# Pigeonhole Principle: Some Applications

PRITHWIJIT DE

The Pigeonhole Principle (PHP) or the Dirichlet Principle is perhaps the easiest theorem that exists in all of Mathematics. It states that if $n+1$ pigeons are put into $n$ pigeonholes, then there is at least one pigeonhole with more than one pigeon. The proof is as easy as the statement. Assume the contrary. Then every pigeonhole has at most one pigeon and therefore the total number of pigeons is at most $n$. A contradiction.

It is natural to ask what is so special about something so trivial. The answer lies in the applications. The goal of this article is to serve up a delectable collection of examples of applications of PHP to the reader.

We start with a couple of simple examples.

(1) *Among any* 13 *people, there are two who have their birthdays in the same month.*

(2) *There are $n$ married couples. How many of the* $2n$ *people must be selected in order to guarantee that one has selected a married couple?*

Think of each couple as a pigeonhole. There are $n$ pigeonholes. If we select $n+1$ people and put each one in the pigeonhole corresponding to the couple to which they belong, then some box will contain two people; that is, we will have selected a married couple. To see that $n+1$ is the smallest, notice that $n$ people may be selected without choosing a married couple, if only the wives or only the husbands are selected.

*Keywords: Pigeonhole principle, pigeons, pigeonhole*

(3) *If $n+1$ integers are chosen from the set $\{1, 2, \ldots, 2n\}$, then there is at least one pair that differ by 1.*

The idea is to pair up the consecutive integers to form $n$ pairs: $(2k-1, 2k)$ for $1 \leq k \leq n$ and in so doing, we reduce it to the previous example.

Try this one:

*If $n+1$ integers are chosen from the set $\{1, 2, \ldots, 2n\}$, then there is at least one pair whose greatest common divisor is 1.*

Interestingly, if $n+1$ integers are chosen from the set $\{1, 2, \ldots, 2n\}$, then there always is a pair $(a, b)$ such that either $a$ divides $b$ or $b$ divides $a$. How do we show this? Note that any positive integer $m$ can be written as $2^k t$ where $k$ is a non-negative integer and $t$ is an odd positive integer. In particular if $m$ is odd, then $k = 0$ and $t = m$. As a consequence of this observation, we see that each positive integer $m \in \{1, 2, 3, \ldots, 2n\}$ can be written as $2^k t$ where $k$ is a non-negative integer and $t \in \{1, 3, 5, \ldots, 2n-1\}$.

Thus there are $n$ different admissible values of $t$. As $n+1$ different numbers are chosen, there are two among these with the same value of $t$, i.e., the same odd part. If the numbers are $a = 2^r t$ and $b = 2^s t$, then it is obvious that $a$ divides $b$ if $r < s$ and $b$ divides $a$ otherwise.

The next example may surprise you:

*Given a set of $m$ integers $\{a_1, a_2, \ldots, a_m\}$, there exists a subset of this set with the property that the sum of all its elements is divisible by $m$.*

As a set with $m$ elements has $2^m$ subsets and therefore $2^m - 1$ nonempty subsets, there are $2^m - 1$ possibilities for the subset we are looking for. For a small enough $m$ (say $m \leq 6$), it is possible to manually enumerate all the possibilities and find the desired subset. But this method fails badly for $m \geq 7$ as the number of possibilities increases rapidly. What do we do then? Let us think a bit. What do we want? We want a subset of $\{a_1, a_2, \ldots, a_m\}$ with the desired property. Consider the $m$ integers

$$a_1, \quad a_1 + a_2, \quad \ldots, \quad a_1 + a_2 + \cdots + a_m.$$

If any one of these sums is divisible by $m$, then we have our required subset. If none is divisible by $m$, then upon division by $m$, each one leaves a remainder that lies between 1 and $m-1$, both inclusive. But then we obtain $m$ remainders, each between 1 and $m-1$. Hence by the PHP, there exist two identical remainders. If the corresponding sums are

$$a_1 + a_2 + \cdots + a_p \quad \text{and} \quad a_1 + a_2 + \cdots + a_q,$$

with $q > p$, then it follows by subtraction that the sum $a_{p+1} + a_{p+2} + \cdots + a_q$ is divisible by $m$; hence the desired subset is $\{a_{p+1}, a_{p+2}, \ldots, a_q\}$. Observe that the proof gives us more than what was asked: it gives us an algorithm to find the subset and quite remarkably the subset consists of some consecutive numbers in the list.

