SJÄLVSTÄNDIGA ARBETEN I MATEMATIK

MATEMATISKA INSTITUTIONEN, STOCKHOLMS UNIVERSITET

Stronger types of continuity

av

Stefan Lilja

2019 - No K17

MATEMATISKA INSTITUTIONEN, STOCKHOLMS UNIVERSITET, 106 91 STOCKHOLM

Stronger types of continuity

Stefan Lilja

Självständigt arbete i matematik 15 högskolepoäng, grundnivå

Handledare: Annemarie Luger

2019

Abstract

In many situations assuming that a function is continuous may not be strong enough but assuming that it is for example continuously differentiable may be too restrictive. The aim of this text is to investigate various notions of smoothness of functions that lie somewhere between continuity and continuous differentiability. These include uniform continuity, Hölder continuity, Lipschitz continuity and absolute continuity. We compile some basic properties of each of them and investigate how these smoothness properties relate to each other. To do this we examine several examples of functions that satisfy some of these conditions, while not satisfying others. We also look at some applications where these types of continuity are useful.

Contents

1 Introduction **3**

2 Continuity and uniform continuity **3**
- 2.1 Continuous functions 5
- 2.2 Uniform continuity 7

3 Hölder continuity and Lipschitz continuity **11**
- 3.1 Hölder continuity 11
- 3.2 Lipschitz continuity 15
  - 3.2.1 Convex functions are locally Lipschitz 17
- 3.3 Inverses and their derivatives 20

4 Two Applications **22**
- 4.1 Initial value problems 22
- 4.2 Limits of Cauchy-type integrals 25

5 Absolute continuity **28**
- 5.1 Basic properties 28
- 5.2 The Lebesgue measure and the Lebesgue integral 33
- 5.3 Bounded variation 35
- 5.4 The fundamental theorem of calculus 37

6 Rectifiable curves **39**

References **44**

1 Introduction

When solving problems in analysis, it is common to make some assumptions that the functions involved are suitably "well behaved" to ensure that solutions exist or that a certain expression converges. Of course, what requirements we need for the function to be "well behaved" will vary greatly from problem to problem. This leads us to formulating many different conditions that functions may or may not satisfy. The most common may be to assume that the function is continuously differentiable a certain number of times. However, sometimes these conditions may not be precise enough leading us to investigating conditions that lie in between.

We begin in chapter 2 by recalling the definition as well as some important properties of continuous functions from elementary calculus. We move on to uniform continuity, a stronger condition which in turn lets us say more of the functions' properties. For example, uniformly continuous functions preserve Cauchy sequences, something continuous functions do not (in general).

In chapter 3 we introduce Hölder continuity and a special case of it called Lipschitz continuity. Among the properties we find is that convex functions satisfy a certain Lipschitz condition. We also find some assumptions that ensure that the inverse of a function satisfies certain Hölder conditions. Chapter 4 is dedicated to two applications of Hölder continuity and Lipschitz continuity. The first involves Lipschitz continuity and initial value problems. We show that if the initial value problem satisfies certain Lipschitz conditions, then the existence as well as uniqueness of the solution is guaranteed. The second uses Hölder continuity in the study of Cauchy-type integrals, a type of complex integral function. We look at the limit of these functions at singularities and present the Sokhotski-Plemelj-formula which details the existence and values of these limits.

In chapter 5 we study absolute continuity and some related properties. We introduce the Lebesgue measure and the Lebesgue integral, and how the Lebesgue integral is connected to absolute continuity through the fundamental theorem of calculus. We also look at another related smoothness property, functions of bounded variation and its basic properties.

Finally, in chapter 6 we study rectifiable curves, curves of finite arc length. We prove a connection between rectifiable curves and the notion of bounded variation and look at examples of both curves that are rectifiable and those that are not. We also introduce an integral formula with which the arc length may be computed.

2 Continuity and uniform continuity

We shall primarily concern ourselves with real valued functions of one real variable. We will denote the domain of a function f by $\mathcal{D}$ or $\mathcal{D}_f$ if multiple functions are involved. Thus $\mathcal{D} \subseteq \mathbb{R}$ unless something else is explicitly specified. The reader is assumed to be familiar with limits, continuity and differentiability in the context of functions from $\mathbb{R}$ into $\mathbb{R}$. Nevertheless many of the concepts and theorems presented in this text can be generalised to for example $\mathbb{R}^n$ using norms.

Definition 2.1. A norm on a real vector space X is a function $\|\cdot\| : X \to \mathbb{R}_{\geq 0}$ satisfying

(i) $\|x\| = 0$ if and only if $x = 0$

(ii) $\|\alpha x\| = |\alpha| \, \|x\|$ for all $\alpha \in \mathbb{R}$ and $x \in X$

(iii) $\|x + y\| \leq \|x\| + \|y\|$ for all $x, y \in X$.

The absolute value function on $\mathbb{R}$ is one example of a norm. By replacing the absolute value with the norm of a space, many of the definitions and theorems below are generalised to that space. It is also possible to extend many of the results into more general metric spaces but that is not the focus of this text.

Another notion we will need is that of Cauchy sequences.

Definition 2.2. A sequence $\{x_n\}$ in $\mathbb{R}$ is called a Cauchy sequence if for every $\epsilon > 0$ there is $N \in \mathbb{N}$ such that

$$m, n > N \Longrightarrow |x_m - x_n| < \epsilon.$$

Cauchy sequences are closely related to convergent sequences. Every convergent sequence is a Cauchy sequence (in any metric space) and every Cauchy sequence in $\mathbb{R}$ converges in $\mathbb{R}$. A space in which all Cauchy sequences converge is called (Cauchy) complete. The main benefit of using Cauchy sequences is that it allows us to look at the sequence without speaking about the limit point. So if the domain $\mathcal{D}$ of a function f is not closed and $\{x_n\}_n$ is a Cauchy sequence in $\mathcal{D}$, then we don't need to worry about whether $\{x_n\}_n$ has a limit in $\mathcal{D}$ when studying the sequence $\{f(x_n)\}_n$.

We end this section with some observations from topology that will be very useful going forward. For further discussion and proofs, we refer to [6, p. 30-40].

Definition 2.3. For subsets $X \subseteq Y \subseteq \mathbb{R}$ we say that X is open in Y if for every $x \in X$ there is $r > 0$ such that $\{y \in Y : |y - x| < r\} \subseteq X$. We also say that X is closed in Y if every limit point of X that lies in Y also lies in X. In the case when $Y = \mathbb{R}$ we shorten this to just saying that X is open (or closed).

Definition 2.4. A family $\{A_i\}_{i \in I}$ of open sets in a metric space M is called an open cover of the set $X \subseteq M$ if $X \subseteq \bigcup_{i \in I} A_i$. We say that X is compact if every open cover of X has a finite subcover, that is if $\{A_i\}_{i \in I}$ is an open cover of X, then there is a finite subset $J \subseteq I$ such that $X \subseteq \bigcup_{i \in J} A_i$.

Theorem 2.5.

(i) A subset $X \subseteq \mathbb{R}$ is compact if and only if it is closed and bounded.

(ii) Every closed subset of $\mathbb{R}$ is complete.

(iii) For subsets $X \subseteq Y \subseteq \mathbb{R}$, X is open in Y if and only if the complement $X^c = Y \backslash X$ is closed in Y.

2.1 Continuous functions

As previously mentioned the reader is assumed to be familiar with continuous functions but we will go through the most important definitions and properties here as we will compare them to other, similar notions. For more details as well as the proofs, we refer to [6, Ch. 4] and [1, Ch. 5.1].

Definition 2.6. A function f is continuous at a point $x_0 \in \mathcal{D}$ if for every $\epsilon > 0$ there is $\delta > 0$ such that if $x \in \mathcal{D}$ and $|x - x_0| < \delta$, then $|f(x) - f(x_0)| < \epsilon$. We say that f is continuous on a subset $I \subseteq \mathcal{D}$ if it is continuous at every point in the set I.

We note that in the above definition the choice of δ will in general depend on ϵ. The number δ may also depend on the point x_0, we shall elaborate on this in the next section.

Definition 2.7. The open ball centered at a point $y \in \mathbb{R}$ with radius $r \in \mathbb{R}$ is defined as the set $\{x \in \mathbb{R} : |x - y| < r\}$ and we will denote it by $B_r(y)$.

This allows us to write the definition above in a more compact way: f is continuous at $x_0 \in \mathcal{D}$ if

$$\forall \epsilon > 0 \, \exists \delta > 0 : x \in \mathcal{D} \cap B_\delta(x_0) \implies f(x) \in B_\epsilon(f(x_0)).$$

The continuity of a function can also be defined in other ways, by the following theorem.

Theorem 2.8. *Let f be a function from $\mathcal{D} \subseteq \mathbb{R}$ into $\mathbb{R}$. Then the following are equivalent:*

(i) f is continuous on $\mathcal{D}$.

(ii) For every $x \in \mathcal{D}$, if $\{x_n\}_n$ is a sequence such that $x_n \to x$ and $x_n \in \mathcal{D}$ for every $n \in \mathbb{N}$, then $f(x_n) \to f(x)$.

(iii) For any set A open in $\mathbb{R}$, the set $f^{-1}(A)$ is open in $\mathcal{D}$.

(iv) For any set A closed in $\mathbb{R}$, the set $f^{-1}(A)$ is closed in $\mathcal{D}$.

Proof. (i) $\Rightarrow$ (ii) : Suppose f is continuous. Let $\{x_n\}_n$ be a sequence such that $x_n \to x$ where $x \in \mathcal{D}$ and $x_n \in \mathcal{D}$ for every $n \in \mathbb{N}$ and let $\epsilon > 0$ be given. Since f is continuous there is $\delta > 0$ such that

$$y \in \mathcal{D} \cap B_\delta(x) \implies f(y) \in B_\epsilon(f(x)).$$

Since $x_n \to x$ there is $N \in \mathbb{N}$ so that $x_n \in \mathcal{D} \cap B_\delta(x)$ for $n \geq N$. But then $f(x_n) \in B_\epsilon(f(x))$ for $n \geq N$ so $f(x_n) \to f(x)$.

(ii) $\Rightarrow$ (i) : Suppose (ii) holds, $x \in \mathcal{D}$ and that $\{x_n\}_n$ is a sequence in $\mathcal{D}$ such that $x_n \to x$. Now suppose for a contradiction that f is not continuous at x, that is

$$\exists \epsilon > 0 \forall \delta > 0 \exists y \in \mathcal{D} \cap B_\delta(x) : f(y) \notin B_\epsilon(f(x)).$$

If we pick such an ϵ, then for every $n \in \mathbb{N}$, $\delta_n = \frac{1}{n}$ satisfies

$$\exists y \in \mathcal{D} \cap B_{\delta_n}(x) : f(y) \notin B_\epsilon(f(x)).$$

For every $n \in \mathbb{N}$ set x_n to be such a value y. We then have $|x_n - x| < \delta_n = \frac{1}{n} \to 0$ when $n \to \infty$ so $x_n \to x$. By (ii) we then have $f(x_n) \to f(x)$. However for every $n \in \mathbb{N}$ we have $|f(x_n) - f(x)| \geq \epsilon$ so we have a contradiction. Thus f must be continuous at x. Since x was chosen arbitrarily in $\mathcal{D}$, f is continuous on all of $\mathcal{D}$.

