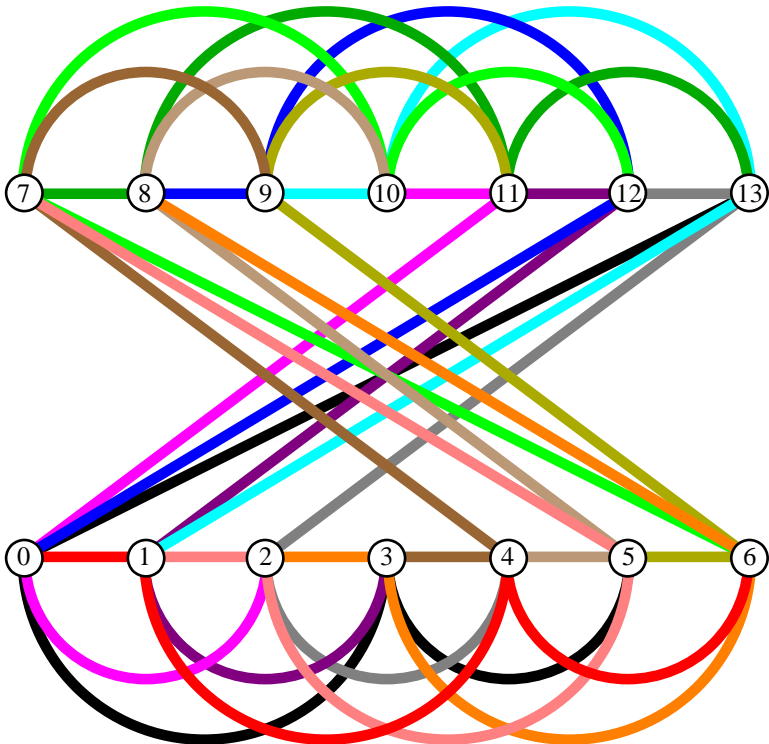


*

ISSN 1350-8539



M500 281



The M500 Society and Officers

The M500 Society is a mathematical society for students, staff and friends of the Open University. By publishing M500 and by organizing residential weekends, the Society aims to promote a better understanding of mathematics, its applications and its teaching. Web address: m500.org.uk.

The magazine M500 is published by the M500 Society six times a year. It provides a forum for its readers' mathematical interests. Neither the editors nor the Open University necessarily agree with the contents.

The Revision Weekend is a residential Friday to Sunday event providing revision and examination preparation for both undergraduate and postgraduate students. For details, please go to the Society's web site.

The Winter Weekend is a residential Friday to Sunday event held each January for mathematical recreation. For details, please go to the Society's web site.

Editor – *Tony Forbes*

Editorial Board – *Eddie Kent*

Editorial Board – *Jeremy Humphries*

Advice to authors We welcome contributions to M500 on virtually anything related to mathematics and at any level from trivia to serious research. Please send material for publication to the Editor, above. We prefer an informal style and we usually edit articles for clarity and mathematical presentation. For more information, go to m500.org.uk/magazine/ from where a LaTeX template may be downloaded.

Mathematics of heredity

Sebastian Hayes

The science of genetics is essentially the work of a single man, Gregor Mendel, a Moravian monk, who received precious little recognition for his painstaking and epoch-making work.

Mendel spent eight years experimenting with peas (some 30,000 specimens). He started off with a parent generation P_1 of homozygous (true-breeding) pairs from two markedly different strains. Peas, like most flowering plants, are hermaphroditic, i.e. bear a set of two reproductive organs on the same plant. The ovaries of such plants are usually hidden at the base of the flower while the sperm cells, contained in the pollen sacks, are held by so-called anthers at the tip of the stamens—they are particularly evident in lilies. Not all hermaphroditic plants are capable of self-pollination but peas are. The characteristics of the offspring of such plants depend on the genes of the parent stock and, in the case of peas, a single pair of alleles (genes that ‘code’ for the same characteristic but which are not necessarily identical) dictates the subsequent colour of the flowers. Other features of the pea such as height, whether the pods are smooth, are wrinkled, and so forth are governed by other pairs of alleles and, if these alleles are situated on different pairs of chromosomes (or on sufficiently distant locations on the same pair of chromosomes), these pairs do not interfere with each other—they obey Mendel’s *Principle of Independent Assortment*. Partly through luck and partly through good judgment, Mendel picked exactly the right sort of plant for his experiments.

Nearly all cells of plants and mammals are diploid, i.e. have pairs of chromosomes that are identical in length and function but which allow some variation at specific loci (human beings have 23 pairs while turkeys, surprisingly, beat us with 41 pairs). During the reproductive process certain cells divide in an unusual way (by meiosis) to produce so-called gametes (sex cells) which are haploid, i.e. have only a single set of chromosomes. Fertilization occurs when two opposed gametes such as spermatozoids and eggs fuse and thus restore the chromosomal number. The important point is that half of the offspring’s chromosomes, and hence its genes, come from the ‘male’ parent and half from the ‘female’—Mendel’s so-called *Principle of Segregation*.

Mendel prevented self-fertilization in P_1 , the parent generation, by removing the pollen bearing stamens of the test plants and himself transferred the pollen (using a small brush) from similar plants. He then enclosed the flower head in a transparent bag to prevent further pollination. Since the

alleles for colour in Mendel's pea cells 'coded' either for purple or white (with nothing in between) and the fertilization process was entirely random, one might expect that some of the first generation, F_1 , would produce purple flowers, some white, while perhaps other flowers again would be an intermediate colour. In fact all specimens (or very nearly all) were purple. Mendel correctly deduced that alleles come in pairs and the allele for purple was 'dominant' over the allele for white in peas—this was nearly a hundred years before the discovery of DNA, and Mendel actually spoke of 'elements of heredity', not of genes and alleles.

Mendel now prevented any cross-fertilization of F_1 and allowed the peas to self-pollinate. This time the result, F_2 , was a collection of plants with purple and white flowers in a 3 : 1 ratio. He concluded that his 'elements' came in pairs with one dominant, one recessive type—his *Principle of Dominance*—and that the recessive element reasserted itself in the offspring of a heterozygous plant (one of mixed ancestry) when paired with a heterozygous plant (either itself or another heterozygous plant of similar ancestry). He continued in this way, allowing F_2 to self-fertilize and produce F_3 and so on, and noted the results.

