41. Shortcomings of Riemann integration

Here we address some deficiencies of Riemann's approach to integration in relation to the requirement of Riemann integrability in Fundamental Theorem of Calculus and boundedness in the setup of the integral.

41.1 Volterra's example.

Recall that the statement of Theorem 40.7 asserting that "the integral of a derivative yields the function" required that the derivative F' be Riemann integrable. As it turns out, this assumption is not at all vacuous:

Theorem 41.1 (Volterra's example) *There exists a function $F\colon \mathbb{R}\to\mathbb{R}$ such that F is differentiable with f' bounded yet not Riemann integrable on $[0,1]$.*

The proof is based on the notion of a "fat Cantor set" which is a generalization of Cantor's ternary set such that it has non-zero length. (Per a wiki page, Volterra seems to have introduced such sets a few years before Cantor, but its first mention is due to Smith who did it even before Volterra.).

Pick a sequence $\{\alpha_n\}_{n\in\mathbb{N}}\in(0,1)^{\mathbb{N}}$. Then set

$$\begin{aligned} C_0 &:= [0,1] \\ C_1 &:= [0,1]\smallsetminus\bigl(\tfrac12(1-\alpha_1),\tfrac12(1+\alpha_1)\bigr) \end{aligned} \tag{41.1}$$

and, generally, assuming that C_n is the union of 2^n disjoint closed intervals of length

$$\ell_n := \frac{1}{2^n}\prod_{i=1}^n(1-\alpha_i) \tag{41.2}$$

let C_{n+1} be the set obtained by removing the open interval of length $\frac12\alpha_{n+1}\ell_n$ in the middle of each of these intervals. Each such interval thus splits into 2 intervals of length

$$\frac12\bigl(\ell_n-\alpha_{n+1}\ell_n\bigr)=\frac12(1-\alpha_{n+1})\ell_n=\ell_{n+1} \tag{41.3}$$

allowing the recursion to proceed. Denote

$$C:=\bigcap_{n\in\mathbb{N}}C_n \tag{41.4}$$

We then have:

Lemma 41.2 ("Fat" Cantor set) *For all $\{\alpha_n\}_{n\in\mathbb{N}}\in(0,1)^{\mathbb{N}}$, the set C is non-empty, compact without isolated points (and thus perfect) and also nowhere dense with $\mathbb{R}\smallsetminus C$ dense in C. Moreover,*

$$\forall\{\alpha_n\}_{n\in\mathbb{N}}\in(0,1)^{\mathbb{N}}\colon\ \sum_{i=1}^\infty\alpha_i<\infty\ \Leftrightarrow\ C\text{ is of non-zero length} \tag{41.5}$$

Proof. The sets $\{C_n\}_{n\in\mathbb{N}}$ are bounded and closed and thus compact. Being also nested, the Cantor Intersection Property ensures that $C\neq\varnothing$ with C compact. Every $x\in C$ is the limit of the endpoints of the intervals in C_n's containing. These endpoints are contained in all C_n's and thus in C and since at least one of them is always different from C, the point x is a limit of a sequence of points from $C\smallsetminus\{x\}$, proving that x is not isolated.

Similarly, every $x \in C$ is within distance at most 2^{-n} of a point in $\mathbb{R} \smallsetminus C_n$ and so x is also a limit point of a sequence from $\mathbb{R} \smallsetminus C$, proving that $\mathbb{R} \smallsetminus C$ is dense in C. The set C_n contains open intervals of length at most 2^{-n} and so C does not contain any open intervals at all. Since C is closed, it is nowhere dense.

Finally, the length of C is the limit of length of C_n's. (Making this precise requires some facts from measure theory.) The length of C_n is easy to compute:

$$\text{length}(C_n) = 2^n \ell_n = \prod_{i=1}^{n} (1 - \alpha_i) \tag{41.6}$$

It is a simple exercise to show that the product tends to a non-zero number if and only if $\{\alpha_i\}_{i \geqslant 1}$ is never equal to one and is summable. $\square$

