A Radical Approach to Real Analysis - David Bressoud Micronotes

Nafis Faisal

March 2025

1 Crisis in Mathematics: Fourier's Series

- Fourier proposed his infinite trigonometric series solution to the steady-state heat propagation problem on a rectangular plate in December 1807:

$$\frac{\partial^2 z}{\partial x^2} + \frac{\partial^2 z}{\partial w^2} = 0 \tag{1}$$

1.1 Background to the Problem

- Fourier found his series by solving for $z(x, w)$ on the $xw-$plane where $z(x, 0) = f(x) = 1$ for $-1 \leq x \leq 1$.
- The solution was found to be,

$$z(x, w) = \frac{4}{\pi} \sum_{i=1}^{\infty} \frac{(-1)^{n-1}}{2n-1} e^{-(2n-1)\pi w/2} cos(\frac{(2n-1)\pi x}{2}) \tag{2}$$

.

- Such infinite trigonometric sums were proposed before by Daniel Bernoulli in 1753 as solutions to the problem of modeling the vibrating string. It was dismissed by Euler citing convergence issues.
- Fourier's solution was also dismissed on the same grounds by Laplace, Lagrange, Lacroix and Monge, summarized by Poisson.
- It was shown that for specific values, e.g $z(0, 0)$, the series converged.

1.2 Difficulties to the Problem

- For Fourier's solution, $z(x + 2, w) = -z(x, w)$, hence it is periodic with period 4, unexpected behavior by the understanding of that time.
- Integrating and differentiating finite sums of functions was allowable without problems.
- Integrating an infinite power series $\sum_i a_i x^i$ term-by-term was dangerous. It was only known then that it could be done within the region of convergence and that doing so might lose convergence at endpoints at most.
- Differentiating term-by-term was even more dangerous. Doing it to Fourier's series led to a series that only converged at even integers although the graph of $z(x, w)$ suggested that the derivative should be zero on intervals and undefined at endpoints of the intervals periodically.

Fourier's infinite trigonometric series solution was another in a long list of things that forced the mathematicians of the time to discard their intuitive understanding of infinite series and develop a more rigorous foundation of infinite series and calculus that could answer the questions posed with newer material.

2 Infinite Summations

- In an infinite series, the symbols $+$ and $\sum$ no longer mean the same thing as they do in a finite series.

2.1 The Archimedean Understanding

- Archimedes introduced the method of exhaustion as a method of finding the area underneath a curve by summing the areas of smaller and smaller inscribed triangles.The name is a misnomer as pointed out by E.J Dijksterhuis since the area is never exhausted, rather only gotten arbitrarily close to.
- Archimedes' argument was to bound the successive partial sums with an upper limit M and a lower limit L and to argue that all partial sums after a certain point remain in the interval (L, M).
- Archimedes' argument also involved setting a target value $T \in (L, M)$ as the limit of the partial sums and arguing that the limit could be neither greater nor lower than T.
- In short, the Archimedean understanding of an infinite series is that, given any open interval (L, M) that contains T, all the partial sums are inside this interval once they have enough terms and if such a value T can be found, it is the value of the infinite series. Uniqueness of T is proved by the Archimedean principle.
- By the early 1800s, it was becoming clearer that the intuitive understanding held by the mathematicians of the time of what was and was not legitimate regarding infinite series was insufficient, as exemplified by the behavior of Fourier's series.
- Infinite series behaved differently from their finite counterparts as they were neither associative nor commutative in general.
- Regrouping and rearrangement of series like,

  $$1 - 1 + 1 - 1 + 1...$$

  and

  $$1 - \frac{1}{2} + \frac{1}{3} - \frac{1}{4} + ...$$

  led to different limiting values of the partial sums.