Mendel now started studying how different traits combined, choosing traits where there were only two contrasting forms, or phenotypes, such as small/tall, smooth/wrinkled and so on. For example, we suppose the cross of two heterozygous plants with one pair of alleles for colour, P and p , and one pair of alleles for a smooth or wrinkled exterior of the grain or fruit, G and g . He found that there were (approximately) four different types of offspring (purple & smooth, purple & wrinkled, white & smooth, white & wrinkled) in (very nearly) a ratio of 9 : 3 : 3 : 1.

The principle can be extended, in theory if not in fact, to any N different loci coding for different traits on pairs of alleles on different chromosomes, so it is time to try to throw this into some mathematical shape. (Incidentally, this problem came up in the course I was doing, *S204*, which is the first biology course I have ever taken at the OU, or elsewhere for that matter.) The whole issue is basically a fairly elementary question in combinatorics but the dominant/recessive dichotomy makes it relatively intractable (to me anyway). The resulting ratios can be obtained from any biology book or the Internet but I wanted to try to solve the problem of the possible genotypes (genes coding for a certain characteristic) and phenotypes (what you actually see) from first principles much in the way Mendel himself must have done.

Mathematicians are supposed to work by deduction but in reality I

suspect that most of them do things the other way round, i.e. examine a mass of numerical data and try to spot a pattern. Glancing at the data for various monohybrid, dihybrid crosses and trihybrid crosses (see the tables at the end of this article), one is struck by the appearance of the binomial coefficients, also the appearance of ascending powers of 3. I remembered, or rather rediscovered, that the two are related since the sum of the n th line of Pascal's Triangle is 2^n while $2^{2^n} = (2^2)^n = (1 + 3)^n$.

This was enough to get started. Although there may be more than one pair of alleles that 'codes' for a particular characteristic such as petal or eye colour (the number varies from species to species), only two choices of alleles can be expressed at equivalent loci on a single pair of chromosomes at the same time, and in any case Mendel was only concerned with such straightforward situations. Supposing alleles R and r , with the capitalized R implying a dominant form and the lower case a recessive one, and an equal chance of getting either, we have two possibilities, R and r , at each locus on a gamete of either type. And for N loci situated on N different pairs of chromosomes—this stipulation is necessary since if the loci of alleles are sufficiently near each other on a single pair of chromosomes they may become linked—the total number of possibilities is 2^N .

Since exactly the same situation applies to the other gamete of the pair and the alleles get thoroughly mixed up during meiosis and the eventual fusion of the successful gametes, each possible combination of alleles is (in theory) equally probable, and so we multiply the possibilities. There are two possibilities at the first locus on the 'male' gamete, two at the equivalent locus of the 'female' gamete, two at the second locus of each gamete, and so on. For N loci there are $(2^N)^2 = 2^{2N}$ possibilities for the genotype of the resulting zygote (diploid offspring).

Examining the tables, one sees the binomial coefficients ${}^{2^n}C_r$ emerging. But should each of these possibilities be considered distinct? That depends what you are looking for. Conventionally, one writes the symbol for the allele from the 'male' first and that from the 'female' second, so that Rr in the offspring means that the dominant allele has come from the male parent and the recessive from the female. The difference of origin may be important for the geneticist who wants to study ancestry and perhaps crucial if the allele involved is linked to some inherited malformation, but usually the order will not matter since there will normally be no difference in the resulting phenotype whether one particular allele comes from the 'male' or the 'female' parent. We thus need to eliminate redundancies. Observation shows that the leading diagonal from the top left hand corner to the bottom

right hand corner has all entries different—different in the sense that one entry cannot be disordered to make another—whereas all entries on the other diagonal are essentially the same, only differing in inversions of the order of certain pairs, e.g. $rRgG$ and $RrGg$. All entries are in fact symmetrical about the main diagonal, but anti-symmetrical about the other diagonal (top right down to bottom left). In effect, the larger square can be subdivided into four squares, 1, 2, 3 and 4 proceeding clockwise, with 1 and 3 distinct and 2 and 4 the same apart from differences of order. For a large square $2^N \cdot 2^N$ each sub-square will have size $2^{N-1} \cdot 2^{N-1}$. Thus, the number of distinct entries will be 2 (number of entries for square $2^{N-1} \cdot 2^{N-1}$) + $1/2$ (2 (number of entries for square $2^{N-1} \cdot 2^{N-1}$)) = 3 (number of entries for square $2^{N-1} \cdot 2^{N-1}$). Since the number of distinct entries for the square $2^2 \cdot 2^2$ is 3, and for $2^3 \cdot 2^3$ is 3^2 , it can easily be proved by mathematical induction that, for N loci on independent distinct chromosomes, there will be $3N$ possibilities if we neglect differences of origin.

As for the distribution of the phenotype possibilities for the phenotype, we are dealing with independent outcomes and so we multiply the probabilities. Also, for any single locus with one pair of alleles, dominant R and recessive r , it is easy to see that we end up with phenotype probabilities in the ratio 3 : 1 in favour of the dominant R . For two loci, we end up with a distribution pattern 9 : 3 : 3 : 1 for combinations RG , Rg , rG and rg respectively. But the second and third are, genetically speaking, the same so we have a distribution $1 \cdot 3^2 + 2 \cdot 3^1 + 1 \cdot 3^0$ i.e. the binomial coefficients 1, 2, 1 multiplying matrix-wise the combinations (2 capitals), (1 capital), (0 capitals). Examination of the table for three loci, a so-called Trihybrid Cross, we see the pattern

$$1 \cdot 3^3 + 3 \cdot 3^2 + 3 \cdot 3^1 + 1 \cdot 3^0$$

emerging for distinct genetic possibilities. That this pattern continues can be proved by mathematical induction as follows.

Suppose we have the phenotype probabilities for $N = k$ loci, with independent pairs of alleles and two possibilities, dominant and recessive,

$$(3R + r)(3G + g) \dots (3T + t) \quad (k \text{ brackets}),$$

giving

$${}^k C_k 3^k \text{ (} k \text{ caps)} + {}^k C_{k-1} 3^{k-1} \text{ (} k-1 \text{ caps)} + \dots + {}^k C_0 3^0 \text{ (0 caps)}.$$

