International Journal of Mathematical Education.
ISSN 0973-6948 Volume 11, Number 1 (2021) pp. 9-15

http://www.ripublication.com/gjpam.htm

A Connection Between the Baire Category Theorem on $\mathbb{R}$ and the Continuity of Certain Functions

Salem Ben Saïd*

Department of Mathematical Sciences, College of Science, United Arab Emirates University, Al Ain Abu Dhabi, UAE.

Abstract

In this note we will ask for the existence of a function that is continuous at each rational number and discontinuous at each irrational one. We will state the answer, which goes back to Volterra, by introducing in a reasonably elementary way senior hight school students to the Baire category theorem.

1. INTRODUCTION

The continuity of functions is one of the most important topics of calculus. In many countries around the world, the continuity of functions is taught in hight school and at the university (first semester). Before the latter part of the nineteenth century, mathematicians, including Newton (1642-1727) and Leibniz (1646-1716), thought of continuous functions in terms of some of the properties they should possess: you can sketch the graph, they do not move nearby points too far from each other. The work of Cauchy (1789-1857) and then Weierstrass (1815-1897) allowed for a precise definition of continuity: the ε-δ definition,

> A function f is continuous at a point a in its domain if and only if for every $\varepsilon > 0$, there exists a $\delta > 0$ so that, if $|x - a| < \delta$, then $|f(x) - f(a)| < \varepsilon$.

In other words, continuous functions are those that have only very small changes in the output when there are small changes in the input. These functions are in contrast to discontinuous functions, where the output changes substantially when the input changes slightly.

*The author is thankful to UAEU for the Start-up Grant No. G00002950.

This note deals with the following challenging problem: On $\mathbb{R}$, let f be that function defined by setting

$$f(x) = \begin{cases} \frac{1}{q} & \text{if } x = \frac{p}{q} \text{ in lowest terms} \\ 0 & \text{if } x \text{ is irrational} \end{cases}$$

The function f is the so-called Thomae function (1840-1921). It is easy to see that f is continuous at every $x \in \mathbb{R} \setminus \mathbb{Q}$ (the set of irrational numbers) yet discontinuous at every $x \in \mathbb{Q}$ (the set of rational numbers). See for instance [4, p. 76]. The main goal of this note is to answer to the following question:

Question 1.1. *Does there exist a function with the reverse property, i.e., a function that is continuous on $\mathbb{Q}$ and discontinuous on $\mathbb{R} \setminus \mathbb{Q}$?*

After spending hours trying to construct a such function, one begins to suspect that no such function can exist. But how to prove it? A beautiful result on set theory due to Baire [1, 2] comes to rescue. More precisely, the answer to the above question will use the Baire category theorem, with attendant of nowhere-dense sets, the concept of pointwise discontinuous functions due to Volterra [7] (see also [6]), i.e., functions whose points of continuity form a dense set, before demonstrating the non-existence of such a function. In order to do so, we will simplify the proof of Baire's category theorem and make it accessible to both last two years hight-school students and university students (the first semester). We believe our presentation provides a fine example of the advantages of studying the history of mathematics; for not only does it give a glimpse into the past but simultaneously satisfies the classroom needs of the present.

2. NOWHERE-DENSE SETS IN $\mathbb{R}$

A set S of real numbers is called dense in $\mathbb{R}$ if for any open interval (a, b) we have $(a, a) \cap S \neq \varnothing$. A set $\not{S}$ fails to be dense if there exists an open interval (a_0, b_0) such that $\not{S} \cap (a_0, b_0) = \varnothing$.

Examples 2.1. (1) *The set $\mathbb{Q}^+$ of positive rational numbers is not dense in $\mathbb{R}$ since, for instance, the open interval $(-2021, 0)$ is free of points of $\mathbb{Q}^+$.*

(2) *The set $\mathbb{Q}$ of rational numbers is dense in $\mathbb{R}$ since every real number has rational numbers that are arbitrarily close to it. The same holds true for $\mathbb{R} \setminus \mathbb{Q}$.*

Definition 2.2. *A set S of real numbers is called nowhere-dense if every open interval (a, b) contains an open subinterval (α, β) such that $(\alpha, \beta) \cap S = \varnothing$.*

The above definition says that even though points of S might be found in a given interval (a, b), there is a subinterval (α, β) with in it that is free of such points. It is worth mentioning that the concept of nowhere-dense is not the logical negation of dense. For instance, as we mentioned above, $\mathbb{Q}^+$ is a non dense set and at the same time is not a nowhere-dense set since the interval $(0, 2021)$ contains no subinterval (a, b) free of points of $\mathbb{Q}^+$.