2.2 Geometric Series

- By the 14th century, scholars in Oxford and Paris like Richard Swineshead and Nicole Oresme were assigning values to infinite series occuring in study of motion. They started with geometric series,

  $$1 + x + x^2 + x^3 + ... = \frac{1}{1-x} \tag{3}$$

- Although letting x equal any value beyond 1 and -1 led to non-sensical answers, Euler accepted such series and used the formula above as a definition for their sums.
- Euler made his view clear in his *"On divergent series"* by stating that the word sum of a series only has meaning for convergent series and the idea of a sum should be given up for divergent series. Thus, for any infinite summation, we need to stretch our definition of sum.
- When an infinite series has a target value in the Archimedean sense, it is called a convergent series.
- Cauchy introduced the error term of a geometric series in 1821 by analyzing finite geometric series sums

  $$1 + x + x^2 + ... + x^{n-1} = \frac{1}{1-x} - \frac{x^n}{1-x} \tag{4}$$

  where the error between the calculated value of summing upto n terms and the target value of $T = \frac{1}{1-x}$ is the remainder term $\frac{x^n}{1-x}$. The series seems to converge only for $|x| < 1$ as the remainder term can be made arbitrarily small by taking more terms in the sum in this case.
- Cauchy showed that equation 3 was valid for $|x| < 1$, a condition that ordinary equalities do not contain. Thus, the equation 3 is about successive approximations where $+$ no longer means the same. This made the Archimedean understanding essential for understanding infinite series.

2.3 Calculating π

- Beginning in the Middle Ages, mathematicians had begun plunging into the infinite, to resurface with treasures. The true power of calculus lied with coupling it to infinite processes. But the intuition of the early explorers regarding the infinite, that is when they could be treated like the finite, provided poor foundation.

- Much initial impetus for using infinite came from the search for a better approximation of π.

- Known to Nilakantha of Kerala, James Gregory, Isaac Newton and Gottfried Leibniz, the archtangent ($tan^{-1}x$) series for $x = 0$ provided a series to calculate π, the same series obtained for $z(0, 0)$ in Fourier's trigonometric series.

- Using Machin's identity on the arctangent series led to a faster converging series for π.

- John Wallis considered integrals of the form

$$\int_0^1 (1 - t^{1/p})^q dt \tag{5}$$

where $p = q = \frac{1}{2}$ led to the area of the unit circle in the first quadrant. Wallis discovered bounds for $\sqrt{\pi/2}$ valid for $n \geq 2$:

$$\frac{2.4.6...(2n-2)\sqrt{2n}}{3.5.7...(2n-1)} > \sqrt{\frac{\pi}{2}} > \frac{2.4.6...(2n-2)(2n)}{3.5.7...(2n-1)\sqrt{2n+1}}, \tag{6}$$

from which we get,

$$\frac{\pi}{2} = \frac{2}{1}.\frac{2}{3}.\frac{4}{3}.\frac{4}{5}.\frac{6}{5}.\frac{6}{7}.\frac{8}{7}... \tag{7}$$

- In 1665 after reading Wallis' account of his discovery, Isaac Newton discovered the binomial series for any real exponent and $|x| < 1$ while trying to generalize

$$\int_0^1 (1 - t^2)^{m/2} dt \tag{8}$$

from even values of m to odd values of m.

- Using the binomial series expansion to integrate term-by-term equation 8 for $m = 1$, another series for $\pi/4$ was obtained by Newton. He also considered the integral and integrated its binomial series term-by-term

$$\int_0^{1/4} (\sqrt{x - x^2}) dx = \int_0^{1/4} x^{1/2}(1 - x)^{1/2} dx = \frac{\pi}{24} - \frac{\sqrt{3}}{32} \text{ (obtained geometrically)} \tag{9}$$

to obtain a series for π.

- Modern calculations of π use faster converging series like the one published by S. Ramanujan in 1915.

2.4 Logarithms and Harmonic Series

- Integrating equation 3 term-by-term after replacing x by $-x$ leads to the following series expansion valid for $|x| < 1$,

$$ln(1 + x) = x - \frac{x^2}{2} + \frac{x^3}{3} - \frac{x^4}{4} + ... = \sum_{i=1}^{\infty} (-1)^{n-1} \frac{x^i}{i} \tag{10}$$

Although discovered independently by Isaac Newton and Nicolaus Mercator in 1667, the latter was the one to first publish it. The series is also valid for $x = 1$ giving the value of the alternating harmonic series as $ln\, 2 = \sum_{i=1}^{\infty} (-1)^{i-1} \frac{1}{i}$.