If we multiply the above by the independent phenotypical probability for a new locus and a new pair of alleles, $(3W + w)$, we obtain

$$\begin{aligned}
& {}^k C_k 3^{k+1} (k+1 \text{ caps}) + {}^k C_{k-1} 3^k (k \text{ caps}) + \cdots + {}^k C_0 3^1 (1 \text{ cap}) \\
& \quad + {}^k C_k 3^k (k \text{ caps}) + {}^k C_{k-1} 3^{k-1} (k-1 \text{ caps}) + \cdots + {}^k C_0 3^0 (0 \text{ caps}) \\
& = {}^k C_k 3^{k+1} (k+1 \text{ caps}) + ({}^k C_{k-1} + {}^k C_k) 3^k (k \text{ caps}) + \cdots + {}^k C_0 (0 \text{ caps}) \\
& = {}^{k+1} C_{k+1} 3^{k+1} (k+1 \text{ caps}) + {}^{k+1} C_k 3^k (k \text{ caps}) + \cdots + {}^{k+1} C_0 3^0 (0 \text{ caps})
\end{aligned}$$

since ${}^n C_r + {}^n C_{r-1} = {}^{n+1} C_r$. But we have, for $k = 2$, the phenotypical probabilities

$$1 \cdot 3^2 (2 \text{ caps}) + 3 \cdot 3^1 (1 \text{ cap}) + 1 \cdot 3^0 (0 \text{ caps}).$$

Generalization for more than two alleles

The preceding is an independent derivation of the Hardy–Weinberg distribution. They consider terms in the expansion of $(p+q)^2$ and map these on to the two allele case, while the three allele case is the trinomial expansion of $(p+q+r)^2$,

$$(p+q+r)^2 = p^2 + q^2 + r^2 + 2pq + 2pr + 2qr.$$

More generally, consider the alleles A_1, \dots, A_i given by the allele frequencies p_1, \dots, p_i , $(p_1 + \cdots + p_i)^2$, giving for all homozygotes $f(A_i A_i) = p_i^2$ and for all heterozygotes $f(A_i A_j) = 2p_i p_j$.

Generalization for polyploidy

The Hardy–Weinberg principle may also be generalized to polyploid systems, that is, for organisms that have more than two copies of each chromosome. Polyploidy—having more than two sets of chromosomes—is extremely rare in animals and infant humans rarely survive for more than a few days, if that. But, surprisingly, it is quite common in plants. A book on genetics informs me that the dandelion, with $n = 8$ chromosomes, can have $2n = 16$, $3n = 24$ or $4n = 32$, i.e. two, three or four sets of chromosomes. Polyploid plants tend to be larger and bear more succulent fruit: Amerindian peoples artificially selected for polyploid maize (which it is thought first arose by chance mutation), while the contemporary cultivated strawberry weighs in at $n = 8$, i.e. actually has eight sets of chromosomes!

The polyploidy case for two alleles is modelled by Hardy and Weinberg on the polynomial expansion of $(p+q)^c$, where c is the ploidy. For example, with tetraploid ($c = 4$) we obtain the following.

Genotype	Polynomial Expansion	Frequency
<i>AAAA</i>	p^4	1
<i>AAAa</i>	p^3q	4
<i>AAaa</i>	p^2q^2	6
<i>Aaaa</i>	pq^2	4
<i>aaaa</i>	q^4	1

For n distinct alleles in c -ploids, the genotype frequencies in the Hardy–Weinberg equilibrium are given by individual terms in the multinomial expansion of $(p_1 + \dots + p_n)^c$:

$$(p_1 + \dots + p_n)^c = \sum_{k_1, \dots, k_n \in \mathbb{N}, k_1 + \dots + k_n = c} \binom{c}{k_1 \dots k_n} p_1^{k_1} \dots p_n^{k_n}.$$

Tables

Two strains, both homozygous (true breeding), one with allele R (dominant) at one locus on one pair of chromosomes, the other with r (recessive) at the equivalent locus.

Genotype P_1	Gamete P_1
Male: RR or rr	R or r
Female: RR or rr	R or r

Zygote Genotype F_1	Phenotype F_1
R r	R r
R RR Rr	R R R
r rR rr	r R r
Ratio: 1 : 2 : 1	Ratio: 3 : 1

Zygote Genotype F_2	Zygote Phenotype F_2
RR Rr rR rr	RR Rr rR rr
RR $RRRR$ $RRRr$ $RRrR$ $RRrr$	RR RR RR RR RR
Rr $RRRr$ $RRrr$ $RrrR$ $Rrrr$	Rr RR Rr RR Rr
rR $rRRR$ $rRRr$ $rrRR$ $rrRr$	rR RR RR rR rR
rr $rRrR$ $rRrr$ $rrrR$ $rrrr$	rr RR Rr rR rr

4 capitals: 1

3 capitals: 4

2 capitals: 6

1 capital: 4

0 capitals: 1

Ratio : 9 : 3 : 3 : 1

3 types : RR , Rr or rR , rr

Homozygous: 2 ($RRRR$ and $rrrr$), heterozygous: 14 (the rest)

Genotypes at Two Loci <i>R/r</i> and <i>G/g</i>					Phenotype			
	<i>RG</i>	<i>Rg</i>	<i>rG</i>	<i>rg</i>				
<i>RG</i>	<i>RRGG</i>	<i>RRGg</i>	<i>RrGG</i>	<i>RrGg</i>	<i>RG</i>	<i>RG</i>	<i>RG</i>	<i>RG</i>
<i>Rg</i>	<i>RRgG</i>	<i>RRgg</i>	<i>RrgG</i>	<i>Rrgg</i>	<i>RG</i>	<i>Rg</i>	<i>RG</i>	<i>Rg</i>
<i>rG</i>	<i>rRGG</i>	<i>rRGg</i>	<i>rrGG</i>	<i>rrGg</i>	<i>RG</i>	<i>RG</i>	<i>rG</i>	<i>rG</i>
<i>rg</i>	<i>rRgG</i>	<i>rRgg</i>	<i>rrgG</i>	<i>rrgg</i>	<i>RG</i>	<i>Rg</i>	<i>rG</i>	<i>rg</i>
4 capitals: <i>RRGG</i> = 1/16					<i>RG:</i>	9/16		
3 capitals: <i>RRGg</i> + <i>RRgG</i> + <i>RrGG</i>					<i>Rg:</i>	3/16		
+ <i>rRrGG</i> = 4/16					<i>rG:</i>	3/16		
2 capitals: <i>RRgg</i> + <i>RrGg</i> + <i>RrgG</i>					<i>rg:</i>	1/16		
+ <i>rRGg</i> + <i>rRgG</i> + <i>rrGG</i> = 6/16								
1 capital: <i>Rrgg</i> + <i>rRgg</i> + <i>rrGg</i> + <i>rrgG</i> = 4/16								
0 capitals: <i>rrgg</i> = 1/16								
Ratio: 1 : 4 : 6 : 4 : 1					Ratio: 9 : 3 : 3 : 1			