$\underline{(i) \Rightarrow (iii)}$: Suppose f is continuous and let A be a set that is open in $\mathbb{R}$. If $f^{-1}(A) = \emptyset$ then it is open. Otherwise, suppose $x_0 \in f^{-1}(A)$. Then $f(x_0) \in A$ and since A is open, there is $\epsilon > 0$ such that $B_\epsilon\big(f(x_0)\big) \subseteq A$. Since f is continuous there is then $\delta > 0$ such that $x \in B_\delta(x_0)$ implies $f(x) \in B_\epsilon\big(f(x_0)\big)$. Thus x_0 is an interior point of $f^{-1}(A)$ and hence $f^{-1}(A)$ is open.

$\underline{(iii) \Rightarrow (i)}$: Suppose that $f^{-1}(A)$ is open in $\mathcal{D}$ for every A open in $\mathbb{R}$. Fix $\epsilon > 0$ and let $x_0 \in \mathcal{D}$. Since $B_\epsilon(f(x_0))$ is open in $\mathbb{R}$, $f^{-1}(B_\epsilon(f(x_0)))$ must be open in $\mathcal{D}$. Thus there is $\delta > 0$ such that $B_\delta(x_0) \subseteq f^{-1}(B_\epsilon(f(x_0)))$. Now if $x \in B_\delta(x_0)$, we have $f(x) \in B_\epsilon(f(x_0))$ so f is continuous at x_0. Since x_0 was chosen arbitrarily in $\mathcal{D}$ it follows that f is continuous on $\mathcal{D}$.

$\underline{(iii) \Leftrightarrow (iv)}$: We note that

$$f^{-1}(A^c) = \{x \in \mathcal{D} | f(x) \in A^c\} = \{x \in \mathcal{D} | f(x) \in A\}^c = \big(f^{-1}(A)\big)^c.$$

By Theorem 2.5 we then have

$$A \text{ open } \Leftrightarrow A^c \text{ closed}$$

(with respect to $\mathbb{R}$) and

$$f^{-1}(A) \text{ open } \Leftrightarrow f^{-1}(A^c) \text{ closed}$$

(with respect to $\mathcal{D}$) and the statement follows. □

Note that (iii) and (iv) above are not the same as saying that continuous functions preserve the openness or closedness of sets, in fact this is not the case. The function e^{-x^2} is continuous on $\mathbb{R}$ which is both open and closed but the image of the function is $(0, 1]$ which is neither open nor closed. Continuity does however preserve compactness, by the following.

Theorem 2.9. *Let $K \subset \mathbb{R}$ be a compact set and $f : K \to \mathbb{R}$ a continuous function. Then $f(K)$ is compact.*

Proof. Let $\{V_\alpha\}_\alpha$ be an open cover of $f(K)$. By the above theorem, the sets $f^{-1}(V_\alpha)$ are open so $\{f^{-1}(V_\alpha)\}_\alpha$ is an open cover of K. Since K is compact there is a finite subcover $\{f^{-1}(V_{\alpha_i})\}_{i=1}^n$ of K. But then

$$f(K) \subseteq \bigcup_{i=1}^n f(f^{-1}(V_{\alpha_i})) \subseteq \bigcup_{i=1}^n V_{\alpha_i}$$

so $\{V_{\alpha_i}\}_{i=1}^n$ is a finite subcover of $f(K)$ which subsequently is compact. □

2.2 Uniform continuity

As was mentioned previously, in the definition of continuity the constant δ may depend on the chosen point x_0. Doing away with this dependence gives us the following definition.

Definition 2.10. A function f is uniformly continuous on $I \subseteq \mathcal{D}$ if for every $\epsilon > 0$ there is $\delta > 0$ such that for all $x, y \in I$,

$$|x - y| < \delta \Longrightarrow |f(x) - f(y)| < \epsilon.$$

Clearly any uniformly continuous function is also continuous in the same subset of the domain. A function that is continuous may be uniformly continuous on some subsets of its domain while not being so on other subsets, as the following example demonstrates.

Example 2.11. Consider the function $f(x) = \frac{1}{x}$ defined on an interval $[a, \infty)$ where $a \in \mathbb{R}_{>0}$. Then for any $\epsilon > 0$ take $\delta = a^2\epsilon$. Then for $x, y \in [a, \infty)$ such that $|y - x| < \delta$ we have

$$|f(x) - f(y)| = \left|\frac{1}{x} - \frac{1}{y}\right| = \frac{|y - x|}{xy} < \frac{\delta}{xy} = \frac{a^2\epsilon}{xy} \leq \epsilon.$$

Thus f is uniformly continuous on $[a, \infty)$. However, if f instead is defined on $(0, \infty)$ then this will not hold. To see this suppose for a contradiction that f was uniformly continuous on $(0, \infty)$. Then for $\epsilon = 1$ there is $\delta > 0$ such that

$$|x - y| < \delta \Longrightarrow \left|\frac{1}{x} - \frac{1}{y}\right| < 1.$$

Now let $x = \frac{1}{n+1}, y = \frac{1}{n}$ for some positive integer n. We then have

$$|x - y| = \left|\frac{1}{n+1} - \frac{1}{n}\right| = \frac{1}{n(n+1)} < \delta$$

if we choose n sufficiently large. On the other hand we have

$$\left|\frac{1}{x} - \frac{1}{y}\right| = |n + 1 - n| = 1$$

which contradicts that f is uniformly continuous on $(0, \infty)$.

There are however some sets where the notions of continuity and uniformly continuity coincide, as detailed in the following theorem. This is one of the most useful methods for showing that a function is uniformly continuous.

Theorem 2.12. *If f is continuous on a compact set $K \subseteq \mathcal{D}$, then f is uniformly continuous on K.*

Proof. Let $\epsilon > 0$ be fixed. By hypothesis, we have for every $x_0 \in K$ a $\delta(x_0) > 0$ such that

$$|x - x_0| < \delta(x_0) \Longrightarrow |f(x) - f(x_0)| < \frac{\epsilon}{2}.$$

For each $x_0 \in K$, $x_0 \in B_{\frac{\delta(x_0)}{2}}(x_0)$ so $\left\{B_{\frac{\delta(x)}{2}}(x)\right\}_{x \in K}$ is an open cover of K. Since K is compact, there is a finite subcover $\left\{B_{\frac{\delta(x_k)}{2}}(x_k)\right\}_{k=1}^{n}$, where $x_k \in K$ for $k = 1, 2, \ldots, n$. Let $\delta = \min_{k=1,\ldots,n}\left(\frac{\delta(x_k)}{2}\right)$. Then for $x, y \in K$ such that $|x - y| < \delta$ we can find $i \in \{1, 2, \ldots, n\}$ such that $x \in B_{\frac{\delta(x_i)}{2}}(x_i)$. Then

$$|y - x_i| \leq |x - y| + |x - x_i| < \delta + \delta(x_i)/2 \leq \delta(x_i)$$

so that $y \in B_{\frac{\delta(x_i)}{2}}(x_i)$. It follows that

$$|f(x) - f(y)| \leq |f(x) - f(x_i)| + |f(y) - f(x_i)| < \epsilon.$$

so f is uniformly continuous on K. □

We saw earlier in Theorem 2.8 that one of the characterisations of continuity is a form of preservation of convergent sequences, under certain conditions. One of these requirements was that the limit of the sequence must be in the domain of the function. If this is not the case, then the function being uniformly continuous will not help us. This can be somewhat remedied using the notion of Cauchy sequences.

One example that shows that continuity is insufficient to preserve Cauchy sequences is the sequence $x_n = \frac{1}{n}$ for $n \in \mathbb{N}$, which is a Cauchy sequence in $(0, 1]$, together with the function $f(x) = \frac{1}{x}$ which of course is continuous on $(0, 1]$. The function gives us a new sequence $f(x_n) = n$ which is not a Cauchy sequence. It is here that we will need uniform continuity.

Theorem 2.13. *Let f be uniformly continuous on $\mathcal{D}$. If $\{x_n\}_n$ is a Cauchy sequence in $\mathcal{D}$ then $\{f(x_n)\}_n$ is a Cauchy sequence in $\mathbb{R}$.*

Proof. If f is uniformly continuous on $\mathcal{D}$ and $\epsilon > 0$ is given, then there is $\delta > 0$ such that, for $x, y \in \mathcal{D}$,

$$|x - y| < \delta \Longrightarrow |f(x) - f(y)| < \epsilon.$$

Since $\{x_n\}_n$ is a Cauchy sequence, there is $N \in \mathbb{N}$ such that

$$m, n \geq N \Longrightarrow |x_n - x_m| < \delta.$$

Thus we have

$$m, n \geq N \Longrightarrow |x_n - x_m| < \delta \Longrightarrow |f(x_n) - f(x_m)| < \epsilon$$

so $\{f(x_n)\}_n$ is a Cauchy sequence in $\mathbb{R}$. □

Another property that is preserved by uniform continuity is boundedness.

Theorem 2.14. *If $\mathcal{D}$ is bounded and f is uniformly continuous on $\mathcal{D}$, then $f(\mathcal{D})$ is bounded.*

Proof. Since f is uniformly continuous on $\mathcal{D}$ we can find $\delta > 0$ such that

$$x, y \in \mathcal{D} \text{ and } |x - y| < \delta \Longrightarrow |f(x) - f(y)| < 1.$$

Now assume for a contradiction that $f(\mathcal{D})$ is not bounded. Then we could find a sequence $\{x_n\}$ in $\mathcal{D}$ such that $|f(x_n)| \geq n$ for all $n \in \mathbb{N}$. Since $\{x_n\}$ is bounded it has a subsequence that converges to some $x \in \overline{\mathcal{D}}$. Thus $B'_{\frac{\delta}{2}}(x)$ contains infinitely many terms of the sequence $\{x_n\}$. Let x_J be one such point in the sequence. Then $|x_n - x_J| < \delta$ is satisfied by infinitely many values of n. Since f is uniformly continuous, $|f(x_n) - f(x_J)| < 1$ is also satisfied by infinitely many values of n. Thus $|f(x_n)| < 1 + |f(x_J)|$ for infinitely many values of n which contradicts that $|f(x_n)| \geq n$ for all $n \in \mathbb{N}$. The function f must therefore be bounded. □

The function $x \mapsto \frac{1}{x}$ defined on the bounded set $(0, 1)$ has image $(1, \infty)$ and thus shows that the function being continuous is not sufficient for preserving boundedness.