The reader may want to try the following problem. It is similar to the one discussed above.

*Prove that, for any $n+1$ integers $a_1, a_2, \ldots, a_{n+1}$, there exist two among them whose difference is divisible by $n$.*

These were examples based on properties of numbers. Let us look at some examples from Geometry.

*Five points are chosen inside a square of side length 2 units. Prove that there are two points which are at most $\sqrt{2}$ units apart.*

One thing has to be made precise before we tackle this problem. By the phrase 'inside a square' we mean the interior as well as the boundary of the square. Observe that the diagonals are $2\sqrt{2}$ units long and half this length is $\sqrt{2}$. Partition the square into four identical squares by drawing two mutually perpendicular lines through the midpoints of the two pairs of opposite sides. The maximum distance between two points inside each of these squares is $\sqrt{2}$. There are 4 squares and 5 points. By PHP one of these squares will have at least two of the chosen points inside it. Here is a similar problem for the reader:

*Determine an integer m such that if m points are chosen within an equilateral triangle of side length* 1 *unit, there are two which are at most* $1/n$ *units apart.* (Naturally, $m$ will depend on $n$, so we can write $m_n$ or $m = f(n)$ instead of just $m$.)

For some applications, a slightly stronger form of PHP is used.

If $n$ objects are placed in $k$ boxes ,where $n = qk + r$, $q$ and $r$ are positive integers and $0 < r < k$, then at least one box contains more than $q$ objects.

The truth of this statement is entirely obvious and the reader may see at once that setting $q = 1$ gives us the form of PHP that was introduced in the beginning. Let us see some applications of the strong form of PHP.

*Each of the vertices of a regular pentagon is coloured either black or white. Both colours are used. Prove that there are three vertices of the pentagon which receive the same colour, and that these form an isosceles triangle.*

One quickly observes that $n = 5$, $k = 2$ and as $5 = 2 \times 2 + 1$, $q = 2$, and concludes that some three vertices must receive the same colour. Let the vertices be named $A$, $B$, and $C$. Now two cases arise.

**Case 1:** *The vertices A, B and C are adjacent.*

In this case, it is easy to see that $AB = BC$ and hence the triangle $ABC$ is isosceles.

**Case 2:** *The vertices A, B are adjacent and C is opposite AB.*

In this case $CA = CB$, so the triangle is isosceles.

Another colouring problem in the same spirit.

*Consider six points in the plane such that no three of them are collinear. Join every pair of points and colour every edge thus obtained either red or blue. If both colours are used, then there must be a triangle whose sides are either all red or all blue.*

Six points joined pairwise give rise to $\binom{6}{2} = 15$ edges. There are two colours. By the strong form of PHP there are at least 8 edges of one colour. But it doesn't lead anywhere. There exist colourings of 8 edges with the same colour without forming a triangle. In fact one can colour 9 of these 15 edges without obtaining a triangle with sides of one colour. What do we do now? The argument is slightly tricky. Consider one of the six points and name it $A$. It is connected to the remaining five points, $B$, $C$, $D$, $E$ and $F$ (say). Each of these five edges receives one of the two colours: red or blue. Thus $n = 5$ and $k = 2$ giving $q = 2$ and by PHP there are at least three edges among these five with the same colour. Without loss of generality assume that the edges $AB$, $AC$ and $AD$ are coloured red. Look at the triangle $BCD$. If all its edges are blue, we have found a 'blue triangle'. If one of the sides of triangle $BCD$ is red, say $BC$ is red, then we have found a 'red triangle' in $ABC$. The reader may verify that the smallest number of points required to ensure the existence of a monochromatic triangle is six by conjuring up counterexamples for $n = 3, 4, 5$. The next example involves a pyramid.