Zygote Genetic Combinations for three Loci (Trihybrid Cross)

	<i>RBC</i>	<i>RBc</i>	<i>RbC</i>	<i>Rbc</i>	<i>rBC</i>	<i>rBc</i>	<i>rbC</i>	<i>rbc</i>
<i>RBC</i>	<i>RRBBCC</i>	<i>RRBBcC</i>	<i>RRBbCC</i>	<i>RRBbcC</i>	<i>RrBBCC</i>	<i>RrBBcC</i>	<i>RrBbCC</i>	<i>RrBbCc</i>
<i>RBc</i>	<i>RRBBcC</i>	<i>RRBBcc</i>	<i>RRBbcC</i>	<i>RRBbcc</i>	<i>RrBBcC</i>	<i>RrBBcc</i>	<i>RrBbcC</i>	<i>RrBbcc</i>
<i>RbC</i>	<i>RRbBCC</i>	<i>RRbBCc</i>	<i>RRbbCC</i>	<i>RRbbCc</i>	<i>RrbBCC</i>	<i>RrbBCc</i>	<i>RrbbCC</i>	<i>RrbbCc</i>
<i>Rbc</i>	<i>RRbBcC</i>	<i>RRbBcc</i>	<i>RRbbcC</i>	<i>RRbbcc</i>	<i>RrbBcC</i>	<i>RrbBcc</i>	<i>RrbbcC</i>	<i>Rrbbcc</i>
<i>rBC</i>	<i>rRBBCC</i>	<i>rRBBcC</i>	<i>rRBbCC</i>	<i>rRBbcC</i>	<i>rrBBCC</i>	<i>rrBBcC</i>	<i>rrBbCC</i>	<i>rrBbCc</i>
<i>rBc</i>	<i>rRBBcC</i>	<i>rRBBcc</i>	<i>rRBbcC</i>	<i>rRBbcc</i>	<i>rrBBcC</i>	<i>rrBBcc</i>	<i>rrBbcC</i>	<i>rrBbcc</i>
<i>rbC</i>	<i>rRbBCC</i>	<i>rRbBCc</i>	<i>rRbbCC</i>	<i>rRbbCc</i>	<i>rrbBCC</i>	<i>rrbBCc</i>	<i>rrbbCC</i>	<i>rrbbCc</i>
<i>rbc</i>	<i>rRbBcC</i>	<i>rRbBcc</i>	<i>rRbbcC</i>	<i>rRbbcc</i>	<i>rrbBcC</i>	<i>rrbBcc</i>	<i>rrbbcC</i>	<i>rrbbcc</i>
6 capitals present: 1		3 capitals present: 20		0 capitals present: 1				
5 capitals present: 6		2 capitals present: 15						
4 capitals present: 15		1 capital present: 6		Total: 64 = (1 + 3) ³				

Offspring Phenotypes for three Loci (Trihybrid Cross)

	<i>RBC</i>	<i>RBc</i>	<i>RbC</i>	<i>Rbc</i>	<i>rBC</i>	<i>rBc</i>	<i>rbC</i>	<i>rbc</i>
<i>RBC</i>	<i>RBC</i>	<i>RBC</i>	<i>RBC</i>	<i>RBC</i>	<i>RBC</i>	<i>RBC</i>	<i>RBC</i>	<i>RBC</i>
<i>RBc</i>	<i>RBC</i>	<i>RBc</i>	<i>RBC</i>	<i>RBc</i>	<i>RBC</i>	<i>RBc</i>	<i>RBC</i>	<i>RBc</i>
<i>RbC</i>	<i>RBC</i>	<i>RBC</i>	<i>RbC</i>	<i>RbC</i>	<i>RBC</i>	<i>RBC</i>	<i>RbC</i>	<i>RbC</i>
<i>Rbc</i>	<i>RBC</i>	<i>RBc</i>	<i>RbC</i>	<i>Rbc</i>	<i>RBC</i>	<i>Rbc</i>	<i>RbC</i>	<i>Rbc</i>
<i>rBC</i>	<i>RBC</i>	<i>RBC</i>	<i>RBC</i>	<i>RBC</i>	<i>rBC</i>	<i>rBC</i>	<i>rBC</i>	<i>rBC</i>
<i>rBc</i>	<i>RBC</i>	<i>RBc</i>	<i>RBC</i>	<i>RBc</i>	<i>rBC</i>	<i>rBc</i>	<i>rBC</i>	<i>rBc</i>
<i>rbC</i>	<i>RBC</i>	<i>RBC</i>	<i>RbC</i>	<i>RbC</i>	<i>rBC</i>	<i>rBC</i>	<i>rbC</i>	<i>rbC</i>
<i>rbc</i>	<i>RBC</i>	<i>RBc</i>	<i>RbC</i>	<i>Rbc</i>	<i>rBC</i>	<i>rBc</i>	<i>rbC</i>	<i>rbc</i>
<i>RBC:</i> 27, <i>RBc:</i> 9, <i>RbC:</i> 9, <i>rBC:</i> 9, <i>Rbc:</i> 3, <i>rBc:</i> 3, <i>rbC:</i> 3, <i>rbc:</i> 1								
Total: 64 = 3 ³ + 3 · 3 ² + 3 · 3 ¹ + 3 ⁰								

Solution 276.9 – Sets

For $n = 1, 2, \dots$, define the set $S(n)$ by

$$\begin{aligned} S(1) &= \{1\}, \\ S(n+1) &= \{2x+1 : x \in S(n)\} \cup \{3x+1 : x \in S(n)\}. \end{aligned}$$

Either show that $S(n)$ always has 2^{n-1} elements, or find a counter-example.

Graham Lovegrove

This is true, not only for 2 and 3, but for most positive numbers. First of all, note that an element x of $S(n)$ can be written as

$$x = 1 + a_1(1 + a_2(1 + \dots + a_{n-1}(1 + a_n)) \dots). \quad (1)$$

We will denote this by $[a_1, a_2, \dots, a_n]$.

