THE BASEL PROBLEM

TIM SMITS

The computation of the value of the infinite series $\sum_{n=1}^{\infty} \frac{1}{n^2}$ was an open problem that gathered interest around 1644. It was attacked by many of the top mathematicians at the time, but none of them were able to conjure up a solution. In 1735, a young mathematican by the name of Euler gave the first solution, making one of his first (of many) important contributions to mathematics. The goal of this handout is to give a solution to the Basel problem using linear algebra.

Definition 0.1. A **metric space** is a set X with a function $d : X \times X \to \mathbb{R}$ called a metric (or distance function) that has the following properties:

(i) $d(x, x) \geq 0$ for all $x \in X$.
(ii) $d(x, y) = d(y, x)$ for all $x, y \in X$.
(iii) $d(x, y) = 0 \iff x = y$.
(iv) (Triangle inequality) $d(x, z) \leq d(x, y) + d(y, z)$ for any $x, y, z \in X$.

When we speak of a metric space, we usually refer to the pair (X, d).

The idea of a metric space is that you have a way of measuring distances between two elements of the set X. Some familiar examples of a metric space are $\mathbb{R}$ with the usual distance between points given by the absolute value $|\cdot|$, and $\mathbb{R}^n$ with the norm $\|\cdot\|$. In a real analysis course, one learns that metric spaces are the natural settings to work in if you want to generalize the ideas of calculus.

Definition 0.2. For a metric space (X, d), a **sequence** is a function $s : \mathbb{N} \to X$. We usually denote the values $s(n)$ as x_n. A sequence **converges** to a point $x \in X$ if for all $\varepsilon > 0$, there is an integer N such that $n \geq N$ means $d(x_n, x) < \varepsilon$. If x_n converges to x, we write $\lim_{n\to\infty} x_n = x$, or $x_n \to x$.

Definition 0.3. A sequence x_n is called **Cauchy** if for any $\varepsilon > 0$, there is an integer N such that if $n, m \geq N$, we have $d(x_n, x_m) < \varepsilon$.

The idea of a Cauchy sequence is that it's a sequence where the terms eventually become arbitrarily close to each other. An important property of convergent sequences is the following:

Proposition 1. *Let (X, d) be a metric space, and suppose that x_n converges. Then x_n is Cauchy.*

One of the important observation that one makes when studying sequences is that the converse is not true: not every Cauchy sequence converges! This is illustrated by the following example:

Example 0.4. Let $X = \mathbb{Q}$ equipped with the metric $d(x, y) = |x - y|$. Define $x_n = (1 + \frac{1}{n})^n$. It turns out that x_n is Cauchy, but x_n does not converge in $\mathbb{Q}$: from calculus, you know that $\lim_{n\to\infty} x_n = e$, but e is not rational.

The metric space $\mathbb{Q}$ has "too many holes" for calculus to work. The way to fix this is to require Cauchy sequences converge.

Definition 0.5. A metric space (X, d) is called **complete** if every Cauchy sequence in X converges to a point in X.

In a first real analysis class, one learns that $(\mathbb{R}, |\cdot|)$ is a complete metric space. This is the key property that allows calculus to be developed. We now apply these ideas in the context of linear algebra.

Proposition 2. *Let $(V, \langle -, - \rangle)$ be an inner product space, and let $\|\cdot\|$ be the norm induced by this inner product. Then with $d(x, y) = \|x - y\|$, (V, d) is a metric space.*

Hopefully this isn't surprising. An inner product on V give us way of measuring lengths in V, which gives us a way of measuring distances!

Definition 0.6. A **Hilbert space** H is an inner product space such that the metric $d(x, y) = \|x - y\|$ makes (H, d) a complete metric space.

Hilbert spaces are the basic object of study in functional analysis (which depending who you ask, is nothing more than linear algebra on infinite dimensional vector spaces). These are vector spaces where we can do calculus. As one might suspect, the most basic example of a Hilbert space is $\mathbb{R}^n$ where the inner product is the usual dot product of vectors. Any finite dimensional inner product space is a Hilbert space, but this requires a bit of work.

