (SRSAs) have been converted to constructing their sentences using Gödel's method. This article is intended to suggest an alternative solution more easily applicable to the search for elegant self-reference. This method is known as the Alternative Gödel Construction (or AGC). When learning about ciphers—or codes as they are often erroneously called—the following system is usually encountered at an early stage:

$$A = 1,\ B = 2,\ \ldots,\ Z = 26 \qquad \text{[All punctuation equals zero]}$$

Using this method, any sentence can be converted into a number. Under AGC, this will be achieved using the method of letter addition, where the values of the letter are added together.

For example: "This sum" $= 20 + 8 + 9 + 19 + 19 + 21 + 13 = 109$

It is now possible to construct sentences that are mathematically self-referential. However it is desirable, in the interests of elegance, to follow certain guidelines:

i) Sentences should be as short as possible.

ii) Sentences should be more or less grammatically correct, and the use of redundant words minimised.

A simple example of a sentence that satisfies these guidelines and that refers to itself is:

$$\text{"Two hundred and fifty one"} = 251 \tag{1}$$

Slightly more adventurous are the following:

$$\text{"One hundred and twelve times three"} = 336 \tag{2}$$

$$\begin{array}{l}\text{"Two hundred and eighty-eight plus} \\ \quad\text{three hundred and forty three"} = 631\end{array} \tag{3}$$

These may be regarded as a little unwieldy, and the brevity of the following may be preferred:

$$\text{"it's five cubed"} = 125 \tag{4}$$

$$\text{“Fifteen squared equals \ldots”} = 225 \tag{5}$$

More of a challenge to SRSA is the Alternative Gödel Construction Of Multiplication (AGCOx) where the letters are not added but multiplied. It was conjectured that under AGCOx the only self-referential sentences were of the general form:

$$\text{“Zero!”} = 0$$

However, two distinct, but related, examples have been discovered:

$$\frac{\text{“Sixty”}}{\text{“by six”}} = \frac{2052000}{205200} = 10 \tag{6}$$

$$\frac{\text{“Seventy”}}{\text{“by seven”}} = \frac{73150000}{7315000} = 10 \tag{7}$$

Returning to AGC, it is clear that dividing sentences (as in the examples above) allows the creation of non-integer self-referential sentences, such as:

$$\frac{\text{“The nth root of”}}{\text{“(four over eight) to the power n”}} = \frac{164}{328} = \frac{1}{2} \tag{8}$$

In the case of irrational numbers it is desirable to introduce further guidelines concerning accuracy.

iii) All approximations to rational numbers should include an indication of their accuracy.

iv) Approximations should be as accurate as considerations of (i) will allow.

Aesthetically, (iv) is the most important factor. Some examples of well known irrationals are:

$$\frac{\text{“Circumference of a circle over (two times its radius)”}}{\text{“To three digits”}} = \frac{499}{159} = 3.14 \tag{9}$$

$$\frac{\text{“The base of the natural logs (in decimal notation)”}}{\text{“To three digits”}} = \frac{432}{159} = 2.72 \tag{10}$$

This accuracy method should also be used for rational number approximations in exceptional circumstances, e.g.

$$\frac{\text{“To three digits”}}{\text{“The cosine of pi by three (in radians)”}} = \frac{159}{316} = 0.50 \tag{11}$$

Finally, the ideal sentence is surely one which is doubly self-referential; i.e. both in the normal sense and under AGC.

A nice example of this is the following haiku (a traditional Japanese poetry form). Not only does it describe its own structure, it includes its own letter sum:

“A haiku? Three lines
In seventeen syllables
Of five, seven, five” = 575.

The Prisoner's Dilemma

Kevin P. Murphy

A PARADOX OF RATIONALITY: THE ONE-SHOT PRISONER'S DILEMMA

Imagine you and an accomplice are thrown into two separate police cells and made the following deal: "If you both stay mum, you'll both be in for two years, since we have some circumstantial evidence against you. But if you squeal on your accomplice, we'll let you out scot free, and he'll get five years." Nervously you ask, "But what if we *both* squeal?" "Then you'll both get four years."

Well, now you're in a pickle! You reason to yourself: if he co-operates with me by remaining silent, I'd be best off defecting by squealing on him; that way I'd get off scot free. But if he squeals, I'd again be better off squealing, because then I'd get 4 instead of 5 years. So I'm better off defecting whatever he does. (This is called a *dominant strategy.*) But the same reasoning goes through the mind of your accomplice—and you both end up spending 4 years in the slammer, when you could have got away with just 2, if only he had co-operated. If only ...

A\B	C	D
C	$3,3 =$ R,R	$0,5 =$ S,T
D	$5,0 =$ T,S	$1,1 =$ P,P

TABLE 1. The payoff matrix for the Prisoner's Dilemma in its canonical form, due to Axelrod. There are two players, **A** (row) and **B** (column). They each choose either to co-operate (C) or defect (D). The payoffs to **A** are listed first. For example, if **A** chooses C and **B** chooses D, **A** gets no years of freedom (= five years in jail) and **B** gets five years free (no years in jail).

This delicious paradox was invented by Melvin Dresher and Merrill Flood, of the RAND Corporation, in about 1951. Albert Tucker wrote the first article on it, and gave it its now-famous name. We can represent it by the payoff matrix shown in Table 1. The size of the actual numbers (or *cardinal utilities*) is not important; all that matters is the *relative values* or *ordinal ranks* of the values. We set

T = temptation to defect
R = reward for co-operation
P = punishment for common defection
S = sucker's payoff

The Prisoner's Dilemma (or PD) is then formally defined by the requirements

$$T > R > P > S \tag{1}$$

and

$$R > \frac{T+S}{2}. \tag{2}$$

Note that the payoffs do not need to be equal for each player, as long as (1) and (2) hold for each player separately. The (P, P) strategy is the only *nash equilibrium*; i.e. a strategy such that either player's deviation from it is disadvantageous to that player. However, it is also *pareto-inferior*, or less than optimal for both sides. Although (R, R) is *pareto-superior*, each player has a temptation to do even better, rendering that strategy unstable. This is the dilemma.

Most of the games people are familiar with are *zero sum*, or at least *constant sum*, where one player wins whatever the other loses (so adding the payoffs always gives the same constant). In a constant sum game, playing to do well for yourself is the same as playing to win—they are games of *total conflict*. The PD is not like that. The sum of the payoffs (if measured cardinally by utilities rather than ordinally by ranks) depends on which strategies are chosen, and hence it is possible to do well without winning and vice versa. (When the PD is played once, it is called *one-shot*.)

A solution to the paradox: Metagame theory

Nigel Howard was the first to use metagame theory to resolve the paradox of the PD.

Consider two players, A and B, playing a one-shot PD. However, we will allow one of the players, A say, four *conditional strategies*

- A_1—choose C regardless of what you think B will choose
- A_2—choose whatever you think B will choose
- A_3—choose the opposite of whatever you think B will choose
- A_4—choose D regardless of what you think B will choose.

B\A	A_1	A_2	A_3	A_4
C	3, 3	3, 3	0, 5	0, 5
D	5, 0	1, 1	5, 0	$*1, 1*$

TABLE 2. Payoff matrix for the metagame: B's scores are listed first. The equilibrium is highlighted. (Rapoport (1967), p. 54.)

The extended payoff matrix is shown in Table 2. The only nash equilibrium is (D, A_4). Let us see why this point is in equilibrium. Given that B has chosen to defect, A can only choose outcomes from the bottom row. Since $1 > 0$, A cannot do better by choosing something other than (D, A_4). Given that A has chosen A_4, B cannot do better than choose (D, A_4). Hence this point is stable for both players. However, it is still pareto-inferior. The paradox remains.

But now let B choose C or D in response to each of A's conditional strategies. B now has $4^2 = 16$ strategies to choose from. B may choose to co-operate regardless of what A does ($CCCC$); B may choose to co-operate only if A plays A_1 ($CDDD$); B may choose to co-operate only if A plays A_2 ($DCDD$); and so on. The new payoff matrix is shown in Table 3.

But look! There are two new nash equilibria, and they are both pareto-superior. It is clear that A should choose A_2. B should choose $DCDD$ (instead of $CCDD$), since that way, if A were to play A_1, B would get 5 instead of just 3. As Rapoport writes,

> Thus the paradox of the Prisoner's Dilemma game was solved. Individual rationality and collective rationality were reconciled. Howard gave the lemma the *coup de grâce* by proving that no new equilibria would appear if metagames of a higher order were constructed. In other words, no further

TABLE 3. Payoff matrix for the meta meta game: B's scores listed first. The equilibria are highlighted. (Based on Rapoport (1967))

B\A	A_1	A_2	A_3	A_4
CCCC	3, 3	3, 3	0, 5	0, 5
CCCD	3, 3	3, 3	0, 5	1, 1
CCDC	3, 3	3, 3	5, 0	0, 5
CCDD	3, 3	*3, 3*	5, 0	1, 1
CDCC	3, 3	1, 1	0, 5	0, 5
CDCD	3, 3	1, 1	0, 5	1, 1
CDDC	3, 3	1, 1	5, 0	1, 1
CDDD	3, 3	1, 1	5, 0	1, 1
DCCC	5, 0	3, 3	0, 5	0, 5
DCCD	5, 0	3, 3	0, 5	1, 1
DCDC	5, 0	3, 3	5, 0	0, 5
DCDD	5, 0	*3, 3*	5, 0	1, 1
DDCC	5, 0	1, 1	0, 5	0, 5
DDCD	5, 0	1, 1	0, 5	1, 1
DDDC	5, 0	1, 1	5, 0	0, 5
DDDD	5, 0	1, 1	5, 0	*1, 1*

extension of the process is necessary: after just two steps, the set of equilibria is closed.†

THE EVOLUTION OF CO-OPERATION: THE DILEMMA ITERATED

The following discussion shows how co-operation can emerge in a world of egoists, without central authority (an example of a benign phenomenon emerging from a non-benign system). Hobbes, for one, argued that this could not happen; that in a world without a central authority people would compete with each other so fiercely that life would be "solitary, poor, nasty, brutish and short".

Axelrod (1984) refutes this, giving many examples, including

- import/export agreements,
- the development of the 'live and let live' system during the trench warfare of World War I,
- examples from biology (see especially Axelrod and Hamilton (1981), and also Dawkins (1989)).

The Iterated Prisoner's Dilemma

Let us get the bad news out of the way first. If the game is played a known finite number of times, there is no rational reason to co-operate. Clearly this is true on the last move, for there is no future to influence (this is a one-shot PD). It is therefore also true on

† Rapoport (1967) p. 55.

the next to last move, since the last move is effectively pre-determined, and hence, by backward induction, defection is always the rational choice. (In practice, people do not always play like this, possibly because of limited information-processing power.†)

When the number of iterations is uncertain, however, co-operation can emerge. We can represent this uncertainty by multiplying future payoffs by a weight, or discount parameter, $0 < w < 1$. This can be interpreted as the probability of the game's continuing, or as the fact that future gains are not valued as highly as current gains. For example, "a bird in the hand is worth two in the bush" implies $w = 1/2$. Now define $V(X \mid Y)$ to be the expected payoff that someone using strategy X gets when playing against someone using strategy Y (the *value* of X given Y). If two people play the *strategy* **ALL C** (always co-operate), they will get

$$V(C \mid C) = R + wR + w^2R + \ldots = R\sum_{j=0}^{\infty} w^j = \frac{R}{1-w}.$$

Using the numerical values in Table 1, and $w = 0.5$, this gives $V(C \mid C) = \frac{3}{1-0.5} = 6$.

We can define a *strategy* as a function from the history of the game so far to a probability of co-operation on the next move. Here is a list of most of the example strategies we shall encounter later (there are of course an infinite number of strategies):

ALL C	always co-operate
ALL D	always defect
DC	alternate D with C (in order)
CD	alternate C with D (in order)
TFT (Tit For Tat)	do whatever the other player did on the previous move; on the first move co-operate
TF2T (Tit For Two Tats)	always co-operate unless the other player makes two consecutive defections, in which case defect once
GRUDGER	co-operate until the other player defects, then play **ALL D**
RANDOM	choose C or D randomly.

What is the best strategy? Unfortunately there is no simple answer to that; it depends on the strategies of the other players.

LEMMA 1. **ALL D** *cannot do better against* **TFT** *than* **TFT** *can if* $w \geq \frac{T-R}{T-P}$.

PROOF.

$$V(\mathbf{TFT} \mid \mathbf{TFT}) = R + wR + w^2R + \ldots = \frac{R}{1-w} \tag{1}$$

$$V(\mathbf{ALL\ D} \mid \mathbf{TFT}) = T + wP + w^2P + \ldots = T + \frac{wP}{1-w} \tag{2}$$

$$V(\mathbf{ALL\ D} \mid \mathbf{TFT}) \leq V(\mathbf{TFT} \mid \mathbf{TFT}) \iff T + \frac{wP}{1-w} \leq \frac{R}{1-w}$$

$$\iff T - wT + wP \leq R \iff w \geq \frac{T-R}{T-P} \qquad \square$$

PROPOSITION 2. *If the discount parameter w is sufficiently large, there is no best strategy independent of the other player's.*

† See Axelrod and Dion (1988) for references.

PROOF. Suppose there were such a best strategy **BEST**. In particular, **BEST** would be the best strategy when playing against **ALL D** *and* **TFT**. Now the best strategy against **ALL D** is **ALL D** (as we have seen previously). But **TFT** is better against **TFT** than **ALL D** is, if w is sufficiently large, by Lemma 1. □

The Computer Tournaments

To say that it is a necessary condition for the evolution of convergence that w is high enough is not the same as saying that this is sufficient. To explore the kinds of strategies that are effective, Axelrod decided to conduct a computer tournament. He had 14 entries from professional game theorists. He conducted a round robin tournament, in which each program played everyone else's, as well as its own twin, and **RANDOM**. The resulting 15×15 matrix of scores is shown in Table 4, using the same utility values as before. There were 200 moves in the game,† so a very good score would be $3 \times 200 = 600$ (for mutual co-operation), and a poor score would be $1 \times 200 = 200$ (for mutual defection), although scores from zero to 1000 are theoretically possible.