The elements of $S(5)$ in order of magnitude are

$$\begin{aligned} [2,2,2,2] &= 31, & [2,2,2,3] &= 39, & [2,2,3,2] &= 43, & [2,3,2,2] &= 45, \\ [3,2,2,2] &= 46, & [2,2,3,3] &= 55, & [2,3,2,3] &= 57, & [3,2,2,3] &= 58, \\ [2,3,3,2] &= 63, & [3,2,3,2] &= 64, & [3,3,2,2] &= 67, & [2,3,3,3] &= 81, \\ [3,2,3,3] &= 82, & [3,3,2,3] &= 85, & [3,3,3,2] &= 94, & [3,3,3,3] &= 121. \end{aligned}$$

There is a clear pattern in the sequences of twos and threes, but the way in which this works is not immediately obvious. The first thing to note is that elements of $S(5)$ with the same number of threes are grouped together in magnitude, and that the smallest in each group has all the twos at the beginning, and the largest has all the twos at the end. Also, within each group, each 3 moves as far left as it can before the next 3 moves. Lastly, notice the transition between $[3, 2, 2, 3] = 58$ and $[2, 3, 3, 2] = 63$. This becomes more obvious when n is larger. If we can explain the relationship between the patterns of twos and threes and the relative magnitudes of the resulting elements of $S(n)$, that should incidentally settle the question of whether any two different patterns can result in the same value. By multiplying out, we can see the following.

Lemma 1

$$[a_1, a_2, \dots, a_n] = 1 + \sum_{i=1}^n \prod_{j=1}^i a_j.$$

So for instance $[2, 3, 3, 2] = 1 + 2 + 2 \cdot 3 + 2 \cdot 3^2 + 2^2 \cdot 3^2$, etc. The question of how the sequencing works remains. It is clear from Lemma 1 that a three

early in the sequence is a multiplier in all the later terms. This suggests an alternative way of writing the sequences, by recording the cumulative number of threes, so we could represent $[2, 3, 2, 3]$ as $[[0, 1, 1, 2]]$. So $S(5)$ would be written

$$\begin{aligned} [[0,0,0,0]] &= 31, & [[0,0,0,1]] &= 39, & [[0,0,1,1]] &= 43, & [[0,1,1,1]] &= 45, \\ [[1,1,1,1]] &= 46, & [[0,0,1,2]] &= 55, & [[0,1,1,2]] &= 57, & [[1,1,1,2]] &= 58, \\ [[0,1,2,2]] &= 63, & [[1,1,2,2]] &= 64, & [[1,2,2,2]] &= 67, & [[0,1,2,3]] &= 81, \\ [[1,1,2,3]] &= 82, & [[1,2,2,3]] &= 85, & [[1,2,3,3]] &= 94, & [[1,2,3,4]] &= 121. \end{aligned}$$

Of course, by emphasizing the cumulative number of threes, this notation ignores the twos altogether. Although this notation doesn't work in quite the same way as normal place-value, this representation does seem to help decide which of two sequences has the bigger value. Reading from right to left we identify the first position where they differ, so for instance $[[0, 1, 2, 2]]$ is bigger than $[[1, 1, 1, 2]]$. To prove that this works, we will show that for two sequences of this second type that differ first in the k th position, e.g. they are of the form $x = [[u_1, u_2, \dots, u_k, u_{k+1}, \dots, u_n]]$ and $y = [[v_1, v_2, \dots, v_k, u_{k+1}, \dots, u_n]]$ with $u_k = v_k + 1$ and the sequences agreeing in all positions after the k th, we have $x > y$ (showing this will also suffice for $u_k > v_k + 1$, since x would then be still larger). To prove this we can restrict ourselves to sequences with the prescribed values in the k th and subsequent positions, but with initial sequences u_1, u_2, \dots, u_{k-1} and v_1, v_2, \dots, v_{k-1} that minimize the value of x and maximize y , so that proving $x - y > 0$ in this case proves it for all cases. Going back to the earlier representation, we know that to maximize the value of a sequence with a given number of twos and threes we must have as many initial threes as possible, and to minimize, we must maximize the number of initial twos. Therefore if $u_k = q + 1$ and $v_k = q$, we can represent x as $k - q - 1$ twos followed by $q + 1$ threes, and y as q threes followed by $k - q$ twos. So we can write

$$\begin{aligned} x &= 1 + 2 + 2^2 + \dots + 2^{k-q-1} + 2^{k-q-1} \cdot 3 + 2^{k-q-1} \cdot 3^2 + \dots \\ &\quad + 2^{k-q-1} \cdot 3^{q+1} + A \\ &= 2^{k-q} - 1 + 2^{k-q-2} \cdot 3 \cdot (3^{q+1} - 1) + A \end{aligned}$$

and

$$\begin{aligned} y &= 1 + 3 + 3^2 + \dots + 3^q + 3^q \cdot 2 + 3^q \cdot 2^2 + \dots + 2^{k-q} \cdot 3^q + A \\ &= \frac{3^{q+1} - 1}{2} + 2 \cdot 3^q \cdot (2^{k-q} - 1) + A, \end{aligned}$$

where A represents the terms resulting from the $k + 1^{\text{th}}$ and subsequent positions, which are the same for both x and y . So

$$\begin{aligned} 2(x - y) &= 2^{k-q+1} - 2 + 2^{k-q-1}(3^{q+2} - 3) - 3^{q+1} + 1 - 3^q(2^{k-q+2} - 4) \\ &= 2^{k-q-1} \cdot 3^q + 2^{k-q-1} + 3^q - 1 > 0 \end{aligned}$$

if $k > 1$, $k \geq q \geq 0$. So this confirms the total order suggested by the second representation and proves that all values of $S(n)$ are distinct, and therefore that $|S(n)| = 2^{n-1}$. It is easy to see from the above that $x - y = 1$ only when $q = 0$ and $k = 1$. In $S(5)$, there are three occurrences of this, for $(x, y) = (57, 58), (63, 64), (81, 82)$.

All this prompts the question: is there anything special about the numbers 2 and 3; will this work with any pair of numbers? The above calculation might suggest this as a possibility. One definite counter-example would be 1 and 2, because $[1, 1, 2, 2] = [2, 1, 1, 1] = 9$.

We generalize the definition of $S(n)$ to the set $S(a, b, n)$, using arbitrary numbers $a, b, b > a$ instead of 2 and 3. Repeating the above calculation, we obtain

$$\begin{aligned} &(a - 1)(b - 1)(x - y) \\ &= (b - a)a^{k-q}b^q \left(\frac{(a - 1)(b - 1) - 1}{a} + \frac{a^{k+q-1} + b^q - 1}{a^{k-q}b^q} \right) \\ &= (b - a) \left((a - 1)(b - 1)a^{k+q-1}b^q - (a^{k+q-1} - 1)(b^q - 1) \right). \end{aligned}$$

If $a, b \geq 2$, then the right-hand side of this equation is certainly greater than zero, since $a^{k+q-1}b^q > (a^{k+q-1} - 1)(b^q - 1)$. So it does work for other values.