Definition 0.7. An orthonormal subset S of H is called **complete** if for $x \in H$, $\langle x, v \rangle = 0$ for all $v \in S$ means that $x = 0$.

We are now ready to state the theorem which we will use for our approach.

Theorem 0.8 (Parseval)**.** *Let H be a Hilbert space, and let $S = \{x_1, x_2, \ldots\}$ be a complete orthonormal subset of H. Then for all $y \in H$, we have $\|y\|^2 = \sum_{n=1}^{\infty} |\langle y, x_i \rangle|^2$.*

Our goal is to use Parseval's theorem to compute the value of $\sum_{n=1}^{\infty} \frac{1}{n^2}$. In order to do this, we need to find a Hilbert space to work in.

The function space $C([-\pi, \pi], \mathbb{C})$ can be given the structure of an inner product space, via $\langle f, g \rangle = \int_{-\pi}^{\pi} f(t)\overline{g(t)}\, dt$. This inner product induces a norm on $C([-\pi, \pi], \mathbb{C})$ defined by $\|f\|_2 = \sqrt{\frac{1}{2\pi} \int_{-\pi}^{\pi} |f(t)|^2\, dt}$. In a real analysis class, one learns that $C([-\pi, \pi], \mathbb{C})$ is *not* a complete metric space with respect to the metric induced by this norm. The way to fix this is to instead look at the larger function space $L^2([-\pi, \pi]) = \{f : [-\pi, \pi] \to \mathbb{C} : \|f\|_2 < \infty\}$. We have the following non-trivial facts:

(i) $L^2([-\pi, \pi])$ *is* complete with respect to the metric induced by $\|f\|_2$, making it a Hilbert space.

(ii) The set $S = \{e^{inx} : n \in \mathbb{Z}\}$ is a complete orthonormal set in $L^2([-\pi, \pi])$.

These two facts form the basis of *Fourier analysis*, which is where our approach draws inspiration from.

Theorem 0.9. $\sum_{n=1}^{\infty} \frac{1}{n^2} = \frac{\pi^2}{6}$.

Proof. Consider $f : [-\pi, \pi] \to \mathbb{C}$ defined by $f(x) = x$. Then $f \in L^2([-\pi, \pi])$. We have $\|f\|_2^2 = \frac{1}{2\pi}\int_{-\pi}^{\pi} x^2\,dx = \frac{\pi^2}{3}$. For $n \neq 0$, $\langle f, e^{inx}\rangle$ is given by $\frac{1}{2\pi}\int_{-\pi}^{\pi} te^{-int}\,dt$. Integrate by parts with $u = t$ and $dv = e^{-int}$ to get $\frac{1}{2\pi}\int_{-\pi}^{\pi} te^{-int}\,dt = \frac{-t}{2\pi in}e^{-int}|_{-\pi}^{\pi} + \frac{1}{in}\int_{-\pi}^{\pi} e^{-int}\,dt = \frac{(-1)^{n+1}}{in}$, so that $|\langle f, e^{inx}\rangle|^2 = \frac{1}{n^2}$. Applying Parseval's theorem says that $\sum_{n=-\infty}^{\infty} \frac{1}{n^2} = 2\sum_{n=1}^{\infty} \frac{1}{n^2} = \frac{\pi^2}{3}$, so that $\sum_{n=1}^{\infty} \frac{1}{n^2} = \frac{\pi^2}{6}$. $\square$

This approach is far from Euler's original solution (which abused notions of convergence of infinite products). However, it should illustrate to you that the ideas of linear algebra are extremely important in the study of function spaces. One benefit of our approach above is that it can be generalized extremely easily. For example, taking $f(x) = x^2$, one can compute that $\sum_{n=1}^{\infty} \frac{1}{n^4} = \frac{\pi^4}{90}$, and in general, you can compute $\sum_{n=1}^{\infty} \frac{1}{n^{2k}}$ for any integer $k \geq 0$. However, whether or not the sum $\sum_{n=1}^{\infty} \frac{1}{n^3}$ has a closed form is still an open question. In fact, we didn't know the value of this sum was irrational until 1978! Why can't the above method be adapted? See if you can find what breaks!