THE PIGEON-HOLE PRINCIPLE

Valentin Goranko

Introduction

The South African Mathematical Society has the responsibility for selecting and training teams to represent South African in the annual International Mathematical Olympiad (IMO).

The process of finding a team to go to the IMO is a long one. It begins with a nationwide Mathematical Talent Search, in which students are sent sets of problems to solve. Their submissions are marked and returned with comments, full solutions and a further set of problems. The principle behind the Talent Search is straightforward: the more problems you solve, the higher up the ladder you climb and the closer you get to selection.

The best students in the Talent Search are invited to attend Mathematical Camps in which specialised problem-solving skills are taught. The students also write a series of challenging Olympiad-level problem papers, leading to selection of a team of six to go to the IMO.

The booklets in this series cover topics of particular relevance to Mathematical Olympiads. Though their primary purpose is preparing students for the International Mathematical Olympiad, they can with profit be read by all interested high school students who would like to extend their mathematical horizons beyond the confines of the school syllabus. They can also be used by teachers and university mathematicians who are interested in setting up Olympiad training programmes and need ideas on topics to cover and sample Olympiad problems.

Titles in the series to date are:

No. 1 *The Pigeon-hole Principle*, by Valentin Goranko

No. 2 *Topics in Number Theory*, by Valentin Goranko

No. 3 *Inequalities for the Olympiad Enthusiast*, by Graeme West

No. 4 *Graph Theory for the Olympiad Enthusiast*, by Graeme West

No. 5 *Functional Equations for the Olympiad Enthusiast*, by Graeme West

No. 6 *Mathematical Induction for the Olympiad Enthusiast*, by David Jacobs

No. 7 *Geometry for the Olympiad Enthusiast*, by Bruce Merry

Details of the South African Mathematical Society's Mathematical Talent Search may be obtained by writing to

Mathematical Talent Search
Department of Mathematics and Applied Mathematics
University of Cape Town
7700 Rondebosch

The International Mathematical Olympiad Talent Search is sponsored by Old Mutual

J H Webb
June 1996

The Pigeon-Hole Principle

Do you know that there are at least two people in Pretoria with the same number of hairs on their heads? (Well, let us ignore bald people to make this statement non-trivial.) I see you would ask me *how i know that.* Of course, you do not believe that I have counted the number of hairs on the head of each inhabitant of Pretoria. All my life would not suffice to accomplish such a task. Then *how*? Very simple: a mixture of little bits of biology, statistics and trivial mathematics tells me that. It is known that the number of hairs on the head of any human being is *less than* 200 000 (biology). Well, you agree that *more than* 200 000 not bald-haired people live in Pretoria (statistics). If every two of them had different number of hairs on their heads, then somebody should have more than 200 000 hairs on his head, because the number of positive integers not exceeding 200 000 is, of course, only 200 000 (trivial mathematics). Isn't that convincing?

1 The basic pigeon-hole principle

The above peculiar argument is a typical example of an extremely simple mathematical principle which, in the same magic way as a conjurer produces a rabbit out of his empty sleeve, implies unbelievably many interesting and deep results which otherwise would require enormous, if not impossible, technical efforts to obtain (as in our example). This principle is often called the *pigeon-hole principle* because a popular version of it reads:

If more than n pigeons fly into n pigeon-holes then at least two pigeons will get into the same pigeon-hole.

Obvious, isn't it? So obvious that if you have to prove it you would wonder what actually is to be proved. Of course you could object that this is not a precise mathematical statement, and therefore no mathematical (i.e. formal) proof is applicable. You would be right. Here is a more formal expression of the same idea.

Let P (for pigeons) and H (for pigeon-holes) be finite sets (i.e. sets with finite numbers of elements; suppose we know what this means) and P has more elements than H. Then for any rule which attaches every element of P to an element of H there are (at least) two elements of P which are attached to the same element of H.

This is what we shall hereafter call the pigeon-hole principle (PHP for short). Try to recognize the original idea in this formal statement. The PHP is also known as *Dirichlet's box principle* after the famous German mathematician Peter Gustav Lejeune-Dirichlet who first obtained serious mathematical results (in number theory) by applying this principle.

And yet, the task to prove the PHP in its formal appearance becomes no simpler. Well, this can be done in various ways, say by mathematical induction. The question however, is whether the principle of mathematical induction is something *more obvious* than the PHP or it should not be proved as well. Now we are wading into deep water and the matter goes beyond the scope of this article. The problem is briefly discussed in the appendix.

Here we shall learn how to apply the PHP in various situations. The art of successful applications of the PHP consists in an appropriate choice of the "pigeons" and "pigeon-holes". Let us start with a paraphrase of the argument at the beginning of the article using our principle. Consider all non bald-headed people from Pretoria as "pigeons" and 200 000 pigeon-holes numbered by the positive integers from 1 to 200 000. Then we let every "pigeon" fly into the pigeon-hole numbered by the number of hairs on the "pigeon's" head. Due to the fact that the number of hairs cannot exceed 200 000 there will be no holeless pigeons. Then, since the "pigeons" are more than the pigeon-holes, according to the PHP at least two "pigeons" will inevitably occur in the same pigeon-hole, i.e. two people will turn out to have the same number of hairs on their heads.

Now, try to apply the PHP to each of the following problems on your own before you read the hints and solutions which are given at the end of the sections. The problems do not involve any mathematics outside the standard school syllabus, except for some basic knowledge about prime numbers and divisibility.

Problem 1. There are 370 students in a school. Show that two of them have birthdays on the same day.

Problem 2. Five points are placed inside an equilateral triangle with sides of length 1. Show that the distance between two of them is less than 0.5.

Problem 3. 25 points are placed in a 6×16 rectangle. Show that there are two between which the distance is not greater than $2\sqrt{2}$.

Problem 4. Prove that for every positive integer n, amongst any $n+1$ integers $m_1, \ldots, m_{n+1}$ there are two, the difference between which is a multiple of n.

Problem 5. Prove that for every positive integer n, and set of n integers $\{m_1, \ldots, m_n\}$ contains a non-empty subset, the sum of whose elements is a multiple of n.

Problem 6. Show that there is a number of the form

$$39913991 \cdots 39910 \cdots 0$$

which is divisible by 1993.

Problem 7. Prove that amongst $n + 1$ different integers $m_1, \ldots, m_{n+1}$, where $1 \leq m_1 < m_2 < \cdots < m_{n+1} \leq 2n$, there are three of which one is equal to the sum of the other two.