The question of what happens when $a = 1$ requires us first to rewrite the equation as

$$(x - y) = (b - a) \left(a^{k+q-1}b^q - \frac{(b^q - 1)}{(b - 1)} \sum_{i=0}^{k+q-2} a^i \right).$$

So for $a = 1$ we have

$$(x - y) = ((b - 1)b^q - (k - q - 1)(b^q - 1)),$$

which is less than or equal to zero for $k \geq b + q + 1$, with equality if $k = b + 2$ and $q = 1$. The expression is, however, positive if $k \leq b + q$. Note however that this deals only with a special choice of successive elements x and y , and

that if these conditions for x to be strictly less than y are satisfied, then it is possible that other possibly non-adjacent elements of the set are equal.

Thus, in general $|S(1, b, n)| < 2^{n-1}$. The first example is $S(1, 2, 4)$, where $[1, 1, 2, 2] = [2, 1, 1, 1] = 9$. For $S(1, 2, 5)$, there are instances of $x - y < 0$, and so the ordering of elements developed above breaks down as predicted. Other related questions come to mind:

1. Does this work for three numbers, say 2, 3 and 4, yielding sets of size 3^{n-1} ?
2. Are there sets of n distinct numbers a_1, a_2, \dots, a_n such that distinct permutations σ generate distinct values of $[a_{\sigma(1)}, a_{\sigma(2)}, \dots, a_{\sigma(n)}]$?

Problem 281.1 – Graph complement

The *complement* of a simple graph, G , is constructed by taking the vertices of G and joining every pair of vertices which are not joined in G . Show that if a graph of v vertices is isomorphic to its complement, then $v \equiv 0$ or $1 \pmod{4}$.

When $v = 4$ we see that the complement of $\bullet - \bullet - \bullet - \bullet$ is $\bullet - \bullet - \bullet - \bullet$, which is indeed isomorphic to $\bullet - \bullet - \bullet - \bullet$. Are there any others?

For $v = 5$, you can verify that the complement of a 5-cycle is a 5-cycle. Can you find all 5-vertex graphs that are isomorphic to their complements?

Find a graph of 8 vertices that is isomorphic to its complement.

Simple means that the graph has no loops or multiple edges, and those edges that it does have are undirected. *Isomorphic* means that as unlabelled graphs they are really the same even if they look different. See, for example, Robin Wilson's *Introduction to Graph Theory* for the technical details.

Thanks to Dick Clements for suggesting the problem at this year's M500 Winter Weekend.

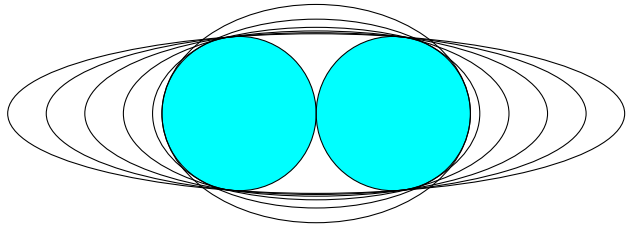
Problem 281.2 – Hours

Jeremy Humphries

Four 'hour positions' are chosen at random on a standard clock face. What is the probability that, taken together, they define (a) a point, (b) a line, (c) a triangle, (d) a quadrilateral?

Solution 278.7 – Two circles and an ellipse

What is the smallest area of an ellipse containing two unit circles?



Dick Boardman

The ellipse will have its centre at the point of contact of the two circles and its major axis will be on the line joining the two centres. The figure will then be symmetrical about both axes; so we only need to worry about the positive quadrant. The parametric form of a unit circle with centre $(1, 0)$ is

$$x = 1 + \cos \theta, \quad y = \sin \theta.$$

The slope of the tangent at the point (x, y) is $-\cot \theta$. Express these parameters in terms of the half-angle $t = -\tan(\theta/2)$:

$$x = 1 + \frac{1 - t^2}{1 + t^2}, \quad y = \frac{2t}{1 + t^2}, \quad \text{slope} = -\frac{1 - t^2}{2t}.$$

The general form of an ellipse centred at the origin is $x^2/a^2 + y^2/b^2 = 1$, which simplifies to

$$\frac{4(b^2 + a^2t^2)}{a^2b^2(1 + t^2)^2} = 1. \quad (1)$$

The slope of the tangent to an ellipse at the point (x, y) is $-x/a^2 \cdot b^2/y$, which simplifies to $-b^2/(a^2t)$. Any ellipse which just encloses the two circles and goes through the point (x, y) will have the same slope at that point. Thus

$$-\frac{1 - t^2}{2t} = -\frac{b^2}{a^2t}. \quad (2)$$

Solving (1) and (2) for positive a and b , gives

$$a = \frac{2}{\sqrt{1 - t^4}}, \quad b = \frac{\sqrt{2}}{\sqrt{1 + t^2}}.$$

The area of the ellipse is πab ; so the area of any ellipse just enclosing the two circles is

$$\frac{2\sqrt{2}\pi}{\sqrt{1 + t^2} \sqrt{1 - t^4}},$$

which attains its minimum of $3\sqrt{3}\pi/2$ at $t = 1/\sqrt{3}$.

Ted Gore

The equation of one of the circles is $(x - 1)^2 + y^2 = 1$. Let the equation for the ellipse be

$$\frac{x^2}{A} + \frac{y^2}{B} = 1.$$

From the circle equation, $y^2 = 2x - x^2$ we wish to find points of intersection of the ellipse with the circle; so, substituting this in the ellipse equation we obtain

$$x = \frac{2A \pm \sqrt{4A^2 - 4(A - B)AB}}{2(A - B)}.$$

If there were two points then part of the circle would fall outside the ellipse. We want just one point so we set $4A^2 = 4(A - B)AB$, which gives

$$B = \frac{A \pm \sqrt{A^2 - 4A}}{2}.$$

Let E be the area of the ellipse and let

$$K = AB = \frac{1}{2} \left(A^2 - \sqrt{A^4 - 4A^3} \right).$$

Then $E = \pi\sqrt{K}$ and

$$\frac{dE}{dA} = \frac{dE}{dK} \times \frac{dK}{dA}, \quad \frac{dK}{dA} = \frac{1}{2} \left(2A - \frac{4A^3 - 12A^2}{2\sqrt{A^4 - 4A^3}} \right).$$

But dK/dA and therefore also dE/dA are zero when $A = 9/2$ and $B = 3/2$. Therefore the minimum value of E is $3\sqrt{3}\pi/2 = 8.1620971$.

Problem 281.3 – Powers of 2

Dave Wild

Show that for any positive integer there is a non-negative power of two which starts with this integer. For example, we have the following.