Example 2.3. Consider the set $S = \{s\}$ with exactly one point. Then S is nowhere-dense. Indeed, if (a, b) is an arbitrary open interval not containing the real number s, then $(a, b) \subseteq (a, b)$ and $(a, b) \cap \{s\} = \varnothing$. Now we assume that $s \in (a, b)$. Then $(a, s) \subseteq (a, b)$ and $(a, s) \cap \{s\} = \varnothing$.

The following two lemmas are useful for later use.

Lemma 2.4. *If S is a nowhere-dense set and $R \subset S$, then R is also a nowhere-dense set.*

Proof. As S is nowhere-dense, then given an open interval (a, b), there exists $(\alpha, \beta) \subseteq (a, b)$ such that $(\alpha, \beta) \cap S = \varnothing$. The assumption $R \subset S$ implies $(\alpha, \beta) \cap R = \varnothing$, i.e., R is a nowhere dense set. □

Lemma 2.5. *The union $S_1 \cup S_2$ of two nowhere-dense sets is a nowhere-dense set.*

Proof. Given that S_1 is nowhere-dense, then for every (a, b) there exists $(\alpha, \beta) \subseteq (a, b)$ such that $(\alpha, \beta) \cap S_1 = \varnothing$. As (α, β) is an open interval and S_2 is nowhere-dense, then we can find an interval $(\gamma, \delta) \subseteq (\alpha, \beta) \subseteq (a, b)$ such that $(\gamma, \delta) \cap S_2 = \varnothing$. In conclusion, $(\gamma, \delta) \cap (S_1 \cup S_2) = \varnothing$ where $(\gamma, \delta) \subseteq (a, b)$. □

From above it follows that a finite union of nowhere-dense sets gives a nowhere-dense set. But what if we take an infinite union of nowhere-dense sets? What use might this be to calculus? This mater was addressed by Baire in his famous category theorem, which we will simplify its presentation and make it accessible to hight school students and university students (first semester).

Henceforth, a set S is called of the first category if

$$S = P_1 \cup P_2 \cup P_3 \cup \cdots \cup P_k \cup \cdots,$$

where every P_k is a nowhere-dense set.

Example 2.6. The set $S = \{a_1, a_2, a_3, \ldots, a_k, \ldots\}$ is of the first category. Indeed, we may rewrite S as $S = \{a_1\} \cup \{a_2\} \cup \{a_3\} \cup \cdots \cup \{a_k\} \cup \cdots$, where, by Example 2.3, each $P_k = \{a_k\}$ is nowhere-dense.

For later use, we prove the following:

Lemma 2.7. *Let S be a set of the first category and let $R \subseteq S$, then R is of the first category.*

Proof. Since S is of the first category, then $S = P_1 \cup P_2 \cup \cdots \cup P_k \cup \cdots$, where every P_k is a nowhere-dense set. Then, $R = R \cap S = (P_1 \cap R) \cup (P_2 \cap R) \cup \cdots \cup (P_k \cap R) \cup \cdots$, where $P_k \cap R \subset P_k$ for all $k \geq 1$. As every P_k is a nowhere-dense set, then Lemma 2.4 implies $P_k \cap R$ is a nowhere-dense set for all $k \geq 1$. □

By mimicking the proof of the above statement, the following lemma would be considered a simple classroom challenge.

Lemma 2.8. *The union of two first category sets is first category.*

The following theorem highlights a particularly important aspect of first category sets. More precisely, a first category set is never sufficient to exhaust an open interval. However, what is indeed important for us in this note is its Corollary 2.10.

Theorem 2.9. (Baire's category theorem) *If $S = P_1 \cup P_2 \cup \cdots \cup P_k \cup \cdots$, where P_k is nowhere-dense for every k, and if (a, b) is an arbitrary open interval, then there exists a point $c \in (a, b)$ such that $c \notin S$.*

Proof. Consider an arbitrary open interval (a, b). As P_1 is nowhere-dense, there exists a closed interval $[a_1, b_1] \subseteq (a, b)$ such that $[a_1, b_1] \cap P_1 = \varnothing$. Now, we consider the open interval (a_1, b_1) and the nowhere-dense set P_2. There exists a subinterval $[a_2, b_2] \subseteq (a_1, b_1) \subseteq [a_1, b_1] \subseteq (a, b)$ where $[a_2, b_2] \cap P_2 = \varnothing$. Continuing in this way, we obtain a sequence of intervals $[a_k, b_k]$ such that $[a_1, b_1] \supseteq [a_2, b_2] \supseteq \cdots \supseteq [a_k, b_k] \supseteq \cdots$, where $[a_k, b_k] \cap P_k = \varnothing$ for every $k \geq 1$. By the nested interval version of the completeness property (Cantor's lemma; see [3]), there exists at least a point c common to all these intervals $[a_k, b_k]$, for $k \geq 1$. Now one needs to show that $c \in (a, b)$ but $c \notin S$. On the one hand, as $c \in [a_k, b_k]$ for every k, then $c \in [a_1, b_1] \subseteq (a, b)$. On the other hand, since $[a_k, b_k] \cap P_k = \varnothing$ and $c \in [a_k, b_k]$ for every $k \geq 1$, it follows $c \notin P_k$ for every k, which implies $c \notin S$. □

Recall that the complement $S^{\complement}$ of a set S is the set of real numbers not belonging to S.