*The base of a pyramid is a convex polygon with 9 sides. Each of the diagonals of the base and each of the edges on the lateral surface of the pyramid is coloured either black or white. Both colours are used. (Note that the sides of the base are not coloured.) Prove that there is a monochromatic triangle.*

The reader may like to have a go at this. Here is another interesting example. In this example, we use upper case letters to denote vertices of a polygon and lower case letters to denote the numbers associated with the vertices. Thus, $A$ is a vertex and $a$ is the number associated with it, etc.

*Consider a regular polygon with 100 vertices. To each vertex a natural number from the set $\{1, 2, 3, \ldots, 49\}$ is assigned. Prove that there are four vertices A, B, C and D which form a parallelogram ABCD and for which $a + b = c + d$.*

A direct application of the strong form of PHP tells us that there is some number that is assigned to at least 3 vertices. But unfortunately this does not lead us anywhere closer to the solution. The beauty of this problem lies in the following simple geometric fact: **The chord joining a given vertex to its farthest neighbour is a diameter of the circumscribing circle.** But how many such diameters are there? 50. If $PQ$ is a diameter, then $0 \leq |p - q| \leq 48$. Thus the difference can take 49 different values. Therefore by PHP there exist at least two diameters for which the absolute values of the difference between the numbers at the endpoints are same. If $AC$ and $BD$ are two such diameters, then $|a - c| = |b - d|$. Without loss of generality we may assume $a \geq c$ and $d \geq b$ to obtain $a + b = c + d$. Note that $ABCD$ is in fact a rectangle.

We now turn to a problem from Algebra which involves PHP. This was asked in the Regional Mathematical Olympiad 2017 conducted in the Maharashtra and Goa region.

*Let $P(x)$ and $Q(x)$ be polynomials of degree 6 and degree 3 respectively, such that*

$$P(x) > Q(x)^2 + Q(x) + x^2 - 6$$

for all real values of $x$.

*If all the roots of $P(x)$ are real numbers, then prove that there exist two roots of $P(x)$, say $\alpha$, $\beta$, such that $|\alpha - \beta| < 1$.*

It is not clear at the outset how this problem is related to PHP. More often than not, Olympiad problems are like that! One wouldn't know right away what will work and what won't. There is an element of surprise in most of the Olympiad problems and this problem is no exception. Let us see how to solve it. However, please bear in mind that the solution presented here need not be the only possible way of solving the problem.

First observe that if $P(x)$ has two identical roots, then the conclusion is obvious. Therefore assume that the roots are distinct. Next a bit of algebraic manipulation is carried out to write the right hand side of the inequality as

$$(Q(x) + 1/2)^2 + x^2 - (5/2)^2.$$

Now observe that the given inequality holds for all real values of $x$. Thus, in particular, it holds for the roots of $P(x)$. Therefore if $u$ is a root of $P(x)$, then

$$0 = P(u) > (Q(u) + 1/2)^2 + u^2 - (5/2)^2,$$

and it follows that $|u| \leq 5/2$. In other words $u \in [-5/2, 5/2]$. Thus all six roots lie in an interval of length 5 units. That's it. Now its over to PHP. Divide the interval into five subintervals of unit length. These are the pigeonholes and the roots are the pigeons. It is evident that there are two or more pigeons in one hole. Thus there exist roots $\alpha$, $\beta$ of $P(x)$ such that $|\alpha - \beta| < 1$.

We end this article with a gem of an example. This can be 'experimentally verified' by the reader in real life.

*Suppose that* $n^2 + 1$ *people are lined up shoulder to shoulder in a straight line. Then it is always possible to choose* $n + 1$ *people from the line to take one step forward so that going from left to right their heights are either non-decreasing or non-increasing.*

The reader may verify this assertion for small values of $n$ and then build up an argument leading to the proof of the statement.

## References


1. Richard A. Brualdi, *Introductory Combinatorics*
2. Kiril Bankov, *Applications of the Pigeonhole Principle*, Mathematical Gazette, Vol. 79, No. 485, pp. 286 - 292



**PRITHWIJIT DE** is a member of the Mathematical Olympiad Cell at Homi Bhabha Centre for Science Education (HBCSE), TIFR. He loves to read and write popular articles in mathematics as much as he enjoys mathematical problem solving. His other interests include puzzles, cricket, reading and music. He may be contacted at de.prithwijit@gmail.com.