Motion and Continuity

Peter J. Olver
School of Mathematics
University of Minnesota
Minneapolis, MN 55455
olver@umn.edu
http://www.math.umn.edu/∼olver

> I don't here consider Mathematical Quantities as composed of Parts *extreamly small*, but as *generated by a continual motion*. Lines are described, and by describing are generated, not by any apposition of Parts, but by a continual motion of Points. Surfaces are generated by the motion of Lines, Solids by the motion of Surfaces, Angles by the Rotation of their Legs, Time by a continual flux, and so in the rest. These *Geneses* are founded upon Nature, and are every Day seen in the motion of Bodies.
>
> Isaac Newton, *Tractatus de Quadratura Curvature*†

Abstract. A fundamental attribute of the macroscopic physical universe is that all objects move continuously. We explain how this fact can be used to motivate the modern topological definition of continuity of functions. This basic definition can then be used to establish a new, fully rigorous foundation for calculus, and, consequently, a promising new approach to the mathematics used to model and understand our world.

The most basic property of motion in the physical universe is that, without exception, it is continuous. In this note, I show how this fundamental fact can be used to motivate and then formulate the modern topological definition definition of continuity, namely a function is continuous if and only if the inverse image of any open set is open, [**15**, **22**]. Moreover, as demonstrated in the author's recent lecture notes, [**23**], a fully rigorous foundation for the entire edifice of calculus can be based on this relatively simple definition,

June 24, 2022

† English translation from [**10**].

thereby completely avoiding the more technically intricate and less intuitive epsilon-delta definitions of limit and continuity that were introduced two centuries ago by Bolzano, Cauchy, and Weierstrass, [**4**, **7**, **8**, **9**], and have, ever since, formed the almost universally accepted basis for rigorous calculus. Moreover, in contrast to the limit-based and other proposed foundations — including nonstandard analysis based on infinitesimals, [**14**, **25**], and the "unlimited" approach of [**19**] — the continuity-based approach can be motivated using just our everyday experience in the physical world.

Philosophical discussions of motion have, by and large, focussed on the nature of velocity; see, for instance, [**5**, **28**] for two recent competing points of view. A key question, dating back to the ancient Greeks, is whether instantaneous velocity has an intrinsic physical meaning. As such, this line of thought became wedded to the notion of limit and the consequential definition of the derivative of the position function, which determines, mathematically, an instantaneous velocity of the body. The continuity-based approach enables one to decouple motion from velocity and limits, and thereby open up new avenues of philosophical inquiry. Indeed, not all physical objects have an instantaneous velocity — a particularly interesting case is *Brownian motion*, where, due to its interactions with the molecules, a dust particle suspended in a fluid follows a fractal path, [**18**], akin to Bolzano's nowhere differentiable function, [**26**], and has no well defined instantaneous velocity anywhere on its stochastic trajectory.

All objects in the universe[†] are in motion. By this statement we include objects at rest, so that "motion" also includes "lack of motion". Indeed, motion is relative: the difference between motion and lack of motion depends on the observer. Two cars moving at the same velocity are in motion relative to a pedestrian watching them go by. But to either driver, the other car is standing still and it is the pedestrian who is in motion, in the opposite direction. As Einstein emphasized, there is no absolute coordinate frame, such as the Aristotelian celestial sphere, that allows one to incontrovertibly assert that a body is motionless; see [**1**, **6**] for the historical antecedents, beginning with Copernicus, Galileo and Descartes. Thus, Zeno's Paradox of the Arrow, [**27**, **28**], is either nonexistent, since all moving bodies can be fixed in place when viewed by a comparably moving observer, or applies equally well to a motionless arrow, which can be placed into motion by any moving observer. If one were to accept Zeno's argument, the only viable conclusion is that an object can only exist for an instantaneous moment in time, since it can neither move nor sit still. But of course this is completely incompatible with our everyday experience of object permanence.

Objects can of course change their shape and size while moving; they can even split apart or combine together. Here one should be thinking not just of solid objects like rocks and people and planets, but also liquids and gases. The motion of objects is subject to

[†] If pressed, I will, for the sake of argument, restrict attention to the macroscopic, non-quantum universe, so that our intuition and experience can be validly relied upon while strange quantum mechanical effects can be ignored. I will also set aside recent theories of a "granular universe", e.g., [**24**], that becomes discrete when one gets down to, say, the Planck length scale. Of course, most of modern physics, including quantum mechanics and relativity, continues to rely on calculus, and so our motivations and constructions remain relevant throughout.

a variety of physical laws, depending on the nature of the object and the environment in which it moves. The motion of a ball depends on its material properties and the forces affecting it, such as its shape and size, whether it is in air or water or outer space, its gravitational mass, its surface roughness, whether or not it is electrically charged, etc.

So given the wide variability of observed motion in our physical universe, are there any universal laws? A moment's reflection should suffice to convince one that there is in fact one behavior that is obeyed by all (macroscopic) motion, without exception. Namely, motion is always continuous. By this we mean that a body in motion must move continuously through space. If it starts at point A and finishes at point B then (assuming it remains intact) it must trace a continuous path that connects A to B so that at each time it is located somewhere on the path and, moreover, passes through the intervening positions in a continuous manner. It cannot suddenly disappear and reappear at a different location. Even the Brownian motion of a dust particle, despite its lack of instantaneous velocity, still obeys this most basic law; no matter how wild its motion, it remains without question continuous. Fans of the classic science fiction TV series Star Trek might dream of teleportation from one world to another, but (as far as we know) this is impossible[†] and the only way to get from here to there is by moving continuously through the intervening space. On the other hand, one can easily accomplish such a nonphysical feat in movies by splicing frames (nowadays, by manipulating movie files on the computer) so that the object suddenly jumps from one location to another — a simple special effect that amazed early film audiences — but we all know that this is physically unrealizable and relies entirely on video postprocessing.