With the fat Cantor set in place, we use it to define a function for which the conclusion will hold. Consider first $h \colon \mathbb{R} \to \mathbb{R}$ defined by

$$h(x) := \begin{cases} (1 - 4x^2)^2 & \text{if } |x| \leqslant 1/2 \\ 0 & \text{if } |x| > 1/2 \end{cases} \tag{41.7}$$

and note that h is differentiable on $\mathbb{R}$ with $h' \neq 0$ on $(-1/2, 0) \cup (0, 1/2)$. Next let $\mathcal{I}_n$ be the set of open intervals constituting $C_n \smallsetminus C_{n+1}$ and, given a sequence $\{\gamma_n\}_{n \in \mathbb{N}} \in (0,1)^{\mathbb{N}}$, define $F \colon \mathbb{R} \to \mathbb{R}$ by

$$F(x) := \begin{cases} \gamma_n (b-a) h\Big(\dfrac{1}{\gamma_n} \dfrac{x - \frac{a+b}{2}}{b-a}\Big) & \text{if } x \in (a,b) \in \mathcal{I}_n \text{ for some } n \in \mathbb{N} \\ 0 & \text{else} \end{cases} \tag{41.8}$$

We then have:

Lemma 41.3 *Let C be the Cantor set above (fat or not). Suppose that $\{\gamma_n\}_{n \in \mathbb{N}} \in (0,1)^{\mathbb{N}}$ defining F is such that $\gamma_n \to 0$. Then F is differentiable on $\mathbb{R}$ with F' bounded and such that*

$$\forall x \in C \colon \;\; F'(x) = 0 \tag{41.9}$$

and

$$\forall n \geqslant 1 \, \forall I \in \mathcal{I}_n \colon \;\; \sup_{x \in I} F'(x) > 1 \tag{41.10}$$

Proof. A direct calculation shows that, whenever $x \in (a,b)$ for some $(a,b) \in \mathcal{I}_n$ and some $n \geqslant 1$, we have

$$F'(x) = h'\Big(\frac{1}{\gamma_n} \frac{x - \frac{a+b}{2}}{b-a}\Big) \tag{41.11}$$

Since $h'(-1/4) = 16\frac{1}{4}(1 - \frac{1}{16}) = \frac{15}{4} \geqslant 3$, we get (41.10). The key point is thus to show differentiability on C. Let $x \in C$ and assume that z is such that $F(z) \neq 0$. Then $z \in (a,b)$ for some $(a,b) \in \mathcal{I}_n$ with $|z - \frac{a+b}{2}| \leqslant \frac{1}{2}\gamma_n(b-a)$. Then $x \notin (a,b)$ gives $|x - \frac{a+b}{2}| \geqslant \frac{1}{2}(b-a)$ and so

$$|x - z| \geqslant \Big|x - \frac{a+b}{2}\Big| - \Big|z - \frac{a+b}{2}\Big| \geqslant \frac{1}{2}(b-a)(1 - \gamma_n) \tag{41.12}$$

by the triangle inequality and some algebra. On the other hand, the fact that h takes values in $[0,1]$ gives

$$\big|F(z) - F(z)\big| = F(z) \leqslant \gamma_n (b-a) \tag{41.13}$$

Combining these two observations we get that $F(z) \neq 0$ for $z \in (a,b)$ for some $(a,b) \in \mathcal{I}_n$ and some $n \geqslant 1$ implies

$$\left|\frac{F(z)-F(x)}{z-x}\right| \leqslant 2\frac{\gamma_n}{1-\gamma_n} \tag{41.14}$$

To complete the proof of differentiability note that, since $\gamma := \sup_{k\in\mathbb{N}} \gamma_k < 1$, the inequalities $|x-z| \geqslant \frac{1}{2}(b-a)(1-\gamma)$ and $|x-z| < \delta$ imply $(b-a) \leqslant 2(1-\gamma)^{-1}\delta$. If $z \in (a,b)$ for some $(a,b) \in \mathcal{I}_n$ and some $n \geqslant 1$, then $|x-z| < \delta$ for $x \in C$ and $F(z) \neq 0$ force n to be so large that $2^{-n}\text{length}(C) < 4\delta$. In short, if $z \to x$ and $F(z) \neq 0$, then n for which $x \in (a,b) \in \mathcal{I}_n$ must diverge and so $\gamma_n \to 0$ forces the ratio on the left of (41.14) tend to zero, proving $F'(x) = 0$. This completes the proof that F is everywhere differentiable with F' bounded and vanishing on C. ☐