- The harmonic series $\sum_{i=1}^{\infty} \frac{1}{i}$ does not converge in the Archimedean sense. It increases unbounded, that is, given any number M, we can pick a number n such that the partial sum of n terms of the harmonic series exceeds M. A series that diverges in this way is said to diverge to infinity.

- A graphical anaysis shows that the partial sum of the harmonic series upto $n-1$ terms exceeds the area underneath the curve $1/x$ from 1 to n or $ln\, n$. The difference between the two, that is, the missing area, approaches a constant value discovered by Euler in 1734, named Euler's constant, γ:

$$\gamma = \lim_{n\to\infty} [(\sum_{i=1}^{n-1} \frac{1}{i}) - ln\, n] \tag{11}$$

- Estimating Euler's constant using the two sequences,

$$x_n = [(\sum_{i=1}^{n-1} \frac{1}{i}) - ln\, n] \tag{12}$$

$$y_n = [(\sum_{i=1}^{n} \frac{1}{i}) - ln\, n] = x_n + \frac{1}{n} \tag{13}$$

we find that the second sequence is decreasing and the first is increasing. Therefore, $x_1 < x_2 < ... < x_n < \gamma < y_n < ... < y_2 < y_1$. As $y_n - x_n = \frac{1}{n}$, they can be brought arbitrarily close. But it is a fundamental question to ask whether any number exists as both sequences narrow in. It was found that this could not be proved, leading to a foundational statement about the real numbers, the **nested interval principle**. Therefore, a foundational assumption was found on which rigorous treatment of calculus can be built. To convert the NIP to a theorem required an equivalent foundational assumption or axiom to be made about the real numbers.

- The **nested interval principle** of real numbers states that for an increasing sequence $\{x_n\}$ and a decreasing sequence $\{y_n\}$ such that, $x_n < y_n$ where their difference can be made arbitarily small, there exists exactly one real number that is greater than or equal to each x_n and less than or equal to each y_n.

2.5 Taylor Series

- Discovery, investigation and utilization of infinite series exploded across the 18th century. In 1748, Euler published his *"Introductio in analysin infinitorum"* as a primer on infinite series as preparation for calculus.

- By the end of 17th century, power series $\sum\limits_{i=0}^{\infty} a_i x^i$ had emerged as one of the primary tools of calculus. Euler solved the differential equation for the vibrations of a circular drumhead,

$$\frac{d^2u}{dr^2} + \frac{1}{r}\frac{du}{dr} + (\alpha^2 - \frac{\beta^2}{r^2})u = 0 \tag{14}$$

using the series $u = r^\lambda \sum\limits_{i=0}^{\infty} a_i r^i$.

- Although many mathematicians of the time like Leibniz, de Moivre, Jean and Johann Bernoulli, Newton knew how to generate the coefficients of the power series given a function, the series is named after the first one to put the method into print in 1715, Brook Taylor. Taylor's expansion formula for $f(x)$ about $x = a$ is given by **Taylor's Theorem**:

$$f(x) = \sum_{i=0}^{\infty} \frac{f^{(i)(a)}}{i!}(x-a)^i \tag{15}$$

provided that all the derivatives exist at $x = a$. The special case for $a = 0$ is known as Maclaurin series. We cannot yet prove that equation 15 is convergent to $f(x)$ in the Archimedean sense.

- d'Alembert was one of the first to study the convergence of infinite series in his paper in 1768, where he studied the binomial expansion

$$\sqrt{1 + \frac{200}{199}} = 1 + (1/2)\frac{200}{199} + \frac{(1/2)(1/2-1)}{2!}(\frac{200}{199})^2 + ... \tag{16}$$

and found that although it started converging to the actual value very quickly, it started diverging from the actual value when the number of terms taken exceeded 300. Comparing the terms with the geometric series, he found that the ratio of two successive terms was greater than 1 when the number of terms taken $n > 300$. Values of divergent series were of important, especially in astronomy.