The fundamental Physical Law of Continuity[‡] that can be gleaned from Newton's initially cited quote was formulated more precisely by Bertrand Russell, [**27**; p. 473]. In the final paragraph of his chapter on motion, Russell states:

> Motion consists merely in the occupation of different places at different times, subject to continuity as explained in Part V. There is no transition from place to place, no consecutive moment or consecutive position, no such thing as velocity except in the sense of a real number which is the limit of a certain set of quotients. The rejection of velocity and acceleration as physical facts (i.e. as properties belonging at each instant to a moving point, and not merely real numbers expressing limits of certain ratios) involves, as we shall see, some difficulties in the statement of the laws of motion; but the reform introduced by Weierstrass in the infinitesimal calculus has rendered this rejection imperative.

According to [**28**], Russell's Law of Continuity has gained almost universal acceptance[§]. In particular, it resolves Zeno's Paradox of the Arrow in that the continuity of the

---

[†] Despite the name, quantum teleportation, [**3**], involves the transfer of information, not physical objects (photons, atoms, molecules).

[‡] Not to be confused with Leibniz's Mathematical Law of Continuity, [**4**, **13**].

[§] On the other hand, as we noted earlier, continuity of velocity is not a universally accepted property of motion. Examples of discontinuous velocities include collisions between solid bodies in motion of the arrow cannot be ascertained from a static view of its location at any instant of time. The continuity of its motion rests not on any limiting behavior or derivative of the function describing the motion (the latter not even existing in the case of a "Brownian arrow"), but, rather, on its behavior over an interval of time. This is the point of view we will develop below. In contrast, to mathematically describe his Law of Continuity, Russell relies on the standard, limit-based epsilon-delta formulation of the definition of continuity of a function, but is forced to admit, [**27**; p. 327]:

> These definitions of continuity and discontinuity of a function, it must be confessed, are somewhat complicated; but it seems impossible to introduce any simplification without loss of rigour.

In this note, I will argue that Russell's final statement is in fact unwarranted. Continuity, as the topologists discovered, [**15**, **17**, **22**], does not require the notion of limit, and hence is the more fundamental concept. Indeed, as we show below, limits (and hence all of calculus) can be rigorously defined in terms of continuity rather than the reverse.

The intent is to motivate the topological definition of continuity purely from a foundational analysis of physical motion. I find the resulting definition to be not only simpler, requiring only the notion of an open subset of the real line[†], and hence easier to assimilate than the standard limit-based version, while, as demonstrated in [**23**], not suffering from any loss of rigor in the consequential development of the mathematical apparatus of calculus. Moreover, one can argue that this approach is more in the spirit of Newton's original concept of a motion-based calculus, as founded on his fluxion apparatus, the latter also dating back to the Greeks. See, for instance, [**12**], which goes on to say

> Both the Greek extremity theory and fluxion theory have been very influential. Not surprisingly we find Isaac Newton (1642-1727) calling upon the motion theory as a support for his fluxional principles of the calculus.

which is then followed by the quote that begins the present paper.

So the question becomes, how does one mathematically formulate the physical requirement of continuity? As noted above, it cannot depend on the position of the object at any one particular time. In the language of the cinema, one cannot predict whether an object will move continuously or not by inspecting a single frame of the movie. To ascertain continuity, one needs to consider the behavior of the object over an interval of time, of some nonzero duration.

---

celestial mechanics, billiards, etc., and Brownian motion. Now it is plausible that at a sufficiently small scale, the bodies experience an abrupt but continuous change in velocity. But much of classical mechanics does not make this assumption and indeed deals with collisions directly as abrupt discontinuities in velocity. And the small scale one must reduce to may be the molecular/quantum level at which point the equations of motion become radically different. The bottom line is that, when modeling macroscopic physical phenomena, one always assumes continuity of motion, but not necessarily continuity of velocity.

[†] One can, of course, generalize to arbitrary topological spaces, but the simplest version will suffice for our purposes.

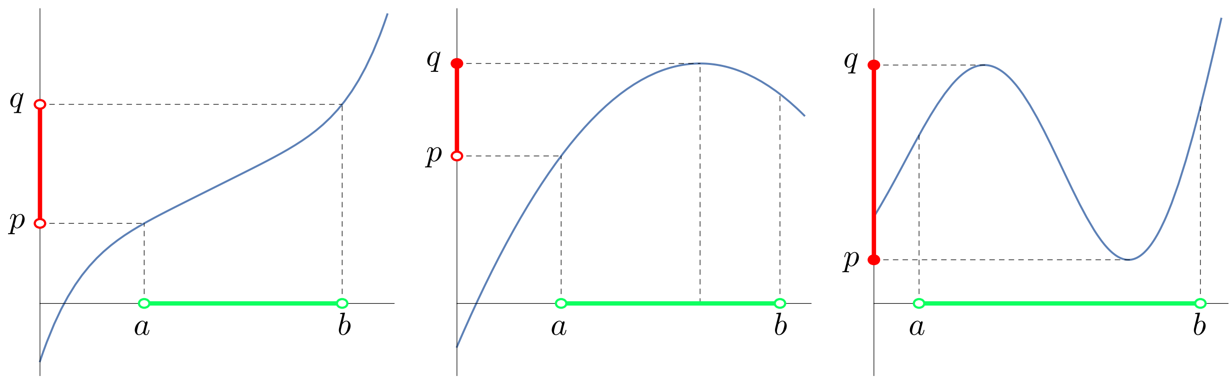

Figure 1. Continuous Motion of a Car.