Integer	1	2	3	4	5	6	7	8	9	10	500
Power of 2	0	1	5	2	9	6	46	3	53	10	1650

This problem was mentioned in the *Coursera* course ‘What is a Proof?’

Solution 259.7 – Admissible numbers

Let k be a positive integer. A positive integer n is called *admissible* if $n(n-1) \equiv 0 \pmod{k(k-1)}$ and $n-1 \equiv 0 \pmod{k-1}$. Show that when k is a prime power the only admissible numbers are $k(k-1)t+1$ and $k(k-1)t+k$, $t = 0, 1, \dots$

Dave Wild

As $n(n-1) \equiv 0 \pmod{k(k-1)}$ and k is a prime power, since $\gcd(n, n-1) = 1$, either k divides n or k divides $n-1$.

Case 1. k divides $n-1$. Since $n-1 \equiv 0 \pmod{k-1}$, both k and $k-1$ divide $n-1$. Since $\gcd(k, k-1) = 1$ and $n > 0$, $n-1 = k(k-1)t$, where t is a non-negative integer. So $n = k(k-1)t+1$, $t = 0, 1, \dots$ satisfies the given conditions.

Case 2. k divides n . Since $n-1 \equiv 0 \pmod{k-1}$, $n-1 = s(k-1)$ for some non-negative integer s . So $n = (s-1)(k-1) + k$. As k divides n and $\gcd(k, k-1) = 1$, k divides $s-1$. As $n > 0$, $n = k(k-1)t+k$, $t = 0, 1, \dots$ also satisfies the given conditions.

Therefore when k is a prime power the only admissible numbers are of the forms found above.

Problem 281.4 – Pythagorean triple generator

Suppose $x^2 + y^2 = z^2$, where x , y and z are integers. Let

$$\begin{bmatrix} X \\ Y \\ Z \end{bmatrix} = \begin{bmatrix} 1 & 2 & 2 \\ 2 & 1 & 2 \\ 2 & 2 & 3 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix}. \quad (1)$$

Show that X , Y and Z are also integers and that they too satisfy $X^2 + Y^2 = Z^2$. For example, if you start with the vector $(1, 0, 1)$ and then repeatedly multiply by the matrix, you get

$(3, 4, 5)$, $(21, 20, 29)$, $(119, 120, 169)$, $(697, 696, 985)$, $(4059, 4060, 5741)$, \dots

Thus (1) preserves Pythagorean tripleness. We wonder if you can solve the problem by some means other than actually computing $X^2 + Y^2$ and Z^2 .

Can you find other 3×3 matrices that have the same property?

Thanks to Rob and Judith Rolfe for suggesting this interesting problem, which I (TF) must admit I have not seen before, and there I was thinking that I knew all there is to know about Pythagorean triples, at this year's M500 Winter Weekend.

Solution 277.5 – Closure and complement

Show that if X is any set of real numbers, then, with the usual topology, the closure of the complement of the closure of the complement of the closure of the complement of the closure of X is equal to the closure of the complement of the closure of X .

For example, if X is the set of rational numbers in the interval $(0, 1)$, you should end up with $(-\infty, 0] \cup [1, \infty)$ in both cases. This problem is actually a special case of Theorem of the Day, number 239, Kuratowski's 14-set Theorem, <http://www.theoremoftheday.org>.

Graham Lovegrove

I don't think this needs to be limited to sets in the real line, as the proof only needs the simple properties of closed sets and closure. I use the normal notations: \overline{A} stands for the closure of A in \mathbb{R} , and \tilde{A} for its complement $\mathbb{R} \setminus A$.

Then $\overline{\overline{\overline{\overline{\overline{X}}}}} \supseteq \overline{\overline{\overline{\overline{\overline{X}}}}} = \mathbb{R} \setminus \overline{\overline{\overline{\overline{\overline{X}}}}}$ since the closure of a set contains the set itself.

So $\overline{\overline{\overline{\overline{\overline{\overline{\overline{\overline{X}}}}}}} \supseteq \overline{\overline{\overline{\overline{\overline{\overline{\overline{\overline{X}}}}}}}$ by taking complements. So $\overline{\overline{\overline{\overline{\overline{\overline{\overline{\overline{X}}}}}}} \supseteq \overline{\overline{\overline{\overline{\overline{\overline{\overline{\overline{X}}}}}}}$ since $\overline{\overline{\overline{\overline{\overline{\overline{\overline{\overline{X}}}}}}}$ is closed, and so contains the closure of any subset.

Conversely, $\overline{\overline{\overline{\overline{\overline{\overline{\overline{\overline{X}}}}}}} \supseteq \overline{\overline{\overline{\overline{\overline{\overline{\overline{\overline{X}}}}}}} = \mathbb{R} \setminus \overline{\overline{\overline{\overline{\overline{\overline{\overline{\overline{X}}}}}}}$ since the closure of a set contains the set itself. So $\overline{\overline{\overline{\overline{\overline{\overline{\overline{\overline{X}}}}}}} \subseteq \mathbb{R} \setminus \overline{\overline{\overline{\overline{\overline{\overline{\overline{\overline{X}}}}}}}$, and so $\overline{\overline{\overline{\overline{\overline{\overline{\overline{\overline{X}}}}}}} \subseteq \overline{\overline{\overline{\overline{\overline{\overline{\overline{\overline{X}}}}}}}$ since $\overline{\overline{\overline{\overline{\overline{\overline{\overline{\overline{X}}}}}}}$ is closed. Therefore $\overline{\overline{\overline{\overline{\overline{\overline{\overline{\overline{X}}}}}}} \supseteq \overline{\overline{\overline{\overline{\overline{\overline{\overline{\overline{X}}}}}}}$ by taking complements, and so we have $\overline{\overline{\overline{\overline{\overline{\overline{\overline{\overline{X}}}}}}} \supseteq \overline{\overline{\overline{\overline{\overline{\overline{\overline{\overline{X}}}}}}}$ since the closure of a set is contained in the closure of any superset.

Hence both $\overline{\overline{\overline{\overline{\overline{\overline{\overline{\overline{X}}}}}}} \supseteq \overline{\overline{\overline{\overline{\overline{\overline{\overline{\overline{X}}}}}}}$ and $\overline{\overline{\overline{\overline{\overline{\overline{\overline{\overline{X}}}}}}} \supseteq \overline{\overline{\overline{\overline{\overline{\overline{\overline{\overline{X}}}}}}}$, and thus $\overline{\overline{\overline{\overline{\overline{\overline{\overline{\overline{X}}}}}}} = \overline{\overline{\overline{\overline{\overline{\overline{\overline{\overline{X}}}}}}}$, as required.