Problem 8. Two integers are called *relatively prime* if they have no common divisor greater than 1. Prove that is m and n are relatively prime positive integers, then there is a positive integer k such that n divides $mk - 1$.

Problem 9. Is there a power of 7 which, when written in decimal notation, ends with 00000001?

Problem 10. Prove that if m and n are relatively prime positive integers, then there is a positive integer k such that n divides $m^k - 1$.

Problem 11. Let A be a set of 101 positive integers, each of which is not greater than 200. Prove that there are two numbers in A, one of which is a divisor of the other.

Problem 12. Show that in every set of 201 positive integers, each less than 301, there are two, the ratio of which is a power of 3.

Problem 13. Let M be a set of 75 positive integers not greater than 100. Show that for each positive integer $k \leq 49$ there are two elements of M which differ by k.

Problem 14. In a college there are 1993 students. Some of them know each other, others not. Assume that if student X knows student Y, then Y knows X too. Show that there are two students who know the same number of fellow students.

Problem 15. Given are $n \geq 2$ points in the plane, some of which are connected by line segments. Prove that there are two points which are ends of the same number of segments.

Problem 16. Let x be a real number and n a positive integer. Show that there exist integers p and q such that

$$1 \leq q \leq n \quad \text{and} \quad \left| x - \frac{p}{q} \right| < \frac{1}{qn}.$$

Hints and Solutions

Solution 1. Let the students be our "pigeons". There are at most 366 days in the year; let them be the "pigeon-holes". By the PHP there must be at least two pigeons in some pigeon-hole, i.e. at least two students are born on the same day. In formal terms: let P be the set of students in the school and H be the set of the days in the year. Our rule attaches every student to tis birthday. Then, according the "formal" PHP two students must be attached to the same day, i.e. must have the same birthday.

Solution 2. The lines connecting the midpoints of the sides of triangle divide it into 4 equilateral triangles each with side of length 0.5. Noe let us take these 4 triangles as "pigeon-holes". Each of the 5 points lie in some (at least one) of them. Therefore, by the PHP, at least two of the points must lie inside the same "pigeon-hole". The the distance between them is less than 0.5.

Hint 3. Partition the rectangle into 24 squares with side 3 and apply the idea of the previous problem.

Solution 4. Note that if a and b are integers then $a - b$ is a multiple of n if a and b leave the same remainder when divided by n. There are exactly n different remainders *modulo* n (i.e. remainders which can be obtained when dividing by n): $0, 1, \ldots, n-1$. Now, if we take the integers $m_1, \ldots, m_{n+1}$ as "pigeons" and the remainders modulo n as "pigeons", by the PHP two of the integers must "fly" into the same "pigeon-hole", i.e. must leave the same remainder modulo n.

Hint 5. Consider the numbers $0, m_1, m_1 + m_2, \ldots, m_1 + \cdots + m_n$ and apply the method of the previous problem.

Hint 6. Consider the numbers 3991, 39913991, ..., $3991 \cdots$ (1994 times) $\cdots 3991$ and apply Problem 4.

Hint 7. Consider the numbers $m_1, \ldots, m_{n+1}, m_2 - m_1, \ldots, m_{n+1} - m_1$, which all lie between 1 and $2n$.

Hint 8. Consider the remainders after dividing m, $2m$, ..., nm by n. They are all different, otherwise n divides $im - jm = (i - j)m$ for some different integers $1 \leq i, j \leq n$, hence n must divide $i - j$ which is not possible since $0 < |i - j| < n$. Then one of the remainders must be 1. For otherwise the n "pigeons" $m, 2m, \ldots, nm$ must "fly" into $n - 1$ different "holes" — the remainders $0, 2, \ldots, n - 1$, hence by the PHP two of the numbers must have the same remainder which is not the case.

Hint 9. Yes. Consider the remainders of the powers of 7 modulo 10^8. Follow the idea of Problem 8. Also show that if 10^8 divides $7^k - 7^m$, then 10^8 divides $7^{k-m} - 1$.

Hint 10. Likewise: consider the remainders after division of $m, m^2, \ldots, m^n$ by n.

Solution 11. Let the numbers from the set A be the "pigeons" and the "pigeon-holes" to be the odd positive integers not greater than 200. Now let each number from A "fly" to its largest odd divisor. For instance $24 = 2^3 \cdot 3$ would fly to 3, $31 = 2^0 \cdot 31$ to 31 and $32 = 2^5 \cdot 1$ to 1. There are exactly 100 pigeon-holes hence two numbers a_1 and a_2 from A must fly into the same pigeon-hole, i.e. $a_1 = 2^i b$ and $a_2 = 2^j b$. Now if $i \leq j$ then a_1 divides a_2, otherwise a_2 divides a_1.

Hint 12. Similar.

Hint 13. Let $M = \{n_1, \ldots, n_{75}\}$ and $P = \{1, 2, \ldots, 150\}$. We attach $n_1, \ldots, n_{75}$ to the numbers $1, 2, \ldots, 75$ and $n_1+k, n_2+k, \ldots, n_{75}+k$ to $76, 77, \ldots, 150$. Thus, a positive integer less than 150 is attached to every element of P. Now apply the PHP and complete the proof on your own.

Solution 14. Of course, there is nothing magic in the number 1993; any other one would do. Let us denote by P the set of students and attach to every student the number of fellow-students she or he knows. Every such number is between 0 and 1992. Thus the set H of numbers attached to the students is included in $\{0, 1, 2, \ldots, 1992\}$, therefore has no more than 1993 elements. This is not very helpful, is it? However, if we think a bit more, we shall realize that if a certain student knows nobody (i.e. the number 0 is attached to him) then nobody knows him, hence no student knows all 1992 other students. Thus H cannot contain both 0 and 1992, hence in fact it has *no more than* 1992 different elements. Therefore by PHP, a certain number from H must be attached to at least two students.

Hint 15. Do you see anything in common with the previous problem?

Solution 16. If z is a real number we denote by $\{z\}$ the fractional part of z, i.e. the unique number from the interval $[0, 1)$ such that $z - \{z\}$ is an integer.