Corollary 2.10. *The complement $S^{\complement}$ of a set S of the first category is dense.*

Proof. By Theorem 2.9, for every interval (a, b), there exists $c \in (a, b)$ such that $c \notin S$. That is $c \in S^{\complement}$. In other words, for every open interval (a, b), we have $(a, b) \cap S^{\complement} \neq \varnothing$, i.e., $S^{\complement}$ is a dense set in $\mathbb{R}$. □

The above result have important consequences for calculus, as the central questions of calculus are about functions. In particular, Corollary 2.10 will play a crucial role in answering to Question 1.1 asked in the Introduction. More precisely, we will use the above result to characterize pointwise discontinuous functions. A function f is called pointwise discontinuous whenever f is discontinuous at infinity many points but still continuous on a dense set.

For a given function f, let us introduce the set

$$P_k = \left\{ \begin{array}{c} x \text{ such that there exists a sequence } (a_n)_n \text{ which converges} \\ \text{to } x \text{ as } n \to \infty \text{ with } |f(x) - f(a_n)| \geq \dfrac{1}{k} \text{ for all } n \geq 1 \end{array} \right\}. \tag{2.1}$$

For instance, let us consider the function

$$f(x) = \begin{cases} \cos(\frac{1}{x}) & \text{if } x \neq 0 \\ 0 & \text{if } x = 0. \end{cases}$$

We claim that $0 \in P_k$ for all $k \geq 1$. Indeed, let $(a_n)_n = \left(\frac{1}{2\pi n}\right)_{n \geq 1}$. Clearly we have $\lim_{n\to\infty} a_n = 0$ and $|f(0) - f(\frac{1}{2\pi n})| = |0 - \cos(2\pi n)| = 1 \geq \frac{1}{k}$ for all $k \geq 1$.

Theorem 2.11. (Baire's characterization of pointwise discontinuous functions) *A function f is pointwise discontinuous if and only if the set S_f of points of discontinuity of f is of the first category.*

Proof. ($\Longrightarrow$) First we will prove that every P_k in (2.1) is nowhere-dense for every $k \geq 1$. Let (a, b) be an open interval. As f is pointwise discontinuous, then f is continuous at some point, say c, in (a, b). That is $\lim_{x\to c} f(x) = f(c)$, and therefore, for $\epsilon = \frac{1}{3k}$, there exists a $\delta > 0$ such that $(c - \delta, c + \delta) \subseteq (a, b)$ and $|f(x) - f(c)| < \frac{1}{3k}$ for all $x \in (c - \delta, c + \delta)$ (see the ε-δ definition of continuity in Introduction). We claim that $(c - \delta, c + \delta) \cap P_k = \varnothing$. To prove so, let us assume the opposite, i.e., there is a $y \in (c - \delta, c + \delta) \cap P_k$. Therefore, there exists a sequence $(a_n)_n$ which converges to y as $n \to \infty$, with $|f(y) - f(a_n)| \geq \frac{1}{k}$ for all $n \geq 1$. Since $y \in (c - \delta, c + \delta)$ and $\lim_{n\to\infty} a_n = y$, there is a subscript N such that $a_N \in (c - \delta, c + \delta)$. By the triangle inequality

$$|u - v| \leq |u - w| + |w - v|, \qquad \text{for every } u, v, w \in \mathbb{R},$$

we obtain

$$\frac{1}{k} \leq |f(a_N) - f(y)| \leq |f(a_N) - f(c)| + |f(c) - f(y)| \leq \frac{1}{3k} + \frac{1}{3k} = \frac{2}{3k}$$

as both a_N and y belong to $(c-\delta, c+\delta)$. Thus we get $\frac{1}{k} \leq \frac{2}{3k}$, which is absurd. Hence $(c-\delta, c+\delta) \cap P_k = \varnothing$. We conclude that there exists a subinterval $(c-\delta, c+\delta) \subseteq (a, b)$ such that $(c-\delta, c+\delta) \cap P_k = \varnothing$, which implies that P_k defined in (2.1) is nowhere-dense for every k, and therefore $P_1 \cup P_2 \cup \cdots \cup P_k \cup \cdots$ is a set of the first category. To finish the proof of the forward direction ($\Longrightarrow$), and in the light of Lemma 2.7, we shall prove that