Much like in the case of continuous functions, uniform continuity is preserved under compositions of uniformly continuous functions. Indeed if g is uniformly continuous on $I \subseteq \mathcal{D}_g$, $g(I) \subseteq \mathcal{D}_f$ and f is uniformly continuous on $g(I)$, then for every $\epsilon > 0$ there is $\delta > 0$ such that for $x, y \in \mathcal{D}_f$,

$$|x - y| < \delta \Longrightarrow |f(x) - f(y)| < \epsilon$$

holds. Furthermore there is then $\gamma > 0$ such that for $z, x \in \mathcal{D}_g$ we have

$$|z - w| < \gamma \Longrightarrow |g(z) - g(w)| < \delta.$$

But then for for $z, w \in \mathcal{D}_g$ we have

$$|z - w| < \gamma \Longrightarrow |f(g(z)) - f(g(w))| < \epsilon.$$

so $f \circ g$ is uniformly continuous.

Uniform continuity is also preserved under algebraic operations, at least under some conditions of boundedness.

Theorem 2.15. *Let f and g be uniformly continuous on I. Then*

(i) *$f + g$ is uniformly continuous on I*

(ii) *if f, g are bounded, then fg is uniformly continuous on I*

(iii) *if f, g are bounded and $\inf |g| > 0$, then $\frac{f}{g}$ is uniformly continuous on I.*

Proof. Let $\epsilon > 0$ be fixed.

(i) Since f and g are uniformly continuous on I, there is $\delta > 0$ such that for any $x, y \in I$ both

$$|x - y| < \delta \Longrightarrow |f(x) - f(y)| < \frac{\epsilon}{2}$$

and

$$|x - y| < \delta \Longrightarrow |g(x) - g(y)| < \frac{\epsilon}{2}$$

hold. Then

$$|x-y| < \delta \Longrightarrow |f(x)+g(x)-f(x)-g(y)| \leq |f(x)-f(y)|+|g(x)-g(y)| < \epsilon$$

so $f + g$ is uniformly continuous.

(ii) Since f and g are uniformly continuous on I, there is $\delta > 0$ such that for any $x, y \in I$ both

$$|x - y| < \delta \Longrightarrow |f(x) - f(y)| < \frac{\epsilon}{2 \sup |g|}$$

and

$$|x - y| < \delta \Longrightarrow |g(x) - g(y)| < \frac{\epsilon}{2 \sup |f|}$$

hold. Then

$$\begin{aligned} &|x - y| < \delta \Longrightarrow |(fg)(x) - (fg)(y)| = \\ &= |f(x)g(x) - f(y)g(x) + f(y)g(x) - f(y)g(y)| \leq \\ &\leq |f(x) - f(y)||g(x)| + |g(x) - g(y)||f(y)| < \\ &< \frac{\epsilon}{2 \sup |g|} \sup |g| + \frac{\epsilon}{2 \sup |f|} \sup |f| = \epsilon \end{aligned}$$

so fg is uniformly continuous.

(iii) Since g is bounded and $\inf |g| > 0$, the function $\frac{1}{g}$ is a composition of uniformly continuous functions and hence itself uniformly continuous. But then $\frac{f}{g}$ is the product of two uniformly continuous functions so it is also uniformly continuous by part (ii).

□

To see that the boundedness in (ii) is necessary, consider the case $f(x) = g(x) = x$. Clearly both functions are uniformly continuous on $\mathbb{R}$ but the product x^2 is not. Note that two points $x \in \mathbb{R}$ and $y = x + \frac{1}{n}$ where n is an integer will satisfy

$$|x - y| = \frac{1}{n} < \delta$$

for any $\delta > 0$ if we just choose n big enough. On the other hand

$$|x^2 - y^2| = |\frac{2x}{n} + \frac{1}{n^2}| > \epsilon$$

for any $\epsilon > 0$ if we just choose x big enough. Hence x^2 is not uniformly continuous on $\mathbb{R}$. To see that the requirement $\inf |g| > 0$ is necessary we need only look at Example 2.11.

Another thing worth mentioning about uniform continuity is its relation to differentiability. While a function being differentiable at a point implies that the function is continuous at that point, a function may be differentiable on a set but not uniformly continuous or uniformly continuous but not differentiable. To see the former, take the function $x \mapsto \frac{1}{x}$ defined on $(0, \infty)$. It is differentiable in its domain but as we saw in Example 2.11 it is not uniformly continuous. To see the latter, look at the function $x \mapsto |x|$ defined on $[-1, 1]$. Since the domain is compact it follows by Theorem 2.12 that it is uniformly continuous but it is not differentiable at $x = 0$.

3 Hölder continuity and Lipschitz continuity

3.1 Hölder continuity

In this chapter we will take a look at Hölder continuity and its properties. The theory is based on [3, Ch. 1].

Definition 3.1. A function f is Hölder continuous on $I \subseteq \mathcal{D}$ if there exist $H \in \mathbb{R}_{\geq 0}$ and $\alpha \in (0, 1]$ such that for any $x, y \in I$,

$$|f(x) - f(y)| \leq H|x - y|^\alpha \tag{1}$$

or equivalently if

$$\sup_{\substack{x,y \in I \\ x \neq y}} \frac{|f(x) - f(y)|}{|x - y|^\alpha} < \infty. \tag{2}$$

Any H that satisfies the relation (1) is called a Hölder constant (or coefficient) of f. If such a number H exists, then the supremum in (2) gives us the smallest constant that satisfies (1). It is called the minimal Hölder constant and we will denote it by $[f]_\alpha$. We also specify that f is Hölder continuous with exponent (or order) α or more succinctly, f is α-Hölder continuous.

Remark 3.2. In some literature the above definition may allow $\alpha > 1$. We choose not to because if $\alpha > 1$ in the above expression, then if I is for example an interval and $x \neq y$ we get

$$\frac{|f(x) - f(y)|}{|x - y|} \leq H|x - y|^{\alpha - 1} \to 0 \text{ when } x \to y$$

so f is differentiable everywhere on I with derivative 0 everywhere, and hence it is constant. Note that this result depends on the structure of $\mathbb{R}$ as a metric space, it does not hold for metric spaces in general. An example is any space equipped with the discrete metric

$$d(x,y) = \begin{cases} 1, & x \neq y \\ 0, & x = y. \end{cases}$$

The corresponding condition for Hölder continuity is

$$d(f(x), f(y)) \leq H d(x,y)^{\alpha}.$$

For simplicity we set $H = 1$. If x and y are equal, then both sides of the equation are zero. If x and y are different then the right hand side is 1 and the left is either 0 or 1. Thus the inequality holds for any function f and any $\alpha > 0$. However there are of course functions that are not constant. We also restrict α to be strictly positive since setting $\alpha = 0$ would reduce the condition to saying that f is bounded on I.

Example 3.3. Consider $f(x) = x^{\beta}$ for some $\beta \in (0,1]$ defined on $[0,\infty)$. Suppose $0 \leq y < x$. We then get

$$|f(x)-f(y)| = x^{\beta} - y^{\beta} = \int_y^x \beta t^{\beta-1} dt \leq \int_y^x \beta(t-y)^{\beta-1} dt = (x-y)^{\beta} = |x-y|^{\beta}.$$

The case where $x < y$ is analogous and the inequality holds trivially if $x = y$. Thus f is β-Hölder continuous with constant 1. Note that for any Hölder constant H of f, setting $y = 0$ and $x = 1$ in Definition 3.1 gives us $1 \leq H$, showing that 1 is the minimal Hölder constant of f.

Hölder continuity is stronger than uniform continuity in the following sense.

Theorem 3.4. *If f is α-Hölder continuous with constant H on $I \subseteq \mathcal{D}$, then f is uniformly continuous on I.*

Proof. Fix an $\epsilon \in \mathbb{R}_{>0}$. Set $\delta = \left(\frac{\epsilon}{H}\right)^{\frac{1}{\alpha}}$. Then for any $x, y \in I$ such that $|x-y| < \delta$ we have

$$|f(x) - f(y)| \leq H|x-y|^{\alpha} < H\delta^{\alpha} = H\frac{\epsilon}{H} = \epsilon$$

□

That the converse does not hold is illustrated by the next example.

Example 3.5. Let

$$f(x) = \begin{cases} 0 \text{ if } x = 0 \\ \frac{1}{\log(x)} \text{ if } x \in (0, \frac{1}{2}] \end{cases}.$$

Since f is the composition of two continuous functions it is continuous on $(0, \frac{1}{2}]$. Furthermore $\lim_{x\to 0^+} f(x) = 0$ so that f is continuous on $[0, \frac{1}{2}]$. Since $[0, \frac{1}{2}]$ is a compact set it follows from Theorem 2.12 that f is uniformly continuous on $[0, \frac{1}{2}]$.

Now we assume that f is Hölder continuous, that is we assume there is $H \geq 0$ and $\alpha \in (0, 1]$ such that $|f(x) - f(y)| \leq H|x - y|^\alpha$ for all $x, y \in [0, \frac{1}{2}]$. Now take $y = 0$ and $x \neq 0$. We get

$$|f(x)| \leq H|x|^\alpha \Longrightarrow \left|\frac{1}{x^\alpha \log(x)}\right| \leq H.$$

However the expression

$$\left|\frac{1}{x^\alpha \log(x)}\right|$$

goes to infinity when x goes to zero for every $\alpha > 0$ so no such numbers α and H can exist, giving us a contradiction.

On bounded sets the Hölder exponents also invoke an order of inclusion on the classes of Hölder continuous functions.

Theorem 3.6. *If f is β-Hölder continuous on $\mathcal{D}$ and $\mathcal{D}$ is bounded, then for every α such that $0 < \alpha < \beta$, f is also α-Hölder continuous on $\mathcal{D}$.*

Proof. If f is β-Hölder continuous, then for some $H \geq 0$,

$$|f(x) - f(y)| \leq H|x - y|^\beta = H|x - y|^{\beta-\alpha}|x - y|^\alpha.$$

Since $\mathcal{D}$ is bounded, the factor $|x - y|^{\beta-\alpha}$ is bounded, say by H'. Setting $M = H \cdot H'$ we then have

$$|f(x) - f(y)| \leq M|x - y|^\alpha$$

so f is α-Hölder continuous. □

In Example 3.3 we saw that $f(x) = \sqrt{x}$ is $\frac{1}{2}$-Hölder continuous on $[0, \infty)$. If $0 < \alpha < \frac{1}{2}$, then f is not α-Hölder continuous. To show this, assume for a contradiction that it is. Then there is $H \geq 0$ such that for all $x, y \in [0, \infty)$

$$|\sqrt{x} - \sqrt{y}| \leq H|x - y|^\alpha.$$

Setting $y = 0$, we get $x^{\frac{1}{2}-\alpha} \leq H$. When x goes to infinity the left hand side also goes to infinity, so we have a contradiction. This demonstrates that the boundedness criterion in the previous theorem is essential.