Player	TFT	T/C	N	Gm	S	S/R	Fr	Da	Gk	Do	Fd	J	T	A	R	Av.
TIT FOR TAT	600	595	600	600	600	595	600	600	597	597	280	225	279	359	441	504
TIDE.&CHIER.	600	596	600	601	600	596	600	600	310	601	271	213	291	455	573	500
NYDEGGER	600	595	600	600	600	595	600	600	433	158	354	374	347	368	464	486
GROFMAN	600	595	600	600	600	594	600	600	376	309	280	236	305	426	507	482
SHUBIK	600	595	600	600	600	595	600	600	348	271	274	272	265	448	543	481
STEIN&RAPOPORT	600	596	600	602	600	596	600	600	319	200	252	249	280	480	592	478
FRIEDMAN	600	595	600	600	600	595	600	600	307	207	235	213	263	489	598	473
DAVIS	600	595	600	600	600	595	600	600	307	194	238	247	253	450	598	472
GRAASKAMP	597	305	462	375	348	314	302	302	588	625	268	238	274	466	548	401
DOWNING	597	591	398	289	261	215	202	239	555	202	436	540	243	487	604	391
FELD	285	272	426	286	297	255	235	239	274	704	246	236	272	420	467	328
JOSS	230	214	409	237	286	254	213	252	244	634	236	224	273	390	469	304
TULLOCK	284	287	415	293	318	271	243	229	278	193	271	260	273	416	478	301
ANON	362	231	397	273	230	149	133	173	187	133	317	366	345	413	526	282
RANDOM	442	142	407	313	219	141	108	137	189	102	360	416	419	300	450	276

TABLE 4. Tournament scores, first round. (Axelrod (1984) p. 194)

As you can see, **TFT**, the simplest of all the rules, won. It was submitted by Anatol Rapoport, and averaged 504 points per game. It turns out that there is a single property that distinguishes the high scoring rules from the low scoring ones. If a rule is *nice*, it is never the first to defect. Each of the top 8 ranking rules is nice. None of the other

† Since all players knew this, some rules started to defect near the end of the game. The very fact that some rules did *not* defect for the *whole* game suggests that the "rational" solution to the finitely repeated PD (namely, **ALL D**) is shaky. End-game variations accounted for very little of the score differences.

entries are. There is even a substantial gap in the scores between the nice and non-nice rules.

Nice rules did well with each other. What distinguished their relative scores was their behaviour with non-nice rules. A rule can be described as *forgiving* if it co-operates after the other player has defected. **TFT** is forgiving, apart from the single round of retaliation. Most rules were considerably less forgiving than this, and so non-nice strategies tended not to do well. **JOSS**, for example, was a variant of **TFT** which, like **TFT**, defects if its opponent defected in the previous game; however, unlike **TFT**, it also defects randomly (with a 10% chance) when it would otherwise have co-operated, in an attempt to sneak some advantage. What actually happens, when it plays against **TFT**, is that both players co-operate until **JOSS** sneaks a defection; then **TFT** retaliates on the next move, **JOSS** retaliates for this retaliation on the subsequent move, and an "echo" is set up, whereby each player alternately defects in retaliation for the rest of the game.† When **JOSS** randomly defects for a second time, it initiates another, interlocking, series of defections, with the result that both players now defect continually until the end of the game. Axelrod comments,

> A major lesson of this tournament is the importance of minimising echo effects in an environment of mutual power. When a single defection can set off a long string of recriminations and counter-recriminations, both sides suffer. ...the real costs may be ... when one's own isolated defections turn into unending mutual recriminations. Without realising it, many of the rules wound up punishing themselves. With the other player serving as a mechanism to delay the self-punishment by a few moves, this aspect of self-punishment was not picked up by many of the decision rules.‡

Axelrod calculated that if **TF2T**, a *more* forgiving version of **TFT**, had been submitted, it would have won.

Axelrod decided to publish the results and to hold a second round, with the hope that people could learn from their mistakes. This time he received 62 entries from 6 countries. This resulted in a massive 63×63 score matrix. End game effects were eliminated by determining the end of each particular game probabilistically. Setting $w = 0.99654$ corresponds to a median game length of 200 moves.

Again *TFT* won. Again nice rules did best. But it now emerged that *retaliatory* rules also did well; these were the rules which "punished" defectors immediately and so could not be taken advantage of. This seemed to be particularly important in this second round, because some people concluded from the first round that "it pays to be nice and forgiving", while others concluded "if others are going to be nice and forgiving, it pays to exploit them". **TF2T** was submitted by John Maynard Smith, an evolutionary biologist, but it only came twenty-fourth—it was exploited by, amongst others, **TESTER**, which could "rip off" **TF2T** and get away with it by never defecting twice in a row.

Axelrod identified a fourth useful attribute of **TFT**: *clarity*. When a player encounters someone using **TFT**, s/he recognises this fact fairly quickly, and realises that the best thing to do is to co-operate. Even rules like **TESTER** will back off once they see that **TFT** will not be exploited. Axelrod warns against being too clever, because strategies which were too complex were indistinguishable from the random, and so gave no incentive to co-operation. Note the sharp difference between this and zero-sum

---

† This is worse than co-operation, since $\frac{S+T}{2} < R$ by definition of the PD.

‡ Axelrod (1984) p. 38.

games like chess, where it is an advantage to conceal your intentions (see also "External information: reputations and labels" on page 30).

An ecological analysis

A repeated tournament was also implemented. Bad rules would be unlikely to be submitted again; successful rules, on the other hand, may be submitted several times (this was explicitly allowed for in the tournament rules). This leads to an ecological perspective, where the number of copies of a rule in a given generation is proportional to the product of the number in the previous generation and its score in the previous generation. It is an *ecological*, as opposed to an *evolutionary*, perspective, since no new rules (mutants) are introduced. The analogy with natural selection ("survival of the fittest") is clear. The first strategies to drop out will be the unsuccessful ones, but then those strategies which survived merely by exploiting those unsuccessful rules will also drop out; those which survive in the end will be both successful and non-exploitative. Equivalent forces act in the human realm—people might copy other people's strategies if they seem successful; people using unsuccessful strategies may get ousted from the game (e.g., go bankrupt or fail to get re-elected). **TFT** also did best in this ecological model.

It would be misleading to suggest that **TFT** is always the best strategy. Indeed, Proposition 1 tells us that there can be no best strategy. We have already seen how **TF2T** would have won the first round, had it been submitted. (It did less well in the second, because of the harsher environment.) Other strategies to beat **TFT** have been developed using techniques from Artificial Intelligence or by Genetic Algorithms, where strategies (considered as "chromosomes") are broken down into component parts ("genes"), and successful strategies allowed to "interbreed"; eventually good strategies arise.

An evolutionary analysis

The most general approach considers all possible strategies (as opposed to just those which were submitted) and asks: which ones are stable in the long run? The concept of an *evolutionary stable strategy* (ESS), which was introduced by Maynard Smith, is important here, though for the time being we shall use the weaker concept which Axelrod calls *collective stability* (CS). The differences are dealt with in "Evolutionary Stability versus Collective Stability" on p. 27.

> Imagine a population of individuals using a certain strategy **B**, and a single mutant using **A**. Strategy **A** is said to *invade* strategy **B** if $V(\mathbf{A} \mid \mathbf{B}) > V(\mathbf{B} \mid \mathbf{B})$, where $V(\mathbf{A} \mid \mathbf{B})$ is the expected payoff **A** gets when playing a **B**, and $V(\mathbf{B} \mid \mathbf{B})$ is the expected payoff a **B** gets when playing another **B**. Since the **B**'s are interacting virtually entirely with other **B**'s, the concept of invasion is equivalent to the single mutant individual being able to do better than the population average. This leads directly to the key concept of the evolutionary approach. A strategy is *collectively stable* (CS) if no strategy can invade it.†

A biological interpretation relates payoffs to Darwinian fitness (survival and fecundity); and hence only a CS strategy can be an equilibrium strategy in the long term. Up until recently, it had not proved possible to specify the necessary and sufficient conditions to be met by *any* CS strategy in the PD. But the insights Axelrod gained

---

† Axelrod (1981) p. 310. A strategy that is CS is in nash equilibrium with itself.

by studying the computer tournaments has now made it possible. The proofs in the following sections are based on Axelrod (1981).

Tit for Tat as a Collectively Stable Strategy

Clearly **TFT** can only be uninvadable if the "shadow of the future" is sufficiently long; that is, if the game is likely to last long enough for retaliation to counteract the temptation to defect.

LEMMA 3. **DC** *cannot invade* **TFT** *if* $w \geq \frac{T-R}{R-S}$.

PROOF. (Sketch)

$$V(\mathbf{DC} \mid \mathbf{TFT}) = \frac{T + wS}{1 - w^2} \tag{3}$$

so by (1), $V(\mathbf{DC} \mid \mathbf{TFT}) \leq V(\mathbf{TFT} \mid \mathbf{TFT}) \iff w \geq \frac{T-R}{R-S}$. □

LEMMA 4. **CD** *cannot invade* **TFT** *if* $w \geq \frac{T-R}{R-S}$.

PROOF. (Sketch)

$$\begin{aligned} V(\mathbf{CD} \mid \mathbf{TFT}) &= \frac{wT + S}{1 - w^2} + R - S \\ V(\mathbf{CD} \mid \mathbf{TFT}) &\leq V(\mathbf{TFT} \mid \mathbf{TFT}) \\ \iff w &\geq \frac{T - R}{R - S} \end{aligned}$$

□

Axelrod seems to think that Lemma 4 is a corollary of Lemmas 1 and 3. He states tersely, "**CD** cannot be better than **DD** and **DC**." However, it can be shown that

LEMMA 4A. $V(\mathbf{CD} \mid \mathbf{TFT}) \geq V(\mathbf{DC} \mid \mathbf{TFT})$ *if* $w \geq \frac{T-R}{R-S}$, and that

LEMMA 4B. $V(\mathbf{CD} \mid \mathbf{TFT}) \geq V(\mathbf{DD} \mid \mathbf{TFT})$ *if* w *is sufficiently large.*

Lemma 4B can be shown numerically. Now we can show

PROPOSITION 5. **TFT** *is CS iff* $w \geq \max\{\frac{T-R}{T-P}, \frac{T-R}{R-S}\}$.

PROOF. For the forward implication: if a strategy is CS, no strategy can invade it; in particular, **ALL D** and **DC** cannot, then by Lemmas 1 and 2, $w \geq \max\{\frac{T-R}{T-P}, \frac{T-R}{R-S}\}$.

For the reverse implication, **TFT** only has a memory of one move. Hence the only possible strategies are repeated sequences of **CC**, **CD**, **DC**, and **DD**. Now $V(\mathbf{CC} \mid \mathbf{TFT}) = V(\mathbf{TFT} \mid \mathbf{TFT})$. By Lemma 1, $V(\mathbf{DD} \mid \mathbf{TFT}) \leq V(\mathbf{TFT} \mid \mathbf{TFT})$ if w is in the specified range. Similarly by Lemma 2, $V(\mathbf{DC} \mid \mathbf{TFT}) \leq V(\mathbf{TFT} \mid \mathbf{TFT})$ and, by Lemma 3, $V(\mathbf{CD} \mid \mathbf{TFT}) \leq V(\mathbf{TFT} \mid \mathbf{TFT})$. Hence, $V(\mathbf{A} \mid \mathbf{TFT}) \leq V(\mathbf{TFT} \mid \mathbf{TFT})$ for all possible strategies **A**. □

The importance of the shadow of the future

Using the previous results, we see that it pays to co-operate with **TFT** if $w \geq \frac{2}{3}$. It pays to play **DC** with **TFT** if $w \geq \frac{1}{2}$. But if $w \leq \frac{1}{2}$, it pays only to play **DD**.

If one player thinks his/her partner will not be around much longer, the perceived value of w drops. One example is where a business is on the verge of bankruptcy:

> Once a manufacturer begins to go under, even his best customers begin refusing payments in merchandise, claiming defects in quality, failure to meet

> specifications, tardy delivery, or what have you. The great enforcer of morality in commerce is the continuing relationship, the belief that one will have to do business again with this customer, or this supplier, and when a failing company loses this automatic enforcer, not even a strong-arm factor is likely to find a substitute.†

Other evidence of this includes the rarity of court enforced settlements between regular trading partners.

The Characterisation of Collectively Stable Strategies

We have seen the requirements for **TFT** to be CS. What about the general case? Intuitively, a CS strategy should only co-operate if it can risk being exploited by the other side and still not be "overtaken". We can say that **B** has a *secure position* over **A** on move n if, no matter what **A** does from move n onwards, $V(\mathbf{A} \mid \mathbf{B}) \leq V(\mathbf{B} \mid \mathbf{B})$, assuming that **B** defects from move n onwards. That is, **B** has a secure position over **A** on move n iff

$$V_n(\mathbf{A} \mid \mathbf{B}) + \frac{w^{n-1}P}{1-w} \leq V(\mathbf{B} \mid \mathbf{B}) \tag{4}$$

where $V_n(\mathbf{A} \mid \mathbf{B})$ is **A**'s discounted cumulative score in the moves before n.

THEOREM *(The Characterisation Theorem).* **B** *is a CSS iff* **B** *defects on move n whenever the other player's score so far is too great, i.e. when $V_n(\mathbf{A} \mid \mathbf{B}) > V(\mathbf{B} \mid \mathbf{B}) - w^{n-1}(T + \frac{wP}{1-w})$.*

PROOF. For the reverse implication, we note that if **B** is CS, $V(\mathbf{A} \mid \mathbf{B}) \leq V(\mathbf{B} \mid \mathbf{B})\ \forall \mathbf{A}$. The proof proceeds by induction. To start, we observe that if $n = 1$, **B** is simply **ALL D**. But $V(\mathbf{A} \mid \mathbf{ALL\ D}) \leq V(\mathbf{ALL\ D} \mid \mathbf{ALL\ D})$ for all **A**. Hence **B** has a secure position over **A** on move 1. For the inductive step, we assume that **B** has a secure position over **A** on move n, and consider two cases:

i) If **B** defects on move n, **A** gets at most P, so

$$V_{n+1}(\mathbf{A} \mid \mathbf{B}) \leq V_n(\mathbf{A} \mid \mathbf{B}) + w^{n-1}P.$$

From (4) we have

$$V_{n+1}(\mathbf{A} \mid \mathbf{B}) \leq V(\mathbf{B} \mid \mathbf{B}) - \frac{w^{n-1}P}{1-w} + w^{n-1}P = V(\mathbf{B} \mid \mathbf{B}) - w^n\frac{P}{1-w}$$

ii) By assumption, **B** will only co-operate on move n when

$$V_n(\mathbf{A} \mid \mathbf{B}) \leq V(\mathbf{B} \mid \mathbf{B}) - w^{n-1}(T + \frac{wP}{1-w}).$$

Since **A** can get at most T on move n, we have

$$V_{n+1}(\mathbf{A} \mid \mathbf{B}) \leq V(\mathbf{B} \mid \mathbf{B}) - w^{n-1}(T + \frac{wP}{1-w}) + w^{n-1}T = V(\mathbf{B} \mid \mathbf{B}) - w^n\frac{P}{1-w}$$

For the forward implication, proof is by contradiction. Suppose that **B** is a CSS and there is an **A** and an n such that **B** does not defect when

$$V_n(\mathbf{A} \mid \mathbf{B}) + w^{n-1}(T + wP(1-w)) > V(\mathbf{B} \mid \mathbf{B}) \tag{5}$$

---

† Mayer, *The Bankers*, quoted in Axelrod (1984), p. 60.