Now let us consider the fractional parts of $0x, 1x, 2x, \ldots, nx$. These are $n+1$ numbers in the interval $[0,1)$. We divide this interval into n equal subintervals, each with length $\frac{1}{n}$: $\left[0, \frac{1}{n}\right)$, $\left[\frac{1}{n}, \frac{2}{n}\right)$, ..., $\left[\frac{n-1}{n}, 1\right)$. Then, by the PHP two of the fractional parts $\{kx\}$ will occur in the same subinterval, i.e. for some k_1, k_2 we have $|\{k_1x\} - \{k_2x\}| < \frac{1}{n}$. Let $k_1 > k_2$, $a = k_1x - \{k_1x\}$ and $b = k_2x - \{k_2x\}$. Then a and b are integers and

$$|(k_1 - k_2)x - (a - b)| < \frac{1}{n}.$$

Now we denote $p = a - b$, $q = k_1 - k_2$ and divide by q.

2 Some combinatorial-arithmetic applications of PHP

Before proceeding further we shall take a brief rest from the pigeon-hole principle with a short combinatorial interlude. The following two problems represent important combinatorial facts which will be used further, but are worth knowing anyway.

Problem 17. Prove that every set A with n elements has exactly 2^n subsets.

Solution 17. An easy way to do that is by induction on n:

- if $n = 1$, i.e. $A = \{a\}$, the only two subsets of A are A and the empty set $\emptyset$;
- let the statement be true for any n-element set and let $A = \{a_1, \ldots, a_n, a_{n+1}\}$. Denote $A' = \{a_1, \ldots, a_n\}$. Then every subset of A is either a subset of A' (if it does not contain a_{n+1}) or is obtained by adjoining a_{n+1} to a subset of A'. By the inductive hypothesis there are 2^n subsets of A'. When we adjoin a_{n+1} to each of them we obtain 2^n new and different subsets of A. Thus the total number of subsets of A is $2^n + 2^n = 2 \cdot 2^n = 2^{n+1}$. The induction is completed.

Problem 18. Show that the number of n-tuples $(\alpha_1, \ldots, \alpha_n)$ where each α_i is either 1 or 0, is 2^n.

Solution 18. At first sight the only common thing between this problem and the previous one is the number 2^n. However, a very useful mathematical trick called "one-to-one correspondence" will show us that in fact both problems mean the same. The trick consists in the following: imagine two sets of arbitrary nature and a rule which attaches to *each* element of the first *exactly one* element of the second one, and that *every* element of the second set is attached to *exactly one* element of the first one. If such a rule is given the two sets are said to be in a *one-to-one correspondence*. Now, the important fact here is that *if two finite sets are in a one-to-one correspondence, they have the same number of elements*. Isn't that clear? In fact, if you think a bit you will realize that *counting* the number of elements of a set is nothing else but establishing a one-to-one correspondence between the set and some set of positive integers $\{1, 2, \ldots, n\}$. (Even long ago, before the positive integers were "invented", the shepherds were aware of that mathematical fact: in order to make sure that no sheep from the flock was lost they used to "count" them with a bag of pebbles called "calculi", just by establishing one-to-one correspondence between the pebbles and the sheep.) Noe we shall establish such a correspondence between the set of subsets of $\{1, 2, \ldots, n\}$ and the set of n-tuples of 0's and 1's. Here it is: to any subset A of $\{1, 2, \ldots, n\}$ we attach the n-tuple $(\alpha_1, \ldots, \alpha_n)$ where

$$\alpha_i = \begin{cases} 1 & \text{if } i \in A \\ 0 & \text{otherwise} \end{cases}$$

Check for yourself that this is a one-to-one correspondence. Now it only remains to apply the previous problem.

The next problem is from the *14th International Mathematical Olympiad in Poland, 1972.*

Problem 19. Let M be an arbitrary set of 10 positive integers not greater than 100. Show that there are two disjoint (i.e. having no common elements) subsets of M which have the same sum of their elements.

Let $p_1, \ldots, p_n$ be n fixed different prime numbers and P be the set of all positive integers the prime divisors of which are among $p_1, \ldots, p_n$. It is known from the fundamental theorem of arithmetic that every number $x \in P$ can be represented in a unique way as a product $p_1^{k_1} \cdots p_n^{k_n}$, where the exponents $k_1, \ldots, k_n$ are non-negative integers.

Problem 20. Show that amongst any $2^n + 1$ numbers from P there are two whose product is a perfect square.

Problem 21. Let the set P be defined as above. Show that any set of $n + 1$ numbers from P contains a subset the product of whose elements is a perfect square.

Problem 22. Let the set P be defined as above. Show that amongst any $3 \cdot 2^n + 1$ numbers from P there are four, the product of which is a perfect fourth power of an integer.

The next problem is from the *26th International Mathematical Olympiad in 1985.*

Problem 23. The set M consists of 1985 positive integers none of which has a prime divisor greater than 26. Prove that there are 4 different numbers from M the product of which is a perfect fourth power of an integer.

Hints and Solutions

Solution 19. First, let us notice that if we find *any* two subsets X and Y of M with the same sum of their elements, after removing the common elements from both of them we shall obtain two *disjoint* subsets of M with the same property. Thus the condition for disjointness is not essential.

Noe, the set A has $2^{10} - 1 = 1023$ non-empty subsets (Problem 17). They will be our "pigeons". Each of them we shall put in the "pigeon-hole" marked by the sum of its elements. Note that every such sum is less than $10 \cdot 100 = 1000$. Thus, the number of the pigeon-holes is less than 1000. Now the PHP completes the proof.

Solution 20. Observe that:

1. a product of numbers from P belongs to P too.
2. a number $x = p_1^{k_1} \cdots p_n^{k_n}$ is a perfect square if and only if each of the exponents $k_1, \dots, k_n$ is an even number,
3. therefore a product of the numbers $p_1^{k_1} \cdots p_n^{k_n}$ and $p_1^{m_1} \cdots p_n^{m_n}$ is a perfect square if and only if k_i and m_i have the same parity (i.e. either both are even or both are odd) for each $i = 1, \dots, n$.

Now we are ready to apply the PHP. As "pigeons" we choose the given $2^n + 1$ numbers from P, To each of them $x = p_1^{k_1} \cdots p_n^{k_n}$ we attach an n-tuple $(\alpha_i, \dots, \alpha_n)$ of 0's and 1's as follows: $\alpha_i = 0$ if k_i is even, otherwise $\alpha_i = 1$ (i.e. α_i is the remainder of k_i when divided by 2.) Now, the number $p_1^{k_1} \cdots p_n^{k_n}$ we put in the "pigeon-hole" $(\alpha_i, \dots, \alpha_n)$ thus defined. Due to observation (3) above, if we succeed to show that two numbers will fall in the same "pigeon-hole", we are done. But we already know that the number of different n-tuples of 0's and 1's is 2^n (Problem 18), whereas the numbers are $2^n + 1$. The PHP completes the proof.

