I know said this in an earlier lecture, but I think I forgot to put it in the notes, and it’s worth putting down anyway.

There are two important notions in a graph, which appear in many applications in computer science and data science: the notion of a maximum induced clique (that is, the largest number of vertices all of which are connected to each other) and the maximum induced independent set (that is, the largest number of vertices none of which are connected to each other.

Figuring out how large the maximum clique or the maximum independent set is in a graph is a well-known problem, with applications ranging from social networks (cliques of friends) to biology; these two problems are very related via the complement graph. A clique in the graph is an independent set in the complement.

There is a clear intuition that in a graph with many vertices, either you ought to have a large clique, or a large independent set. This is more or less what Ramsey graph theory (with 2 colors) is investigating, and here is the connection.

Take any graph $G$ on $n$ vertices, and represent it as a subset of the complete graph with the same number of vertices, by coloring the edges present in $G$ with Red and the ones that are not present in $G$ with Blue. Existence of an all-Red $K_k$ signifies presence of that clique in $G$, while existence of an all-Blue $K_l$ signifies presence in $G$ of an $l$-large independent set.

In other words, what Ramsey theory establishes is that for any $k, l \geq 2$, a graph with sufficiently many vertices either has a $k$-clique or an $l$-independent set. Of course, $k$ and $l$ are then lower bounds for the maximum clique and respectively independent sets that actually might be present.

Suprisingly, Ramsey theory has a continuous version, which involved all the points in the plane (or space). It also has applications in arithmetic and sequences! And some of the results we have shown are suboptimal. Here are a few examples.

**Example 1.** Suppose that the plane is colored with two colors. There exists a unit segment with both endpoints of the same color.

The proof is trivial: take an equilateral triangle with length 1, by Pigeonhole, two of the three points have to have the same color.

**Example 2.** The above remains true if we allow the plane to be colored with 3 colors.

We construct the proof in two steps: first, we show that if this weren’t true, then all line segments of length $\sqrt{3}$ have same-color endpoints; then we take a circle of length $\sqrt{3}$ around a point, and show there are two points on it at distance 1.

Pick three points in the plane which form the vertices of an equilateral triangle. Either two of the vertices are colored with the same color (in which case we’re done), or all three vertices have different colors. Now reflect one of the points $A$ across the line formed by the other two $BC$; this fourth point $D$ also forms an equilateral triangle of side 1 with $B$ and $C$. Thus, if it has $B$’s or $C$’s color, we’re done. Else it has the same color as $A$. The distance $AD = \sqrt{3}$.

Similarly, we show that every point in the circle of radius $\sqrt{3}$ centered at $A$ must have the same color as $A$, or else we’ve shown there is a line segment of length 1 with both endpoints of the same
But a circle of radius $\sqrt{3}$ has plenty of points at distance 1 on it. So there must be two points of the same color as $A$ at the same distance on the circle.

**Sums of members.** Suppose that we color each positive integers with one of $k$ colors. Show that there exists some $N(k)$ such that for all $n \geq N(k)$, there exist 3 integers $a, b, c$ of the same color so that $a + b = c$ (we could have $a = b$).

Suppose we place the integers as labels on a $K_n$, and we color the edge $(i, j)$ with the color assigned to the integer $|i - j|$. By letting $n$ be larger than $R(3, 3, \ldots, 3)$ (with $k$ 3s), we know that the coloring of $K_n$ will have a monochromatic triangle formed by $i, j, k$ (assume wlog $i < j < k$). Then by letting $a = j - i$, $b = k - j$, $c = k - i$, we have the statement of the problem.

**Probabilistic Method.** To understand more fully what this method is about, we need to do a very brief recap of discrete probability. We will do this next time, but for now it suffices to outline the (main) central idea.

**Definition** Given some finite set of objects called a *sample space* $\Omega$, and a property $A$ compatible with the objects in $\Omega$, suppose we want to figure out whether some object in $\Omega$ has property $A$. Pick a *random* object $X$ in $A$ (here we will think of randomness mostly as being uniform, that is, every object will get the same probability of being picked; this is unnecessary, as all we will really need is that the distribution we choose over $\Omega$ assigns to each object some non-zero probability). If we can establish that the probability that $X$ has property $A$ is *strictly positive*, it follows that some object in $\Omega$ has property $A$.

More, next time.