Define $\mathbf{A}'$ as the same as $\mathbf{A}$ on the first $n-1$ moves, and as $\mathbf{D}$ thereafter. $\mathbf{A}'$ gets T on move n (since $\mathbf{B}$ co-operated then, by assumption), and at least P thereafter. So,

$$V_n(\mathbf{A}' \mid \mathbf{B}) \geq V_n(\mathbf{A} \mid \mathbf{B}) + w^{n-1}(T + wP(1-w))$$

and from (5), $V_n(\mathbf{A}' \mid \mathbf{B}) > V(\mathbf{B} \mid \mathbf{B})$. Hence $\mathbf{A}'$ invades $\mathbf{B}$, contrary to the assumption that $\mathbf{B}$ is CS. □

COROLLARY. **ALL D** *is always CS.*

PROOF. **ALL D** always defects, and hence always defects when required by the Characterisation Theorem. □

It also follows that a strategy has the flexibility to either defect or co-operate and still be CS, as long as its opponent has not scored too highly already. This explains why there are many CS strategies. Moreover, nice rules have the most flexibility, since they do so well when playing against each other; this means that they can afford to be generous towards invaders. They do not, however, have unlimited flexibility, as is shown by

PROPOSITION 6. *For a nice strategy to be CS, it must be **provoked** by the first defection of the other player, i.e. on some later move, the rule must have a non-zero chance of retaliating with a defection of its own.*

PROOF. If a nice strategy were not provoked by a defection on move n, then it could be invaded by a rule that only defected on move n. □

It was shown earlier that **TFT** is CS iff w is sufficiently large. This is, in fact, a general property.

PROPOSITION 7. *Any rule* **B** *which may be the first to co-operate is only CS when w is sufficiently large.*

PROOF. If **B** co-operates on the first move, $V(\mathbf{ALL\ D} \mid \mathbf{B}) \geq T + wP/(1-w)$. But for any **B**, $r/(1-w) \geq V(\mathbf{B} \mid \mathbf{B})$ since R is the best **B** can do with another **B** by the assumption that $R > P$ and $R > (S+T)/2$. Therefore $V(\mathbf{ALL\ D} \mid \mathbf{B}) > V(\mathbf{B} \mid \mathbf{B})$ is so whenever $T + wP/(1-w) > R/(1-w)$. This implies that **ALL D** invades a **B** which co-operates on the first move whenever $w < (T-R)/(T-P)$. If **B** has a positive chance of co-operating on the first move, then the gain of $V(\mathbf{ALL\ D} \mid \mathbf{B})$ over $V(\mathbf{B} \mid \mathbf{B})$ can only be nullified if w is sufficiently large. Likewise, if **B** will not be the first to co-operate until move n, $V_n(\mathbf{ALL\ D} \mid \mathbf{B}) = V_n(\mathbf{B} \mid \mathbf{B})$ and the gain of $V_{n+1}(\mathbf{ALL\ D} \mid \mathbf{B})$ over $V_{n+1}(\mathbf{B} \mid \mathbf{B})$ can only be nullified if w is sufficiently large. □

Invasion of a population of "meanies" by cooperative strategies

If a nice individual enters a world of "meanies" it will not survive, because it has no-one to reciprocate its co-operation with. However, suppose a cluster of individuals using strategy **A** arrives. Let the proportion of their interactions with each other be p; the proportion of their interactions with the natives using **B** is clearly $1-p$. **A**s are considered to contribute a negligible fraction of the environment of the **B**s, so any given **B** will interact almost entirely with other **B**s (assuming a non-random distribution of the **A**s). Then a *p-cluster of* **A** *invades* **B** if

$$pV(\mathbf{A} \mid \mathbf{A}) + (1-p)V(\mathbf{A} \mid \mathbf{B}) > V(\mathbf{B} \mid \mathbf{B})$$

i.e., if

$$p > \frac{V(\mathbf{B} \mid \mathbf{B}) - V(\mathbf{A} \mid \mathbf{B})}{V(\mathbf{A} \mid \mathbf{A}) - V(\mathbf{A} \mid \mathbf{B})}. \tag{6}$$

The striking thing is just how easy invasion of **ALL D** by clusters can be. Consider, for instance, the value of p needed for a **TFT** cluster to invade an **ALL D** world if $w = 0.9$. We have $\mathbf{A} = \mathbf{TFT}$, $\mathbf{B} = \mathbf{ALL\ D}$, $V(\mathbf{B} \mid \mathbf{B} = P(1-w) = 10$, $V(\mathbf{A} \mid \mathbf{B}) = S + wP/(1-w) = 9$ and $V(\mathbf{A} \mid \mathbf{A}) = R/(1-w) = 30$, so $p > 1/21$. If the median game length is 200, corresponding to $w = 0.99654$, then 1 interaction in 1000 with a like-minded **TFT**er is enough for the strategy to invade an **ALL D** world.

What kind of strategies can best invade an **ALL D** world? We can call a strategy *maximally discriminating* if it will eventually co-operate even if the other side has never co-operated yet, and once it co-operates it will never co-operate again with **ALL D**, but will always co-operate with another player using the same strategy. (For example, **TFT** is maximally discriminating.)

PROPOSITION 8. *The strategies which can invade* **ALL D** *in a cluster with the smallest value of p are those which are maximally discriminating.*

PROOF. To be able to invade **ALL D**, a rule must have a positive chance of co-operating first. Stochastic co-operation is not as good as deterministic co-operation with another player using the same rule since stochastic co-operation yields equal probability of S and T, and $(S+T)/2 < R$ in the PD. Therefore, a strategy which can invade **ALL D** with the smallest p must co-operate first on some move, n, even if the other player has never co-operated yet. (The value of n does not make any difference.) Equation (6) shows that the rules which invade $\mathbf{B} = \mathbf{ALL\ D}$ with the lowest values of p are those which have the lowest values of $\hat{p}$, where

$$\hat{p} = \frac{V(\mathbf{B} \mid \mathbf{B}) - V(\mathbf{A} \mid \mathbf{B})}{V(\mathbf{A} \mid \mathbf{A}) - V(\mathbf{A} \mid \mathbf{B})}.$$

The value of $\hat{p}$ is minimised when $V(\mathbf{A} \mid \mathbf{A})$ and $V(\mathbf{A} \mid \mathbf{B})$ are maximised (subject to the constraint that **A** co-operates for the first time on move n), since $V(\mathbf{A} \mid \mathbf{A}) > V(\mathbf{B} \mid \mathbf{B}) > V(\mathbf{A} \mid \mathbf{B})$. $V(\mathbf{A} \mid \mathbf{A})$ and $V(\mathbf{A} \mid \mathbf{B})$ are maximised subject to this constraint only if **A** is a maximally discriminating rule. □

How vulnerable are nice strategies to invasion? Not very.†

PROPOSITION 9. *If a nice strategy cannot be invaded by a single individual, it cannot be invaded by any cluster of individuals either.*

PROOF. For a cluster of rule **A** to invade a population of rule **B**, there must be a $p \leq 1$ such that $pV(\mathbf{A} \mid \mathbf{A}) + (1-p)V(\mathbf{A} \mid \mathbf{B}) > V(\mathbf{B} \mid \mathbf{B})$. But if **B** is nice, then $V(\mathbf{A} \mid \mathbf{A}) \leq V(\mathbf{B} \mid \mathbf{B})$. Let $V(\mathbf{A} \mid \mathbf{A}) = \alpha V(\mathbf{B} \mid \mathbf{B})$, $0 < \alpha \leq 1$. Then

$$\begin{aligned}
\alpha p V(\mathbf{B} \mid \mathbf{B}) + (1-p)V(\mathbf{A} \mid \mathbf{B}) &> V(\mathbf{B} \mid \mathbf{B}) \\
(1-p)V(\mathbf{A} \mid \mathbf{B}) &> V(\mathbf{B} \mid \mathbf{B})(1-\alpha p) \\
V(\mathbf{A} \mid \mathbf{B}) &> V(\mathbf{B} \mid \mathbf{B})\beta
\end{aligned}$$

where

$$\beta = \frac{1-\alpha p}{1-p}, \quad \beta \geq 1$$

† But see "Evolutionary stability versus Collective Stability".

i.e., $V(\mathbf{A} \mid \mathbf{B}) > V(\mathbf{B} \mid \mathbf{B})$. But that is equivalent to **A** invading as an individual. □

Thus nice strategies do not show the same vulnerability to invasion as **ALL D** does. Once co-operative strategies have invaded a world of unconditional defection, they will thrive (as we saw in the tournaments), and they can protect themselves from invasion by less co-operative strategies. "Thus the gear wheels of social evolution have a ratchet", as Axelrod puts it.†

Evolutionary Stability versus Collective Stability

Maynard Smith introduced the concept of evolutionary stability (ES), which he described like this:‡

> Suppose that a population contains a small fraction p of behavioural "mutants" adopting strategy **M**, and that the remainder of the population q $(= 1 - p)$ adopts the strategy **S**. If the total Darwinian fitness (expected number of surviving offspring) of the members of the population before a series of contests is C, then after the contest
>
> $$W(S) = C + q\,V(\mathbf{S} \mid \mathbf{S}) + p\,V(\mathbf{S} \mid \mathbf{M})$$
> $$W(M) = C + q\,V(\mathbf{M} \mid \mathbf{S}) + p\,V(\mathbf{M} \mid \mathbf{M})$$
>
> where $V(\mathbf{S} \mid \mathbf{M})$ is the expected payoff (change in fitness) to an individual employing strategy **S** in a contest with an individual employing strategy **M**, $W(S)$ is the total increased fitness acquired by employing strategy **S**, and so on.
>
> If **S** is an ESS, then $W(S) > W(M)$ for any mutant strategy **M**. Hence either $V(\mathbf{S} \mid \mathbf{S}) > V(\mathbf{M} \mid \mathbf{S})$ or $V(\mathbf{S} \mid \mathbf{S}) = V(\mathbf{M} \mid \mathbf{S})$ and $V(\mathbf{S} \mid \mathbf{M}) > V(\mathbf{M} \mid \mathbf{M})$.

It is clear that ES is a stronger concept than CS since ES⇒CS but not, in general, vice versa. Axelrod says, "I have used the new definitions to simplify the proofs and to highlight the difference between the effect of a single mutant and a small number of mutants". He claims that all his propositions still hold good if "ES" is substituted for "CS", except for the Characterisation Theorem, where the condition becomes necessary but not sufficient.

Clearly ES ⇔ CS if $V(\mathbf{S} \mid \mathbf{S}) > V(\mathbf{M} \mid \mathbf{S})$ for all **M**. Axelrod claims that, if **S** is nice, this is always true. But if **M** is nice, $V(\mathbf{S} \mid \mathbf{S}) = V(\mathbf{M} \mid \mathbf{S})$ and so it is not always true. This can be understood intuitively: if a nice mutant "wanders into" a nice population, it will be indistinguishable from the natives.

Whether **S** can be invaded by a nice strategy **M** depends on how it does relative to **S** with rare non-nice strategies.§

Consider, for instance, a population of **TFT** in which one-way mutation maintains *suspicious tit for tat* (**STFT**), which defects on the first move and then plays **TFT**.|| **TF2T** can invade **TFT** whenever $V(\mathbf{TF2T} \mid \mathbf{STFT}) > V(\mathbf{TFT} \mid \mathbf{STFT})$. "When w is large enough, this inequality is always satisfied because the more tolerant **TF2T**

---

† Axelrod (1990), p. 21.

‡ Maynard Smith (1978), p. 141.

§ If **M** is nice, invasion in the Axelrod sense is the same as invasion in the Maynard Smith sense, namely $V(\mathbf{M} \mid \mathbf{S}) > V(\mathbf{S} \mid \mathbf{S})$

|| The rest of this section is based on Boyd and Lorberbaum (1987).

induces **STFT** to cooperate, whereas **TFT** becomes involved in an endless series of reprisals. In fact, the only population configuration that is locally stable is a mixture of **TF2T** and **STFT**." Let p be the frequency of **TF2T** at this equilibrium. Then

$$\begin{aligned} &pV(\mathbf{TF2T} \mid \mathbf{TF2T}) + (1-p)V(\mathbf{TF2T} \mid \mathbf{STFT}) \\ &= pV(\mathbf{STFT} \mid \mathbf{TF2T}) + (1-p)V(\mathbf{STFT} \mid \mathbf{STFT}). \end{aligned}$$

Now,

$$\begin{aligned} V(\mathbf{TF2T} \mid \mathbf{TF2T}) &= R + wR + w^2R + \ldots = \frac{R}{1-w} \\ V(\mathbf{TF2T} \mid \mathbf{STFT}) &= S + wR + w^2R + \ldots = \frac{R}{1-w} - R + S \\ V(\mathbf{STFT} \mid \mathbf{STFT}) &= P + wP + w^2P + \ldots = \frac{P}{1-w} \\ V(\mathbf{STFT} \mid \mathbf{TF2T}) &= T + wR + w^2R + \ldots = \frac{R}{1-w} - R + T \end{aligned}$$

so

$$\begin{aligned} \frac{pR + (1-p)(R + (S-R)(1-w))}{1-w} &= \frac{p(R + (T-R)(1-w)) + (1-p)P}{1-w} \\ p &= \frac{(1-w)S + wR - P}{(1-w)(T+S-R) + wR - P} \qquad (7) \\ \lim_{w \to 1} p &= 1 \end{aligned}$$

For Axelrod's values, $p = 0.9863$. Hence, a population of **TFT**, although CS, can be invaded by a mixture of a nice, tolerant morph (**TF2T**) and a non-nice, somewhat exploitative morph (**STFT**). Worse, **TF2T** can itself be invaded by **DC**.† We must, however, remember that "invasion" is being used in a stronger sense than in the previous sections.