We are now ready to conclude:

Proof of Theorem 41.1. Suppose that the sequence $\{\alpha_n\}_{n\in\mathbb{N}} \in (0,1)^{\mathbb{N}}$ defining C is summable and $\{\gamma_n\}_{n\in\mathbb{N}} \in (0,1)^{\mathbb{N}}$ defining F tends to zero. By Lemma 41.3, F is differentiable, vanishing on C. Yet, since every $x \in C$ is, for each $n \geqslant 1$, within 2^{-n} of all $z \in (a,b)$ for at least one $(a,b) \in \mathcal{I}_{n+1}$, the bound (41.10) implies that F' has a discontinuity at $x \in C$. Lemma 47.3 tells us that C is of non-zero length and so Theorem 39.3 lets us conclude that F' is NOT Riemann integrable on $[0,1]$. ☐

We remark that Volterra's example has been a source of motivation for the creation of Lebesgue's theory of integration; indeed, in this theory an everywhere differentiable function with a bounded derivative does satisfy the conclusion of (40.16). (This is because the derivative of an everywhere differentiable function is necessarily *measurable*, which is the type of regularity required by the Lebesgue integral.) Voltera's function is Lipschitz continuous and so we have shown that there are differentiable Lipschitz functions that do NOT arise as Riemann integrals. (As an aside we note that all such functions do arise as Lebesgue integrals.)

41.2 Improper integrals.

One unpleasant limitation of the Riemann integral is the requirement of boundedness, both for the integrated function and the underlying domain. This prevents us from integrating functions with divergences at the endpoints of integration domain directly, or to integrate over unbounded domains. This is usually fixed by taking additional limits leading us to the following concept:

Definition 41.4 (Improper integral) *Let $f\colon (a,b] \to \mathbb{R}$ be Riemann integrable on $[c,b]$ for every $c \in (a,b)$. The* improper Riemann integral *of f on $[a,b]$ is then defined by*

$$\int_a^b f(x)\mathrm{d}x := \lim_{c\to a^+} \int_c^b f(x)\mathrm{d}x \tag{41.15}$$

whenever the limit exists. A similar definition applies to the upper limit of integration and the limits $a \to -\infty$ and $b \to +\infty$. (Only one limit is to be taken at each time.)

Using this concept we can evaluate integrals

$$\int_0^1 \frac{1}{\sqrt{x}}\mathrm{d}x := \lim_{a\to 0^+} \int_a^1 \frac{1}{\sqrt{x}}\mathrm{d}x = \lim_{a\to 0^+} 2\sqrt{x}\Big|_a^1 = \lim_{a\to 0^+} 2[1-a] = 2 \tag{41.16}$$

where we use the FTCII along with the fact that $x \mapsto 2\sqrt{x}$ is an antiderivative of $x \mapsto \frac{1}{\sqrt{x}}$, as well as integrals such as

$$\begin{aligned}\int_0^1 \frac{\sin(1/x)}{x}\mathrm{d}x &:= \lim_{a\to 0^+}\int_a^1 x\frac{\mathrm{d}}{\mathrm{d}x}\cos(1/x)\mathrm{d}x \\ &= \lim_{a\to 0^+} x\cos(1/x)\big|_a^1 - \lim_{a\to 0^+}\int_a^1 \cos(1/x)\mathrm{d}x \\ &= \cos(1) - \int_0^1 \cos(1/x)\mathrm{d}x\end{aligned} \tag{41.17}$$

where we used integration by parts and then the fact that $x \mapsto \cos(1/x)$ is Riemann integrable on $[0,1]$. While the example (41.16) is completely fine for Lebesgue's theory of integration, the example (41.17) is out of consideration there as well. (This is due to the fact that the Lebesgue integral requires that the integral of at least one of the positive and negative part are finite, which fails for the function integrated in (41.17).)