Example 3.7. We return to the function $f(x) = x^\beta$ from Example 3.3 but restricting the domain to $[0, 1]$. If $\alpha \in (0, \beta]$, then the previous theorem tells us that f is also α-Hölder continuous on [0,1] since the interval is bounded. However if $\beta < \alpha \leq 1$ then f is not α-Hölder continuous. To see this, assume for a contradiction that there is $H > 0$ such that

$$|x^\beta - y^\beta| \leq H|x - y|^\alpha$$

for $x, y \in [0,1]$. Let $y = 0$ and $x > 0$. We then have

$$|x^\beta| \leq H|x|^\alpha \implies x^{\beta-\alpha} \leq H.$$

However, since $\beta - \alpha < 0$, for any $H \geq 0$, we can choose x sufficiently close to 0 so that $x^{\beta-\alpha} > H$. Thus we have a contradiction.

Having found some Hölder continuous functions, we might ask if combining them will yield new Hölder continuous functions. We first take a look at compositions, which do indeed preserve the property of being Hölder continuous though the Hölder exponent may change in the process.

Theorem 3.8. *If g is β-Hölder continuous on $I \subseteq \mathcal{D}_g$, $g(I) \subseteq \mathcal{D}_f$ and f is α-Hölder continuous on $g(I)$, then the composition $f \circ g$ is $\alpha\beta$-Hölder continuous on I and*

$$[f \circ g]_{\alpha\beta} \leq [f]_\alpha([g]_\beta)^\alpha.$$

Proof. We have

$$|f(g(x)) - f(g(y))| \leq [f]_\alpha|g(x) - g(y)|^\alpha \leq [f]_\alpha([g]_\beta)^\alpha|x - y|^{\alpha\beta}$$

for all $x, y \in I$, proving the theorem. □

Hölder continuity is also preserved under algebraic operators, at least if the functions have the same Hölder exponent and meet certain boundedness conditions. If the functions have different Hölder exponents, then Theorem 3.6 may be of use to find a common Hölder exponent.

Theorem 3.9. *Let f and g be α-Hölder continuous on I. Then*

(i) *$f + g$ is α-Hölder continuous on I and $[f + g]_\alpha \leq [f]_\alpha + [g]_\alpha$*

(ii) *if f, g are bounded, then fg is α-Hölder continuous on I and $[fg]_\alpha \leq [f]_\alpha \sup|g| + [g]_\alpha \sup|f|$*

(iii) *if f, g are bounded and $\inf|g| > 0$, then $\frac{f}{g}$ is α-Hölder continuous on I and $\left[\frac{f}{g}\right]_\alpha \leq \frac{[f]_\alpha \sup|g| + [g]_\alpha \sup|f|}{(\inf|g|)^2}$.*

Proof.

(i) For $x, y \in I$ we get

$$\begin{aligned} &|(f + g)(x) - (f + g)(y)| \leq |f(x) - f(y)| + |g(x) - g(y)| \leq \\ &\leq [f]_\alpha|x - y|^\alpha + [g]_\alpha|x - y|^\alpha = ([f]_\alpha + [g]_\alpha)|x - y|^\alpha. \end{aligned}$$

(ii) For $x, y \in I$

$$\begin{aligned} &|(fg)(x) - (fg)(y)| = |f(x)g(x) - f(y)g(x) + f(y)g(x) - f(y)g(y)| \leq \\ &\leq |f(x) - f(y)||g(x)| + |g(x) - g(y)||f(y)| \leq \\ &\leq ([f]_\alpha \sup|g| + [g]_\alpha \sup|f|)|x - y|^\alpha \end{aligned}$$

where the suprema are finite since f and g are bounded.

(iii) For $x, y \in I$

$$\begin{aligned}
&\left|\frac{f}{g}(x) - \frac{f}{g}(y)\right| = \frac{|f(x)g(y) - f(y)g(x)|}{|g(x)g(y)|} = \\
&= \frac{|f(x)g(y) - f(x)g(x) + f(x)g(x) - f(y)g(x)|}{|g(x)g(y)|} \leq \\
&\leq \frac{|f(x) - f(y)||g(x)| + |g(y) - g(x)||f(x)|}{|g(x)g(y)|} \leq \\
&\leq \frac{[f]_\alpha \sup|g| + [g]_\alpha \sup|f|}{(\inf|g|)^2}|x-y|^\alpha
\end{aligned}$$

where again the suprema are finite since f, g are bounded and $\frac{1}{\inf|g|}$ is finite since $\inf|g| > 0$.

□

To motivate the boundedness condition in (ii) above, consider $f(x) = g(x) = x$ which are both 1-Hölder continuous on $\mathbb{R}$. Their product $h(x) = x^2$ however is not 1-Hölder continuous. Indeed, if $H \geq 0$ were a Hölder constant of h, then setting $y = 0$ in Definition 3.1 we would get

$$|x^2| \leq H|x| \implies |x| \leq H$$

for every $x \in \mathbb{R}$, which clearly is not true.

For the quotient in (iii) to make sense we of course need to assume that $g \neq 0$. This is however not strong enough. Consider $f(x) = 1$ and $g(x) = x$ defined on $(0, 1)$. They are both 1-Hölder continuous, bounded, and $g \neq 0$ but the quotient $\frac{1}{x}$ is as we have seen earlier (Example 2.11) not even uniformly continuous on (0,1) and hence cannot be Hölder continuous.

3.2 Lipschitz continuity

A case of Hölder continuity that is of special interest is when $\alpha = 1$. We therefore give it its own name.

Definition 3.10. A function f is Lipschitz continuous on $I \subseteq \mathcal{D}$ if there exist $L \in \mathbb{R}_{\geq 0}$ such that for any $x, y \in I$,

$$|f(x) - f(y)| \leq L|x - y|.$$

Much like för Hölder continuity, if L satisfies the relation above it is known as a Lipschitz constant of f, and the smallest such L (denoted by $[f]_1$) is known as the minimal Lipschitz constant. If $[f]_1 < 1$ we call f a contraction. We say that a function is locally Lipschitz on a set A if it is Lipschitz continuous on every compact subset of A.

Since it is a special case it of course holds that all Lipschitz continuous functions are also Hölder continuous and the theorems about Hölder continuous functions we have proven also hold for Lipschitz continuous functions. That not all Hölder continuous function are Lipschitz continuous is illustrated by Example 3.7 by choosing $\beta < 1$.

Example 3.11. Consider the function $f(x) = |x|$. Using the triangle inequality we find that

$$|x| = |y + (x - y)| \leq |y| + |x - y|$$
$$|y| = |x + (y - x)| \leq |x| + |y - x|$$

for all $x, y \in \mathbb{R}$. Combining these we find that $\big||x|-|y|\big| \leq |x-y|$ so f is Lipschitz continuous on $\mathbb{R}$ with constant 1. To see that this is in fact the minimal Lipschitz constant, assume L is a Lipschitz constant and set $x = 1$ and $y = 0$ to get the inequality $1 \leq L$.

The absolute value function is not differentiable at 0, showing that Lipschitz continuous functions are not necessarily differentiable. However if we know a function is differentiable then we can use the following theorem to determine whether it is Lipschitz continuous.

Theorem 3.12. *If f is a differentiable function on a non-degenerate interval $I \subseteq \mathbb{R}$, then it is Lipschitz continuous on I if and only if f' is bounded on I. Furthermore, in the case where f' is bounded and f is Lipschitz continuous, we have* $\sup_I |f'| = [f]_1$.

Proof. Assume first that f is Lipschitz continuous. Then for every $x, y \in I$

$$\frac{|f(x) - f(y)|}{|x - y|} \leq [f]_1.$$

Taking the limit as x goes to y we get $|f'(y)| \leq [f]_1$ and so f' is bounded on I.

Next we assume that f' is bounded on I by a constant L in the sense that $|f'| \leq L$. Then by the mean value theorem there is, for all $x, y \in I$, a point $z \in I$ between x and y satisfying

$$f(x) - f(y) = f'(z)(x - y).$$

Taking the absolute value on both sides gives

$$|f(x) - f(y)| = |f'(z)||x - y| \leq \sup_I |f'||x - y|$$

so f is Lipschitz continuous on I with Lipschitz constant $\sup_I |f'|$. Hence $[f]_1 \leq \sup_I |f'|$ and so we must have $[f]_1 = \sup_I |f'|$. □

A consequence of this theorem is that if f is continuously differentiable, then it is locally Lipschitz continuous. This follows from the fact that if f' is continuous, then it is bounded on any compact subset of its domain (Theorem

2.9). As an example consider $f(x) = e^x$. Since it is continuously differentiable on $\mathbb{R}$, we can conclude that it is locally Lipschitz on $\mathbb{R}$. However, f is not Lipschitz continuous on $\mathbb{R}$ since if it were, setting $y = 0$ in Definition 3.10 we would have

$$\left| \frac{e^x - 1}{x} \right| \leq L$$

for $x \in \mathbb{R}, L \geq 0$ but the left hand side goes to infinity when x goes to infinity. We can conclude that locally Lipschitz continuous functions are not necessarily Lipschitz continuous. We do however have the following slightly weaker result.

Theorem 3.13. *If f is locally Lipschitz continuous on an open set $\mathcal{D} \subseteq \mathbb{R}^n$, then f is continuous on $\mathcal{D}$.*

Proof. For any $x \in \mathcal{D}$, since $\mathcal{D}$ is open we can find $\epsilon > 0$ such that $\overline{B_\epsilon(x)} \subseteq \mathcal{D}$. Since $\overline{B_\epsilon(x)}$ is compact, f is Lipschitz continuous on $\overline{B_\epsilon(x)}$. This implies that f is continuous at every point of $\overline{B_\epsilon(x)}$, and in particular at the point x. □

3.2.1 Convex functions are locally Lipschitz

In this section we will see how Lipschitz continuity relates to the notion of convexity. The theory and proof is based on [8]. We start by introducing the notion of convex sets.

Definition 3.14. A set $E \subseteq \mathbb{R}^n$ is convex if

$$tx + (1-t)y \in E$$

whenever $x, y \in E$ and $t \in [0, 1]$.

This means that the line segment between the points x and y also lies in E. For a subset of $\mathbb{R}$ convexity then means that it is an interval. In $\mathbb{R}^n$ we have for example open balls that are convex. Along with this definition we also need the convexity of functions.

Definition 3.15. A function f from a convex subset $\mathcal{D} \subseteq \mathbb{R}^n$ into $\mathbb{R}$ is convex if

$$f(tx + (1-t)y) \leq tf(x) + (1-t)f(y)$$

for $x, y \in \mathcal{D}$ and $t \in [0, 1]$.

For example the absolute value function is convex, since by the triangle inequality

$$|tx + (1-t)y| \leq |tx| + |(1-t)y| = t|x| + (1-t)|y|$$

whenever $x, y \in \mathbb{R}$ and $t \in [0, 1]$. Before we get to the main result we will prove the following lemma.