It was shown, in the Corollary to the Characterisation Theorem on page 24, that **ALL D** is CS. But it too can be invaded by a mixture of strategies. If w is large enough, **ALL D** can be invaded without population structure. For example, **STFT** can invade a population in which **ALL D** is common and in which **TF2T** is maintained by one-way mutation. Neither **ALL D** nor **STFT** are ever the first to co-operate; when paired, they both defect during every interaction. Thus, their relative fitness depends on how each fares against **TF2T**. For w large enough, **STFT** has a larger payoff than **ALL D**, since **STFT** can be induced to co-operate whereas **ALL D** cannot. Thus **STFT** can invade **ALL D**. The only stable equilibrium in this population is the mixture of **TF2T** and **STFT** given in equation (7). We have seen already that **TFT** and **ALL D** are not ES. In fact, no pure strategy can be ES if w is sufficiently large.

LEMMA 10. *Suppose a strategy* **M** *plays* **ALL D** *against* **B** *and* **ALL C** *against* **A**. *The minimum score difference between* **A** *and* **B** *is given by*

$$V(\mathbf{A} \mid \mathbf{M}) - V(\mathbf{B} \mid \mathbf{M}) = \frac{R-P}{1-w}.$$

† Axelrod and Dion (1988) p. 1389.

PROOF. Clearly **A** can earn either R or T every move. Similarly, **B** can earn P or S. But $\min\{R-P, R-S, T-P, T-S\} = R-P$. To see this, remember that $T > R > P > S$, by definition of the PD. Clearly R and P are "closest together". More formally, we can let $P = S + \alpha$, $R = P + \beta$, $T = R + \gamma$, $\alpha, \beta, \gamma > 0$. Hence, $\min\{R-P, R-S, T-P, T-S\} = \min\{\beta, \beta+\alpha, \beta+\gamma, \alpha+\beta+\gamma\} = \beta$. Hence the minimum score difference is $(R-P)(1+w+w^2+\ldots)$. □

It can also be shown that

PROPOSITION 11. *No non-random, pure strategy can be ES if*

$$w > \min\{\frac{T-R}{T-P}, \frac{P-S}{R-S}\}.$$

This result can be understood in terms of Axelrod's insight that if w is large enough, no strategy can be best against all opponents. When two strategies interact with each other in the same way that they do with themselves, their relative fitness depends on their interactions with other strategies. Because neither strategy can be best against every possible third strategy, no pure strategy can resist invasion by every combination of strategies. This means that the "ratchet" which Axelrod referred to can be undone.

Playing to win versus playing to do well

All of the previous conclusions have assumed that the aim is to get as many points as possible, rather than to win at all costs. However, sometimes you *do* want to win at any cost. Such games are called by Shubik "pure status" games. Many border disputes, where the land being squabbled over is of no real value, are of this form. (One thinks of the Falklands, although there establishing a reputation was probably at least as important a goal—see the next section.)

Behr (1981) has calculated which strategies would have won the first round of Axelrod's tournament had the criterion for winning been the greatest number of victories (or, alternatively, the greatest margin of victory). It turns out that those rules which previously did well, notably **TFT**, now did poorly, and those that previously did poorly now did well: **JOSS** and **FELD** were the winners. Why such a dramatic turnaround? "Because maximising scores requires maximized cooperation, and achieving victories requires at least some willingness to defect, [and hence] it seems impossible for any single decision rule to be eminently successful at both."† Clearly one must decide one's goals before choosing one's strategies.

Interestingly, when people play the IPD in practice, they usually tend to play to win. They cannot resist trying to do better than their opponents, even at their own expense. An interesting real-world example is pay-bargaining: why should what another occupation earns matter to you? Rationally, it should make no difference, but it is a fact that many of the recent pay disputes in the UK have been caused by "jealousy" of this kind. Another example comes from the US Congress; members co-operate with members from other states, to their mutual benefit, but "play to win" against legislators from their own State, who may challenge them at future elections.

† Behr (1981), p. 299.

External information: reputations and labels

Up until now, we have assumed that the only information available to each player is the past history of their interactions with the other. In the real world, however, one usually has access to other sources of information, for example, observing the behaviour of a player with someone else. This leads to the establishment of reputations.

How useful is it to know the other player's strategy? That depends on what it is. It would be very useful to know if it were **TF2T**, for example, since then you could easily do very well (and win) by alternating defection and co-operation. How useful is it for you to have your strategy known? Again, it depends. If you were playing **TFT**, it would be very useful, because then the other player would realise that the best solution is to co-operate right from the start (assuming, of course, that w is sufficiently large).

> Having a firm reputation for using **TFT** is advantageous to a player, but it is not actually the best reputation to have. The best reputation to have is the reputation for being a bully. The best kind of bully to be is the one who has a reputation for squeezing the most out of the other player while not tolerating any defections at all from the other. The way to squeeze the most out of the other is to defect so often that the other player just barely prefers cooperating all the time to defecting all the time. And the best way to encourage cooperation from the other is to be known as someone who will never cooperate again if the other defects even once.
>
> Fortunately, it is not easy to establish a reputation as a bully. To become known as a bully, you have to defect a lot, which means you are likely to provoke the other player into retaliation. Until your reputation is well established, you are likely to have to get into a lot of very unrewarding contests of will. For example, if the other player defects even once, you will be torn between acting as tough as the reputation you want to establish requires and attempting to restore amicable relations in the current interaction.
>
> What darkens the picture even more is that the other player may also be trying to establish a reputation, and for this reason may be unforgiving of the defections you use to try to establish your own reputation. When two players are each trying to establish their reputations for use against other players in future games, it is easy to see that their own interactions can spiral downward into a long series of mutual punishments.
>
> Each side has an incentive to pretend not to be noticing what the other is trying to do. Both sides want to appear to be untrainable so that the other will stop bullying them.
>
> The PD tournament suggests that a good way for a player to appear untrainable is for the player to use the strategy of **TFT**. The utter simplicity of the strategy makes it easy to assert as a fixed pattern of behaviour. And the ease of recognition makes it hard for the other player to maintain an ignorance of it. Using **TFT** is an effective way of holding still and letting the *other* player do the adapting. it refuses to be bullied, but does not do any bullying of its own. If the other player does adapt to it, the result is mutual cooperation. In fact, deterrence is achieved through the establishment of a reputation.†

Sometimes one makes *assumptions* about a person's behaviour on the basis of some fixed, observable characteristics known as a *label.*

> One of the most interesting but disturbing consequences of labels is that they can lead to self-confirming stereotypes. To see how this can happen,

† Axelrod (1984), p. 152.

suppose that everyone has either a Blue label or a Green label. Further, suppose that both groups are nice to members of their own group and mean to members of the other. For the sake of concreteness, suppose that members of both groups employ **TFT** with each other and **ALL D** with members of the other group. And suppose that the discount parameter w is high enough to make **TFT** a CSS. Then a single individual, whether Blue or Green, can do no better than to do what everyone else is doing and be nice to one's own type and mean to the other type.

This incentive means that stereotypes can be stable, even when they are not based on any objective differences. The Blues believe that the Greens are mean, and whenever they meet a Green, they have their beliefs confirmed. The Greens think that only other Greens will reciprocate cooperation, and they have their beliefs confirmed. If you try to break out of the system, you will find that your own payoff falls and your hopes will be dashed. So if you become a deviant, you are likely to return, sooner or later, to the role that is expected of you.†

It is clear that everyone is doing less well than they could. What is perhaps not so clear is that the minority do relatively worse.

To see why, suppose that there are 80 Greens and 20 Blues in a town, and everyone interacts with everyone else once a week. Then for the Greens, most of their interactions are within their own group and hence result in mutual cooperation. But for the Blues, most of their interactions are with the other group (the Greens), and hence result in punishing mutual defection. Thus the average score of the minority Blues is less than the average score of the majority Greens.

Labels can also serve to support status hierarchies. Suppose that everyone could be ranked on the basis of some characteristic such as height or skin tone. For simplicity we ignore the possibility of ties.

Suppose everyone uses the following strategy when meeting someone beneath them: alternate defection with cooperation unless the other player defects even once, in which case never cooperate again. This is being a bully in that you are often defecting, but never tolerating a defection from the other player. And suppose that everyone uses the following strategy when meeting someone above them: cooperate unless the other defects twice in a row, in which case never cooperate again. This is being meek in that you are tolerating being a sucker on alternating moves, but it is also being provokable in that you are not tolerating more than a certain amount of exploitation.

This pattern of behaviour sets up a status hierarchy based on the observable characteristic. The people near the top do well because they can lord it over nearly everybody. Conversely, the people near the bottom are doing poorly because they are being meek to almost everyone. It is easy to see why someone near the top is happy with the social structure, but is there anything someone near the bottom can do about it acting alone?

Actually there isn't. The reason is that when the discount parameter is high enough, it would be better to take one's medicine every other move from the bully than do defect and face unending punishment.

To see this, being meek gives you $S + wR + w^2R + \ldots = (S + wR)/(1 - w^2)$. If you revolt, you might as well defect all the time, which gives $P + wP + w^2P + \ldots = P/(1 - w)$.

† Axelrod (1984), p. 148.

So there is no incentive to revolt whenever $(S+wR)/(1-w^2) > (P+wP)/(1-w^2)$ or $w > (P-S)/(R-P)$. Using the usual values, it does not pay to revolt if $w > 1/2$.

> The futility of isolated revolt is a consequence of the immutability of the other player's strategies. A revolt by a low-status player would actually hurt *both* sides. If the higher-status players might alter their behaviour under duress, then this fact should be taken into account by a lower-status player contemplating revolt.

Other theories of co-operation: kin selection and reciprocal altruism

The image many people have of natural selection is summed up in Tennyson's wonderfully evocative phrase, "Nature red in tooth and claw". This image is further enhanced by such phrases as "survival of the fittest" and "the selfish gene". It is, however, an empirical fact that altruism exists, and not only among humans. One attempt to explain this is known as kin selection, and was invented by W. D. Hamilton, although it had precursors in the work of early geneticists, and has since been extended by many others. Basically it says: the reason a mother looks after her children is that she is programmed to do so by her genes;† and the reason they do that is to ensure their survival—each child gets half its genes from its mother, half from its father (we say the "coefficient of relatedness", r, is $1/2$), and so it is in the genes' interest to ensure that the child survives and reproduces in turn.‡ Similar reasoning applies to other kin. For example, $r = 1/8$ for first cousins, so a gene for suicidally saving eight first cousins would just pay. Perhaps the greatest success of this theory was in explaining why the (female) workers in the class of insects known as Hymenoptera—including ants, bees and wasps—should care for the offspring of the queen. The answer is that the workers are more closely related to their sisters ($r = 3/4$ for the queen's offspring) than to their own children ($r = 1/2$). This phenomenon is known as haplo-diploidy, and is caused by males only having one set of chromosomes (hence "haploid") and females having the usual double set (hence "diploid"). The details can be found in Dawkins (1989).

One problem with this theory is that it doesn't explain why non-related individuals co-operate, except by "accident". However, the preceding sections on the IPD have offered an alternative explanation, which is known as "reciprocal altruism", first developed by R. L. Trivers. As Morris points out, however, it strictly ought not to be called "altruism", since it is in our *own* interests. "True" altruism would be helping someone at a cost to oneself, with no expectation of future reward.§ Thus from the viewpoint of the mother, helping her children would be regarded as altruistic; but from the viewpoint of her genes it would be regarded as "selfish".

It must not be thought that reciprocal altruism can only evolve when the participants can calculate expected payoffs, etc. (A similar charge has been laid at the feet

---

† "We are survival machines—robot vehicles blindly programmed to preserve the selfish molecules known as genes," as Richard Dawkins says.

‡ Of course, genes cannot calculate. This is just a shorthand way of saying: those mothers who look after their children are likely to have more surviving offspring, and hence any genes which contribute to this behaviour will spread in the gene-pool. Actually, half of *all* the mother's genes will be passed on, not just those that contribute to increased fitness. This fact has been used to explain the presence of "junk" DNA—DNA fragments which do not code for proteins. This "selfish DNA" is regarded as a parasite, "hitching a ride" for free off the backs of the other genes.

§ If one regards the promise of Heaven as a reward, then even religiously motivated altruism becomes "selfish".

of kin selection, and is equally fallacious.) The example of the bacteria should gainsay any such impression. How do they do it? Well, their strategies can be "hardwired" by their genetic "program", and the need to recognise other players can be reduced by *site tenacity* (always "playing" in the same place, and hence (on the whole) with the same players), or completely eliminated by only "playing" with one individual, as in many mutualisms.

We have seen how co-operative, and indeed altruistic, behaviour can evolve in a world of selfish individuals. This is not to imply that humans *are* fundamentally selfish in nature—after all, software (memes) can always override the dictates of the hardware (genes)—but I for one find it comforting to know that we do not *need* to rely on the hypothesised "morality" of humans to achieve a humane society. As Morris puts it,

> In this manner it becomes possible to explain the biological bases for man's seemingly altruistic behaviour. This is in no way intended to belittle such activities, but merely to point out that the more usual, alternative explanations are not necessary. For example, it is often stated that man is fundamentally wicked and that his kind acts are largely the result of the teachings of moralists, philosophers and priests; that if he is left to his own devices he will become increasingly violent, savage and cruel. The confidence trick involved here is that if we accept this viewpoint we will attribute all society's good points to the brilliant work of these great teachers. The biological truth appears to be rather different.†

BIBLIOGRAPHY

Axelrod, R. (1981). "The Emergence of Co-operation among Egoists", *American Political Science Review* 75:306-18.

Axelrod, R. (1984). *The Evolution of Co-operation.* Basic Books, 1984. Penguin 1990.

Axelrod, R.; Dion, D. (1988). "The Further Evolution of Co-operation", *Science* 242:1385-90

Axelrod, R.; Hamilton, W.D. (1981). "The Evolution of Co-operation", *Science* 211:1390-96.