Hint 21. Let the chosen numbers be $x_1, \dots, x_{n+1}$. The product of the numbers in any non-empty subset of $\{x_1, \dots, x_{n+1}\}$ belongs to P, hence it is of the kind $p_1^{k_1} \cdots p_n^{k_n}$. Every such a product we attach to an n-tuple $(\alpha_i, \dots, \alpha_n)$ of 0's and 1's as in the previous problem. The number of non-empty subsets of $\{a_1, \dots, a_{n+1}$ is $2^{n+1} - 1$ which is more than the number of n-tuples $(\alpha_i, \dots, \alpha_n)$. hence there are two subsets the products of which are attached to the same n-tuple. We can regard these subsets disjoint (why?). Now, consider the product of these two products.

Solution 22. Since $3 \cdot 2^n + 1 > 2^n + 1$ we can choose two of them the product of which is a perfect square, as stated in Problem 18. For the remaining $3 \cdot 2^n - 1$ we can apply the same argument and find another pair the product of which is a perfect square. This argument can be repeated $\frac{(3 \cdot 2^n + 1) - (2^n - 1)}{2} = 2^n + 1$ times and thus we obtain $2^n + 1$ pairs (x_i, y_i) $(i = 1, \dots, 2^n + 1)$, such that each $x_i y_i$ is a perfect square. Therefore the numbers $\sqrt{x_i y_i}$, $(i = 1, \dots, n)$ are positive integers which belong to P. Now, applying again Problem 18 we obtain that the product of two of them, say $\sqrt{x_i y_i}$ and $\sqrt{x_j y_j}$, is a perfect square. Then $x_i y_i x_j y_j$ is a perfect fourth power.

Solution 23. There are 9 prime numbers not greater than 26: $2, 3, 5, 7, 11, 13, 17, 19, 23$. Since $3 \cdot 2^9 + 1 = 1537 < 1985$ we can apply Problem 22.

3 Finite generalizations of the PHP

To be more precise let us call the principle which we have exploited so far the *"shortage" principle of pigeon-holes* to distinguish from its dual, the *"surplus" principle of pigeon-holes*:

If n pigeons fly into $n + 1$ pigeon-holes then some pigeon-hole will remain empty.

This is as obvious as the "shortage" version, but still can be deduced from it: suppose that each pigeon-hole is occupied and exchange the placed of the pigeons and the pigeon-holes in the original PHP. As a result you will find that some pigeon has simultaneously flown in two different pigeon-holes which is impossible.

Try to formulate the formal version of the "surplus" PHP.

There are various generalizations of the PHP. Here we shall consider some of the most important and useful ones.

PHP':*Let H be a finite set and $P_1, \ldots, P_n$ be subsets of H such that the sum of the numbers of elements of $P_1, \ldots, P_n$ is greater than the number of elements of H. Then at least two of these subsets have a common element.*

For instance, if out of 20 students 13 have passed the test in physics and 8 have passed the test in mathematics, then at least one student has passed both tests.

Problem 24. Prove the original PHP using the above statement.

Solution 24. Let $P = \{p_1, \ldots, p_n\}$ be the set of "pigeons" and the rule attaches every p_i to an element h_i from the set of "pigeon-holes" H. Thus we obtain the singleton sets $\{h_1\}, \ldots, \{h_n\}$ which are subsets of H and the sum of the numbers of their elements is $n \times 1 = n$ which is greater than the number of elements of H. Therefore two of these singleton sets must have an element in common, i.e. two different elements of P must have the same element attached.

Problem 25. Formulate the "surplus" version of the PHP'.

Here is another generalization of the PHP, let us call it *congested pigeon-holes principle.*

If (at least) $kn + 1$ pigeons fly into n pigeon-holes then some pigeon-hole will have to accommodate at least $k + 1$ pigeons.

For instance if 5 pigeon-holes have to accommodate 16 pigeons, then at least 4 pigeons will share the same pigeon-hole. The reason is that if no more than 3 pigeons fly in each pigeon-hole then the total number of pigeons in the 5 pigeon-holes would be no more than

$$5 \text{ pigeon-holes } \times 3 \text{ pigeons } = 15\text{pigeons}.$$

And the formal statement:

CPHP:*Let P and H be finite sets with respectively m and n elements and k be a positive integer such that $m > kn$. Then for any rule which attaches every element of P which are attached to the same element of H.*

Note that when $k = 1$ in CPHP we obtain the original PHP.

Problem 26. 101 points are placed in the plane in such a way that amongst every three of them there are two the distance between which is less than 1. Show that some 51 of these points can be covered by a circle of radius 1.

Problem 27. 606 points are placed in a square with side 1. Show that at least 6 of them can be covered by a circle of radius $1/15$.

Problem 28. Each of nine lines divides a square into two trapeziums the ratio of the areas of which is 19:93. Show that at least three of the lines are concurrent.

Problem 29. Show that amongst any $mn+1$ positive integers there are either $n+1$ equal or $m+1$ pairwise different.

Problem 30. There are 13 blue, 10 red, 8 green and 6 yellow balls in a box. How many of them must be taken out at random so that amongst them there will be:

1. at least 5 of the same colour;
2. at least 3 green;
3. at least one of each colour?

Problem 31. In a school there are 370 students in 17 groups. Show that either there is a group with at least 25 students or there are two groups with at least 24 students each?

Problem 32. Let n be an odd positive integer. Show that amongst any $(n-1)^2+1$ integers n can be chosen whose sum is divisible by n.

Another way to express the CPHP is the following:

Let P and H be finite sets with respectively m and n elements and k be a natural number such that $m > kn$. Then for any rule which attaches to each element of H no more than k elements of P there is an element of P which is not attached to any element of H.

Verify that this statement is equivalent to the CPHP.

Here is a version of a very popular problem (in graph theory) which is solved by a simple application of the CPHP.

Problem 33. In a company of 6 people every two of them either like or hate each other. Show that amongst these 6 people there are either three who like one another, or three who hate one another.