To understand what is going on, and thus motivate the eventual mathematical definition, let us look at a specific example. Consider a car driving along a road. At each time t the car is at a certain location x along the road, measured in some chosen units of time and distance (seconds, minutes; kilometers, miles, etc.). We assume the car maintains its overall shape (no accidents!) — although it does change its material form in that it is burning gasoline and emitting the resulting exhaust as it travels. To be precise, let us measure the position of the car by choosing a single fixed point on it — for example, the middle of the front bumper. We remark that choosing the center of mass is trickier due to the above mentioned change in composition. However, if one goes down to the molecular level, even before quantum effects become manifest, the choice of a single fixed point on the car can become problematic due to thermal motion and the detachment/attachment of individual atoms and molecules from its body. Anyone who is troubled by such small scale physical details can replace our "car" by a mathematically idealized "point particle" that travels along a straight line with a well defined position at each time in what follows, keeping in mind that the aim is to motivate the definition of continuity, and not be sidetracked into a detailed discussion of the underlying physics.

We use functional notation, and write $x(t)$ for the position of the car (or point particle) along the road at time t. Thus, at each time t in the time interval under consideration there is one and only one location of the car, namely $x(t)$. This means that our car is not allowed to disappear or leave the road during the times that we observe it. See the accompanying figures for plots of some representative motions.

We will allow the car to move in any fashion whatsoever that is, at least in principle, physically realizable, i.e., we ignore dynamical and material constraints, like top speed, maximum acceleration, braking ability, etc. The car can speed up; it can slow down; it can stop; it can reverse direction; and any combination thereof is allowed. However, by the Physical Law of Continuity, it is not allowed to instantaneously jump from one position on the road to another. If at time $t = c$ it is at position $p = x(c)$ and at a later time $t = d$ it appears at position $q = x(d)$, then in the intervening times it must have visited every intermediate position $p < x < q$ at least once, and it must have done so in a continuous fashion without any smaller instantaneous jumps. However, this "Intermediate Value Property", while a consequence of continuity, is, in and of itself, insufficient to

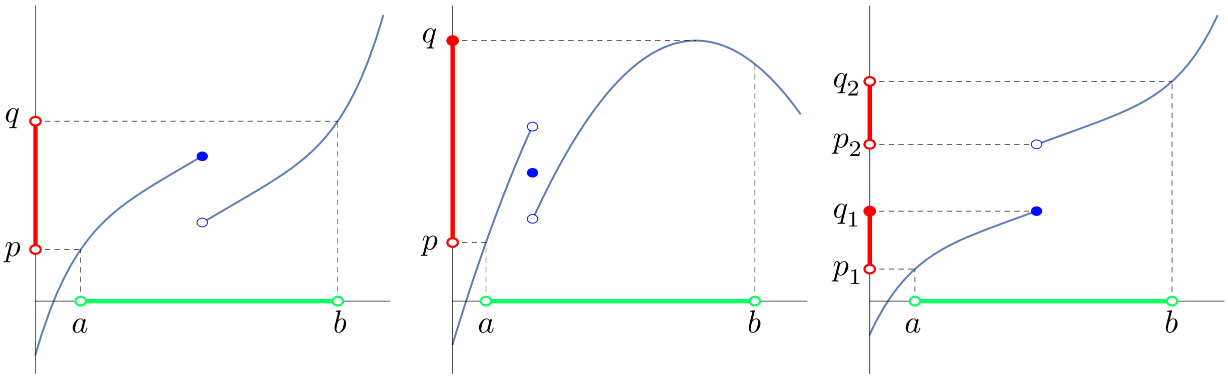

Figure 2. Discontinuous Motions of a Car.

produce a mathematically precise definition of continuity, and we need to understand the possible motions in more detail.

To this end, let us look at what happens to the car on some time interval. We will focus on an open interval $a < t < b$ meaning that we do not include the endpoints a, b. This is just for specificity and later we will discuss the case of closed intervals $a \le t \le b$ that include both endpoints. Consider the part of the road that the car has visited; we will refer to this as the *range* of the car over the given time interval, which is, mathematically, the range of the function $x(t)$ for $a < t < b$. Well, if the car only moves in a single direction, as in the first plot[†] in Figure 1, then the road segment that the car has travelled along, i.e., its range, will also be an open interval, say $p < x < q$. However, if the car reverses direction once at some intermediate time then the range will be a half open interval, say $p < x \le q$ if it goes forward and then reverses, as illustrated in the second plot, or $p \le x < q$ if it does the opposite. If it reverses direction more than once, the range could even be a closed interval $p \le x \le q$ as in the third plot.

However, the same can be said for a variety of non-physical motions of the car, as illustrated by a few representative plots[‡] in Figure 2. The range can be either an open, half open, or closed interval, or it could be more general, e.g., the union of two or more disjoint intervals, the latter case not allowed for a continuously moving physical car. So this approach to the problem is not going to help us unambiguously distinguish continuous from discontinuous motion.

Now comes the key insight. Let's switch our point of view, and look instead at what happens on an open interval of the road. We will investigate the collection of times for which the car is located within the prescribed road interval. In other words, consider an open interval $p < x < q$. (The endpoints p, q could be indicated by placing flags by the side of the road.) Our question is: what values of t will see the car situated between positions p and q? Suppose first that when $a < t < b$, the car drives in only one direction. If it

---

[†] In the plots, time intervals are colored green, and the corresponding road intervals are colored red. Open circles indicate open intervals; closed circles indicate closed intervals; an interval with one of each is half open.

[‡] The solid blue circle on the graph indicates the position of the car at the discontinuity.