Lemma 3.16. *If $a < b < c$ are points in the open interval $\mathcal{D} \subseteq \mathbb{R}$ and f is convex on $\mathcal{D}$, then*

$$\frac{f(b) - f(a)}{b - a} \leq \frac{f(c) - f(a)}{c - a} \leq \frac{f(c) - f(b)}{c - b}.$$

Proof. Since $a < b < c$ we can find $t \in (0, 1)$ such that $b = ta + (1 - t)c$, or equivalently $t = \frac{c-b}{c-a}$. Since f is convex we get

$$f(b) = f(ta + (1 - t)c) \leq tf(a) + (1 - t)f(c)$$

which is equivalent to

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

which we can write

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

proving one of the desired inequalities. Setting $s = 1-t$ we have $b = sc+(1-s)a$ or equivalently $s = \frac{b-a}{c-a}$ and

$$f(b) \leq sf(c) + (1 - s)f(a).$$

Similar to before we get

$$f(b) - f(a) \leq \frac{b - a}{c - a}(f(c) - f(a)) \Leftrightarrow \frac{f(b) - f(a)}{b - a} \leq \frac{f(c) - f(a)}{c - a}$$

establishing the other inequality. □

Before we state the main result of this section we want to generalise a definition from earlier. If a real-valued function f is defined on $\mathcal{D} \subseteq \mathbb{R}^n$, then we say that it is locally Lipschitz continuous if, for any compact subset $K \subseteq \mathcal{D}$ there is $L \geq 0$ such that for every $x, y \in K$

$$|f(x) - f(y)| \leq L \|x - y\|$$

holds. Here $\|\cdot\|$ is the standard norm in $\mathbb{R}^n$.

Theorem 3.17. *If f is a real-valued, convex function defined on an open and convex subset $\mathcal{D}$ of $\mathbb{R}^n$, then f is locally Lipschitz continuous in $\mathcal{D}$.*

Proof. The proof is by induction on the dimension n of $\mathbb{R}^n$. First, we consider the case $n = 1$. Then $\mathcal{D}$ is an open interval in $\mathbb{R}$. Let K be a compact subset of $\mathcal{D}$ and let $x, y \in K$ satisfying $x < y$. Since $\mathcal{D}$ is open, we can find $a < b < c < d$ in $\mathcal{D}$ such that for every $x, y \in K$ satisfying $x < y$, $a < b < x < y < c < d$ holds. By applying Lemma 3.16 to $a < b < x$ we get

$$\frac{f(b) - f(a)}{b - a} \leq \frac{f(x) - f(b)}{x - b}$$

and by applying it to $b < x < y$ we get

$$\frac{f(x) - f(b)}{x - b} \leq \frac{f(y) - f(x)}{y - x}.$$

Putting these together we have

$$\frac{f(b)-f(a)}{b-a} \leq \frac{f(y)-f(x)}{y-x}$$

and by a similar argument on $x < y < c < d$ we find that

$$\frac{f(y)-f(x)}{y-x} \leq \frac{f(d)-f(c)}{d-c}.$$

Now set

$$L = \max\left(\left|\frac{f(b)-f(a)}{b-a}\right|, \left|\frac{f(d)-f(c)}{d-c}\right|\right).$$

If $\frac{f(y)-f(x)}{y-x}$ is positive, then

$$\left|\frac{f(y)-f(x)}{y-x}\right| \leq \left|\frac{f(d)-f(c)}{d-c}\right| \leq L$$

and if $\frac{f(y)-f(x)}{y-x}$ is negative, then

$$\left|\frac{f(y)-f(x)}{y-x}\right| \leq \left|\frac{f(b)-f(a)}{b-a}\right| \leq L$$

so f is Lipschitz continuous with constant L on K, and hence locally Lipschitz continuous on $\mathcal{D}$. This concludes the case $n = 1$.

Now assume that the theorem holds for $n = k-1$, $k \geq 2$, and let f be a real-valued, convex function defined on a convex and open subset $\mathcal{D}$ of $\mathbb{R}^k$ and let K be a compact subset of $\mathcal{D}$. We will proceed by finding compact sets X and Y satisfying

(i) $K \subseteq X \subseteq Y \subseteq \mathcal{D}$

(ii) $K, \partial X$ and ∂Y are pairwise disjoint

(iii) X and Y are finite unions of k-dimensional boxes with edges parallel to the coordinate axes.

Here ∂X and ∂Y refers to the boundaries of X and Y respectively. Since K is compact and $\mathcal{D}$ is open, we can find $r > 0$ such that any point of $\mathbb{R}^k$ whose distance to K is less than r must lie in $\mathcal{D}$. For k-tuples of integers $(m_1, \ldots, m_k)$, define the box

$$B(m_1, \ldots, m_k) = \left\{(x_1, \ldots, x_k) \in \mathbb{R}^k : \frac{m_i r}{10\sqrt{k}} \leq x_i \leq \frac{(m_i+1)r}{10\sqrt{k}},\ i = 1, \ldots, k\right\}.$$

Note that these boxes cover all of $\mathbb{R}^k$ and that they have side lengths $\frac{r}{10\sqrt{k}}$ and hence diagonal $\frac{r}{10}$. Now let X be the union of all such boxes with distance to K smaller than $\frac{r}{5}$ and let Y be the union of all such boxes with distance to

K smaller than $\frac{4r}{5}$. By construction X and Y satisfy (i) and (iii) above (the unions are finite since K is bounded).

Let $x \in K$. Then the closest point in any box adjacent to one that x lies in (x may lie in more than one box if it lies on a boundary) is a distance smaller than $\frac{r}{10}$ (one diagonal length) from x. Hence any box that is adjacent to a box containing x must be included in X and subsequently in Y since $X \subseteq Y$. As a consequence $x \notin \partial X$ and $x \notin \partial Y$.

Now let $x \in \partial X$. Then x is at a distance smaller than $\frac{r}{5} + \frac{r}{10} = \frac{3r}{10}$ from K. Thus all boxes adjacent to a box that x lies in are included in Y and so $x \notin \partial Y$. We can conclude that $K, \partial X$ and ∂Y satisfy (ii) above.

We note that ∂X and ∂Y are unions of $(k-1)$-dimensional faces of k-dimensional boxes. Let H be a hyperplane that intersects some box in one of its faces. The restriction of f to $H \cap \mathcal{D}$ is then a convex function on an open, convex space of dimension $k-1$. By the induction hypothesis f is then locally Lipschitz on $H \cap \mathcal{D}$ and thus, by Theorem 3.13, f is also continuous on $H \cap \mathcal{D}$. For every $x \in \partial X, y \in \partial Y$ the function

$$Q(x, y) = \frac{|f(x) - f(y)|}{\|x - y\|}$$

is then continuous. Note that since ∂X and ∂Y are compact and disjoint, the denominator is bounded away from zero. Since Q is continuous on the compact set $(\partial X) \times (\partial Y)$ it attains some maximal value M therein.

Let $x, y \in K$ such that $x \neq y$. Then there is a unique line l that passes through x and y. If we travel along l starting at x and going away from y, then we must leave X, passing through some point $b \in \partial X$ and later leave Y, passing though a point $a \in \partial Y$. Going in the opposite direction we also find $c \in \partial X$ and $d \in \partial Y$ so that the order of the points on l is a, b, x, y, c, d. We can now consider the restriction of f to $l \cap \mathcal{D}$ as a convex function on an open interval of $\mathbb{R}$. By an analogous argument to the one used in the case $n = 1$ above, we find that

$$\frac{|f(x) - f(y)|}{\|x - y\|} \leq \max\left(\frac{|f(c) - f(d)|}{\|c - d\|}, \frac{|f(b) - f(a)|}{\|b - a\|}\right) \leq M$$

so f is Lipschitz continuous on K and so locally Lipschitz continuous on $\mathcal{D}$. By induction the theorem then holds for any $n \geq 1$. □

3.3 Inverses and their derivatives

While it is true that the inverse (assuming it exists) of a continuous function is also continuous, the same does not hold for Hölder continuity. An example of this is the function $f(x) = x^3$ defined on [0,1]. Since f has a bounded derivative, it is Lipschitz continuous by Theorem 3.12 and then by Theorem 3.6 it is also $\frac{1}{2}$-Hölder continuous. The inverse of f is $f^{-1}(x) = x^{\frac{1}{3}}$, which as we saw in Example 3.7 has maximum Hölder exponent $\frac{1}{3}$ and hence cannot be $\frac{1}{2}$-Hölder continuous.

The continuity of the derivative will however allow us to say something about the inverse.

Theorem 3.18. *Let f be a differentiable function defined in an open interval $\mathcal{D}$. Also assume that f' is continuous in $\mathcal{D}$, that $|f'|$ is bounded and that $\inf |f'| > 0$. Then f is invertible and f^{-1} is Lipschitz continuous in the image of f with*

$$[f^{-1}]_1 \leq \frac{1}{\inf |f'|}.$$

Proof. Since $\inf |f'| > 0$ and $\mathcal{D}$ is an interval, f is strictly monotonous in $\mathcal{D}$ and thus it is invertible. The derivative of the inverse is

$$Df^{-1}(x) = \frac{1}{f'\left(f^{-1}(x)\right)}.$$

Now, since $\inf |f'| > 0$, Df^{-1} is bounded and so by Theorem 3.12 f^{-1} is Lipschitz continuous and

$$[f^{-1}]_1 = \sup_{f(\mathcal{D})} |Df^{-1}| = \sup_{f(\mathcal{D})} \frac{1}{|f'\left(f^{-1}(x)\right)|} \leq \frac{1}{\inf_{\mathcal{D}} |f'|}.$$

□

Furthermore, Hölder continuity of the derivative allows us to assert the Hölder continuity of the derivative of the inverse.

Theorem 3.19. *Let f be a differentiable function defined in an open interval $\mathcal{D}$. Assume that f' is α-Hölder continuous in $\mathcal{D}$ and that $\inf |f'| > 0$. Then f is invertible, and the derivative of the inverse is α-Hölder continuous with*

$$[Df^{-1}]_\alpha \leq \frac{[f']_\alpha}{(\inf |f'|)^{2+\alpha}}.$$

Proof. We have $\inf |f'| > 0$ so f is strictly monotone, and hence invertible. For x, y in $f(\mathcal{D})$ we have

$$\begin{aligned} \left|Df^{-1}(x) - Df^{-1}(y)\right| &= \left|\frac{1}{f'(f^{-1}(x))} - \frac{1}{f'(f^{-1}(y))}\right| = \\ &= \frac{|f'(f^{-1}(x)) - f'(f^{-1}(y))|}{|f'(f^{-1}(x))f'(f^{-1}(y))|} \leq \frac{[f']_\alpha |f^{-1}(x) - f^{-1}(y)|^\alpha}{(\inf |f'|)^2} \end{aligned}$$

and then, by the preceding theorem

$$\begin{aligned} \frac{[f']_\alpha |f^{-1}(x) - f^{-1}(y)|^\alpha}{(\inf |f'|)^2} &\leq \frac{[f']_\alpha ([f^{-1}]_1)^\alpha |x - y|^\alpha}{(\inf |f'|)^2} \leq \\ &\leq \frac{[f']_\alpha}{(\inf |f'|)^{2+\alpha}} |x - y|^\alpha \end{aligned}$$

which completes the proof. □

4 Two Applications

4.1 Initial value problems

In this section we will see one application of Lipschitz continuity in the context of initial value problems. More specifically we will study problems on the form

$$\begin{cases} x'(t) = f(t,x) \\ x(t_0) = x_0 \end{cases}$$

where f is a real valued function defined in some subset of $\mathbb{R}^2$ that satisfies certain continuity conditions. This section is a simplification of [7, p. 33-38], the original source also handles the case where $x(t)$ is vector valued. First we will have to say a few things about normed spaces and fixed points of functions.