Another formulation of this problem, in terms of graph theory is: every two of six points are connected by a line segment coloured either red or blue; prove that there is either a red triangle or a blue triangle with vertices among these points.

Problem 34. Every two of 17 points are connected by a line segment coloured either in red, blue or green. Prove that there is a triangle with vertices among these points and all sides of the same colour.

Problem 35. State and prove an analogue to the previous problem about 4 colours.

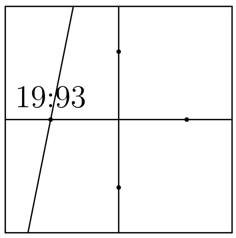
Hints and Solutions

Solution 25. Let H be a finite set and $P_1, \ldots, P_n$ be subsets of H such that the sum of the numbers of elements of $P_1, \ldots, P_n$ is less than the number of elements of H. Then at least one element of H does not belong to any of these subsets.

Solution 26. Let A be one of the points. If all other points lie within a circle C_A of radius 1 and centre A er are done. Suppose that some point B lies outside C_A and consider the circle C_B of radius 1 and centre B. Then each of the remaining 99 points lies in at least one of C_A and C_B (why?). Therefore, by the PHP, one of these circles contains, besides its centre, at least 50 of the given points.

Solution 27. Let us partition the square into 11^2 equal squares with a side $1/11$ each. At least 6 points must be in the same small square since $121 \times 5 = 605$. These points will be covered by the circumcircle of the square, the radius of which is $\frac{1}{11\sqrt{2}} < \frac{1}{15}$.

Solution 28. The area of a trapezium equals the product of the height and the length of the segment connecting the midpoints of the non-parallel sides. Each line cuts a pair of opposite sides of the square and divides it into two trapeziums with the same height, therefore it divides the segment connecting the midpoints of the other two opposite sides of the square in the ratio 19:93. There are exactly 4 points in the square, two on each of the two halving segments. Thus, each of the nine lines passes through one of these 4 points. Therefore at least three of the lines pass through the same point.



Solution 29. Suppose that amongst the given numbers there are no more than n different, i.e. each of them is equal to some of $a_1, \ldots, a_k$ for $k \leq n$. Take k pigeon-holes, mark them by $a_1, \ldots, a_k$ and put all given numbers in their pigeon-holes. Then, by the congested PHP, some of $a_1, \ldots, a_k$ will occur at least $n+1$ times in the given numbers.

Hint 30. The answers are

$$\text{(a) } 4 \times 4 + 1 = 17; \text{(b) } 13 + 10 + 3 + 6 = 32; \text{(c) } 13 + 10 + 8 + 1 = 32.$$

Hint 32. By Problem 29, amongst any $(n-1)^2 + 1$ integers either there are n with different remainders modulo n or there are n with the same remainder modulo n. In either case the sum of those n numbers is divisible by n. (Why?)

Solution 33. Choose any of the people, call him X. He either likes at least three of the other five or hates at least three of them. (Here we applied the PHP.) Suppose the former is the case. If any two of the three liked by X like each other then they, together with X, form a friendly triangle. If not, these three form a hostile triangle.

Solution 34. Choose one of the points. Among the 16 segments with an end at that point there are at least 6 of the same colour, say green. If there is a green segment with ends among these 6 points we are done. If not, apply to them the problem for two colours.

Hint 35. If the line segments between $1 + (1 + 4(17 - 1)) = 66$ points are coloured in one of four colours there will always be a monotone triangle with vertices amongst them.

4 Infinite generalizations of the PHP

Now, try to imagine *infinitely many* pigeons. What will happen if they fly into *finitely many* pigeon-holes. Of course, at least one of the pigeon-holes will have to accommodate infinitely many pigeons. Formally:

Let P be an infinite set and H be a finite set. Then for any rule which attaches every element of P to an element of H there are infinitely many elements of P which are attached to the same element of H.

Problem 36. Prove the infinitary version of the PHP.

Hint 36. Suppose the contrary and apply the CPHP to get a contradiction.

The next problem is from the *17th International Mathematical Olympiad in Bulgaria, 1975*

Problem 37. Let $a_1, a_2, \ldots$ be a strictly increasing sequence of positive integers. Prove that infinitely many terms of the sequence can be represented in the form

$$a_n = xa_p + ya_q$$

where x and y are positive integers and $p \neq q$.

Problem 38. Let k be an arbitrary positive integer. Prove that there exists a prime number p and a strictly increasing sequence of positive integers $a_1, a_2, \ldots$ such that all terms of the sequence $p + ka_1, p + ka_2, \ldots, p + ka_n, \ldots$ are prime numbers.

Problem 39. A man walks with a step 1 along a line on which infinitely many ditches, each d wide, are dug with their centres at a distance $\sqrt{2}$ from each other. Prove that the man will certainly step into some ditch.

Solution 39. Let us "coil" the line on a circumference with length $\sqrt{2}$Then all ditches will coincide. Since $\sqrt{2}$ is an irrational number, no two steps of the man will coincide on the circumference, otherwise $p\sqrt{2} = q \cdot 1$ for some integers p and q. Let N be a positive integer greater than $\frac{\sqrt{2}}{d}$. Let us partition the circumference into N equal arcs each with length $\frac{\sqrt{2}}{N} < d$. After $N + 1$ steps of the man two steps will get into the same arc. let them be the k-th and the m-th steps where $k > m$. Then if we "snap" the positions of the man at the moments m, $m+(k-m)$, $m+2(k-m)$, $\ldots$, we shall see that he is moving on the circumference in the same direction with a step less than $\frac{\sqrt{2}}{N} < d$. Then he will certainly fall into the ditch.

Note that the same proof works whenever the man walks with a step x and the ditches are at a distance y from each other, if x/y is an irrational number.

Problem 40. A coordinate system is fixed in the plane and trees with the same diameter d are planted at the points with integer coordinates. A marksman stands at the origin and shoots in an arbitrary direction. Assuming that the bullet can fly infinitely far, prove that is till certainly hit a tree.

Let X be a set of real numbers. X is said to be *dense on the real line* if for every two real numbers a and b, $a < b$, there is an element x from X such that $a < x < b$. For instance the set of all real numbers is dense: given a and b, take $x = (a+b)/2$. A non-trivial example of a dense set is the set of rational numbers. Try to show that.

