



Abel Prize Laureate 2012 Endre Szemerédi

Arithmetic progressions

An arithmetic progression is a sequence of natural numbers of constant difference. The set of natural numbers contains a lot of arithmetic progressions, but if we consider smaller proper subsets of the natural numbers, it is not obvious that there are any arithmetic progressions. A general problem is to decide how small a randomly chosen subset can be, and still contain arithmetic progressions.

In primary school children learn by heart the two times table, $2, 4, 6, \dots$, the three times table, $3, 6, 9, 12, 15, \dots$, and so on. The mathematical term for a times table is *arithmetic progression*. The concept of arithmetic progression also includes finite sequences like $10, 13, 16, 19$. This is an arithmetic progression of length 4, of (constant) difference 3 and initial value 10. The two times table is an arithmetic progression of infinite length, of difference 2 and initial value 2. An arithmetic progression is determined by its length, difference and initial value. If you are asked to write up the arithmetic progression of length 5, difference 7 and initial value 23, you should answer $23, 30, 37, 44, 51$. Notice that $4, 5, 6, 7, 8, 9$ is also an arithmetic progression, the difference in this case is 1.

The Dutch mathematician **Pierre Joseph Henry Baudet** formulated in 1921 the following conjecture: *If one divides the natural numbers $1, 2, 3, \dots$ ad infinitum into a random number of boxes, then there is nevertheless always at least one box which contains an arithmetic progression of arbitrary length.* Baudet died shortly after, at the age of 30. The conjecture was proved in 1927 by another Dutch mathematician, **Bartel Leendert van der Waerden**.

A strengthening of van der Waerden's result was conjectured by **Pál Erdős** and **Pál Turán** in 1936. They believed that the reason for the existence of

Colouring

By colouring the natural numbers using only two colours, say red and blue, it is easy to see that 9 consequent numbers e.g. $1, 2, \dots, 9$, are needed to ensure that there is an arithmetic progression of length 3. Why? Let us try to prove the opposite, so suppose the sequence $1, 2, \dots, 9$ does not contain arithmetic progressions of length 3. For this reason, number 1, 5 and 9 cannot be all equally coloured. So assume first that 1 and 5 are red, and 9 is blue. Since 1 and 5 are red, 3 has to be blue. But 9 is also blue, so 6 has to be red. Now 5 and 6 are red, forcing 4 and 7 to be blue. Number 8 has to be red since 7 and 9 are blue, and number 2 has to be red since 3 and 4 are blue. But then 2, 5 and 8 are all red, which is a contradiction. The case where 1 and 9 are red and 5 is blue is treated similarly. On the other hand, the sequence of length 8, given by **R****B****R****B**-**B****R****B****R** has no arithmetic progressions of length 3. Thus 9 is a sharp bound for this property. This means that the so-called *van der Waerden number* $W(2, 3) = 9$.



arithmetic progressions is that some colour occupies a set of natural numbers of strict positive upper density. For a subset A of the natural numbers, the *upper density* is defined as follows: For each natural number N we intersect the set A with the set $\{1, 2, \dots, N\}$, count the number of integers in the intersection and divide by N . This rational number between 0 and 1 measures the size of A compared to all integers between 1 and N . We do this for increasing numbers N . If the fraction for huge N never exceeds a certain number, we say that this number is an upper bound for the fractions. The smallest possible upper bound for huge N is called the *upper density* for A .

Notice that there is also a concept of *lower density*, given as the greatest lower bound for the fraction, when N is a huge number.

Upper density

Example 1. Let A be the set of even numbers. For a given N the set of even numbers between 1 and N has cardinality $N/2$ if N is an even number and $(N-1)/2$ if N is odd. Thus the fraction we are looking for is $1/2$, which is the upper density for the even numbers.

Example 2. Now let A be some finite set, say all natural numbers from 1 to 100 . For N less than 100 , the fraction is 1 , but for N greater than 100 the fraction will decrease and eventually tend to zero. So the upper density is 0 .

Example 3. We consider the set of powers of 10 , $A = \{10, 100, 1000, \dots\}$. If we compare this set to the set $\{1, 2, 3, \dots, 10^k\}$ for some natural number k , it is easy to see that the fraction will be $k/10^k$. As k grows, this fraction will tend to 0 , and again the upper density is 0 .

In 1953, **Klaus Friedrich Roth** proved that any subset of the integers with positive upper density contains an arithmetic progression of length 3 . In 1969, **Endre Szemerédi** proved that the subset must contain an arithmetic progression of length 4 , and then in 1975 proved that any subset with positive upper density must contain arithmetic progressions of arbitrary length, known as Szemerédi's theorem.

Erdős formulated in 1973 a stronger version of the Erdős-Turán conjecture: *Let A be a subset of the natural numbers such that the sum of their reciprocals exceeds any natural number. Then A must have arithmetic progressions of arbitrary length.* Erdős offered a prize of US\$ 3000 for a proof of this conjecture at the time. The problem is currently worth US\$ 5000. One can prove that the sets of natural numbers of positive upper density necessarily have divergent reciprocal sums. Thus Erdős' conjecture implies Szemerédi's theorem. It is also known that the set of primes have divergent reciprocals, first proved by **Leonhard Euler** in 1737. The theorem of **Ben Green** and **Terence Tao** from 2004 about the existence of arithmetic progressions of arbitrary length in the set of prime numbers is a special case of this conjecture.

The results concerning existence of arithmetic progressions are based on the interplay between size, randomness and structure. The bigger the sets are, the more likely it is that they have arithmetic progressions. Szemerédi states that positive upper density is a sufficient condition. For even smaller sets, of zero density, some additional structure is needed. The set of prime numbers has zero density, but Green and Tao show that this set nevertheless has some structural similarities to the natural numbers, enough to prove the existence of arithmetic progressions, based on Szemerédi's theorem.