- In 1813 in the revised edition of *"Theorie des fonctions analytiques"*, Lagrange gave the means to estimate the difference between partial sums of the Taylor series expanstion of a function and the actual functional value as stated in **Lagrange's Remainder Theorem**:

$$D_n(a,x) = f(x) - [\sum_{i=0}^{n-1} \frac{f^{(i)}(a)}{i!}(x-a)^i] = \frac{f^{(n)}(c)}{n!}(x-a)^n,\ c \in (a,x) \tag{17}$$

.

- Using Lagrange's Remainder Theorem and **Stirling's Formula**:

$$\lim_{n\to\infty} \frac{n!}{(n/e)^n\sqrt{2\pi n}} = 1 \tag{18}$$

we can prove that exponential, sine and cosine functions have Taylor expansions valid for any real value of x. For the binomial expansion studied by d'Alembert in equation 16, d'Alembert's findings are confirmed.

- The true significance of Lagrange's Remainder Theorem lies in its precise definition of the difference between a function and the Taylor polynomial approximating it. It acts as a step towards the Archimedean understanding of power series.

2.6 Emerging Doubts

- Newton's fluxions - velocities and rates of change, Leibniz's differentials both proved to be weak foundations for calculus as they both used infinity in a cavalier fashion without defining it properly as pointed out by George Berkeley in his *"The Analyst"* in 1734.

- Even the Eulerian way of taking the definition of divergent sums to be given by equation 3 was problematic as different algebraic manipulations of the same series $1 - 1 + 1 - ...$ would return different sums. Thus, the same diverging series can have different values depending upon the context in which it arises.

- In 1747, d'Alembert published the differential equation for a vibrating string of length l:

$$\frac{\partial^2 y}{\partial x^2} = \frac{1}{c^2}\frac{\partial^2 y}{\partial t^2} \tag{19}$$

which can be solved for all time t given an initial function $y(x) = y(x,0)$ with boundary conditions $y(0,t) = y(l,t) = 0$. Like heat propagation, if the initial position was linear combination, $\sum_n a_n sin(n\pi x/l)$, then the general solution could be $\sum_n a_n sin(n\pi x/l)cos(n\pi ct/l)$. Daniel Bernoulli suggested that the string would be able to accommodate infinitely many harmonics as an initial state, and so any initial state could be written as $\sum_n^\infty a_n sin(n\pi x/l)$. But Euler rejected this possibility by stating that the infinite sum initial state is periodic with period $2l$ and could not handle initial states that are not periodic functions of x. Such a rejection was valid to Euler and his contemporaries because they understood a function as an expression borne from polynomials, exponential, logarithm, trignometric functions – valid for all values of x it could be defined for. Even piecewise defined functions were assumed to be many different functions joined together.

- Charles Babbage, John Heschel, George Peacock would champion Lagrange's viewpoint that a function was defined at all values of x from its value and the values of all its derivatives at some point $x = a$ through its Taylor expansion. Functions with continuous derivatives upto order p are said to belong to the class of functions C^p. If all derivative exist, it belongs to C^∞. But all C^∞ functions do not have a unique power series representation about any point, as shown by Cauchy in 1821 through the function

$$f(x) = \begin{cases} 0,\ x = 0 \\ e^{-1/x^2},\ \text{elsewhere} \end{cases}$$

All derivatives of $f(x)$ at $x = 0$ is 0 therefore the power series expansion of this function about $x = 0$ is the same as that of the function 0. It was later understood that a subset of C^∞ could be written as power series about any point, which would be called **analytic functions**.