The next problem is an important theorem proved by the German mathematician L. Kronecker, which extracts the mathematical essence from the last two problem. Try to prove it in two ways: using Problem 38, and applying the PHP directly.

Problem 41. Let α be an irrational number. Prove that the set A of all real numbers $m\alpha + n$ where m and n are integers is dense on the real line.

Problem 42. Given that $\log_2 10$ is an irrational number, show that for some integer n, the decimal expression of 2^n begins with 1993, i.e. $2^n = 1993\cdots$.

Hints and Solutions

Solution 37. Let m be any integer such that $0 \leq m < a_1$ and A_m be the set of those terms of the sequence a_k which divided by a_1 give a remainder m. Since every term of the sequence gets into one of $A_0, \ldots, A_{a_1-1}$ then at least one of there sets is infinite. Let $A_m = \{a_{k_1}, a_{k_2}, \ldots\}$ where $k_1 < k_2 < \cdots$ and $a_{k_i} = p_i a_i + m$ $(i = 1, 2, \ldots)$. Then $a_{k_i} - a_{k_1} = (p_i - p_1)a_1$ hence for every $i = 1, 2, \ldots$, we have $a_{k_i} = x a_{k_1} + y a_1$ where $x = 1$ and $y = p_i - p_1 > 0$ and $k_1 \neq 1$ since $k_1 > 1$.

Solution 38. Let P be the set of all prime numbers. For every $i = 0, 1, \ldots, k-1$ we denote by P_i the set of primes which divided by k give a remainder i. Every prime gets into one of the sets P_0, P_1, $\ldots$, P_{k-1}. Then one of them is infinite since the set of all prime numbers is infinite. Suppose that P_i is infinite and let p be its least element. Let $x_1, x_2, \ldots$ be the elements of P_i in an increasing order. For every x_n we denote $a_n = \frac{x_n - p}{k}$. Then $a_1, a_2, \ldots$ is a sequence of positive integers (why?) with the desired property.

Solution 40. Let the angle between the direction of shooting and the x-axis be α. The the bullet will pass through all points (x, y) for which $x \geq 0$, and $y = x\tan\alpha$. If $\tan\alpha$ is rational then y will be an integer for some integer x, hence the bullet will hit the tree at (x, y) in the middle. Now let $\beta = \tan\alpha$ be an irrational number. We shall apply the previous problem. Let ditches with length d be dug along the y-axis, at a distance 1 from each other and midpoints at $(0,0), (0,1), (0,2), \ldots$. Let a man walk from the point $(0,0)$ northwards with a step β. Then for some integer x, after x steps the man will step into a ditch, which means that the point $(x, x\beta)$ from the trajectory of the bullet is at a distance less than $d/2$ from a point with integer coordinates, hence the bullet will hit the tree planted at that point.

Solution 41. Let a and b be real numbers, $a < b$, and $N > \frac{1}{b-a}$. Let ud partition the interval $[0,1)$ into N equal subintervals $\left[0, \frac{1}{N}\right), \left[\frac{1}{N}, \frac{2}{N}\right), \ldots, \left[\frac{N-1}{N}, 1\right)$. For every integer m we choose the integer $n = \{m\alpha\} - m\alpha$ where $\{m\alpha\}$ is the fractional part of $m\alpha$. Since $[0,1)$ is partitioned into *finitely many* subintervals, there are two different integers m_1 and m_2, such that $\{m_1\alpha\}$ and $\{m_2\alpha\}$ get into the same subinterval. At that $\{m_1\alpha\}$ and $\{m_2\alpha\}$ are different, otherwise $\alpha = \frac{n_2-n_1}{m_2-m_1}$, where $n_1 = \{m_1\alpha\} - m_1\alpha$ and $n_2 = \{m_2\alpha\} - m_2\alpha$, is a rational number. Let $m_2\alpha + n_2 > m_1\alpha + n_1$. Then $0 < (m_2-m_1)\alpha + (n_2-n_1) < \frac{1}{n} < b-a$. Therefore some of the multiples of $(m_2-m_1)\alpha + (n_2-n_1)$, which too are of the type $m\alpha + n$, will be between a and b (why?).

Hint 42. Show that there are integers m and n such that $\log_{10} 1993 < n \log_{10} 2 - m < \log_{10} 1994$.

5 Geometric (measure-theoretic) generalizations of the PHP

It is intuitively clear that the idea behind the PHP is more general than the "pigeon" version we have used so far. Let us look at the following practical problem which I sometimes face. While I am working on my desk I want to have all books which I am using open and readily available, i.e. not overlapping with each other. The area of the surface of my desk is $2m^2$The area of the surface which each open book occupies on the desk is $0.077m^2$. Now, do you see my problem? *I cannot use more than* 25 *books at a time.* If I open 26 books on my desk the total area which they will occupy is $0.077 \times 26 = 2.002m^2$ which is more than the area of the surface of my desk, so that at least two of them must overlap.

Do you see in that argument anything which resembles the PHP? You should. If not, let us try to explain it again with pigeons. Imagine a cage with a volume of $1ft^3$ and 300 pigeons, each of which occupying a space with a volume $0.0034ft^3$. We have to put all these pigeons in the cage. All this sounds rather cruel, but let us forget for a moment that each pigeon needs some space full of air to breathe. The only concern now is to arrange them somehow in the cage. Still, do we have any chance to succeed? Does it make any sense to try this wild experiment at all? Fortunately not, simply because:

1. no two pigeons can overlap in the space; and
2. the sum of the volumes of the pigeons is $0.0034 \times 300 = 1.02ft^3$, i.e. *more than* the volume of the cage.

As you see, these two situations are very similar. Let us extract the mathematical moral from them. First, what is common between area of a surface and volume of a solid, and, let us add, length of a curve? All these are different examples of *measures of geometric figures*. Loosely speaking, a *measure* is a rule which attaches a non-negative real number to each member of a certain class of sets called *measurable sets*. Each measure determines its own class of measurable sets, of instance the class of space figures (regarded as sets of points in the space) which have a volume, or the class of surfaces which have an area, or the class of curves which have a length, etc. I shall not risk making this notes difficult supplying a general definition of measure and measurable sets; this is a rather advanced branch of mathematics. By measure you may think of any of the above mentioned particular examples of geometric measures. The only fact we need about measures is the following property (we suppose that a measure and hence its class of measurable sets are fixed):

If H is a measurable set and $P_1, \dots, P_n$ are measurable subsets of H such that the sum o f the measures of $P_1, \dots, P_n$ is greater than the measure of H then at least two of these subsets have a non-empty intersection.