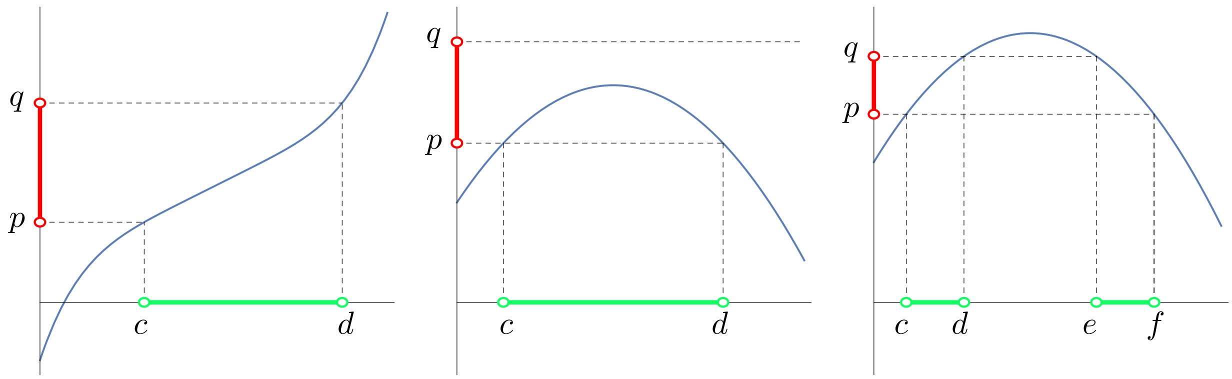

Figure 3. Continuous Motions of a Car.

starts out past $p < x(a)$ and finishes up short of $q > x(b)$ then the answer is the entire open time interval $a < t < b$. On the other hand, if it starts out before the first flag, so $x(a) < p$ and finishes further down the road, so $x(b) > q$, then, by continuity of the motion (the Intermediate Value Property), it must be at p at some time $a < c < b$, so $x(c) = p$, and, at some later time $c < d < b$ (keep in mind that here the car is always moving in the positive direction), it is at $q = x(d)$, as in the first plot in Figure 3. Therefore, again by the continuity of motion, it is in the open road interval $p < x < q$ at all times t belonging to the open time interval $c < t < d$ for some $a \leq c < d \leq b$. Similar considerations apply to the other two cases of unidirectional motion, where it starts in the road interval and finishes outside, or when it starts outside and finishes inside. In all cases the time interval when it is in the chosen road interval is an open time interval of the form $c < t < d$. We claim that this fact is, with one amplification noted below, true no matter how the car moves! For example, suppose the car enters the interval at position p at time c, reverses direction and leaves through the same point at a later time $d > c$, so $x(c) = x(d) = p$, while $p < x(t) < q$ for all $c < t < d$. Again, the times the car is in our chosen road interval form an open time interval $c < t < d$; see the second plot in Figure 3. Indeed, the same holds no matter how many times the car changes direction within the road interval, and no matter at which endpoint it enters or leaves, or even whether or not it starts and/or ends in the road interval. Even a "Brownian car" would obey the same rules.

The one complication is when the car exits the chosen road interval, then reverses direction (or uses a alternative route to return to the other end), and enters it again, perhaps doing this more than once, where we continue to allow for the possibility that the car reverses direction whilst inside the road interval. In such cases, the times when the car is in the given section of road form a disjoint union of several open time intervals; see the third plot in Figure 3. For example, the car might have position $p < x(t) < q$ when either $c < t < d$ or $e < t < f$ and so on, depending upon how many times it visits our chosen section of road; at the endpoints of the given time intervals, $c, d, e, f, \ldots,$ the car would be passing through one of the road interval endpoints, p or q. (If $d = e$, it would reverse direction immediately upon reaching an endpoint. If $a = c$ the car could start out at a point within the road interval, and similarly if the final endpoint $f = b$ it could end up at another such point.) Let us call a union of one or more open time intervals an *open set*, in accordance with standard topological terminology. Then the claim is that, *no matter how the car moves along the road, the set of times when it is in any chosen open interval of road is open*, meaning that it is the union of a collection of open time intervals. It could even be empty if the car always stays outside the road interval under consideration, and so we also deem the empty set to be open. The set of times is still open if the car stops during its travels, or even remains stopped for the entire time.

On the other hand, we claim that this observation is *not* valid if the car performs a non-physical motion. In other words, if the motion is discontinuous, we can find a road interval $p < x < q$ with the property that the times at which the car is located there *do not* form an open set. Let's start with a very simple case. Suppose that the car is stopped at position $x = y$ for some time, say $x(t) = y$ for $a < t < c$ and then immediately appears at a different location $x = z$, so that $x(t) = z$ for $c < t < b$. To be specific, let's assume $y < z$, meaning that the car has been teleported further down the road. (A very useful property for winning races!) Now at time $t = c$, the car has to be somewhere — as before, we do not allow physical objects to disappear completely. Mathematically this means that the domain of the location function $x(t)$ is the entire interval $a < t < b$. Suppose $x(c) = y$. Choose any locations $p < y < q < z$, so p is located in front of where the car was originally sitting and q is some point between there and where it ends up. Consider the road interval $p < x < q$. At what times is the car located inside this interval? Well, using our assumption on p and q, this would be when $a < t \leq c$ since the car sits at position y up until and including the moment when it teleports, but not any later since it is then at location z which is outside the road interval under consideration. Thus, the specified time interval is half open, not fully open as it was in the continuous case; see the first plot in Figure 4. A similar argument, left to the reader, applies when $x(c) = z$, in which case we look at the road interval $p < x < q$ where $y < p < z < q$. What about if $x(c) = w$ where w is different from both y and z? This corresponds to a car sitting at position y, then instantaneously appearing at a different location w before jumping to position z where it remains for all times thereafter. Well, since $w \neq y$ and $w \neq z$, we can find a small road interval $p < x < q$ such that w belongs to it, so $p < w < q$, but both y and z lie outside it. But then at what times is the car in this particular interval? The answer is only when it is at position $x = w$, which only occurs at time $t = c$, as illustrated in the second plot in Figure 4. Thus, the set of times when the car sits in the given road interval $p < x < q$ is a single point, $t = c$, and not an open interval or union thereof. Note that this also applies when $y = z \neq w$, i.e., the car is sitting at one location suddenly for one instant appears elsewhere before immediately returning to its original position, as in the third plot.