Behr, R.L. (1981). "Nice Guys Finish Last—Sometimes", *Journal of Conflict Resolution* 25.2:289-300

Boyd, R.; Lorberbaum, J.P. (1987). "No Pure Strategy is Evolutionarily Stable in the Repeated prisoner's Dilemma Game", *Nature* 327:58-59

Dawkins, R. (1989). *The Selfish Gene.* First edition 1976. This edition has two new chapters, and extensive commentary on the first edition. OUP.

Maynard Smith, J. (1978). "The Evolution of Behaviour", *Scientific American*, Sept. 1978.

Morris, D. (1978). *Manwatching. A field guide to human behviour.* Trid Panther.

Rapoport, A. (1967). "Escape from paradox." *Scientific American* June 1967

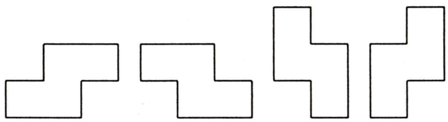
† Morris (1978), p. 154.

Tetris

Richard Tucker

In *Eureka* **50**, Adam Chalcraft considered the game of Tris, a simplified form of Tetris, and proved that it was a win for the player. (For the very few who haven't yet met Tetris, his article gives a concise summary of the game.) He conjectured that Tetris was a win for the daemon dropping the pieces, but that if the daemon had to announce in advance the sequence of pieces it was going to drop, the player could win. Here, we demonstrate that in fact the daemon can announce the sequence in advance, allow the player to choose the dimensions of the board, and still win.

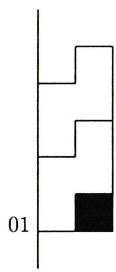
First, some terminology. The daemon's only going to use the two pieces called S and Z. When rotated, we refer to them as S' and Z':



S $\quad Z \quad S' \quad Z'$

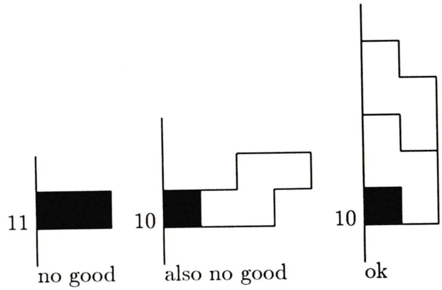
We also have some names for various sorts of row: a 11-row is a row that has a block in its leftmost square and another block in the square next to that, and anything you like in the other squares. Similarly, a 10-row begins with a block followed by a space, followed by anything at all, a 01-row goes space, block, anything, and a 00-row has nothing in the first two columns.

Suppose the player is in the middle of a game of Tetris involving only pieces S and Z. Then the number of 01-rows cannot ever decrease, since the only way of eliminating one is to drop a Z' into it, thereby creating another 01-row higher up. So if we further suppose that we have a strategy that will beat the daemon (i.e. we can play forever) there must be some point after which the number of 01-rows remains constant. Now consider the highest row which is not a 00-row. If it is a 01-row, then the only piece we can drop in either of the two leftmost columns without increasing the number of 01-rows is a Z', and so we must continue to fill up the left column with Z's forever:



If it is a 11-row, there is nothing at all we can drop in either of the first two columns without creating a 01-row, so we're stuck, and if it's a 10-row, we can either drop an S'

or an S in it, except if we drop an S we'll end up with a 11-row which we've just said is useless. So in fact we have to keep dropping S's into it:



So eventually, we must either fill the leftmost two columns with an S'-*tower* or a Z'-*tower*; therefore we can ignore these columns and look at the rest of the board. We label the rows 00, 01, 10 and 11 as before, but this time depending on what they have in their third and fourth squares from the left, and by the same argument we deduce that the next two columns must also be filled up by a tower. By induction therefore, the whole width of the board (which is of course finite) must eventually be taken up by such towers. From this we know that the board must be an even number of columns wide (otherwise no such strategy could exist) and that the pieces S and Z must be falling in an appropriate long-term ratio to enable all the towers to be built with equal speed.

Now I can reveal the daemon's winning sequence of moves: simply pick any irrational number x between 0 and 1 and drop pieces S and Z in that ratio. (If you want to be specific, first of all drop a Z and then drop whichever piece brings the ratio of the number of pieces S so far to the number of pieces Z so far closer to x). Whatever size board the player chooses, and however the columns are made into S'- and Z'-towers, the pieces will be needed in a rational ratio, but they are being dropped in an irrational ratio, so the player cannot possibly win. $\square$

Unanswered questions:

1. On a board of given size, for what n can the daemon stop the player from completing more than n lines?
2. For what subsets of the Tetris pieces can the daemon win a game involving only those pieces?

Miss Warren's Profession

Graham Nelson

This is a (hopefully constructive) protest article about the Mathematical Tripos, or rather the terminus it currently leads to. The reader who finds subjective polemic of rather limited interest is encouraged to stop reading now, with my apologies.

In classic soap-box style, one should begin by quoting a famous but long dead socialist. The title arises from the following passage from Act I of Bernard Shaw's play 'Mrs Warren's Profession', an Unpleasant play which was immediately censored when it was written in 1894: its message is that women become prostitutes because of hopeless poverty and not, as was conventionally considered, wanton immorality. Mrs Warren's 'respectably' brought up daughter Vivie, the other main character, has just made the acquaintance of Praed, a middle-aged architect:

> PRAED: ... It was perfectly splendid, your tieing with the third wrangler. Just the right place, you know. The first wrangler is always a dreamy, morbid fellow, in whom the thing is pushed to the length of a disease.

To Praed's shock, Vivie then explains that she only did so because her mother had bribed her to do so for £50, and this was certainly too little.

> PRAED: But surely it's practical to consider not only the work these honors cost, but also the culture they bring.
>
> VIVIE: Culture! My dear Mr Praed: do you know what the mathematical tripos means? It means grind, grind, grind for six to eight hours a day at mathematics, and nothing but mathematics. I'm supposed to know something about science; but I know nothing except the mathematics it involves.

People often look back on dreadful old Cambridge customs as quaintly romantic, like the annual reading of the results ceremony in the Senate House: great fun if you can close your ears to the sound of crying. (Although that's not nearly as bad as the gathering of would-be research students in DPMMS common room immediately after, where ashen-faced people walk one by one into an office and silently emerge, with or without a blue piece of paper.) The Tripos has been reformed a little in the intervening ninety-six years; Cambridge has finally conceded that women may be capable of thought, and nowadays the pretence is maintained that nobody is told exactly where they came in the exams. But, as I hope to demonstrate, the old profession of wrangling for places is very much alive.

1. How Part III works at the moment

At present, after the three-year Tripos is complete, there is a fourth year called Part III, which is exceptionally funded directly by the State so that quite a large number of people are able to read it. It is primarily intended as a preparation for research, although quite a number of people take it in order to spend another comfortable year at Cambridge (despite the miscellaneous selection of diplomas ideal for that purpose). Students register with either the Pure or Applied departments, which have very little to do with each other (so that pure and applied lectures on dynamical systems may clash, for example!). This distance is such that, having been Pure, I have hardly any idea of the situation in the Applied department apart from the general impression that it is more organised.

Teaching is done by offering a choice of a couple of dozen lecture courses on specialist subjects, usually so that members of the department give courses on the subjects they are experts in. Some courses have a few example classes as well, and these can be fairly helpful, but there aren't any supervisions and in effect lectures are the only source. There is little contact with the departments; Part III students don't often go to seminars, for example, as would be thought quite normal for first-year graduates at most universities. In effect, for someone going on to a doctorate, this year is the most isolated and cut off of all seven or so; which is unfortunate given how easy it is to be at a loose end during it. Having said this, students are provided with careers advice and are usually made welcome when they go and talk to people in the departments.

At the end of the time, candidates choose six of these courses to take examinations in, although they have the option (which in practice almost everyone exercises) of writing an extended essay in place of one of them. This can be quite rewarding and at least feels like real mathematics.

Unfortunately, many of the courses are very specialized, and the degree to which you have to focus on them is such that you never get time to look at anything else. The effect is like looking through a telescope in five or six different directions and pretending to have seen a landscape. What is perhaps worse is that after this process, you may have the mistaken impression that you are now too specialized to be able to look at any of the things missed out: in Dr Körner's memorable observation, in a few years you find yourself at the top of an ever narrowing gum tree.

But the really dreadful thing about Part III is the examination itself. The courses are (rightly) about difficult things and most of them are lectured at great speed so as to pack a large amount of material in. The result is that most reasonable questions that could be posed about them might take a research student a week to work out, so that it's quite hard to set questions which test understanding. In consequence, most of the examinations test rote learning on a grand scale; of the papers I took, only small parts of two questions on one of the papers had answers which did not consist of recital of the appropriate passage from the lecture notes. The hardest part of programming a computer to pass the course would be teaching it to read the rubric and work out what combinations of questions are legal. There are some papers which call for actual thought, but they are a somewhat dangerous proposition since you are bound to get less marks on them than others will get on the memory tests: which matters a great deal if you want to stay in Cambridge, since research students are selected solely on the basis of the mark ordering. (Other universities have been known to treat the results as meaningful, as well.) In general the papers vary greatly in difficulty which must make it difficult for the examiners to normalize.

In any event, the amount of theory required is such that there is no way to stop and try to think how to derive things, any more than school Latin and a rough knowledge of the plot is enough to be able to translate the Aeneid at writing speed. So you need to memorize about 120 lectures of writing, basically symbol by symbol. It might be thought that this is so large a task that you would be forced to understand it pretty thoroughly. There is some truth in this; then again, you also end up learning exactly where every lemma lies on the page in your notes. A year later, I could still leaf through these pages in the mind's eye. Also, you understand what is in the notes, but little else; it is mostly theory you have no motivation for. To give an extreme example, immediately before the Commutative Algebra paper, several of us who could all have recited all the main theorems and proofs of dimension theory of rings had an argument about whether there were or were not any rings whose dimension was not 1.

Miss Warren's estimate of the amount of time this takes is definitely on the conservative side. Her modern counterparts resort to devices such as tieing themselves to their chairs with string, wearing earplugs to prevent them from hearing any conversation they would be tempted to join, keeping appalling graphs of their daily twelve-hour workloads, and so on. (In the end people really do become dreamy and morbid; I can distinctly remember holding a toothbrush on the night before one of the papers and not being entirely sure what to do with it.) When it is over, you tend to recoil so much that it may be months before you are prepared to look a maths book in the contents page again.

2. Reform

The above notwithstanding, Part III is a way of learning some useful maths and effectively getting an extra year's funding before beginning research, and it has some good features, such as the essay. Also, it must not be thought that the Cambridge departments are indifferent to the lot of their students. Anyone who has ever attended the (variably effectual) meetings of the Faculty Board will realise that many lecturers are genuinely concerned about reform. At present, a large-scale reform of the undergraduate Tripos is under way, which aims to remove some of the arbitrary division of teaching between Pure and Applied mathematics (vector spaces and differentiation are Pure, but distributions and vector fields are Applied, and so forth) and to make it accessible to more A-level students. Much time and effort has gone into this thoroughly worthwhile aim. It was interesting to note that in the early stages, many people were very concerned that at almost any cost, Part III had to be maintained, that being the jewel in the crown of the Cambridge Tripos.

Also, two years ago an attempt was made by the Pure department to reform the problem of overspecialized courses. It was decided that the old system of having many courses which only one or two people would study, would have to go; a broad range of introductory courses would be added. (This did not stop them from quietly allowing a few people to take specially set second papers on combinatorics and category theory, however; although as the existence of these papers was not directly communicated even to those who had attended the first courses, this number might have been artificially few.) This was partially successful, although many of the lecture courses seemed to be still fundamentally about specific theories in great detail without much reference to their relevance to each other. The problem of overspecialization in effect remains, except that now there are fewer courses available, it is no longer always possible for this specialization to be in subjects people intend to work in.

Having said this, complaining to the powers that be about Part III is like the guests at Fawlty Towers complaining about Manuel to Basil Fawlty, who replies "I know, he's awful, isn't he?" Owing to the rather odd election rules for student representatives on the Faculty Board, it is essentially impossible for a Part III student to be one (if you are an undergraduate but graduate part-way through the time, you must step down; if you are a Part III student, you have no certainty of still being in Cambridge for any more than one term from when the elections take place); so it is a cause with no obvious spokesman. But at least there will be an excellent opportunity to do something about it when the first intake to the reformed Tripos reach Part III.

As for how to reform it, at this point I slide yet further into subjectivity. In many ways a useful way to spend the year would be to take another selection of Part II courses passed over in the year before (for example, General Relativity and Linear Analysis is not that unlikely a combination of courses, but sheer pressure on timetables usually clashes them), but this is unrealistic. At the other extreme, moving over to a full MPhil course would result in the exceptional funding status of Part III being lost, which would mean that very few people could take it and it would not be feasible to offer such a range of courses. I propose that the balance between course-work (i.e. on dissertations of some kind) and examinations be drastically shifted, possibly so that the examination part could be just on a few of the 'core' courses; other lecture courses would be to provide background, understanding of which would be necessary to work on the dissertations part of it, as well as being useful to research students. Ideally, students would work on several different essays, but on different subjects.

Shaw would probably have considered these proposals dismally unradical, but the Faculty Board (or, more importantly, the staff meetings) may well disagree. Most of my argument, which has perhaps itself been pushed to the point of a disease, has been to propose reform just to make the course more humane for students, but the departments have just as much to gain: a more sensible system for selecting research students and (more optimistically) some integration of all these potential future mathematicians into the institutions they nominally belong to.

Free Market Maths

Paul Norridge

In recent years the problem of funding in academic circles has become one of considerable size. Obviously, in maths, partnerships with industry are not always feasible. Many of the other conventional solutions are also difficult to implement. So what other possibilities are there in this area? Where can we turn to combat this result of current economic policy?

One answer which I think has been seriously overlooked is that of theorem sponsorship. For instance, think of the money that could be generated by the Coca-Cola Theory of Relativity, or the Pepsi–Bolzano–Weierstrass theorem. Organisations could be charged a small amount for each lecture or paper involving their chosen subject. This would have the advantage of generating capital for the study of those areas which are most often used. This fits very nicely with the government's current view on the free market and would no doubt be widely welcomed.

In fact, the idea has more possibilities than may be obvious at first glance. For example, companies could sponsor those subjects which are particularly relevant to their trade—"Kelvin's Economy 7 Circulation Theorem", the "Dulux White-with-a-hint-of-Green's function", or the irresistable "Remington's Fuzzaway Fuzzy Subsets".