- Fourier asserted in 1807 that both Daniel Bernoulli and Euler were right. That any initial position could be written as a linear combination of trigonometric functions and that it would be periodic, except outside the $(0, l)$ interval it was defined in. That is, the description of the function within $(0, l)$ could tell us nothing about the function outside it. Fourier's assertion was another piece in the longstanding controversy in mathematics of that time. Assuming there had to be a flaw, Lagrange and others who reviewed Fourier's script assumed that the series would not converge. In succeeeding years, Fourier showed this was not true, forcing a critical revolution of the meaning of function, infinite series, derivative and the notion of infinity. It was Cauchy who realized that the only true foundation was a return to the Archimedean understanding.

3 Differentiability and Continuity

- The goal of this chapter is to compress roughly fifty years of struggle to understand differentiability and continuity, and to prove Lagrange Remainder theorem.

- Modern interpretations of continuity and differentiability were developing from the early 1800s. Fr. Bernhard Bolzano (1781-1848) gave a somewhat modern definition although did not have much impact, Gauss had insight into this, as seen in his notebooks in 1814, but did not publish. Augustin Louis Cauchy (1789-1857) is usually credited for the current interpretation along with the books he wrote for his courses in 1820s, especially *"Cours d'analyse de l'Ecole Royale Polytechnique"*. The modern standards of rigor came into analysis by 1850s and 1860s due to the efforts of Karl Weierstrass (1815-1897) and Bernhard Riemann (1826-1866).

3.1 Differentiability

- Lagrange Remainder theorem will be proved by induction. So the base case for $n = 1$ needs to be proven, which is usually known as **Mean Value Theorem**.

  Theorem 1 (**Mean Value Theorem**). *Given a function f that is differentiable $\forall x \in (a, b)$ and continuous on $[a, b]$, there exists a real number $c \in (a, b)$ such that,*

$$\frac{f(b) - f(a)}{b - a} = f'(c)$$

- To define the limits using the Archimedean principle:

  Definition 1 (**Archimedean definition of a Derivative**). *The derivative of a function f at a point a, denoted by $f'(a)$ is a value such that we can force*

$$L < \frac{f(x) - f(a)}{x - a} < M$$

  by taking x sufficiently close to but not equal to $f(a)$ for all $L < f'(a)$ and $M > f'(a)$.

- Cauchy introduced the following definition in 1823 by noting that L and M can be replaced by $f'(a) - \epsilon$ and $f'(a) + \epsilon$ respectively without any loss of generality:

  Definition 2 (**Cauchy's definition of a Derivative**). *The derivative of a function f at a point a, denoted by $f'(a)$ is a value such that for any real number $\epsilon > 0$, there exists a real $\delta > 0$ such that,*

$$0 < |x - a| < \delta \implies |E(x, a)| = |f'(a) - \frac{f(x) - f(a)}{x - a}| < \epsilon$$

  If such a value does not exist, f is not differentiable at a.

- The $\epsilon - \delta$ definition of the derivative is the result of much struggle and proves to work in those gray areas where differentiation does not seem to work as expected.

- This definition of the derivative is not of much use if it cannot tell us something we did not know before. So we look at the question of differentiating infinite series of functions. For a finite series of n differentiable functions $\{f_i\}$, the error can be split up among each of the functions. So if the bound required to keep the error term $|E_i(x, a)| = |f_i'(a) - \frac{f_i(x) - f_i(a)}{x - a}| < \epsilon_i = \frac{\epsilon}{n}$ is δ_i, then taking $\delta = min(\delta_1, \delta_2, ..., \delta_n)$ ensures that all of the errors remain within bound. Thus, the total error $|E(x, a)| = \sum |E_i(x, a)| < \epsilon$ and so any finite sum of differentiable functions is itself differentiable. But for an infinite series, the sum of the errors might not converge to the chosen ϵ for any choice of δ.

- Using this definition, it can be proven that at $x = 0$ the derivative of $f(x) = x^2 sin(x^{-2}), f(0) = 0$ is 0 and all order derivatives of $g(x) = e^{-1/x^2}, g(0) = 0$ is 0. Thus, the Taylor series of $g(x)$ at $x = 0$ is exactly equal to that of the constant function $h(x) = 0$ since all the coefficients are 0.