With a little more thought, one discovers that the same occurs for any non-physical discontinuous motion of the car. If this happens, and assuming the car is located somewhere along the road at every time in some open interval $a < t < b$, then one can find a road interval $p < x < q$ such that the set of times where the car is located inside this road interval is not open, meaning that it is not an open interval nor a union of open time intervals. This applies to *all* other types of discontinuous motion beyond just suddenly jumping from one location to another; see the Appendix for an interesting example.

We have thus landed upon the topological formalization of continuity. It relies on the definition of an open subset of the real line, denoted by $\mathbb{R}$, and which, as we remarked

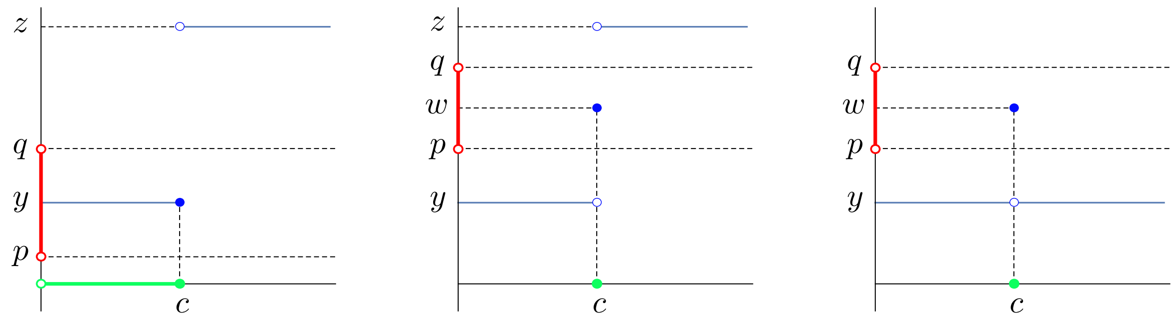

Figure 4. Discontinuous Motions of a Car.

earlier, can be taken to be either the empty set, or an open interval, or a union — finite or infinite — of open intervals. We also introduce the following notation. Given a scalar function $x(t)$ and a subset $S \subset \mathbb{R}$ — think of S being the road interval — we define its inverse image $x^{-1}(S)$ to be the set of all t for which $x(t)$ lies in S:

$$x^{-1}(S) = \{\, t \in \mathbb{R} \mid \; x(t) \in S \,\} .$$

With this in hand we can now state the basic definition of continuity.

Definition 1. A function $x(t)$ is called *continuous* if, whenever $S \subset \mathbb{R}$ is open, then $x^{-1}(S) \subset \mathbb{R}$ is also open.

Note that the open set S is, in general, a union of open intervals, $S = \bigcup I_j$, and, by our preceding observations, each individual $x^{-1}(I_j)$ is a union of open intervals, and hence so is $x^{-1}(S) = \bigcup x^{-1}(I_j)$.

Observe that no limits are required to formulate this definition of continuity. On the other hand, it is a useful mathematical exercise to prove that the usual limit-based definition and the above topological definition are equivalent. But here, we discard the former entirely, and take Definition 1 as our starting point.

Remark: According to Moore, [**20**, **21**], Definition 1 can be traced back to Hausdorff, [**11**], modulo his nonstandard definition of topological space, and, in its fully modern form, to Lefschetz, [**17**].

And it is this foundation upon which one can build the entire edifice of calculus, following the unorthodox approach taken in the author's notes, [**23**]. These notes demonstrate that this basic definition can be used as a starting point to rigorously develop all of calculus, in both one and several variables, including the intermediate value, extreme value, and mean value theorems, etc., often with conceptually simpler proofs that do not require the introduction of any epsilons and deltas. The definition can even be adapted to treat discrete limits, Cauchy sequences, uniform convergence, and much more. However, I will not delve into the full details here, all of which can be found in the above-referenced lecture notes.

Furthermore, I am of the opinion that, in light of the preceding motivational arguments, this continuity-based approach to the subject is (a) more directly tied to everyday experience, and, thus, (b) easier for the novice to grasp — in contrast to the usual limit-based approach, that relies on the unintuitive and technically challenging epsilon-delta construction. Moreover, it completely avoids any "ghosts of departed quantities" that underpin Berkeley's pointed critique, [**2**], of Newton's and Leibniz's methods; see also [**4**, **8**]. It further has the pedagogical advantage of introducing the beginning student to some elementary point set topology, which plays an essential role in their later mathematical education. On the other hand, I have not, as yet, had the opportunity to test this opinion in a classroom setting, and so I use this occasion to invite the motivated reader to consider this possibility.

Going back to the beginning of our automobile analogy, one may have asked: why concentrate on open intervals of time and space? What happens if one looks at closed intervals, say $a \leq t \leq b$? What we find by considering the various scenarios is, unlike open intervals, if the car moves continuously then its range over a closed time interval is also a closed interval of the road: $p \leq x \leq q$. We leave it to the reader to convince themselves of this fact by looking at particular cases. While it is, in fact, a general property of continuous functions, this criterion is *not* sufficient to distinguish continuous motions from discontinuous motions. For example, consider the car that sits at position $x = y$ and then suddenly jumps to $x = z$. Then the range over any closed time interval is either one or the other or both points, and hence is itself a closed set, albeit not necessarily a closed interval. It is a little harder to construct a discontinuous motion where the image of every closed time interval is a closed road interval, but this can be done; one example can be found in the Appendix.