Problem 281.5 – Integral

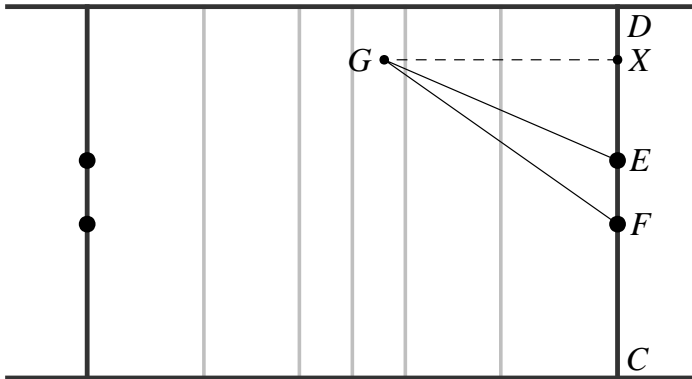
Show that

$$\int_{-\infty}^{\infty} \frac{\cos x}{1+x^6} dx = \frac{\pi (1 + \sqrt{e} \cos(\sqrt{3}/2) + \sqrt{3e} \sin(\sqrt{3}/2))}{3e} \approx 1.63477.$$

Just in case it might be relevant, recall that in Problem 242.1 we asked you for a proof that $\int_{-\infty}^{\infty} \cos x/(1+x^2) dx = \pi/e$.

Solution 278.6 – Rugby conversions

On the rugby pitch below a try is scored by touching the ball down at X . Where is G to maximize the angle EGF subject to GX orthogonal to the goal line CD ?



Chris Pile

Let the goal width be $EF = 2$ in some units. Let t be the distance from the point of touch down, X , to the mid point between the goal posts. Let d be the distance walked to the kicking point, G , from X . Let $A = \angle XGF$, $B = \angle XGE$. Then

$$\tan A = \frac{t+1}{d}, \quad \tan B = \frac{t-1}{d},$$

$$\tan(a - B) = \frac{\tan A - \tan B}{1 + \tan A \tan B} = \frac{2d}{d^2 + t^2 - 1}.$$

To maximize the angle EGF we can maximize $\tan(A - B)$ by differentiating and equating to 0:

$$\frac{2(d^2 + t^2 + 1) - (2d)^2}{(d^2 + t^2 - 1)^2} = 0; \text{ hence } t^2 - 1 - d^2 = 0.$$

Therefore $d = \sqrt{t^2 - 1}$.

By intuition, we would expect the angle to be about 45 degrees, as indeed it is, approximately, for t not too small. If t is less than about 2, it would be necessary to walk further for the ball to gain height to go over the bar.

M500 Mathematics Revision Weekend 2018

The forty-fourth M500 Revision Weekend will be held at

**Yarnfield Park Training and Conference Centre,
Yarnfield, Staffordshire ST15 0NL
from Friday 18th to Sunday 20th May 2018.**

The standard cost, including accommodation (with en suite facilities) and all meals from dinner on Friday evening to lunch on Sunday is £265 for single occupancy, or £230 per person for two students sharing in either a double or twin bedded room. The standard cost for non-residents, including Saturday and Sunday lunch, is £150.

Members may make a reservation with a £25 deposit, with the balance payable at the end of February. Non-members must pay in full at the time of application and all applications received after 28th February 2018 must be paid in full before the booking is confirmed. Members will be entitled to a discount of £15 for all applications received before 18th April 2018. The Late Booking Fee for applications received after 18th April 2018 is £20, with no membership discount applicable.

A shuttle bus service will be provided between Stone station and Yarnfield Park on Friday and Sunday. This will be free of charge, but seats will be allocated for each service and must be requested before 1st May.

There is free on-site parking for those travelling by private transport. For full details and an application form, see the Society's web site:

www.m500.org.uk.

The Weekend is open to all Open University students, and is designed to help with revision and exam preparation. We expect to offer tutorials for most undergraduate and postgraduate mathematics OU modules, subject to the availability of tutors and sufficient applications.

Please note that the venue is not the same as last year. We go back to the 2016 location.

Problem 281.6 – Primitive Pythagorean triples

A *Pythagorean triple* is three non-negative integers (a, b, c) such that $a^2 + b^2 = c^2$, and *primitive* means $\gcd(a, b, c) = 1$.

Show that a primitive Pythagorean triple contains exactly one multiple of 3, exactly one multiple of 4 and exactly one multiple of 5. For example, in the triple $(119, 7080, 7081)$ the unique multiples of 3, 4 and 5 are respectively 7080, 7080 and 7080.

Thanks to Judith and Rob Rolfe for suggesting the problem at this year's M500 Winter Weekend. See page 14 in this issue for another problem about Pythagorean triples.

Mathematics of heredity	
Sebastian Hayes	1
Solution 276.9 – Sets	
Graham Lovegrove	8
Problem 281.1 – Graph complement	11
Problem 281.2 – Hours	
Jeremy Humphries	11
Solution 278.7 – Two circles and an ellipse	
Dick Boardman	12
Ted Gore	13
Problem 281.3 – Powers of 2	
Dave Wild	13
Solution 259.7 – Admissible numbers	
Dave Wild	14
Problem 281.4 – Pythagorean triple generator	14
Solution 277.5 – Closure and complement	
Graham Lovegrove	15
Problem 281.5 – Integral	15
Solution 278.6 – Rugby conversions	
Chris Pile	16
M500 Mathematics Revision Weekend 2018	17
Problem 281.6 – Primitive Pythagorean triples	17

Front cover A group divisible design with block size 4 and group type 2^7 . The 7 groups of size 2 are arranged vertically, $\{0, 7\}$, $\{1, 8\}$, ..., $\{6, 13\}$. There are 14 blocks, which are shown in various shades of grey (colours) and are generated from the grey (red) block $\{0, 1, 4, 6\}$ by $x \mapsto x + 1 \pmod{14}$. The reader can verify that each pair of points occurs in precisely one group or in precisely one block but not both.

We are again running low. I have had to hang on to a certain amount of material, including solutions to problems that appeared in issue 279, to make sure there is something to print in 282. As usual, apart from problems and their solutions, we desperately want articles of, say, 2–8 pages on mathematical topics that can be readily understood by first-year undergraduates taking mathematics courses. Please see ‘Advice for Authors’ on the M500 Web site. — TF