In other words:

If H is a measurable set and $P_1, \dots, P_n$ are pairwise disjoint measurable subsets of H (i.e. every two of them have an empty intersection) then the sum of the measures of $P_1, \dots, P_n$ cannot be greater than the measure of H.

We shall call either of these two equivalent statements a *general pigeon-hole principle*.

If we take the class of finite sets and measure each of them by the number of tis elements, the general PHP yields precisely the generalization PHP' of the original PHP.

You should be convinced in the validity of the general PHP for each of the mentioned geometric measures. Here is the general HP for lengths of segments:

If several line segments with a sum of their lengths more than x are placed on a line segment with a length x, then at least two of them have a common point.

And the dual version

If several line segments with a sum of their lengths less than x are placed on a line segment L with a length x then there is a point of L which is not covered by any of the small segments.

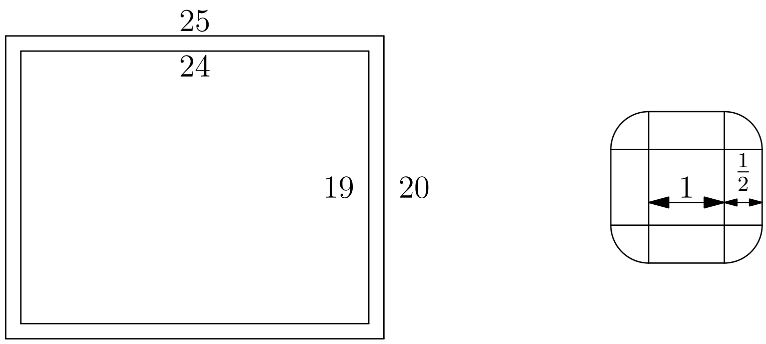
It is a good exercise to formulate the general PHP for areas and for volumes. Another particular case of the general PHP is:

If several arcs with a sum of their lengths more than 2π are placed on a circumference of radius 1, then at least two arcs have a common point.

We continue with a series of problems which employ various versions of the general PHP.

Problem 43. 120 squares with a side 1 are placed on a rectangle with sides 20 and 25. Prove that a circle of diameter 1 can be placed in the rectangle without intersecting any of the squares.

Solution 43. If a circle of diameter 1 lies within a 20×25 rectangle then the centre of the circle lies within a smaller 24×19 rectangle with area 456. Now, if the circle does not overlap with a square with a side 1 then the centre of the circle lies outside a region around the square, with boundary at a distance $1/2$ from the square, as shown. The area of every such region is $S = 1 + 4 \cdot 1 \cdot \frac{1}{2} + \pi \left(\frac{1}{2}\right)^2 = 3 + \frac{\pi}{4} < 3.79$. Then the sum of the areas of the regions around all 120 squares is $120 \times 3.79 = 454.8 < 456$. Therefore there exists a point O in the inner rectangle not covered by any of the regions. The circle with centre O and diameter 1 is what we need.



Problem 44. Show that a circle of radius S/P can be placed inside every convex quadrilateral with perimeter P and area S.

Problem 45. Several chords are drawn in a circle of radius 1. If every diameter cuts at most k chords, show that the sum of lengths of the chords is less than $k\pi$.

Problem 46. A point O is fixed in the plane. Are there 4 circles not covering O such that any ray with a beginning O intersects (has a common point with) at least two of the circles?

Problem 47. $4n$ line segments each of length 1 are placed inside a circle of radius less than n. Show that there is a line parallel to one of the axes in the plane and intersecting at least 2 of the segments.

Problem 48. Several circles with circumferences of total length 10 area placed inside a square with a side 1. Show that there is a line which intersects at least 4 of the circles.

Problem 49 (Blihfeld's Theorem)**.** Prove that is a figure in the plane has an area more than 1, then there are two points from the interior of the figure with coordinates (x_1, y_1) and (x_2, y_2) respectively, such that $x_1 - x_2$ and $y_1 - y_2$ are integers.

Problem 50. The plane is chequered by vertical and horizontal lines at a distance 1 from each other. A figure is given with an area less than 1. Show that the figure can be placed on the plane without covering any of the intersection points.

Hints and Solutions

Solution 44. On each of the sides of the quadrilateral as a base, we draw a rectangle with height S/P expanding into the quadrilateral. The total area of these rectangles is $P \times S/P = S$ and some of them overlap (why?). Therefore there is a point O in Q which is not covered by any of the rectangles. Therefore a circle of centre O and a radius S/P will lie inside the quadrilateral.

Solution 45. Assume that the sum of lengths of the chords is at least $k\pi$. Then the sum of the lengths of the small arcs adjacent to these chords is more than $k\pi$. If we consider those arcs together with the diametrically opposite ones, then the sum of the lengths becomes more than $2k\pi$. Therefore, by the general PHP, there is a point on the circle which belongs to at least $k+1$ arcs. The diameter through that point must cut at least $k+1$ chords, which is not the case. Therefore our assumption is wrong.

Solution 46. No. For every circle not covering O the angle between the two tangents through O is less than $180°$. Then the sum of these angles for any 4 circles not covering O is less than $720°$. Therefore there is a point in the plane covered by at most one circle.

Solution 47. Denote the segments by $s_1, \ldots, s_{4n}$ and the projections of s_i on the x-axis and y-axis by a_i and b_i respectively. Then $a_i + b + i \geq 1$ (why?), hence $(a_1 + \cdots + a_{4n}) + (b_1 + \cdots + b_{4n}) \geq 4n$. Suppose that $(a_1 + \cdots + a_{4n}) \geq (b_1 + \cdots + b_{4n})$. Then $(a_1 + \cdots + a_{4n}) \geq 2n$. The projections of all these segments on the x-axis lie within the projection of the circle, which is a segment with length less than $2n$. Therefore two projections have a common point. The line through that point parallel to the y-axis will intersect the respective segments.

Hint 48. Project all circles on one of the sides of the square and compare the total length of the projections with the length of the side.