On the other hand, one can verify that under a continuous motion, the set of times when the car is located in a closed road interval, say $p \leq x \leq q$, is itself a closed interval or, more generally, a closed subset. Indeed, this follows immediately from Definition 1 using the fact that the complement of a closed set is an open set and vice versa. By the same reasoning, this does not happen when the motion is discontinuous. So this could be used as an alternative characterization of continuity, although in point set topology, one traditionally concentrates on the version based on open sets, [**15**, **22**].

Now that we have developed the mathematical basis of continuity, as inspired by the continuity of physical motion, let consider the issue of velocity. To understand it at a basic level, let us return to our car driving along a road. If the car starts out at position $p = x(a)$ at an initial time $t = a$ and, at a subsequent time $t = b$, ends up at position $q = x(b)$, then its average velocity is, by definition, the distance covered divided by the time elapsed:

$$v_{av} = \frac{q - p}{b - a} = \frac{x(b) - x(a)}{b - a}\,. \tag{1}$$

Note that this does not require that the car's motion from point p to point q be monotone; it could for instance overshoot q and subsequently come back to it. Of course, this would require a higher average *speed* to accomplish, but the average velocity remains the same no matter how fast it travels. Since velocity has a sign, the same formula (1) holds whatever the two positions are, so q could be greater than, less than, or equal to p. It also holds when the times are taken in reverse order; in other words, we could allow $b < a$ and the same argument would go through. What we are not allowed to do as yet is to take $b = a$ because both numerator and denominator of (1) would vanish, producing an indeterminate quotient. Here is where the notion of an instantaneous velocity arises. But rather than try to use a limiting procedure — after all we have (purposefully) not even attempted to define what a limit is yet — we use continuity to define the instantaneous velocity, when it exists. The resulting continuity-based definition of derivative dates back to Carathéodory, cf. [**16**, **23**], albeit in a slightly different formulation.

Definition 2. Let $x(t)$ be continuous. We say that $x(t)$ has a *derivative* at $t = a$ if there exists a real number v^* such that the *difference quotient* function

$$v(t) = \begin{cases} \dfrac{x(t) - x(a)}{t - a}, & t \neq a, \\ v^*, & t = a. \end{cases} \tag{2}$$

is continuous. In this case, we write $x'(a) = v(a) = v^*$.

Given that algebraic combinations of continuous functions remain continuous, where defined, [**15**, **22**, **23**], since $x(t)$ is continuous, the average velocity function $v(t)$ is continuous when $t \neq a$, and hence the only potentially problematic time is $t = a$. In the case when $x(t)$ represents the motion of an object, the derivative v^*, when it exists, can be viewed as the object's instantaneous velocity at $t = a$. Its existence is purely a consequence of the continuity of the amended average velocity function (2) — there is no need to appeal to limits over vanishingly small time intervals! On the other hand, not all continuous functions have a derivative, i.e., not all motions allowed in the physical universe have an instantaneous velocity. For example, the absolute value function $x(t) = |\,t\,|$ corresponding to a car suddenly reversing direction at time $t = 0$ does not have a derivative there. Indeed, the difference quotient function (2) in this case is the sign function,

$$v(t) = \frac{x(t) - x(0)}{t - 0} = \frac{|\,t\,|}{t} = \operatorname{sign} t = \begin{cases} 1, & t > 0, \\ -1, & t < 0, \end{cases}$$

and there is no value v^* that will make it continuous everywhere including $t = 0$. On the other hand, the absolute value function does have a derivative at any $t \neq 0$, namely $\operatorname{sign} t$. A more extreme example is provided by Brownian motion, which is genuinely physical (as opposed to the sudden reversal of a car, which no doubt violates various physical constraints), but does not have a derivative at any point, [**18**].

We note that one can readily adapt the above discussion to functions whose domain is an open subinterval or a more general open subset of the real line. Adaptations to non-open domains are extremely interesting, but that would take us too far field; see [**23**] for extensions that include limits of sequences and the Cauchy convergence criterion.

Finally, we should explain how limits enter into this continuity-based formulation of calculus. The answer is that they are defined using continuity rather than the traditional reverse order. Indeed, we say that a function $x(t)$ has the limit L at a point $t = a$ where it is not necessarily defined if the function

$$\widetilde{x}(t) = \begin{cases} x(t), & t = a, \\ L, & t = a, \end{cases}$$

is continuous, in which case we write

$$\lim_{t \to a} x(t) = L.$$

In particular, if $x(t)$ is continuous and defined at $t = a$, then, almost tautologously,

$$\lim_{t \to a} x(t) = x(a).$$

Again, the latter fact is a *consequence* of the topological definition of continuity, and is emphatically *not* a defining property. Thus, one can write the instantaneous velocity v^* in (2) as the limit of the average velocities:

$$v^* = \lim_{t \to a} v(t) = \lim_{t \to a} \frac{x(t) - x(a)}{t - a}.$$

But, in the present framework, this is merely an alternative way of writing the continuity-based Definition 2. As before, no epsilons and deltas are required!

To me, this new continuity-based foundation for calculus — and hence mechanics and physics — represents a compelling alternative to the 200 year old Bolzano–Cauchy–Weierstrass limit-based tradition, and one that successfully overcomes Russell's above quoted qualms. As we have demonstrated, it is founded on very basic observations of the physical world, and avoids complicated and unintuitive mathematical artifices such as epsilons and deltas or, for that matter, the hyperreals and infinitesimals. Moreover, this novel approach may inspire new philosophical lines of inquiry into the nature of motion, velocity, and so on, that have heretofore been forced to rely on the traditional definitions of limit and continuity. It would also be of great interest to see if this alternative foundation has any pedagogical impacts on the teaching of calculus.