Of particular interest to us is the space $C(I)$ of continuous functions defined on some compact subset I of $\mathbb{R}^n$. Defining the operations pointwise and defining

$$\|x\| = \sup_{t\in I} |x(t)|$$

for $x \in C(I)$, this is a normed vector space. This norm is called the supremum norm and a sequence in $C(I)$ that converges with respect to this norm is said to converge uniformly. Furthermore $C(I)$ is a complete normed space as established by the following lemma. We refer to complete normed spaces as Banach spaces.

Lemma 4.1. *For any compact set $I \subset \mathbb{R}^n$, $C(I)$ equipped with the norm $\|x\| = \sup_{t\in I} |x(t)|$ is complete.*

Proof. Let $x_n(t)$ be a Cauchy sequence in $C(I)$. For any fixed $t \in I$, $x_n(t)$ forms a Cauchy sequence in $\mathbb{R}$ and by the completeness of $\mathbb{R}$ this sequence has a limit. For every $t \in \mathbb{R}$, let $x(t)$ denote this limit. We have found a limit function $x(t)$ of our sequence, it remains to show that $x_n(t)$ converges to $x(t)$ uniformly and that $x(t) \in C(I)$.

Fix $\epsilon > 0$. Since $x_n(t)$ is a Cauchy sequence we can find $N \in \mathbb{N}$ such that for all $m, n \geq N$

$$\|x_n - x_m\| \leq \epsilon$$

and so for all $t \in I$

$$|x_n(t) - x_m(t)| \leq \epsilon.$$

If we now let $m \to \infty$, we get

$$|x_n(t) - x(t)| \leq \epsilon$$

for every $t \in I$ and hence

$$\|x_n - x\| \leq \epsilon$$

so x_n converges to x uniformly.

It remains to show that x is continuous. Fix $\epsilon > 0$ and $t \in I$. By the above we can find $m \in \mathbb{N}$ so that $\|x_m - x\| < \frac{\epsilon}{3}$ and since x_m is continuous we can find

$\delta > 0$ so that $|t-s| < \delta$ implies $|x_m(t) - x_m(s)| < \frac{\epsilon}{3}$. Then whenever $|t-s| < \delta$, we have

$$
\begin{aligned}
|x(t) - x(s)| &\leq |x(t) - x_m(t)| + |x_m(t) - x_m(s)| + |x_m(s) - x(s)| < \\
&< \frac{\epsilon}{3} + \frac{\epsilon}{3} + \frac{\epsilon}{3} = \epsilon
\end{aligned}
$$

so x is continuous at t. Since t was arbitrary in I we can conclude that $x \in C(I)$. □

In the more general setting of a normed space X we say that a mapping $f : \mathcal{D} \subseteq X \to X$ is Lipschitz continuous if there is $L \geq 0$ such that

$$\|f(x) - f(y)\| \leq L \|x - y\|$$

for all $x, y \in \mathcal{D}$. Just like in definition 3.10 we also say that f is a contraction if there is $L < 1$ satisfying the above inequality. Since we have established that $C(I)$ is a Banach space, the following result is applicable.

Theorem 4.2 (Contraction principle)**.** *Let A be a non-empty closed subset of a Banach space X and let $f : A \to A$ be a contraction with Lipschitz constant $L < 1$. Then f has a unique fixed point in A, that is there is a unique $x \in A$ such that $f(x) = x$.*

Proof. To show the existence of such a point, choose an arbitrary point $x_0 \in A$ and for $n = 1, 2, \ldots$ define $x_n = f(x_{n-1})$. We get

$$\|x_{n+1} - x_n\| \leq L \|x_n - x_{n-1}\| \leq \cdots \leq L^n \|x_1 - x_0\|$$

and hence for $n > m$, using the triangle inequality,

$$\|x_n - x_m\| \leq \sum_{i=m}^{n-1} \|x_{i+1} - x_i\| \leq \sum_{i=m}^{n-1} L^i \|x_1 - x_0\| \leq \frac{L^m}{1-L} \|x_1 - x_0\| .$$

If we now let $m \to \infty$, we see that x_n is a Cauchy sequence. Since A is closed and a subset of a Banach space we can conclude that x_n converges to a point $x \in A$. Now since f is continuous we have

$$f(x) = f\left(\lim_{n\to\infty} x_n\right) = \lim_{n\to\infty} f(x_n) = \lim_{n\to\infty} x_{n+1} = x$$

so x is a fixed point of f. To see that the fixed point is unique note that if x and x' are fixed points, then

$$\|x - x'\| = \|f(x) - f(x')\| \leq L \|x - x'\| .$$

Since $L < 1$ we must then have $\|x - x'\| = 0$ and so $x = x'$. □

We are now ready to formulate the main result of this section.

Theorem 4.3 (Picard-Lindelöf). *Suppose $f \in C(U)$, where U is an open subset of $\mathbb{R}^2$ and $(t_0, x_0) \in U$. If f is locally Lipschitz continuous in the second argument and uniformly continuous in the first, then there is a unique local solution $\overline{x}(t) \in C(I)$ to the IVP*

$$\begin{cases} x'(t) = f(t,x) \\ x(t_0) = x_0 \end{cases} \tag{3}$$

where $I = [t_0, t_0 + T_0]$ for some $T_0 > 0$.

By saying that f should be Lipschitz continuous in the second argument we mean that we can find $L > 0$ independent of t that satisfies

$$|f(t,x) - f(t,y)| \leq L \, \|x - y\|$$

for every $(t,x), (t,y) \in U$. By saying that there is a local solution we mean that it is possible to find T_0 so that there is a solution in $C([t_0, t_0 + T_0])$. Having found such a T_0, this solution is then unique in the sense that it is the only solution in $C([t_0, t_0 + T_0])$.

The theorem can be generalised in several ways. It can be extended to vector valued functions of several real variables. Changing a few details of the proof we get a similar result when the domain I of the solution is on the forms $I = [t_0 - T_0, t_0]$ or $I = [t_0 - T_0, t_0 + T_0]$.

Proof. The main idea of the proof is to use the contraction principle to find the solution to the IVP as the unique fixed point of some contraction. By integrating both sides of the equation in (3) and using the initial condition we see that the IVP is equivalent to the integral equation

$$x(t) = x_0 + \int_{t_0}^{t} f(s, x(s)) ds.$$

We now define a mapping from $C([t_0, t_0 + T])$ into $C([t_0, t_0 + T])$, for some suitable value of $T > 0$, by

$$(Kx)(t) := x_0 + \int_{t_0}^{t} f(s, x(s)) ds.$$

We note that if $\overline{x}$ is a solution to the IVP then it is a fixed point of the mapping K in the sense that $(K\overline{x})(t) = \overline{x}(t)$ for all $t \in [t_0, t_0 + T]$. By Lemma 4.1, $C([t_0, t_0 + T])$ is a Banach space for any $T > 0$. To use the contraction principle we also need a (non-empty) closed set, which we shall call A, on which K is a contraction self-mapping. To find such a set, note first that since U is open we can choose $T > 0$ and $\delta > 0$ such that $V = [t_0, t_0 + T] \times \overline{B_\delta(x_0)} \subseteq U$. Let

$$M = \max_{(t,x) \in V} |f(t,x)|$$

which exists since f is continuous and V is compact. Now if the graph of $x(t)$ lies in V, then

$$|(Kx)(t) - x_0| \leq \int_{t_0}^{t} |f(s, x(s))| ds \leq M(t - t_0).$$

Choosing T_0 so that $0 < T_0 \leq \min\left(T, \frac{\delta}{M}\right)$ (if $M = 0$ we interpret this as $\frac{\delta}{M} = \infty$ and so the condition is reduced to $T_0 \leq T$), we then have $T_0 M \leq \delta$ and so the graph of $(Kx)(t)$ restricted to $[t_0, t_0 + T_0]$ also lies in V. Since $T_0 \leq T$, M will also be a bound of $|f|$ on $V_0 = [t_0, t_0 + T_0] \times \overline{B_\delta(x_0)} \subseteq V$. Thus if we choose $X = C([t_0, t_0 + T_0])$ with norm $\|x\| = \max_{t \in [t_0, t_0 + T_0]} |x(t)|$ as our Banach space and use the subset $A = \{x \in X | \, \|x - x_0\| \leq \delta\}$, then the image of K restricted to A is a subset of A, or in other words K is a self-mapping on A.

Finally we must show that K is a contraction. Now using that f is locally Lipschitz continuous in the second argument, we get

$$|(Kx)(t) - (Ky)(t)| \leq \int_{t_0}^{t} |f(s, x(s)) - f(s, y(s))| ds \leq$$

$$\leq L \int_{t_0}^{t} |x(s) - y(s)| ds \leq L(t - t_0) \sup_{t_0 \leq s \leq t} |x(s) - y(s)|$$

for $x, y \in A$. Taking the maximum over $t \in [t_0, t_0 + T_0]$, we get

$$\|Kx - Ky\| \leq LT_0 \|x - y\| .$$

If we add to the definition of T_0 that it should satisfy that $T_0 < L^{-1}$, then K is indeed a contraction. We can now apply the contraction principle to furnish us a unique solution $\overline{x}$ to the IVP, thereby proving the theorem. □

While this theorem proves the existence of a solution it is not much help in finding this solution in general. In the proof of the contraction principle we constructed a sequence whose limit is the desired solution by successively applying a function. In this case the function is an integral function and in many cases this iteration of integrals, known as Picard iteration, will not be possible to do analytically. However in some cases numerical methods can be used to find the solution.

4.2 Limits of Cauchy-type integrals

In this section we will study Cauchy-type integrals, that is functions on the form

$$F(z) = \frac{1}{2\pi i} \int_\gamma \frac{f(t)}{t - z} dt$$

where γ is some curve in $\mathbb{C}$, f is a function defined on γ and $z \in \mathbb{C} \backslash \gamma$. In particular we will study what happens if we let z approach some point $t_0 \in \gamma$ and what properties f needs to satisfy for these limits to exist. In this text we will restrict ourselves to the case when the curve γ is some interval $[a, b]$ in the real line.

