18.103. Lecture 1, Wed, Sept 4, 2013: Introduction.

One of the main goals this course is to establish rules for the limiting behavior of functions so that we can deal with functions with as much confidence as we do real or complex numbers. Today we give a preview, without any proofs.

Part 2. Fourier Analysis (starting about week 7). A complex-valued, periodic function $f(x)$ of period 2π is represented by the **Fourier series**

$$f(x) = \sum_{n=-\infty}^{\infty} c_n e^{inx}, \quad c_n \in \mathbb{C} \tag{1}$$

The family

$$e^{inx}, \quad n \in \mathbb{Z} = \{0, \pm 1, \pm 2, \dots\}$$

should be viewed as a basis for the infinite-dimensional vector space of periodic functions. The remarkably simple and practical formula of Fourier (from the early 1800s) for the coefficients c_n is

$$c_n = \frac{1}{2\pi} \int_{-\pi}^{\pi} f(x) e^{-inx} dx \tag{2}$$

Formula (2) giving the correspondence $f \mapsto c_n$ is the companion to Formula (1) which computes $c_n \mapsto f$.

Consider the partial sum

$$S_N(x) = \sum_{|n| \le N} c_n e^{inx}$$

The main question in the *second half* of the course is whether and in what sense

$$\lim_{N\to\infty} S_N = f \ ?$$

The cleanest answer (among many) is that convergence works in the following average sense. Assume that

$$\sum_{n=-\infty}^{\infty} |c_n|^2 < \infty$$

then

$$\lim_{N\to\infty} \int_{-\pi}^{\pi} |f(x) - S_N(x)|^2 \, dx = 0$$

This conclusion is based on remarkably simple formulas based on the Pythagorean theorem. Define the length of a function in terms of its Fourier coefficients as

$$\|f\| = \sqrt{\cdots + |c_{-1}|^2 + |c_0|^2 + |c_1|^2 + \cdots}$$

This gives a definition of the length of f because the correspondence given by (2) and (1) between f and the sequence of coefficients c_n is one-to-one. This length can also be expressed directly in terms of f as follows.

Theorem (Parseval's formula) $\displaystyle\sum_{n\in\mathbb{Z}} |c_n|^2 = \frac{1}{2\pi}\int_{-\pi}^{\pi} |f(x)|^2 dx$

Because the series has a finite sum, its tail $\displaystyle\sum_{|n|>N} |c_n|^2$ tends to zero. Hence

$$f - S_N = \sum_{|n|>N} c_n e^{inx} \implies \|f - S_N\|^2 = \sum_{|n|>N} |c_n|^2 \longrightarrow 0 \quad \text{as } N \to \infty$$

In other words, by Parseval's formula, the square mean distance from f to S_N tends to zero:

$$\int_{-\pi}^{\pi} |f(x) - S_N(x)|^2 dx = 2\pi \sum_{|n|>N} |c_n|^2 \longrightarrow 0, \quad \text{as } N \to \infty$$

Part 1. Lebesgue measure and integrals. The task in the *first half* of the course is to introduce *Lebesgue measure* and establish properties of the *Lebesgue integral.* Our textbook (Adams and Guillemin) introduces Lebesgue measure using motivation and examples from probability theory. An equally important motivation (that will only become clear in the second half) is that the systematic study of Fourier series requires the Lebesgue integral. The square mean convergence of Fourier series and Parseval's formula cannot be stated accurately in proper generality without the Lebesgue integral and Lebesgue integrable functions $f(x)$.

Probability theory. A *Bernoulli sequence* is an infinite sequence of outcomes

$$HTTHTTTH\cdots$$

of coin tosses with H representing heads and T representing tails. Assume that heads and tails are equally likely and that the tosses are independent of each other. Then the 8-letter initial sequence displayed has probability 2^{-8} as does each of the 2^8 possible words of length 8 in the letters H and T. The first paradox we face is that the probability of any single infinite string is zero (the limit of 2^{-n} as $n \to \infty$), whereas if we add up all possibilities we get

$$\sum 0 = 1 \quad ??$$

It turns out that this paradox is a real contradiction that has to be avoided. The collection of outcomes is uncountably infinite. (This is proved by the Cantor diagonal argument and is the subject of our first homework exercise.) We will have to give up on sums of probabilities with uncountably many terms.

Given that some operations are illegal, our job will be to figure out what operations are legal and give meaningful probability values. Fortunately there are many meaningful questions we can answer. For $n \geq 0$, consider

$$S_n = \text{number of heads minus number of tails in first } n \text{ tosses}$$

The trajectory of S_n is known as a random walk (in one dimension, that is, on the integers $\mathbb{Z}$). I graphed this for the example $HTTHTTTH$: $S_0 = 0$, $S_1 = 1$, $S_2 = 0$, $S_3 = -1$, etc. A statement referring to full, infinite strings can only make sense using a correct formulation of probability on the uncountable collection of Bernoulli sequences. Here is an example that does make sense and has a coherent answer.

Strong Law of Large Numbers.

$$\text{With probability 1,} \quad \lim_{n\to\infty} \frac{S_n}{n} = 0$$

Combining probability and Fourier analysis. After we have developed probability theory on Bernoulli sequences, using a correspondence with Lebesgue measure on the unit interval, we will discuss the Lebesgue integral and some Fourier analysis. Then we will use some Fourier analysis to prove more theorems in probability. By the end of the semester we will have all the tools to discuss the continuum limit of a (suitably scaled) random walk, namely *Brownian motion.*

The first rigorous formulation of Brownian motion was given by Norbert Wiener (in the math department at MIT in the 1920s) before probability theory itself was on a fully rigorous footing! Brownian motion starting from $B(0) = 0$ is given on $0 \le t \le \pi$ by

$$B(t) = a_0 t + \sum_{k=1}^{\infty} a_k \frac{\sin kt}{k}$$

where the a_k, $k = 1, 2, \dots$ are independent, standard (mean 0, variance 1) normal random variables. (The coefficient a_0 is also a normal random variable, but with a different variance, which we will figure out later. In lecture, for simplicity, I omitted a_0 initially. Without the $a_0 t$ term, $B(t)$ returns to 0 at $t = \pi$ and is known as a Brownian bridge.)

Wiener showed that with probability 1, $B(t)$ is continuous but not differentiable. Curiously, with the help of Fourier series representations, even non-differentiable functions can be differentiated. The catch is that the answer is not a function. The derivative

$$dB/dt = \sum_{k=0}^{\infty} a_k \cos kt$$

is known as "white noise." With probability 1 the series on the right is not convergent. Although with probability 1 this series does not represent a function (much less a continuous function), it nevertheless has a good interpretation as a so-called "generalized function." You have already seen one such generalized function in 18.03, namely the delta function. We will discuss generalized functions much later in the course.

In Lecture 2, we will begin the systematic theory of Lebesgue measure, starting with the case of the unit interval, in tandem with probability theory for Bernoulli sequences, following the Adams-Guillemin text.