Solution 49. Through every point in the plane with integer coordinates we draw lines parallel to the axes. Thus the plane is chequered into squares with side 1 each. Now we shift every such square, by means of a translation $x \mapsto x - m, y \mapsto y - n$, where (m, n) are the coordinates of the bottom left vertex of the square, to the square with vertices (0,0), (0,1), (1,0), (1,1). Thus the parts of Φ lying in the different squares will gather in this square. Since the area of the square is 1 while the sum of areas of all parts of Φ is more than 1, there will be a point (x_0, y_0) covered by two different parts, i.e. two different points from Φ, (x_1, y_1) and (x_2, y_2),a re translated onto (x_0, y_0). Then $x_1 - x_0$ and $y_1 - y_0$ are integers and likewise for $x_2 - x_0$ and $y_2 - y_0$. Therefore $x_1 - x_2$ and $y_1 - y_2$ are integers, too.

Hint 50. Let us place the figure arbitrarily and translate all squares into one as in Problem 49. The sum of the areas of the pieces of the figure is less than the area of the square, hence it will not be covered by the pieces. Let us choose an uncovered point A in the check and, fixing the figure placed as originally, translate the plane in such a way that the translations of all intersection points get into A. Then the figure will not cover any intersection points.

6 Additional problems

We shall finish with a miscellany of problems involving the pigeon-hole principle. They are left without solutions. Some of them are either easy or use ideas from the previous problems, others are really challenging.

Problem 51. 101 points are placed in a square with a side 1. Show that at least 5 of them can be covered by a circle of radius $1/7$.

Problem 52. Φ is a figure on a sphere, which has area greater than half of the area of the sphere. Prove that there are two diametrically opposed points in Φ.

Problem 53. Show that there exists a positive integer divisible by 3991 whose decimal expression ends with 1993.

Problem 54. 50 points are placed in a circle of diameter 12. Show that there are two points at a distance not greater than 2 apart.

Problem 55. 11 points are placed in a ball with volume V. Prove that a portion with volume $V/8$ can be cut from the ball, which does not contain in its interior any of the points.

Problem 56. Prove that amongst any $2m+1$ different integers with absolute values not greater than $2m-1$, there are three whose sum is zero.

Problem 57. Let $k \geq 1$ and $n \geq 1$ be positive integers and A be a set of $(k-1)n+1$ integers each of which is not greater than kn. Prove that some element of A can be represented as a sum of k (not necessarily different) elements of A.

Problem 58. 15 points are placed inside an equilateral triangle with side 15. Prove that three of the points can be covered by a circle of diameter $\sqrt{3}$.

Problem 59. Show that amongst every 11 different real numbers from the interval $[1, 1000]$ two can be chosen, say x and y, which satisfy the inequalities

$$0 < x - y < 1 + 3\sqrt[3]{xy}.$$

Problem 60. Let $x_1, \ldots, x_k$ be real numbers and n a positive integer. Show that there exist integers $p_1, \ldots, p_k$ and q such that $1 \leq q \leq n^k$ and $\left|x_i - \frac{p_i}{q}\right| < \frac{1}{qn}$ for $i = 1, \ldots, k$.

Problem 61 (18-th IMO, Austria, 1976)**.** In the system of p equations for $q = 2p$ unknowns

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1q}x_q &= 0 \\ &\cdots \\ a_{p1}x_1 + a_{p2}x_2 + \cdots + a_{pq}x_q &= 0 \end{aligned}$$

all coefficients are either -1, 0 or 1. Prove that there exists a solution $(x_1, \ldots, x_q)$ of the system, such that:

1. all x_j $(1 \leq j \leq q)$ are integers;

2. at least one x_j $(1 \le j \le q)$ is non-zero;

3. for every j $(1 \le j \le q)$ $|x_j| \le q$.

Problem 62 (28th IMO, Cuba, 1987)**.** Let $x_1, x_2, \ldots, x_n$ be real numbers and $x_1^2 + x_2^2 + \cdots + x_n^2 = 1$. Prove that for every integer $k \ge 2$ there are integers $a_1, a_2, \ldots, a_n$ not all zero, such that $|a_i| \le k-1$, $i = 1, 2, \ldots, n$ and the following inequality holds:

$$|a_1x_1 + a_2x_2 + \cdots + a_nx_n| \le \frac{(k-1)\sqrt{n}}{k^n - 1}.$$

Appendix

It can be shown that in a precise mathematical sense the pigeon-hole principle and the principle of mathematical induction *are equivalent*, and thus, to verify one of them by using the other would be like trying to get out of a quagmire by pulling up your own hair. If we then start trying to verify the PHP directly, very soon we shall realize in desperation that *we cannot do so* because there are *infinitely many* finite sets about which we must check the validity of the PHP. Well, this task is much more difficult than to count the number of hairs on the head of each inhabitant of Pretoria. It is simply *impossible*. Thus, we have no choice but to *believe* or to *assume* that, according to our understanding of the notions of *natural number* and *finite set*, the PHP *must be valid*. Actually, in the same way the mathematicians have proceeded with the principle of mathematical induction.

The South African Committee for the IMO

Professor J H Webb (University of Cape Town) - Chairman
Professor N J H Heideman (Rhodes University)
Dr L R le Riche (University of Stellenbosch)
Professor P Persens (University of Western Cape)
Dr G P West (University of the Witswatersrand)

Regional Organizers

Mr R Anguelov (Vista University-Welkom)
Dr W Bartoszek (University of South Africa)
Professor G L Booth (University of Transkei)
Dr F Bullock (University of South Africa)
Dr M H Burton (Rhodes University)
Dr J de Klerk (University of Potchefstroom)
Professor J C Engelbrecht (University of Pretoria)
Mr D Fish (University of Zululand)
Dr W Greybe (University of Pretoria)
Professor M J Glencross (University of Transkei)
Mr H Glover (University of Port Elizabeth)
Mr D Jacobs (University of Cape Town)
Mr P Jeganathan (University of Transkei)
Professor P J Laridon (University of the Witswatersrand)
Professor D P Laurie (University of Potchefstroom)
Dr J H Meyer (University of the Orange Free State)
Professor J Moori (University of Natal, Pietermaritzburg)
Professor P Pillay (University of Durban-Westville)
Professor J N Ridley (University of the Witswatersrand)
Professor T Schulz (University of the North-West)
Dr L C Slammert (University of the Western Cape)
Dr A J M Snyders (University of Port Elizabeth)

Tutors

Mr M Berman, Mr D Fraser, Mr I A Webb (University of Cape Town)
Ms E Mitove, Mr A Skeen (University of the Witswatersrand)

Administration

Mrs C Sher (University of Cape Town)