Perhaps the most profitable area would be politics itself, surely Chaos is ideal for this. Or maybe we could have the Labour Left-half plane and the Conservative Right-half plane (though no doubt the latter would soon be privatised.)

Only one problem really springs to mind—Could Sainsbury's be prosecuted by the Advertising Standards Authority for sponsoring the infinite server queue ...?

Problems Drive 1991

Michael Earnshaw and Tim Wilkins

1. Consider the following game ...

You have to find a mystery number between 00 and 99. A turn consists of picking a number, which is given a score as follows:

1 mark for a correct digit in the wrong place,

3 marks for a correct digit in the correct place.

After typing in the correct number, you are told how many attempts you required to discover it. Let x_s be the number of attempts required to find x using strategy s. Find:

$$\text{a)}\quad \min_s\left(\max_x x_s\right)$$

$$\text{b)}\quad \max_x\left(\min_s x_s\right)$$

2. Define $f(\theta) = \sum_{n=1}^{\infty} \frac{1}{n} \cos n\theta \cos^n \theta$

i) Find some $\theta \in (0, \frac{1}{2}\pi)$ such that $f(\pi - \theta) = f(\theta)$ and $f(2\theta) = f(\theta)$.

ii) Find some θ such that $f(\theta) = 1$.

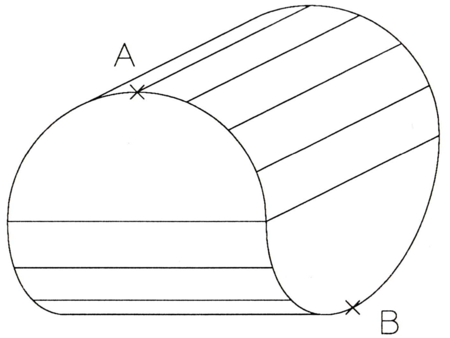
3. You are sent an encoded message: b_j: $0, 4, 30, 32, 26, 34, 12, 18$

$$b_j = a_j + \alpha + \beta j^3 \pmod{36} \qquad \alpha, \beta \text{ are integers; } j = 1, 2, \ldots, 8$$

You know that the original message (a_j) consists of the numbers $7, 8, 11, 17, 22, 24, 29, 30$ in some order. Find the original message.

Hints: $6^3 = 36 \times 6$, $\frac{1}{4} \times 8^2 \times (8+1)^2 = 36 \times 36$

4. The surface shown is formed by dividing a cylinder of radius 1, height 2, into two equal pieces, rotating one half by 90° and reattaching the halves. Find the shortest distance from A to B along a path on the surface.



5. Fill in the following number cube. The answers are in base 5, the centre is 0, and no answer begins with a zero.

g	h	i
d	e	f
a	b	c

o	p	q
m		n
j	k	l

x	y	z
u	v	w
r	s	t

m	to	n	
n	to	p	a pythagorean triple including 29
n	to	k	
g	to	c	sum to 51
i	to	a	
d	to	h	perfect numbers
o	to	q	
x	to	z	(m to n)
h	to	y	α^2
s	to	y	(m to n) + (k to n)
r	to	z	β^3
j	to	l	$\alpha\beta^2$
x	to	a	a prime
b	to	f	$\beta!$
x	to	r	a fifth power
t	to	z	sum of all the digits in the cube

6. A farmer has a triangular field with equal area and perimeter in suitable units. He tethers a goat so that it is free to graze within a circle in the field, whilst maximising its grazing area. Find this area in these units.

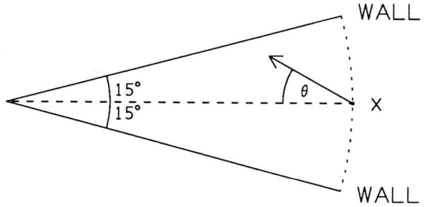
$$\text{Hint: Consider } \frac{\text{area of triangle}}{\text{radius of circle}}$$

7. There are seven people A,B,C,D,E,F and yourself at an Archimedeans' talk. Of the other six, two are clinically insane, two are committee members, and two are from Trinity College (these pairs are not necessarily disjoint). Sane people are always truthful, whereas mad ones are unreliable. Identify the committee members and the mad pair from the following statements:

i) C: "Exactly one of the Trinity people is mad."
ii) D: "Statement (i) is true."
iii) F: "C is not mad."
iv) E: "If I am on the committee then F is not."
v) B: "E and myself are from Trinity."
vi) E: "B and myself are from Trinity."
vii) C: "A and E are sane."
viii) A: "C is not on the committee, but exactly one mad person is."
ix) F: "D is sane. B and F are the committee."
x) E: "The committee members are next to each other alphabetically."

8. Place a small puck at X. Fire at angle θ as shown. θ is uniformly distributed on $(0°, 360°)$. The bounces off the two walls shown are elastic. Find:

i) $\mathbb{P}$(precisely 3 bounces occur)
ii) $\mathbb{P}$(precisely 6 bounces occur)
iii) $\mathbb{P}$(precisely 8 bounces occur)



9. Find the upper and lower limits on positive x such that $x^{x^{x^{\cdot^{\cdot^{\cdot}}}}}$ converges.

10. Let $S = \mathbb{Z}^3 \cup \{(x, y, z) \colon (x - \frac{1}{2}, y - \frac{1}{2}, z - \frac{1}{2}) \in \mathbb{Z}^3\}$, and $\mathbf{u} \in S$.
Define $T(\mathbf{u}) = \{\mathbf{v} \in \mathbb{R}^3 \colon |\mathbf{u} - \mathbf{v}| \leq |\mathbf{v} - \mathbf{w}| \ \forall \mathbf{w} \in S\}$.
Consider $T(\mathbf{0})$ as a solid body, with volume V.

a of its faces are regular polygons with b sides,
c of its faces are regular polygons with a sides,
Find a, b, c and V.

11. Find the next number in each of the following sequences:

i) $3, 5, 11, 17, 31, \ldots$
ii) $1, \frac{\pi}{4}, \frac{2}{3}, \frac{3\pi}{16}, \frac{8}{15}, \ldots$
iii) $1, 6, 15, 28, 45, \ldots$
iv) $11, 12, 13, 21, 21, 23, \ldots$

12. The following diagrams represent positions on solitaire boards. In each case decide whether it is possible to get from position A to position B using only legal solitaire moves (i.e. one piece jumps over an adjacent piece into an empty space, taking the piece it jumps over; no diagonal jumps).

N.B. There are no marks for partially correct answers.

```
              A                          B
i)          × × ×                      × × ×
            o × o                      o × o
        × × × o × × o              o × × o o o o
        × × o × × × o              o × o × × × o
        o × o o × × o              o o o o o × o
            o × o                      × × o
            × × ×                      × × ×

ii)         o × o                      o o o
            × o ×                      o o o
        o × o × o × o              o o o o o o o
        × o × o × o ×              o o o o o o o
        o × o × o × ×              o o o o o o ×
            × o ×                      o o o
            o × o                      o o o
```

iii)

			A								B			
		o	×	o						o	o	o		
		×	o	×						o	o	o		
o	×	o	×	o	×	o		o	o	o	o	o	o	×
×	o	×	o	×	o	×		o	o	o	o	o	o	o
o	×	o	×	o	×	×		o	o	o	o	o	o	o
		×	o	×						o	o	o		
		o	×	o						o	o	o		

iv)

		o	×	o						o	o	o		
		o	×	×						o	o	o		
o	×	o	×	o	o	o		o	o	×	o	o	o	o
×	×	×	o	×	o	×		o	o	o	o	o	o	×
o	o	×	×	×	×	o		o	o	×	o	o	o	o
		×	×	o						o	o	o		
		o	×	o						o	×	×		

THE RULES

Candidates compete in teams of two. Each team is given a new question every five minutes, and each question is removed after ten minutes; an extra ten minutes is allowed at the end for scrabbling answers together, inspired guesswork and so forth. Marks are awarded according to a bizarre algorithm, by tradition the total for each question being one mark, subdivided as appropriate. The winning team this year consisted of Vin de Silva and Oliver Riordan of Trinity College; they set the questions for the 1992 Problems Drive. The wooden spoon was awarded to Timothy Luffingham, who unwisely entered by himself.

The solutions are on page 55.

Clerihews

A. Nomet

Alexander Kolmogorov
Had a headache, but it wore off.
He didn't discover the test that statis-
ticians learn of
Until he discovered Smirnoff.

I find Maclaurin
Very borin'
And Taylor's series
Also wearies.

G.H. Hardy
Was rather tardy
At patching up his quarrel
With Stan Laurel.

Bertrand Russell
Once choked on a mussel
Which is why he says nothing exciting
about shellfish in his writing.

Some Unsolved Problems

In Geometry, Number Theory and Combinatorics

Paul Erdős

GEOMETRY

1. Let there be given n points in the plane, no five on a line. An old conjecture of mine states that if $h(n)$ denotes the maximum number of lines which contain four points,
£75. then $h(n)/n^2 \to 0$. I offer £75 for a proof or disproof of this conjecture. Grűnbaum constructed a set of points showing that $h(n) > cn^{3/2}$; it is not impossible that in fact $h(n) < c'n^{3/2}$.

Let us now drop the condition that no five of our points are on a line. George Purdy and I considered the following problem. Denote by $g(n)$ the maximum number of distinct lines which contain four or more of our points. The lattice points show that $g(n) > cn^2$. Perhaps the exact determination of $g(n)$ is hopeless, but we could not even determine $\lim_{n\to\infty} g(n)/n^2$. Perhaps if there are at least cn^2 distinct lines with at least four points, then there is a line which contains "many" of our points. We could not even prove that there is a line which contains five points, but perhaps there is a line which must contain n^ε or even more points.

2. Let $x_1, \ldots, x_n$ be n points in the plane. Denote by $f(n)$ the largest integer so that the points determine at least $f(n)$ distinct distances. An old and no doubt very difficult conjecture of mine states that

$$f(n) > n^{c/\log\log n}\,. \tag{1}$$

£300. I offer £300 for a proof or disproof of (1).

Let now $d_1, \ldots, d_k$ be the distinct distances determined by $x_1, \ldots, x_n$. Assume that the distance d_i occurs s_i times. Clearly

$$\sum_i s_i = \binom{n}{2}\,.$$

I conjecture that there is a c_1 for which

$$\sum_i s_i^2 < n^3(\log n)^{c_1}\,. \tag{2}$$

£300. The lattice points show that both (1) and (2), if true, are best possible. I offer £300 for a proof or disproof of (2).

Assume now that the points $x_i, 1 \le i \le n$, form a convex polygon. I conjecture that

$$\sum_i s_i^2 < cn^3\,. \tag{3}$$

Perhaps inequality (3) is easy, so I offer £25 for a proof or disproof of it. £25.

Fishburn and I conjecture that for $n > 8$, $\sum s_i^2$ is maximal for the regular polygon. I further conjectured that

$$\max s_i < cn\,, \tag{4}$$

and Fishburn and I conjectured that $\max s_i < 2n$ which, if true, is nearly best possible since Edelsbrunner and Peter Hajnal showed that $\max s_i = 2n - 7$ is possible. Füredi proved

$$\max s_i < cn \log n$$

which is the current record. I conjectured that in every convex n-gon there is a vertex which does not have four vertices equidistant from it. If true this is very much stronger than (4). I offer £100 for a proof or disproof of this assertion. Originally I conjectured this with three instead of four, but this was disproved by Danzer. £100.

I conjectured and Altman proved that if $x_1, \dots, x_n$ is convex then the points determine at least $\lfloor n/2 \rfloor$ distinct distances. The regular polygon shows that $\lfloor n/2 \rfloor$ is best possible. Denote by $f(x_i)$ the number of distinct distances from x_i. I conjectured that

$$\max_{1 \le i \le n} f(x_i) \ge \frac{n}{2}\,, \tag{5}$$

and Fishburn conjectured that

$$\sum_{i=1}^{n} f(x_i) \ge \binom{n}{2}\,. \tag{6}$$

Inequality (6) is, of course, much stronger than (5), and the regular polygon shows that it is best possible, if true.

Let $g(x_i)$ denote the largest number of points equidistant from x_i. Following Fishburn, I conjectured that

$$\sum_{i=1}^{n} (g(x_i)) < 4n \tag{7}$$

which is, of course, much stronger than (4), and even stronger than $\min_{1 \le i \le n} g(x_i) < 4$. Inequality (7), if true, is best possible, since 4 can not be replaced by $4 - \varepsilon$. In fact the example of Edelsbrunner and Hajnal shows that

$$\sum_{i=1}^{n} g(x_i) > 4n - c$$

is possible.

3. In a note entitled "Some Problems on Elementary Geometry" published in Australian Math. Gazette 2 (1975), 2–3, I published a few problems which seemed interesting to me; since the Gazette is almost unknown outside Australia it might be worthwhile to repeat some of them here.

Let $x_1, \dots, x_n$ be n distinct points in the plane (they do not have to be in general position). Consider all the circles which go through at least three of our points. Denote by $f(n)$ the largest integer for which there are $f(n)$ distinct unit circles which go through at least three of our points. I conjectured that

$$\frac{f(n)}{n^2} \to 0 \text{ and } \frac{f(n)}{n} \to \infty \,.$$

Elekes found a very simple and ingenious proof of the inequality

$$f(n) > cn^{3/2}\,.$$

£100. Perhaps $f(n) < c'n^{3/2}$ also holds. I offer £100 for a proof or disproof of this. Suppose now that our points are in general position (no four on a circle and no three on a line); to what extent does this change the value of $f(n)$?

Let $h(n)$ be the largest integer so that among any n points in general position there are at least $h(n)$ circles of different radii passing through three of our points. Determine or estimate $h(n)$ as well as you can.

Is it true that to every k there is an n_k so that among any n_k points in general position there are always k so that all the $\binom{k}{3}$ circles determined by our k points have different radii? Prove that n_k exists and determine it as well as you can.