Appendix A. A Non-Elementary Example of a Discontinuous Function.

The following construction is not essential to our text, but may be of interest to the more mathematically inclined reader.

We first remark that a commonly used mathematical example of a discontinuous function $x(t)$ that satisfies both the Intermediate Value Property and the property that the image of any closed interval is a closed interval is the trigonometric function $x(t) = \sin(1/t)$ for $t \neq 0$ and $x(0) = 0$. However, since we wish to avoid assuming any familiarity with trigonometry, let us construct a more elementary example, and then show that it has the stated properties†.

---

† Bolzano's and Weierstrass' nowhere differentiable functions, [**18**, **26**], are similarly related; the former is entirely elementary, whereas the latter requires trigonometry.

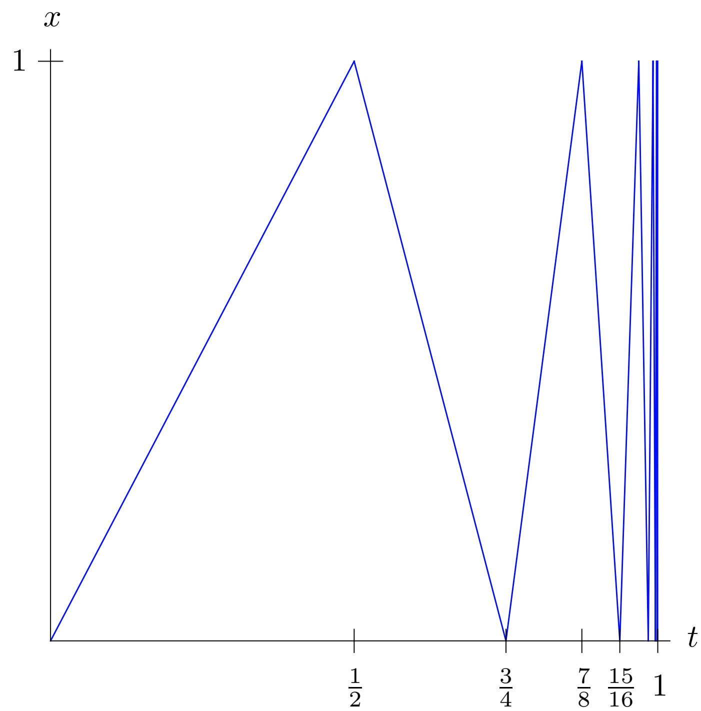

Figure 5. A Discontinuous Motion.

We again appeal to the automobile, and work on the road interval $0 \leq x \leq 1$. Let our car (or, if you prefer, point particle) be at rest at position $x(t) = 0$ outside of the time interval $0 \leq t \leq 1$. At time $t = 0$, the car starts out at $x(0) = 0$ and then drives with constant speed to reach the other endpoint at time $t = 1/2$, so that $x(1/2) = 1$. It then instantaneously† reverses direction and drives back to the start, again at constant (much faster) speed, reaching it at time $t = 3/4$, so $x(3/4) = 0$. It then reverses direction again and drives back to the far endpoint, at an even faster constant speed, reaching it when $t = 7/8$, so $x(7/8) = 1$. This process of driving back and forth repeats ad infinitum on shorter and shorter time intervals, the car reaching astronomical speeds as the time $t \to 1$, as it becomes, to the observer, a blur racing back and forth infinitely often between the endpoints of the road. Finally, at the end of this amazing journey it ends up back where it began, so $x(1) = 0$ and stays there, exhausted, from then on. More explicitly, for any natural number $k = 0, 1, 2, 3, \ldots$, we set

$$x\,(1 - 2^{-2k}) = 0, \qquad x\,(1 - 2^{-2k-1}) = 1,$$

and hence, by our assumption of constant speed between the instants when it reverses

† It is possible to smooth out the motion to avoid discontinuities in the velocity, acceleration, etc., at these points, but the formulas become more complicated, and, for such purposes, it is perhaps preferable to return to the smooth trigonometric example instead.

direction,

$$x(t) = \begin{cases} 2^{2k+1}\, t - 2^{2k+1} + 2, & 1 - 2^{-2k} \le t \le 1 - 2^{-2k-1}, \\ -2^{2(k+1)}\, t + 2^{2(k+1)} - 1, & 1 - 2^{-2k-1} \le t \le 1 - 2^{-2(k+1)}, \qquad k = 0, 1, 2, \ \ldots \ . \\ 0, & t = 1, \end{cases}$$

The car's motion is depicted in Figure 5 and is, of course, completely unphysical, the car eventually violating relativistic limits by traveling faster than the speed of light. And, while this motion is continuous on any open time interval that does not include $t = 1$, it is not continuous, either physically or mathematically, at $t = 1$.

To prove that this function does not satisfy the topological Definition 1 of continuity, consider the open interval $S = \{-1 < x < 1\}$. The car is located in S except at the times $t_k = 1 - 2^{-2k-1}$, $k = 0, 1, 2, \ldots$, when it is at the far end of its range: $x(t_k) = 1$. Thus, $x^{-1}(S)$ consists of the open time intervals $\{t < t_0 = 1/2\}$ and $\{t_k < t < t_{k+1}\}$ for all $k \ge 0$, along with the closed interval $\{t \ge 1\}$. The union of all but the last is open, but the entire set $x^{-1}(S)$ is not open: it cannot be written as the union of open intervals. This proves the claim, since to verify that a function is not continuous, it suffices to find one open set S for which $x^{-1}(S)$ is not open.