3.2 Cauchy and the Mean Value Theorem

- Cauchy set out to prove the Lagrange's form of the Taylor theorem but first needed to prove the mean value theorem. The proof of the mean value theorem seen in most textbooks is by Ossian Bonnet (1819-1892) anf first published in Joseph Alfred Serret's calculus text of 1868 *"Cours de Calcul Differentiel et Integral"*.

- Cauchy's first proof of the theorem had some problems. It invoked the intermediate value property, assumed the same δ to work for all points in the interval and also assumed that the derivative was bounded, which would be true if it was continuous, but the continuity of the derivative was assumed later in the proof when it was not needed.

  Definition 3 (**Intermediate Value Property**). *A function f is said to have the intermediate value property on the interval $[a, b]$ if for all $x_1, x_2 \in [a, b]$ such that for some N,*

  $$f(x_1) < N < f(x_2), \exists c \in (x_1, x_2) \text{ such that, } f(c) = N$$

- Cauchy's second proof of the theorem actually proved the much more powerful Generalized Mean Value theorem and proved Theorem 1 as a special case ($F(x) = x$). It appeared in an appendix to *"Resume des Lecons"*. But this proof assumed that a function with a non-negative derivative in an interval is monotonically increasing in that interval. This was not proven then and the proof required a deeper understanding of continuity.

  Theorem 2 (**Generalized Mean Value Theorem**). *If f and F are both continuous at every point in $[a, b]$ and differentiable at every point in the open interval (a, b) and if $F' \neq 0$ in this interval, then*

  $$\frac{f(b) - f(a)}{F(b) - F(a)} = \frac{f'(c)}{F'(c)}$$

  for at least one point $c \in (a, b)$.

3.3 Continuity

- Continuity of a function was such an intuitive geometric concept that any rigorization of it did not appear until 1817 when Bernhard Bolzano first published the modern definition of continuity in the Proceedings of the Prague Scientific Society under a title that translates to *"Purely analytic proof of the theorem that between any two values that yield results of opposite sign there will be at least one real root of the equation"*. He was trying to prove that any continuous function has the intermediate value property (Definition 3).

- It seems strange that although Bolzano was trying to prove that continuity implies the intermediate value property (IVP), he did not use this property to define continuity. It would be in line with the geometric intuition about continuity that any function having it was an unbroken curve achieving all the values between any two value. There are several reasons for this. Firstly, a function that had the IVP was not necessarily bounded (e.g. $f(x) = x^{-1}sin(1/x), f(0) = 0$). Secondly, the sum of two functions having IVP. $g(x) = cos^2(1/x)$, $h(x) = sin^2(1/x)$, $g(0) = h(0) = 0$ could not have it as $g(x) + h(x) = 1$, $g(0) + h(0) = 0$. Furthermore, continuity would also mean that if x was close to a then $f(x)$ would be close to $f(a)$. But, $f(x) = sin(1/x)$, $f(0) = 0$ did not have this property at $x = 0$ although it satisfies IVP no matter how $f(0)$ is defined as long as $|f(0)| \leq 1$. This was damning.

- For defining continuity of a function at a point a, we want to be able to keep $f(x)$ as close to $f(a)$ as possible by keeping x close to a. This led to the definition of continuity as stated by Bolzano in 1817 and by Cauchy in 1821 in his *"Cours d'analyse"*. Cauchy used the Archimedean understanding.

  Definition 4 (**Continuity**). *A function f is continuous at a point a if for any real $\epsilon > 0$, there exists a real $\delta > 0$, such that,*

  $$0 < |x - a| < \delta \implies |f(x) - f(a)| < \epsilon$$

  If a function is continuous at every point in an interval I, then it is continuous on I.

- This definition does it all. It proves IVP, boundedness of continuous functions and achievement of those bounds and also that sum, product or composition of continuous functions are continuous.

3.4 Consequences of Continuity

-

3.5 Consequences of the Mean Value Theorem