To demonstrate integrals with infinite limits, we introduce the *Fresnel integral*

$$\begin{aligned}\int_0^\infty \sin(x^2)\mathrm{d}x &:= \lim_{b\to\infty}\int_0^b \sin(x^2)\mathrm{d}x = \lim_{b\to\infty}\int_0^b \frac{\sin(t)}{2\sqrt{t}}\mathrm{d}t \\ &= \lim_{N\to\infty}\sum_{n=0}^N \int_{2\pi n}^{2\pi(n+1)} \frac{\sin(t)}{2\sqrt{t}}\mathrm{d}t\end{aligned} \tag{41.18}$$

where use use the Substitution rule and then rewrote the general limit $b \to \infty$ as the limit along the naturals (this uses that the function decays to zero). Using the the fact that $\sin(t+\pi) = -\sin(t)$ we get

$$\int_{2\pi n}^{2\pi(n+1)} \frac{\sin(t)}{\sqrt{t}}\mathrm{d}t = \int_0^\pi \sin(t)\Big[\frac{1}{\sqrt{2\pi n + t}} - \frac{1}{\sqrt{2\pi n + \pi + t}}\Big]\mathrm{d}t \tag{41.19}$$

Elementary manipulations now show

$$\Big|\frac{1}{\sqrt{2\pi n + t}} - \frac{1}{\sqrt{2\pi n + \pi + t}}\Big| \leqslant \frac{2\pi}{2(2\pi n)^{3/2}} = \frac{1}{\sqrt{8\pi}}\frac{1}{n^{3/2}} \tag{41.20}$$

thus bounding the integral in (41.19) by $\frac{1}{\sqrt{8\pi}}n^{-3/2}$. It follows that the limit on the right of (41.18) exists. (The limit value of the integral is actually known to be $\sqrt{\pi/8}$.)

There is also a suitable limit procedure to address functions that have singularities inside the integration domain. One example of this is:

Lemma 41.5 (Cauchy principal value) *Let $a < 0 < b$ and suppose $f\colon [a,b] \to \mathbb{R}$ is continuous and such that $f'(0)$ exists. Then*

$$\int_a^b \frac{f(x)}{x}\mathrm{d}x := \lim_{\epsilon\to 0^+}\int_a^b \frac{f(x)}{x}1_{[a,-\epsilon]\cup[\epsilon,b]}(x)\mathrm{d}x \tag{41.21}$$

exists. (Here $1_A(x)$ equals one when $x \in A$ and zero when $x \notin A$.)

Proof. Assuming $a = -b$ for simplicity, we rewrite the integral under the limit as

$$\int_{-b}^b \frac{f(x)}{x}1_{[-b,-\epsilon]\cup[\epsilon,b]}(x)\mathrm{d}x = \int_\epsilon^b \frac{f(t) - f(-t)}{t}\mathrm{d}t \tag{41.22}$$

Since $f'(0)$ exists, the integrand is now easily checked to be a continuous function on $[0, a]$ and so the limit $\epsilon \to 0^+$ of the integral exists. $\square$

The computation using the limit generalizes beyond the case of f differentiable at the singularity. Indeed, if we know that f is α-Hölder at zero, the difference $f(t) - f(-t)$ would be at most a constant times t^α and the limit would still exist by a calculation similar to (41.16). One can also apply the same to other functions than $x \mapsto f(x)/x$; the key point is to use a well-articulated way of taking the limit near the singularity.

To finish the discussion of integrals over infinite domain, we note the following fact:

Lemma 41.6 *Let $f\colon [0, \infty) \to \mathbb{R}$ be non-increasing, non-negative with $\lim_{t\to\infty} f(t) = 0$. Let $\varphi\colon \mathbb{R} \to \mathbb{R}$ be continuous and, given some $L > 0$, such that $\forall t \in \mathbb{R}\colon \varphi(t + L) = -\varphi(t)$. Then the improper integral*

$$\int_0^\infty f(t)\varphi(t)\mathrm{d}t := \lim_{a\to\infty} \int_0^a f(t)\varphi(t)\mathrm{d}t \tag{41.23}$$

exists. (This includes that $t \mapsto f(t)\varphi(t)$ is Riemann integrable on $[0, a]$, for all $a > 0$.)

The proof of this fact uses a similar idea as invoked in the calculation in (41.18) which is itself reminiscent of the Alternating-series test. A typical example of φ is $\varphi(t) = \sin(t)$. We leave the proof to a homework exercise.