Let us next explain why, even though $x(t)$ is discontinuous, the range of every closed time interval is a closed road interval. Let the time interval be $c \le t \le d$. If $d \le 0$ or $c \ge 1$, then the range is a single point, namely $x = 0$, which is a closed road interval. If $c = d$, the range is also a single point, namely $x(c)$. Otherwise, if $c < d < 1$ then the continuity of $x(t)$ for $t < 1$ immediately implies the result, although it is not hard to check this directly. The range of most "sufficiently long" time intervals is the entire road $0 \le x \le 1$ because the car has had time to traverse from one end to the other. Short closed time intervals can have shorter ranges, but they are always short closed road intervals. On the other hand, if $c < 1 \le d$, no matter how close c is to 1, the car has been able to go back and forth along the road infinitely often on the intervening time interval $c \le t \le d$, and hence its range is the entire closed road interval $0 \le x \le 1$. This completes our demonstration of the claimed fact.

We further note that $x(t)$ also satisfies the Intermediate Value Property even though it is not continuous. In other words, given any two times $a < b$ when the car is at respective locations $p = x(a)$, $q = x(b)$, and any location z lying between p and q, there exists a time $a < c < b$ such that $x(c) = z$. We leave the verification of this property, which requires checking a few different cases, as an exercise for the reader.

Acknowledgments: I thank Harvey Brown, Jeremy Butterfield, and Samuel Fletcher for references and corrections, and for their many thought-provoking comments on earlier versions. I also thank the referee for commentary that led to several improvements of the text.

References

[**1**] Barbour, J.B., *Absolute of Relative Motion? Vol. 1. The Discovery of Dynamics*, Cambridge University Press, Cambridge, 1989.

[**2**] Berkeley, G., *The Analyst*, J. Tonson, London, 1734.

[**3**] Bouwmeester, D., Pan, J.-W., Mattle, K., Eibl, M., Weinfurter, H., and Zeilinger, A., Experimental quantum teleportation, *Nature* **390** (1997), 575–579.

[**4**] Boyer, C.B., *The History of the Calculus and its Conceptual Development*, Dover Publ., New York, 1949.

[**5**] Butterfield, J., Against *pointillisme* about mechanics, *Brit. J. Phil. Sci.* **57** (2006), 709–753.

[**6**] DiSalle, R., Space and Time: Inertial Frames, *in*: *The Stanford Encyclopedia of Philosophy*, E.N. Zalta, ed., 2020.
`https://plato.stanford.edu/archives/sum2020/entries/spacetime-iframes/`

[**7**] Felscher, W., Bolzano, Cauchy, epsilon, delta, *Amer. Math. Monthly* **107** (2000), 844–862.

[**8**] Grabiner, J.V., Who gave you the epsilon? Cauchy and the origins of rigorous calculus, *Amer. Math. Monthly* **90** (1983), 185–194.

[**9**] Grattan-Guinness, I., Bolzano, Cauchy and the "New Analysis" of the early nineteenth century, *Arch. Hist. Exact Sci.* **6** (1970), 372–400.

[**10**] Harris, J., *Lexicon Technicum*, vol. 2, D. Brown, London, 1710.

[**11**] Hausdorff, F., *Grundzüge der Mengenlehre*, Veit, Leipzig, Germany, 1914.

[**12**] Johnson, D.M., The problem of the invariance of dimension in the growth of modern topology. I, *Arch. Hist. Exact Sci.* **20** (1979), 97–188.

[**13**] Katz, M.G., and Sherry, D.M., Leibniz's laws of continuity and homogeneity, *Notices Amer. Math. Soc.* **59** (2012), 1550–1558.

[**14**] Keisler, H.J., *Elementary Calculus: An Infinitesimal Approach*, 3rd ed., Dover Publ., New York, 2012.

[**15**] Kelley, J.L., *General Topology*, Graduate Texts in Mathematics, vol. 27, Springer–Verlag, New York, 1975.

[**16**] Kuhn, S., The derivative á la Carathéodory, *Amer. Math. Monthly* **98** (1991), 40–44.

[**17**] Lefschetz, S., *Algebraic Topology*, American Math. Soc. Colloquium Publ., vol. 27, New York, 1942.

[**18**] Mandelbrot, B.B., *The Fractal Geometry of Nature*, W.H. Freeman, New York, 1983.

[**19**] Marsden, J., and Weinstein, A.J., *Calculus Unlimited*, Benjamin/Cummings, Menlo Park. Calif., 1981.

[**20**] Moore, G.H., The evolution of the concept of homeomorphism, *Historia Math.* **34** (2007), 333–343.

[**21**] Moore, G.H., The emergence of open sets, closed sets, and limit points in analysis and topology, *Historia Math.* **35** (2008), 220–241.

[**22**] Munkres, J.R., *Topology*, 2nd ed., Prentice–Hall, Inc., Upper Saddle River, N.J., 2000.

[**23**] Olver, P.J., *Continuous Calculus*, Lecture Notes, University of Minnesota, 2020, `http://www.math.umn.edu/∼olver/ln_/cc.pdf`

[**24**] Rovelli, C., and Vidotto, F., *Covariant Loop Quantum Gravity*: *An Elementary Introduction to Quantum Gravity and Spinfoam Theory*, Cambridge University Press, Cambridge, 2014.

[**25**] Robinson, A., *Non-standard Analysis*, Princeton University Press, Princeton, N.J., 1974.

[**26**] Russ, S., Bolzano's analytic programme, *Math. Intelligencer* **14**(3) (1992), 45–53.

[**27**] Russell, B., *The Principles of Mathematics*, 2nd ed., W.W. Norton & Co., New York, 1938.

[**28**] Tooley, M., In defense of the existence of states of motion, *Philosophical Topics* **16** (1988), 225–254.