At first one might consider

$$\frac{1}{2\pi i} \int_a^b \frac{f(t)}{t - t_0} dt$$

as the limit when z approaches t_0. However since the integrand is, in the general case, unbounded, this integral will not converge. However if we exclude some small neighbourhood of t_0 from our domain of integration the integral should converge. This leads us to define the principle value of the integral, denoted by

$$\text{p.v.} \int_a^b \frac{f(t)}{t - t_0} dt$$

as the limit

$$\lim_{\epsilon \to 0^+} \left(\int_a^{t_0 - \epsilon} \frac{f(t)}{t - t_0} dt + \int_{t_0 + \epsilon}^b \frac{f(t)}{t - t_0} dt \right)$$

if it exists. The principle value is not the whole story however.

Example 4.4. In this example we will look at the simple case when $f = 1$ and so

$$F(z) = \frac{1}{2\pi i} \int_a^b \frac{dt}{t - z}$$

for $z \notin [a, b]$. Let t_0 be a point in the interval (a, b). Let us first look at the principle value.

$$\frac{1}{2\pi i} \text{p.v.} \int_a^b \frac{dt}{t - t_0} = \frac{1}{2\pi i} \lim_{\epsilon \to 0^+} \left([\log |t - t_0|]_a^{t_0 - \epsilon} + [\log |t - t_0|]_{t_0 + \epsilon}^b \right) =$$

$$= \frac{1}{2\pi i} \lim_{\epsilon \to 0^+} (\log |\epsilon| - \log |a - t_0| + \log |b - t_0| - \log |\epsilon|) = \frac{1}{2\pi i} \log \frac{|b - t_0|}{|a - t_0|}$$

If we instead write $z = x + iy$ where $x, y \in \mathbb{R}$, we get

$$F(z) = \frac{1}{2\pi i} \int_a^b \frac{dt}{t - z} = \frac{1}{2\pi i} \int_a^b \frac{t - x}{(t - x)^2 + y^2} dt + \frac{1}{2\pi} \int_a^b \frac{y}{(t - x)^2 + y^2} dt$$

We now have two real integrals. The first evaluates to

$$\frac{1}{2\pi i} \int_a^b \frac{t - x}{(t - x)^2 + y^2} dt = \frac{1}{2\pi i} \left[\frac{1}{2} \log \left((t - x)^2 + y^2 \right) \right]_a^b =$$

$$\frac{1}{2\pi i} \left[\log \left(\sqrt{(t - x)^2 + y^2} \right) \right]_a^b = \frac{1}{2\pi i} [\log |t - z|]_a^b = \frac{1}{2\pi i} \log \frac{|b - z|}{|a - z|}$$

If we now let $z \to t_0$, then this converges to $\frac{1}{2\pi i} \log \frac{|b - t_0|}{|a - t_0|}$ which we recognise as the principal value above. The other integral evaluates to

$$\frac{1}{2\pi} \int_a^b \frac{y}{(t - x)^2 + y^2} dt = \frac{1}{2\pi} \left[\arctan \left(\frac{t - x}{y} \right) \right]_a^b =$$

$$\frac{1}{2\pi} \left(\arctan \left(\frac{b - x}{y} \right) + \arctan \left(\frac{x - a}{y} \right) \right).$$

We can rephrase z approaching the point t_0 as x approaching t_0 and y going to zero. When x gets close to t_0, then the enumerators $b-x$ and $x-a$ will remain positive and bounded. On the other hand, the denominators y goes to zero. If we approach the point t_0 from above in the complex plane ($y > 0$), then the quotient goes to $+\infty$ and the limit of the arctan function is $\frac{\pi}{2}$. If we on the other hand approach t_0 from below ($y < 0$), then the quotient goes to $-\infty$ and so the limit of the arctan is $-\frac{\pi}{2}$. If we denote the limits as $z \to t_0$ from above and below by $F^+(t_0)$ and $F^-(t_0)$ respectively, then what we have found is that

$$F^+(t_0) = \frac{1}{2\pi i} \log \left| \frac{b - t_0}{a - t_0} \right| + \frac{1}{2}$$

and

$$F^-(t_0) = \frac{1}{2\pi i} \log \left| \frac{b - t_0}{a - t_0} \right| - \frac{1}{2}.$$

Not only is the limit not given by the principle value of the integral, we get two different limits depending on from which side of the line $[a, b]$ we approach t_0. This will also be the case in general. We may now wonder when these limits exist and exactly how they depend on f. As it turns out, f being Hölder continuous is a sufficient condition for these limits to exist and their values are given by the so called Sokhotski-Plemelj-formula.

Theorem 4.5. *If f is Hölder continuous on $[a, b]$, then the principal value*

$$\text{p.v.} \int_a^b \frac{f(t)}{t - t_0} dt$$

exist finitely. If we also define

$$F(z) = \frac{1}{2\pi i} \int_a^b \frac{f(t)}{t - z} dt$$

for $z \notin [a, b]$, then the limits of F as $z \to t_0 \in (a, b)$ exist and are given by

$$F^+(t_0) = +\frac{f(t_0)}{2} + \frac{1}{2\pi i} \text{p.v.} \int_a^b \frac{f(t)}{t - t_0} dt$$

and

$$F^-(t_0) = -\frac{f(t_0)}{2} + \frac{1}{2\pi i} \text{p.v.} \int_a^b \frac{f(t)}{t - t_0} dt.$$

For a proof we refer to [5, p. 23].

In example 3.3 we saw that $f(t) = \sqrt{t}$ is Hölder continuous in the interval $[0, 1]$. We can use the theorem to find the limits of

$$F(z) = \frac{1}{2\pi i} \int_0^1 \frac{\sqrt{t}}{t - z} dt$$

as $z \to \frac{1}{2}$. To do so we first need to calculate the principal value

$$\text{p.v.} \int_0^1 \frac{\sqrt{t}}{t-\frac{1}{2}} dt = \text{p.v.} \int_0^1 \frac{\sqrt{t}-\frac{1}{\sqrt{2}}}{t-\frac{1}{2}} dt + \frac{1}{\sqrt{2}} \text{p.v.} \int_0^1 \frac{dt}{t-\frac{1}{2}}$$

The second term we have already seen in the previous example, we get

$$\text{p.v.} \int_0^1 \frac{dt}{t-\frac{1}{2}} = \log \frac{|1-\frac{1}{2}|}{|0-\frac{1}{2}|} = 0.$$

For the other one we get

$$\text{p.v.} \int_0^1 \frac{\sqrt{t}-\frac{1}{\sqrt{2}}}{t-\frac{1}{2}} dt = \text{p.v.} \int_0^1 \frac{1}{\sqrt{t}+\frac{1}{\sqrt{2}}} dt = \left[2\sqrt{t} - \sqrt{2} \log \left(\sqrt{t} + \frac{1}{\sqrt{2}} \right) \right]_0^1 =$$
$$= 2 - \sqrt{2} \log \left(1 + \frac{1}{\sqrt{2}} \right) + \sqrt{2} \log \left(\frac{1}{\sqrt{2}} \right) = 2 - \sqrt{2} \log \left(2 + \sqrt{2} \right)$$

We can then conclude that

$$F^+ \left(\frac{1}{2} \right) = \frac{1}{2\sqrt{2}} + \frac{1}{2\pi i} \left(2 - \sqrt{2} \log \left(2 + \sqrt{2} \right) \right)$$

and

$$F^- \left(\frac{1}{2} \right) = -\frac{1}{2\sqrt{2}} + \frac{1}{2\pi i} \left(2 - \sqrt{2} \log \left(2 + \sqrt{2} \right) \right).$$

5 Absolute continuity

5.1 Basic properties

We will start with defining absolute continuity and make some basic observations.

Definition 5.1. A function $f : [a, b] \to \mathbb{R}$ is said to be absolutely continuous on $[a, b]$ if for every $\epsilon > 0$ there is $\delta > 0$ such that

$$\sum_{i=1}^{n} |f(y_i) - f(x_i)| < \epsilon$$

whenever $\{[x_i, y_i] : i = 1, \ldots, n\}$ is a finite collection of mutually disjoint subintervals of $[a, b]$ satisfying

$$\sum_{i=1}^{n} |y_i - x_i| < \delta.$$

If we let $n = 1$ in the definition above, the statement reduces to the definition of uniform continuity (on a set of the form $[a, b]$) so absolute continuity is stronger than uniform continuity. The converse however, is not true as the following counterexample shows.

Example 5.2. Let

$$f(x) = \begin{cases} x\sin\left(\frac{1}{x}\right), & \text{for } x \in (0,1] \\ 0, & \text{for } x = 0. \end{cases}$$

This function is continuous on a compact set, and hence uniformly continuous. We now consider subintervals given by $x_i = \frac{2}{(2i+1)m\pi}, y_i = \frac{1}{im\pi}$ for $i = 1, \ldots, n$, where m is a positive integer. Now given some $\epsilon > 0$ we have, for every $\delta > 0$ that

$$\sum_{i=1}^{n} |y_i - x_i| = \frac{1}{m\pi} \sum_{i=1}^{n} \frac{1}{i(2i+1)} < \delta$$

if we choose m sufficiently big. However in this case we get

$$\sum_{i=1}^{n} |f(y_i) - f(x_i)| = \frac{1}{m\pi} \sum_{i=1}^{n} \frac{2}{2i+1} > \epsilon$$

if we choose n sufficiently big. Thus f is not absolutely continuous.

On the other hand, absolute continuity is weaker than Lipschitz continuity.

Theorem 5.3. *If f is Lipschitz continuous on $[a,b]$, then f is absolutely continuous on $[a,b]$.*

Proof. Let L be a Lipschitz constant of f. Given $\epsilon > 0$ we set $\delta = \frac{\epsilon}{L}$. Then whenever $\{[x_i, y_i] : i = 1, \ldots, n\}$ is a finite collection of mutually disjoint subintervals of $[a,b]$ satisfying

$$\sum_{i=1}^{n} |y_i - x_i| < \delta.$$

we have

$$\sum_{i=1}^{n} |f(y_i) - f(x_i)| \leq \sum_{i=1}^{n} L|y_i - x_i| < L\delta = \epsilon$$

so f is absolutely continuous on $[a,b]$. □

That the converse does not hold is illustrated by the function $f(x) = \sqrt{x}$ defined on $[0,1]$. The function f is not Lipschitz continuous (see Example 3.7). Given $\epsilon > 0$, if we set $\delta = \epsilon^2$ we find that whenever$\{[x_i, y_i] : i = 1, \ldots, n\}$ is a finite collection of mutually disjoint subintervals of $[0,1]$ satisfying $\sum_{i=1}^{n} |y_i - x_i| < \delta$, we get the estimate

$$\sum_{i=1}^{n} |\sqrt{y_i} - \sqrt{x_i}| = \sum_{i=1}^{n} \int_{x_i}^{y_i} \frac{dt}{2\sqrt{t}} \leq \int_{0}^{\sum_{i=1}^{n} |y_i - x_i|} \frac{dt}{2\sqrt{t}}$$

since the integrand is monotone decreasing and positive. The right hand side evaluates to

$$\sqrt{\sum_{i=1}^{n} |y_i - x_i|} < \sqrt{\delta} = \epsilon$$

so f is absolutely continuous.

