Chapter I. THE PIGEON-HOLE PRINCIPLE

1.1 The Basic Pigeon-Hole Principle

Theorem 1.1. (Pigeon-hole principle)

Let n and k be positive integers such that $n > k$. Suppose we have to place n balls (pigeons) into k boxes (pigeon holes). Then there will be at least one box in which we place at least two balls.

Proof: We pove this theorem in an indirect way (prove by contradiction), that is, we assume that its contrary is true, and deduce a contradiction from that assumption.

Assume that there is no box with at least two balls. Then each of k boxes has either 0 or 1 ball in it. It follows that

$$n = \#\text{ balls} \leq \#\text{ boxes} = k.$$

This contradict to our condition that $n > k$. Therefore, our assumption that there is no box with at least two balls must have been false. □

The "proof by contradiction" method follows from a basic fact in logic as follows.

The statement **A is true** is equivalent to the statement **(Not A) is false.**

We will show that our simple pigeon-hole principle is in fact a **very** power tool.

Example 1.1. There is an element in the sequence

$$3, \ 33, \ 333, \ 3333, \ 33333, \ldots$$

that is a multiple of 2017.

Proof. We prove an even stronger statement is true, in fact, "one of the first 2017 elements of the sequence is divisible by 2017."

Assume that the above fact is not true. Take the first 2017 elements of the sequence and divide each of them by 2017. As none of them is divisible by 2017, they will have a remainder that is at least 1 and at most 2016. As we have 2017 remainders, and only 2016 possible values for the remainders, there are two elements whose have the same remainder. Assume that these are the ith and the jth elements of the sequence $(1 \le i < j \le 2017)$.

We have

$$\begin{array}{rl} 33\ldots333\ldots3 & j \text{ digits} \\ - \qquad\qquad 33\ldots3 & i \text{ digits} \\ \hline \underbrace{33\ldots3}_{j-i}\underbrace{00\ldots0}_{i} & \end{array}$$

This means that

$$(\underbrace{33\ldots3}_{j-i}) \times 10^i \text{ is divisible by } 2017$$

However, 10^i and 2017 are relatively prime. It follows that $\underbrace{33\ldots3}_{j-i}$ must be a multiple of 2017. This contradicts to our assumption that all the first 2017 elements of the sequence are not divisible by 2017 ($1 \leq j-i \leq 2016$). $\square$

In this example:

"**Boxes**" are "**1, 2, ..., 2016**"

"**Balls**" are "**The remainders**".

Example 1.2* Let a be an irrational number (a real number that can not be written in the form p/q, where p, q are integers). Prove that there exist (infinitely many) rational numbers $r = \frac{p}{q}$ (p, q are integers) such that

$$|a - r| < \frac{1}{q^2}$$

Proof. Let Q be a positive integer. Without loss of generality, we can assume that $a > 0$. Consider the "factional part" $\{0\}, \{a\}, \{2a\}, \ldots, \{Qa\}$ of the first $(Q+1)$ multiples of a.

By the pigeon-hole principle, there are two of these in one of Q semi-open intervals:

$$[0, \tfrac{1}{Q}), [\tfrac{1}{Q}, \tfrac{2}{Q}), \ldots, [\tfrac{Q-1}{Q}, 1)$$

In other words, there are integers s, q_1, and q_2 such that:

$$\{q_1 a\} \text{ and } \{q_2 a\} \in [\tfrac{s}{Q}, \tfrac{s+1}{Q})$$

Taking $q = |q_1 - q_2|$ we obtain some integer p

$$|qa - p| < 1/Q.$$

Dividing by q, we get

$$\left| a - \frac{p}{q} \right| < \frac{1}{Qq} \leq \frac{1}{q^2} \quad (\text{Since } 0 < q \leq Q).$$

To prove that the number of such pairs (p,q) is infinite we need extra argument as follows.

Assume on the contrary that only for a finite numbers of $r_i = \frac{p_i}{q_i}$, $i = 1, 2, \ldots, N$, $|a - r_i| < \frac{1}{q_i^2}$.

Since non of the differences is exactly 0, there exists an integer Q such that $|a - r_i| > \frac{1}{Q}$ for all $i = 1, \ldots, N$. Apply our above argument for this Q to produce $r_0 = \frac{p_0}{q_0}$ such that

$$|a - r_0| < \frac{1}{Qq_0} \leq \frac{1}{Q}.$$

Hence r_0 can not be one of the r_i $(i = 1, \ldots, N)$.
On the other hand, as before, $|a - r_0| < \frac{1}{q_0^2}$

contradicting the assumption that the fractions r_i $(i=1,N)$ were all the factions with this property. $\square$

Example 1.3. A chess tournament has n participants, and any two players play exactly one game against each other. Then it is true that in any point of the time, there are two players who have finished the same number of games.

Proof. First we could think that the pigeon-hole principle will not be applicable here as the number of players ("balls") is n, and the number of possibilities for the number of games finished by any one of them ("boxes") is also n. Indeed, a player could finish either no games, or one game, or two games, ... and so on, up to and including $n-1$ games.

However, if there is a player who plays no games, then there are no players who played against all other ones, i.e. who play $n-1$ games. (and vice versa). It means that the number of possibilities of the number of games is always at most $n-1$. In particular, this set is either $\{0,1,\ldots,n-2\}$ or $\{1,\ldots,n-1\}$.

By pigeon-hole principle, there are two players who has finished the same number of games. $\square$

Example 1.4. 181 distinct points have been chosen on a circle, all with integer number of degrees. Prove that there are at least one pair of points that are 180° apart.