$$S_f \subset P_1 \cup P_2 \cup \cdots \cup P_k \cup \cdots .$$

Let $d \in S_f$ be an arbitrary point of discontinuity of f. Then there exists $\epsilon > 0$ such that for every $\delta > 0$, we can find x such that $|x - d| < \delta$ with $|f(x) - f(d)| \geq \epsilon$. Choose an integer k_0 such that $\frac{1}{k_0} < \epsilon$ and let δ equal, in tern, $1, \frac{1}{2}, \frac{1}{3}, \frac{1}{4}, \cdots$ to generate a sequence $a_1, a_2, a_3, a_4, \cdots$ which converges to d, with $|f(a_n) - f(d)| \geq \frac{1}{k_0}$. Hence $d \in P_{k_0}$. In conclusion, $S_f \subset P_1 \cup P_2 \cup \cdots \cup P_{k_0} \cup \cdots$. By Lemma 2.7 we deduce that S_f is of the first category.

Now we turn our attention to proof of the reverse direction ($\Longleftarrow$). We shall prove that if S_f is of the first category, then f is pointwise discontinuous. As S_f is of the first category, by Corollary 2.10, its complement $S_f^{\complement}$ is dense, which is precisely the required condition for f to be pointwise discontinuous. $\square$

We close this note by answering to Question 1.1 asked in the Introduction.

Theorem 2.12. (Volterra's theorem revised) *There do not exist two pointwise discontinuous functions on an open interval* (a, b) *for which the continuity points of one are the discontinuity points of the other, and vice versa.*

Proof. Assume there exist two pointwise discontinuous functions f and g such that the continuity points of f are the discontinuity points of g, and vise versa. By the previous result, S_f and S_g are two sets of the first category. Therefore, Lemma 2.8 implies $S_f \cup S_g$ is also a set of the first category. Using Corollary 2.10, it follows that the complement $(S_f \cup S_g)^{\complement}$ is dense. However, $(S_f \cup S_g)^{\complement} = S_f^{\complement} \cap S_g^{\complement}$ represents the set of common points of continuity for f and g, which is empty. Thus we reached a contradiction. $\square$

With these reasonably elementary arguments—a tribute to the genius of Baire and Volterra —we can use some of yesterday's mathematics to answer today's questions. Students should be both satisfied and, one hopes, impressed. And in this instance the history of mathematics comes alive by proving its value in the contemporary classroom.

3. CONCLUDING REMARKS

(1) After presenting the content of this note to students, it could be very tempting to challenge them to try to think about the function defined by

$$f(x) = \begin{cases} x & \text{if } x \text{ is irrational} \\ p\,\sin(1/q) & \text{if } x = p/q \text{ in lowest terms.} \end{cases}$$

(2) Students may use the fact that a first category set cannot exhaust an open interval (Theorem 2.9) to give a short and elementary proof for the known Cantor's theorem [5]:

> *If $\{x_n\}_n$ is a sequence of distinct real numbers, then any open interval (a, b) contains a point not included among the sequence $\{x_n\}_n$.*

(3) Finally, a challenging question for students clould be an extension of Volterra's revised theorem to a sequence of pointwise discontinuous functions defined on a common interval. Indeed, one can show that there exists a point (indeed, a dense set of points) at which all of these are simultaneously continuous.

REFERENCES

[1] R. Baire, *Sur les fonctions discontinues qui se rattachent aux fonctions continues*, Comptes Rendus Hebdomadaires des Séances de l'Académie des Sciences, Paris, vol. 129 (1898), 1621–1623.

[2] R. Baire, *Sur les fonctions de variables réelles*, Annali di Matematica Pura ed Applicata, vol. 3 (1899), 1–123.

[3] R. G. Bartle and D. R. Sherbert, *Introduction to Real Analysis*, John Wiley & Sons Inc, 2011.

[4] C. Clark, *The Theoretical Side of Calculus,* Krieger, 1978.

[5] G. Cantor, *Gesammelte Abhandlungen mathematischen und philosophischen Inhalts*, Edited by Ernst Zermelo with biography of Cantor by Abraham A. Fraenkel. Berlin: Springer-Verfag,1932/Hildesheim: Georg Olms, 1962.

[6] H. Royden, *Real Analysis,* 2nd ed., Macmillan, 1968.

[7] V. Volterra, *Alcune osservasioni sulle funzioni punteggiate discontinue*, in Opere Matematiche, vol. 1 (Rome; Accademia Nazionale dei Lincei, 1954), (originally in Giornale di Matematiche 19 (1881), 76–86).