4. Let $x_1, \ldots, x_{2n}$ be $2n$ points in $\mathbb{R}^n$. Denote by $f_r(n)$ the largest integer for which these points determine at least $f_r(n)$ distinct distances. It is a simple exercise to prove that

$$f_r(n) > n^{\varepsilon_r}$$

for some $\varepsilon_r > 0$. The exact determination of ε_r does not concern us; we need only that $\varepsilon_r > 0$. Now let $g_r(n)$ be the largest integer so that if $y_1, \ldots, y_n$ and $z_1, \ldots, z_n$ are two sets of n points then there are at least $g_r(n)$ distinct distances $d(y_i, z_j)$, i.e. there are at least $g_r(n)$ distinct distances between the y's and the z's. An example of Lenz shows that $g_r(n) = 1$ for $r \geq 4$. It suffices to take the y's on a circle and the z's on a circle orthogonal to it.

Is it now true that for $r = 2$ and $r = 3$ we have $g_r(n)/f_r(n) \to 0$? Perhaps the
£50. answer will be different for $r = 3$ and $r = 2$. I offer £50 for a solution of this question.

5. Let $x_1, \ldots, x_n$ be n points in the plane. Denote by $d(x_i, x_j)$ the distance between x_i and x_j. Assume that $d(x_i, x_j) \geq 1$, $1 \leq i \leq j \leq n$ and further assume that if two distances differ, then they differ by at least 1. Is it then true that the diameter $D(x_1, \ldots, x_n)$ is greater than cn? In fact perhaps for $n > n_0$

$$D(x_1, \ldots, x_n) \geq n - 1\,;$$

clearly, equality holds if the points are $0, 1, \ldots, n-1$. In any case it would be of some interest to determine $\min D(x_1, \ldots, x_n)$ as well as possible for small values of n.

NUMBER THEORY

1. Here is a completely forgotten old problem of Surányi and myself, published in a Hungarian paper†. Let $f(n)$ be the smallest integer for which if $a_1 < a_2 < \cdots < a_n$ are any n integers then among any a_n consecutive integers one can always select $f(n)$ of them whose product is a multiple of $\prod_{i=1}^{n} a_i$. In our paper we show that $f(n) > (2-\varepsilon)n$ for every $\varepsilon > 0$ if $n > n_0(\varepsilon)$. Is it true that $f(n) < (2+\varepsilon)n$, or perhaps $f(n) < 2n$?

It is easily seen that $f(2) = 2$; in our paper we show that $f(3) = 4$.

† Mat. Lapok 10 (1959), 39–48

2. Denote by $p(n)$ the least prime factor of n. Is it true that for every $n > n_0$ there is a composite m for which $m > n$ but

$$m - p(m) < n? \tag{8}$$

I am afraid that (8) is hopeless. It would be of some interest to try to determine the largest n for which (8) does not hold. Also, put

$$f(n) = \min_{m>n}(m - p(m)).$$

Perhaps a table of $n - f(n)$ would be of some interest. The order of magnitude of $n - f(n)$ seems interesting. Is it true that $(n - f(n))/\sqrt{n} \to c$, $c > 0$?

3. Let $f(n)$ be a number theoretic function. Call m a barrier of $f(n)$ if $m + f(m) \le n$ whenever $m < n$.

Denote by $\omega(n)$ the number of distinct prime factors of n and by $\Omega(n)$ the number of prime factors of n, multiple factors counted multiply. It seems certain that both $\omega(n)$ and $\Omega(n)$ have infinitely many barriers, but I am afraid that it will not be easy to decide these questions. The only way I can think of attacking these problems are sieve methods and these are probably not yet sufficiently powerful. I could not even prove that there is an ε for which there are infinitely many integers n for which for every $m < n$, $m + \varepsilon\omega(m) \le n$.

Denote by $q(n)$ the number of squares dividing n. It is easy to prove that $q(n)$ has infinitely many barriers. Selfridge and I tried to find integers n for which for every $m < n$

$$m + d(m) \le n + 2, \tag{9}$$

where $d(m)$ is the number of divisors of m. $n = 24$ is perhaps the largest integer satisfying (9). I offer £25 for another such integer. I am being rather stingy but we old people are stingy. I expect that £25.

$$\lim_{n\to\infty} \max_{m<n}(m + d(m) - n) = \infty.$$

4. Denote by $M(n;k)$ the least common multiple of $n+1, n+2, \ldots, n+k$. The equation $M(n;k) = M(m;l)$ seems to have very few solutions for $m \ge n+k$, $l > 1$. The only solutions I know are $M(4;3) = M(13;2) = 210$, and $M(3;4) = M(19;2) = 420$. Put $A(m;k) = \prod_{i=1}^{k}(n+i)$. There are probably only finitely many solutions of $A(n;k) = A(n;l)$, $m \ge n+k$, $l > 1$.

Let now $k \ge 3$. It is easy to see that for every k and $m \ge n+k$

$$M(n;k) > M(m;k) \tag{10}$$

has infinitely many solutions. Let n_k be the smallest solution of (10). It is not hard to prove that $n_k/k \to \infty$ but I have no good upper or lower bounds for n_k. I wondered if

$$M(n;k) > M(m;k+1)$$

has infinitely many solutions. I could not find any, but Selfridge found two, namely $M(96;7) > M(102;8)$ and $M(132;7) > M(139;8)$.

COMBINATORICS

1. In our paper with Ko and Rado, Quarterly J. Math. Oxford (2) **12** (1961, 313–320), many problems were posed, and all but one of which has been solved. Let me state the problem which is still open. Let S be a set with $4n$ elements and let $A_1, A_2, \ldots, A_{f(n)}$ be subsets of S with $2n$ elements each, such that $|A_i \cap A_j| \geq 2$ for all $i \neq j$. Is it then true that

$$f(n) \leq \frac{1}{2}\left(\binom{4n}{2n} - \binom{2n}{n}^2\right)? \tag{11}$$

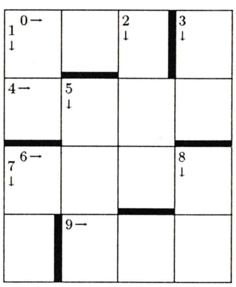
It is easy to see that (11), if true, is best possible. To see this let S be the set of integers $1, 2, \ldots, 4n$ and let the A's be the sets which contain at least $n+1$ integers at most $2n$. It is immediate that these sets satisfy $|A_i \cap A_j| \geq 2$ and their number is $\frac{1}{2}\left(\binom{4n}{2n} - \binom{2n}{n}^2\right)$.

£250. I offer £250 for a proof or disproof that (11) is best possible.

Cross-Number

Ampenditus

No answer begins with a zero, and all are written in base ten. The lights are clued in terms of one another, with one clue each, but unfortunately the clues have been mixed up:



$5+8+1$
7×7
3
1×3
$8+8-1-1-3$
1
$5+8$
$2+2+2$
$2-5$
$(7-3) \times 9 - 3$

The answer to the first clue is not three digits long, and the answer to the last clue is not two digits long.

What is Mac Lane missing?

A. R. D. Mathias

A sociologist observing the 1989/90 Logic Year at the Berkeley Mathematical Sciences Research Institute would have judged it to be a typical gathering of mathematicians, exchanging ideas, running seminars to chip away at current problems, and writing papers and books. But there was one speaker who from time to time would tell the others that they were working on the wrong problems in the wrong subject. This was not the result of a momentary aversion: Professor Mac Lane has for at least twenty years been saying that "set theory is obsolete", that "measurable cardinals are bizarre", and so on, and he has written one large book and many articles in order to present his view of mathematics.

It is the purpose of this essay to examine his stance, and to suggest that in as far as Mac Lane urges the unity of mathematics, he is to be supported, but in as far as he secretly desires the uniformity of mathematics, he is to be opposed.

Perhaps I should begin with a few reflections on the psychology of mathematics. One of the remarkable things about mathematics is that I can formulate a problem, be unable to solve it, pass it to you; you solve it; and then I can make use of your solution. There is a unity here: we benefit from each other's efforts. In this regard mathematicians interact much more than do (say) historians or composers.

But if I pause to ask *why* you have succeeded where I have failed to solve a problem, I find myself faced with the baffling fact that you have thought of the problem in a very different way from me: and if I look around the whole spectrum of mathematical activity the huge variety of styles of thought becomes even more evident.

Is it desirable to press mathematicians all to think in the same way? I say not: if you take someone who wishes to become a set theorist and force him to do (say) algebraic topology, what you get is not a topologist but a neurotic. Uniformity is not desirable, and an attempt to attain it, by (say) manipulating the funding agencies, will have unhealthy consequences.

The purpose of foundational work in mathematics is to promote the unity of mathematics: the larger hope is to establish an ontology within which all can work in their different ways.

What, then, is Mac Lane's ontology? This seems to admit a clear answer. In his book *Mathematics: Form and Function*, he urges the claims of a system he calls ZBQC, which initials stand for Zermelo with Bounded Quantification and Choice, to supply all that he needs to do the mathematics he wants to do.

The axioms of this system are those of the more familiar ZFC, weakened by dropping Replacement and only allowing Comprehension (Separation) for Δ_0 formulae; that is, ones in which quantifiers appear only in formulae of the form, $\forall x \in A \ldots$ or $\exists x \in A \ldots$ with A fixed—whence "Bounded".

This system provides for the existence of the real numbers, and for ω-types over them, thus yielding the complex numbers, functions from reals to reals, functionals and so on.

That this system represents a natural portion of mathematics may be seen from the way in which it keeps reappearing, first as the simple theory of types, and more recently as topos theory, with each of which it is equiconsistent. (That is, it has a model—or "universe" of sets which it describes—if and only if either of the latter has a model.) A natural model for it is $V_{\omega+\omega}$ (a set described by the usual axioms of ZF).

It is plain from Mac Lane's book that this system indeed supports a large amount of mathematics, more than I shall ever learn. Why then need we go outside it?

I suggest that an area ill supported by Mac Lane's system ZBQC is that of iterative constructions. We know from the work of Cantor onwards that there are processes which need more than ω steps to terminate; of which examples may be found even within traditional areas of mathematics. For example, within the space of continuous functions on [0,1], the class of differentiable functions forms a set which is not a Borel set but is naturally expressible as the union of $\aleph_1$ Borel sets; and this has implications for the problem of building the anti-derivative of a given function.

So therefore let us look for a moment at abstract recursion theory and ask how easily it sits within Mac Lane's system.

A well-established axiomatic framework for abstract recursion theory is the system KP of Kripke and Platek. It turns out that if ZBQC is consistent then so is ZBQC+KP: the idea behind the proof is that one can "code" a model for the latter system inside a model for the former. With slightly more trouble one can indeed establish that the consistency of ZBQ (that is, ZBQC without the axiom of choice) implies that of ZBQC+KP.

Thus ZBQC has via suitable coding a reasonable capacity for recursive constructions; and this would support Mac Lane's thesis that it is a reasonable basis for much of mathematics. However it will, as is clear from the work of Harvey Friedman, fail to support many constructions: it will not be able to prove Borel determinacy, a famous theorem about recursion which requires the iteration of the power set operation through all countable ordinals; similarly it will not be able to prove Borel diagonalisation.

Set theory is so rich a theory that it has been claimed for much of this century to be the foundation of mathematics. In ontological terms this claim is not unreasonable; but Mac Lane resists. I would guess that his reason is not so much that he objects to the ontology of set theory but that he finds the set-theoretic cast of mind oppressive and feels that other modes of thought are more appropriate to the mathematics he wishes to do.

One must acknowledge that ideas from category theory provide a smooth way to handle a large amount of material. However, to reject a claim that set theory supplies a universal mode of mathematical thought and of mathematical existence need not compel one to declare set theory entirely valueless.

Let us therefore set aside set theory's claim to be a foundation of the whole of mathematics, it being misguided to define the worth of a subject solely in terms of its serviceability to other areas of mathematics. Instead let us define set theory to be the study of well-foundedness. As such, it is a worthy object of study; and it can scarcely be said that this is a subject of little content!

From this point of view, Mac Lane's view that "measurable cardinals are bizarre" becomes hard to defend. May we suppose him to mean that he sees no need to think about them and therefore resents a suggestion that he should think about them?

In terms of the study of well-foundedness, measurable cardinals are natural objects: just as ZBQC has resurfaced in many forms, so do measurable cardinals keep bobbing up in unexpected contexts. The hypothesis that they exist, or the hypothesis that in some inner model there are measurable cardinals, may be construed as saying that in certain circumstances the direct limit of well-founded structures is well founded. Other large cardinal axioms may also be interpreted as assertions of this general kind. These hypotheses seem worthy of study: well-foundedness is important, being central to the general enterprise of constructing objects by recursion, and it is natural to ask when well-foundedness is preserved under direct limits. These questions are interesting in their own right.

This might be a good moment to challenge one of Mac Lane's opinions, which I believe to rest on a misconception. On page 359 of his book he writes, after reflecting on the plethora of independence results, that "For these reasons 'set' turns out to have many meanings, so that the purported foundation of all of Mathematics upon set theory totters." Elsewhere, on page 385, he remarks that "the Platonic notion that there is somewhere *the* ideal realm of sets, not yet fully described, is a glorious illusion."

I would suggest a contrary view: independence results within set theory are generally achieved either by examining an inner model of the universe (an inner model being a transitive class containing all ordinals) or by utilising forcing to build a larger universe of which the original one is an inner model. The conception that begins to seem more and more reasonable with the advance of the inner model program on the one hand and a deeper understanding of iterated forcing on the other is that within one enormous universe there are many inner models, and the various "independence arguments" may be reworked to give positive information about the way the various inner models relate to each other. Far from undermining the unity of the set-theoretic view, the various techniques available for building models actually promote that unity.

In a more diplomatic mood, Mac Lane writes on page 407:

> Neither organization is wholly successful. Categories and functors are everywhere in topology and in parts of algebra, but they do not as yet relate very well to most of analysis. Set theory is a handy vehicle, but its constructions are artificial. ... We conclude that there is as yet no simple and adequate way of conceptually organizing all of Mathematics.

Let me now consider briefly whether there can be a unified foundation for Mathematics. In probing this question I have found myself coming to a view that can be traced back certainly to Plato, namely that that there are *two* primitive mathematical intuitions; which might be called the geometrical and the arithmetical; or, alternatively, the spatial and the temporal.

Plato did not have the advantage of modern research into the functions of the left and right half of the brain; this work suggests that the temporal mode (which would include recursive constructions) is handled in the left brain, whereas the spatial mode is handled in the right.

What can each mode of thought contribute to the understanding of the other? I believe, a lot.