Proof. Denote our set of point by A.
Let B be the set of "antipoles" of the points in A, i.e. B consists of points that are 180° apart from some point in A.
Assume that set A contains no pair of antipoles. This assumption leaves $A \cap B = \phi$. This means that 199 (distinct) points in B are at 179 integer-degree positions outside A.
By pigeon-hole principle, there are two points of B at the same position, a contradiction. $\square$

Example 1.5. Given any sequence of 3000 integers (positive or negative, not necessarily all different), some consecutive subsequence has the property that the sum of the members of the subsequence is a multiple of 3000.

Proof. Assume that our sequence is

$$a_1, a_2, \dots, a_{3000}.$$

Denote by

$$S_n := a_1 + a_2 + \cdots + a_n$$

for $n = 1, 2, \dots, 3000$.
If there exist a S_n divisible by 3000, then

we are done. Otherwise, $s_1, \dots, s_{3000}$ are not divisible by 3000. Similar to Example 1.1, there are two s_i and s_j that have the same remainder when divided by 3000 $(1 \le i < j \le 3000)$.

This implies that

$$s_j - s_i = a_{i+1} + a_{i+2} + \cdots + a_j$$

is divisible by 3000.

$\square$

Example 1.6. In every polyhedron there is at least one pair of faces with the same number of sides.

Proof. Let N be the greatest number of sides in a face of a given polyhedron. Then there are N adjacent faces each having the number of sides between 3 and N. By pigeon-hole principle, there are two such faces having the same number of sides.

$\square$

Example 1.7. Let n be an odd number, and $a_1, \dots, a_n$ be a "permutation" of $\{1, 2, \dots, n\}$ (i.e. $\{a_1, \dots, a_n\} = \{1, 2, \dots, n\}$). Prove that the product

$$\prod_{i=1}^{n} (a_i - i) \quad \text{is even.}$$

Hint. Consider odd's a_i and odd j. There are $n+1$ of them (since n is odd). ← balls

Boxes = n terms $(a_i - i)$.

1.2 The Generalized Pigeon-Hole Principle

It is easy to generalized the pigeon-hole principle to the following theorem:

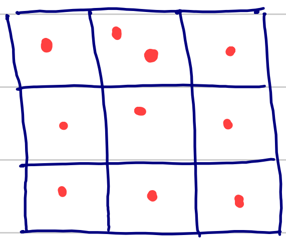
Theorem 1.2. (Pigeon-hole principle, general version)

Let n, m and r be positive integers so that $n > rm$. Let us put n balls into m boxes. Then there will be at least one box into which we place at least $r+1$ balls.

Proof.
Similar to the proof of Theorem 1.1, we assome the contrary statement that each box can hold at most r balls. Thus then all m boxes can hold at most $rm < n$ balls, which contradicts to the requirement that we distribute n balls. □

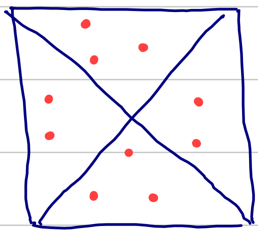
Example 1.8. Ten points are given within a unit square. Then we can always find two of them that are closer to each other than 0.48, and we can find three which can be cover by a circle of radius 0.5.

Proof.
Let us split the unit square into 9 small squres of side $\frac{1}{3}$ as in the picture.



As we have 10 points given inside the 9 small squares there will be at least one small square containing two of the 10 points. The distance between these two points is at most the length of the diagonal of the small square (of side $\frac{1}{3}$), which equals $\frac{\sqrt{2}}{3} < 0.48$, so the first part follows.

To prove the second statement, we divide the unit square into four parts by the diagonals



Theorem 1.2 implies that atleast one part contains atleast 3 of the 10 points. As the radius of the circumcirle of these parts (isoceles right triangles) is 0.5, the second part follows. $\square$

Example 1.9. (Quantum Magazine)

(a) Given 9 points in a triangle (inside or on the edges) of area 1. Prove that there are three points forming a triangle of area less than or equal $\frac{1}{4}$.

(b) Prove the same fact for only 7 points.

Hint. (a) Divide the triangle into four small ones whose areas equal $\frac{1}{4}$.

(b) "Any triangle inside a parallelogram of area S has area at most $S/2$"

Example 1.10. (Quantum Magazine)

The **Integer Lattice** is the collection of all points on the plane with both coordinates integer. The latter points are called the "**integer points**".

Given a "convex pentagon" whose vertices are all integer points. Prove that it contains at least one more integer point.

Is the statement is still true if our pentagon is concave?

Hint / Defn: If M_1 and M_2 are in a convex shape P, then any points of the line segment M_1M_2 are also in P.

Remark A more formal statement of the pigeonhole principle is:

" For two finite set A and B, there exists a 1-1 correspondence $f: A \to B$ if and only if $|A| = |B|$."

Example 1.11. Among six people, there are always three who know each other or three who are complete strangers.

Is the statement still true for only 5 people?

Example 1.12. Consider the Fibonacci sequence defined by

$$a_1 = a_2 = 1, \quad a_{n+1} = a_n + a_{n-1}, \quad n > 1.$$

Prove that, for any n, there is a fibonacci number ending with n zeros.

Example 1.13. The positive integers $1, 2, \ldots, 101$ are written down in any order. Prove that you can cross 90 of these number, such that a motonic sequence remains.

Proof. L_m = the length of the longest monotonically increasing with the last element m.

R_m = the length of the longest monotonically decreasing seq. starting with m.

claim: $L_m \neq L_K$ or $R_m \neq R_K$. for $m \neq K$.

So all pairs (L_m, R_m) are distinct.

Assume that no such monotonic subseq of length 11 exists. $L_m, R_m \in \{1, 2, \ldots, 10\}$. $\Rightarrow$ we have 100 boxes, and 101 balls. $\Rightarrow \exists\, m \neq K$ s.t.

$$(L_m, R_m) = (L_K, R_K) \quad >< .$$

$\square$