Can either be reduced to the other? I should say not; certain formal translations exist, but the underlying intuitions do not translate; and these obstructions show themselves as paradoxes such as the Banach–Tarski paradox. This notorious result states that a solid ball in $\mathbb{R}^3$ may be decomposed as the union of finitely many sets (in fact rather a small number of them), which may then be 'reassembled'—after translations and rotations in space—to form two solid balls the same size as the original.

Let me refer to my contention that there are these two modes, neither reducible to the other, as positing an essential *bimodality* of mathematical thought.

In earlier pieces I have remarked how Mac Lane's choice of axioms agrees with that made by Bourbaki, at least initially; it has recently been suggested that Bourbaki's choice of topics was made by consideration of the needs of physicists.

This suggests a speculative question: what need is there for a theory of recursion in physics? There is certainly a need for a theory of recursion in mathematics. The recursion theorem itself is the heart of logic; it is the watershed where processes become objects. In descriptive set theory it takes the shape of the Coding Theorem of Moschovakis, and is thus the source of the strength of the axiom of determinacy.

My sense of the bimodality of mathematics is such that to suppress the ordinals or other frameworks on which to carry out recursions is to suppress half one's mathematical consciousness. I wonder therefore what physicists might be missing by using only the Bourbaki–Mac Lane portion of mathematics in their modelling. Might it be that time might fruitfully be modelled with ordinals rather than reals, so that a leap to the next limit ordinal corresponds to some discontinuous event?

Such speculation prompts a further question: is it necessary for all the mathematical concepts invoked in physical explanation to have a direct physical meaning? Or might it be desirable to have abstract concepts which have the merit of making the physics easier to understand without having a perceptible physical interpretation?

But physics aside, the unity of mathematics is a desirable aim; and I would suggest as a modest first step that working in ZBQC + KP rather than ZBQC would encourage awareness of the temporal side of mathematics as well as the spatial side.

Mac Lane's set theory is weak in constructive power, but strong in manipulating the objects naturally arising in geometry. The reverse, as I expect Mac Lane would agree, is true of set theory. I suggest that category theory is as natural a framework for spatial mathematics as set theory is for temporal. I suggest therefore that we should seek an organisation of mathematics that will allow both the fundamental intuitions room to develop and to interact; in doing so, we should move away from the regrettable situation so pithily described by Augustus de Morgan over a century ago and still, sadly, to be found today:

> We know that mathematicians care no more for logic than logicians for mathematics. The two eyes of exact science are mathematics and logic: the mathematical sect puts out the logical eye, the logical sect puts out the mathematical eye; each believing that it sees better with one eye than with two.

Book Reviews

More Mathematical Byways

by **Hugh ApSimon** Reviewed by **Stephen Turner**

This book, the seventh in the OUP's *Recreations in Mathematics* series and the sequel to *Mathematical Byways in Ayling, Beeling and Ceiling*, is recommended to anyone who enjoys solving those Braintwisters, Brainteasers and Enigmas that appear in certain newspapers and magazines. For those not acquainted with such things I should say that they tend to be short arithmetical or geometrical puzzles ending with some such question as "How many sheep were there?" or "How far is the church from the tree?" and being best approached without using (messy) calculus or trigonometry. In other words they are accessible to the numerate layman, which will be a disappointment to a few Cambridge mathematicians and a relief to most. Mr ApSimon presents fourteen such problems of his own, some previously published and some new, and examines them. This is where the book becomes worthwhile. Rather than just giving the solution to the particular problem in each case, ApSimon also looks at the general problem of which the particular problem is an example and at extension questions, some of which he must leave unanswered, as well as telling how he invented and developed the puzzle. These for me were the most interesting parts of the book. I found some of the problems frustrating as they often rely on finding, for example, some neat geometrical construction which simplifies the problem enormously and which is very beautiful once you have found it. Or got fed up and turned to the solution. To learn a little of how a professional puzzle setter works is in my opinion far more satisfying, and the extension questions often looked more interesting than the particular problem.

As for the price, £15 for fourteen problems does seem a little dear. By all means borrow this book from the library, but if you want to buy it (for yourself or for your numerate layman friends) wait until it comes out in paperback.

More Mathematical Byways, by Hugh ApSimon, OUP hardback £14.95.

The Penguin Dictionary of Curious and Interesting Geometry

by **David Wells** Reviewed by **Vin de Silva**

Euclidean geometry is one of those subjects which, despite having held an apparently impregnable position in mathematics syllabuses for centuries, is now considered by mathematicians to be of merely circumstantial interest, and of little relevance to any worthwhile branch of mathematics. The common perception of the subject is even less generous, and is easily explained by a glance at any one of the many dreary, now defunct textbooks that purported to teach geometry to school-children: that it is rather dull. While David Wells' new book, *The Penguin Dictionary of Curious and Interesting Geometry*, a sequel to his similarly-titled volume on Numbers, may do little to alter the professional view on the importance of Eucliden geometry, it should certainly play its part in correcting the preconceptions held by the vast majority of people, that it is intrinsically uninteresting.

Like its predecessor, the book is set out in the form of an illustrated dictionary, with alphabetically-ordered paragraphs on diverse topics that come under the umbrella of "geometry". Although there is no restriction specifically to the Euclidean domain—numerous examples are taken from the study of Fractals (not to mention Chaos), Topology, Projective Geometry, Optimisation,Tesselations (including hinged and hyperbolic varieties), Knot Theory, and so on—it is strong on material of the kind the Greeks would have recognised and been proud of: theorems about circles and triangles. It is a feast of wonderful facts and historical anecdotes, and its layout encourages the dilettante to dip in and browse. We find explicit constructions for the regular 17-gon (using compass and straightedge) and the regular 7-gon (using toothpicks); we learn of the mathematician Steiner who discovered countless theorems by using the trick of circular inversion, but did not reveal his secret, so as to continue impressing his colleagues; we are told that the Emperor Napoleon was no mean mathematician, and may well have discovered the theorem traditionally known as Napoleon's Theorem.

The depth and thoroughness of the research makes the book invaluable as a reference work, and it is very well-written. A special mention must go to the beautiful diagrams and illustrations by John Sharp which complement the text. Geometry thrives on pretty pictures, and here we are lavishly furnished with them. These alone make a strong recommendation, but the book as a whole is indispensable. I welcome its appearance wholeheartedly.

The Penguin Dictionary of Curious and Interesting Geometry, by David Wells,
Penguin paperback £10.99.

More Clerihews

A. Nomet

Stephen Hawking
Is always talking
Of that stitch in Time
Which saves nine.

David Hilbert
Was often mistaken for Gilbert.
"I did NOT write Trial by Jury!"
He would say, in fury.

Alan Turing
Needed reassuring
That a Turing machine made of
papyrus
Was immune to almost every virus.

Faraday†
Is not easy to parody.
He would take legal action
If you said "Proof by induction".

Fleischmann and Pons
Confounded their fellow dons
By demonstrating fusion in a football
stadium
And later at the London Palladium.

Professor John Coates
Mistyped his lecture notes;
But "Useful for Bankers" had more
takers
Than the intended: "Useful for
Bakers".

† This one requires a Scottish accent.

Solutions to Problems Drive

1. Need to try e.g. 98, 76, 54, 32, 10 (i.e. distinct pairs), and then consider cases. In the worst case seven tries might be needed, so (a) is 7, and (b) is 1 trivially.

2. (i)$f(\pi-\theta)=f(\theta)$ since (-1)'s square up, so try $\pi-\theta=2\theta$. Then $\theta=\pi/3$.
(ii) Let $z=\cos\theta\, e^{i\theta}$, so Re $z^n=\cos^n\theta\cos n\theta$. Hence $-$Re $\ln(1-z)=f(\theta)=$ Re $(z+\frac{1}{2}z^2+\frac{1}{3}z^3+\ldots)$, so $f(\theta)=-\ln|1-z|$. But $1-z=\sin\theta(\sin\theta-i\cos\theta)$, so $f(\theta)=-\ln|\sin\theta|$. $f(\theta)=1\Rightarrow|\sin\theta|=1/e$, so $\theta=\pm\sin^{-1}(1/e)+n\pi$.

3. 17, 30, 29, 22, 7, 24, 11, 8. Sum the numbers to get $8=\sum b_j-\sum a_j=8\alpha+36n$ ($n\in\mathbb{Z}$). No β dependence, since $\sum_1^8 j^3=\frac{1}{4}.8^2.(8+1)^2=0 \pmod{36}$, by the hint. Get four possible values of α. Now consider b_6, since $6^3=0 \pmod{36}$. This gives α. Note β is odd, so we know which a_i are even; add them to deduce β, and the rest follows.

4. Cut open the solid to get the figure shown (over). The path must be straight, and the two dashed lines give $l^2=\pi^2/4+\pi+2$ or 8, so $l=(\pi^2/4+\pi+2)^{1/2}$, by comparison.

5.

1	1	1
1	0	4
1	4	0

1	0	3
1	0	4
3	1	0

1	0	4
1	2	3
2	4	1

6. Grazing region is the inscribed circle (see over). Let B and P be the area and perimeter of the triangle respectively. Then we get $B=\frac{1}{2}bc\sin\alpha$ and $r=x\tan\frac{1}{2}\alpha=xt$, so $r^2=x^2t^2$. By the cosine rule on α and some algebra, we get that $\frac{B}{r}=\frac{P}{2}$, so $r=2\left(\frac{B}{P}\right)$. Therefore grazing area $=\pi r^2=4\pi$ here, since $B/P=1$.

7. If C is mad, so is F by (iii), so everyone else is sane. Specifically (ii) and (v) are true, so one of B and E is also mad, a contradiction. So C is sane, whence so are A and E by (vii). (x) contradicts (ix) so F is mad. And (vi) is true, so by (i), B is mad. Now (viii), (x) and (iv) give the committee members as A and B.

8. By reflective symmetry on bouncing (see over), the path is a straight line. (Note $\alpha=30°$.) Then by simple geometry (i) $\mathbb{P}(3)=\frac{1}{12}$, (ii) $\mathbb{P}(6)=\frac{1}{24}$ and (iii) $\mathbb{P}(8)=0$. (On the diagram, ϕ shows the ranges of angle giving 3 bounces.)

9. Let $y=x^{x^{x^{\cdot^{\cdot^{\cdot}}}}}$, so we require $y=x^y=f(y)$. Iterate: $y_{n+1}=x^{y_n}=f(y_n)$ (x fixed). We get convergence if $\left|\frac{\mathrm{d}f}{\mathrm{d}y}\right|<1$ at y. I.e. require the derivative to have modulus 1 at the limits of the range.

$x^y\ln x=\pm1$ so $y\ln x=\pm1$, but $\ln x=(\ln y)/y$ so $\ln y=\pm1$ and therefore $y=e$ or $1/e$. So we get a lower limit of e^{-e} and an upper limit of $e^{1/e}$.

10. $a=4$, $b=4$, $c=8$ as in the diagram (over). (This is the Wigner–Seitz unit cell for a body-centred cubic lattice, otherwise known as a splatt.)

We expect 8 of something from chopping the corners of the unit cube, and 6 of something from the 6 faces. So we get either 6 octagons and 8 triangles, or 6 squares and 8 hexagons. It must be the latter since no edges of the unit cube are in $T(0)$. Consider small spheres at each point in S. The unit cube contains 2 spheres (1 from

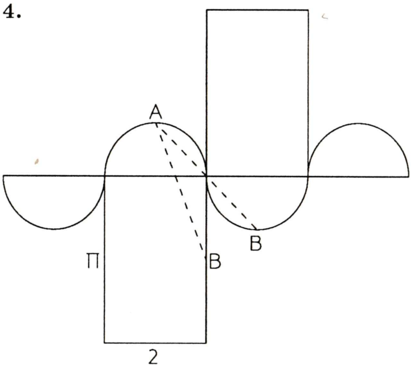
(0,0,0) and $\frac{1}{8}$ from each vertex). But $T(\mathbf{u})$ tesselate, and each contains just one sphere at the centre. Therefore $V = \frac{1}{2}$.

11.

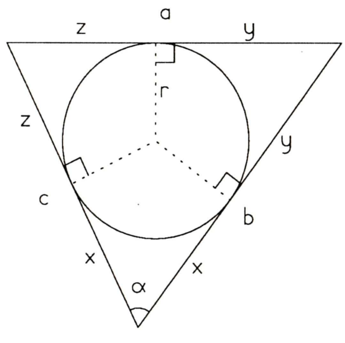
i)	41	$a_n = p_{p_n}$ (2nd, 3rd, 5th, ... prime)
ii)	$5\pi/32$	$a_n = \int_0^{\pi/2} \sin^n x \, dx$
iii)	66	alternate triangle numbers
iv)	23	$a_n =$ the n^{th} prime in base n, for n at least 2.

12. (i) No; (ii) Yes; (iii) No; (iv) No. (ii) and (iii) are obvious. For (i) and (iv) consider the four invariant sets, i.e. sets of points such that marbles cannot move between the sets.

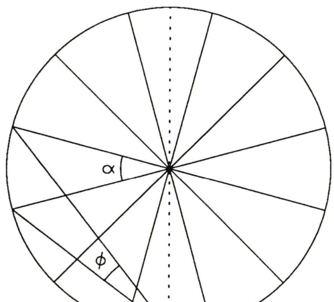
4.



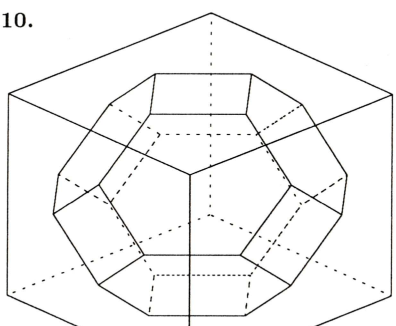
6.



8.



10.



Limerick

Eureka, you'll note, is delayed;
And perhaps you'll be rather dismayed.
It's a bit of a shame,
But I know who's to blame:
It's the Editor's fault, I'm afraid.

ERRATUM SLIP

The art of printing has advanced considerably this century; nevertheless the gremlins have adapted themselves remarkably well to the new technology. The following are the material mistakes that have come to light since printing.

p. 3, para. 4, line 5: *for* Tashkent, *read* Peking
. line 6: *for* Salt Lake City, *read* Poheng
p. 4, para. 1, line 5: *for* nocturnal, *read* but satisfying
p. 15, para. 3, line 3: *for* rational, *read* irrational
p. 46, para. 3, line 2: *for* ${}^{(k}_{3)}$, *read* $\binom{k}{3}$.