The above theorem allows us to quickly find many absolutely continuous functions. If f is continuous on $[a, b]$, then the function $F(x) = \int_a^x f(t)dt$ where $x \in [a, b]$, has the derivative $F'(x) = f(x)$ by the Fundamental Theorem of Calculus. Since f is continuous on a compact set it is bounded, and so by Theorem 3.12 F is Lipschitz continuous and hence absolute continuous. Absolute continuity is actually very closely related to the Fundamental Theorem of Calculus which we will explore further in a later section.

We have now established that absolute continuity lies between Lipschitz continuity and uniform continuity, and earlier we have seen that Hölder continuity also lies between these two notions. It is then natural to ask, is there some implication between these two? The answer to this question is no, neither of them implies the other. The next example demonstrates that absolute continuity does not imply Hölder continuity.

Example 5.4. Consider $f(x) = x^\beta$ where $0 < \beta < 1$, defined on $[0, 1]$. In Example 3.7 we saw that this function is not α-Hölder continuous for $\alpha > \beta$. It is however absolutely continuous. For any $\epsilon > 0$ set $\delta = \epsilon^{\frac{1}{\beta}}$ and suppose $\{[x_i, y_i] : i = 1, \ldots, n\}$ is a finite collection of pairwise disjoint subintervals of $[0, 1]$ satisfying

$$\sum_{i=1}^{n} |y_i - x_i| < \delta.$$

We then have

$$\sum_{i=1}^{n} |f(y_i) - f(x_i)| = \sum_{i=1}^{n} |y_i^\beta - x_i^\beta| = \sum_{i=1}^{n} \int_{x_i}^{y_i} \beta t^{\beta-1} dt = \int_{\bigcup_{i=1}^{n}[x_i, y_i]} \beta t^{\beta-1} dt.$$

Now since $t^{\beta-1}$ is monotone decreasing and positive we get the estimate

$$\int_{\bigcup_{i=1}^{n}[x_i, y_i]} \beta t^{\beta-1} dt \leq \int_0^{\sum_{i=1}^{n} |y_i - x_i|} \beta t^{\beta-1} dt = \left(\sum_{i=1}^{n} |y_i - x_i| \right)^\beta < \delta^\beta = \epsilon$$

which shows that x^β is absolutely continuous. Thus for any $0 < \alpha < 1$ we can find a function that is absolutely continuous but not α-Hölder continuous.

For an example of functions that are Hölder continuous but not absolutely continuous we consider Weierstrass functions, which are functions on the form

$$f(x) = \sum_{i=0}^{\infty} a^n \cos(b^n \pi x)$$

where $0 < a < 1$, $b > 1$ and $ab > 1$. The graph in the case $a = \frac{1}{2}, b = 5$ is shown in figure 1.

It can be shown that Weierstrass functions are Hölder continuous with exponent $-\frac{\log a}{\log b}$ but not differentiable anywhere (see [4] for the proof). We shall

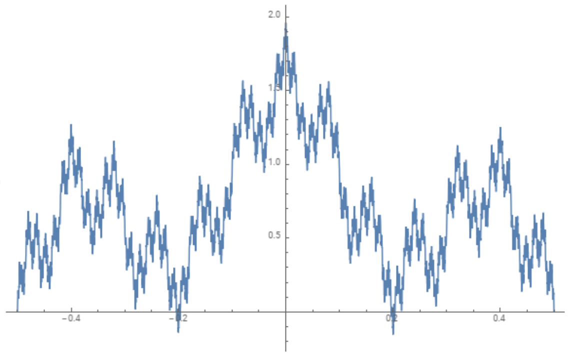

Figure 1: A Weierstrass function

see later that a function that is nowhere differentiable cannot be absolutely continuous.

Compositions of absolutely continuous functions are not necessarily absolutely continuous. For an example, consider the functions

$$g(x) = \begin{cases} x^2 \sin^2\left(\frac{1}{x}\right), x \in (0,1] \\ 0, x = 0 \end{cases}$$

and $f(x) = \sqrt{x}$. As we have seen in Example 5.4, f is absolute continuous on $[0, 1]$. The derivative of g is

$$g'(x) = 2x \sin^2\left(\frac{1}{x}\right) - 2\sin\left(\frac{1}{x}\right)\cos\left(\frac{1}{x}\right)$$

for $x \in (0, 1)$. This derivative is continuous and

$$|g'(x)| \leq \left|2x\sin^2\left(\frac{1}{x}\right)\right| + \left|2\sin\left(\frac{1}{x}\right)\cos\left(\frac{1}{x}\right)\right| \leq 4.$$

Thus for any subinterval $[x, y] \subseteq [0, 1]$ we can find, by the mean value theorem, a point $z \in (x, y)$ such that

$$|g(y) - g(x)| = |g'(z)||y - x| \leq 4|y - x|.$$

Now given $\epsilon > 0$ we set $\delta = \frac{\epsilon}{4}$. Then whenever $\{[x_i, y_i] : i = 1, \ldots, n\}$ is a finite collection of mutually disjoint subintervals of $[0, 1]$ satisfying

$$\sum_{i=1}^{n} |y_i - x_i| < \delta$$

we have

$$\sum_{i=1}^{n} |g(y_i) - g(x_i)| \leq \sum_{i=1}^{n} 4|y_i - x_i| < 4\delta = \epsilon$$

so g is absolutely continuous on $[0, 1]$. The composition $h = f \circ g$ however, is not an absolutely continuous function. The function h is very similar to the function studied in Example 5.2. An analogous argument using the same subintervals can be used to show that h is not absolutely continuous.

If we make the stronger assumption that f is Lipschitz continuous and g absolutely continuous on $[a, b]$, then the composition $f \circ g$ will be absolutely continuous. Indeed, for any $\epsilon > 0$ we can find $\delta > 0$ so that whenever $\{[x_i, y_i] : i = 1, \ldots, n\}$ is a finite collection of mutually disjoint subintervals of $[a, b]$ satisfying

$$\sum_{i=1}^{n} |y_i - x_i| < \delta.$$

we have

$$\sum_{i=1}^{n} |g(y_i) - g(x_i)| < \frac{\epsilon}{[f]_1}$$

and so

$$\sum_{i=1}^{n} |f(g(y_i)) - f(g(x_i))| \leq [f]_1 \sum_{i=1}^{n} |g(y_i) - g(x_i)| < \epsilon$$

which shows that $f \circ g$ is absolutely continuous on $[a, b]$.

Absolute continuity is also preserved under basic algebraic operations.

Theorem 5.5. *Let $f, g : [a, b] \to \mathbb{R}$ be absolutely continuous. Then*

(i) $f + g$ is absolutely continuous

(ii) fg is absolutely continuous

(iii) if we assume that $\inf |g| > 0$*, then $\frac{f}{g}$ is absolutely continuous.*

Proof. Let $\epsilon > 0$ be fixed.

(i) Since f and g are absolutely continuous, there is $\delta > 0$ so that if $\{[x_i, y_i] : i = 1, \ldots, n\}$ is a finite collection of mutually disjoint subintervals of $[a, b]$ satisfying $\sum_{i=1}^{n} |y_i - x_i| < \delta$, then both

$$\sum_{i=1}^{n} |f(y_i) - f(x_i)| < \frac{\epsilon}{2}$$

and

$$\sum_{i=1}^{n} |g(y_i) - g(x_i)| < \frac{\epsilon}{2}$$

hold. Then

$$\sum_{i=1}^{n} |(f+g)(y_i) - (f+g)(x_i)| \leq \sum_{i=1}^{n} |f(y_i) - f(x_i)| + \sum_{i=1}^{n} |g(y_i) - g(x_i)| < \epsilon$$

so $f + g$ is absolutely continuous.

(ii) Since f and g are absolutely continuous, there is $\delta > 0$ so that if $\{[x_i, y_i] : i = 1, \ldots, n\}$ is a finite collection of mutually disjoint subintervals of $[a, b]$ satisfying $\sum_{i=1}^{n} |y_i - x_i| < \delta$, then both

$$\sum_{i=1}^{n} |f(y_i) - f(x_i)| < \frac{\epsilon}{2 \sup |g|}$$

and

$$\sum_{i=1}^{n} |g(y_i) - g(x_i)| < \frac{\epsilon}{2 \sup |f|}$$

hold. Since f and g are continuous on a compact set they are also bounded, which guarantees that the suprema are finite. We now get

$$\begin{aligned}
&\sum_{i=1}^{n} |(fg)(y_i) - (fg)(x_i)| = \\
&= \sum_{i=1}^{n} |f(x_i)g(x_i) - f(x_i)g(x_{i-1}) + f(x_i)g(x_{i-1}) - f(x_i)g(x_{i-1})| \leq \\
&\leq \sum_{i=1}^{n} |f(x_i)||g(x_i) - g(x_{i-1})| + \sum_{i=1}^{n} |g(x_{i-1})||f(x_i) - f(x_{i-1})| < \\
&< \sup |f| \frac{\epsilon}{2 \sup |f|} + \sup |g| \frac{\epsilon}{2 \sup |g|} = \epsilon
\end{aligned}$$

so fg is absolutely continuous.

(iii) If $\inf |g| > 0$, then the image of g is contained in a closed interval that does not include zero. On such an interval, the function $x \mapsto \frac{1}{x}$ is Lipschitz continuous since its derivative $-\frac{1}{x^2}$ is bounded. Then the composition $\frac{1}{g}$ is absolutely continuous by the discussion above this theorem. Finally, $\frac{f}{g}$ is a product of f and $\frac{1}{g}$, both absolutely continuous, and so is also absolutely continuous by part *(ii)*.

□

For part *(iii)* note that if g could take values arbitrarily close to zero, then $\frac{1}{g}$ would not be bounded, and hence it could then not be absolutely continuous.

5.2 The Lebesgue measure and the Lebesgue integral

The Lebesgue integral is a generalisation of the Riemann integral in the sense that it allows us to integrate a larger set of functions. To do so we first need to introduce a few key concepts. For a more in depth look at the theory with proofs we refer to [9, Ch. 4.3,Ch. 6.1]. First is the Lebesgue measure, which lets us calculate the size of certain subsets of $\mathbb{R}$. Let M be the smallest (with respect to inclusion) set whose elements include all open subsets of $\mathbb{R}$ and is closed under complements and countable unions. The elements of M are known as the Lebesgue measurable sets. For our purposes we note that in particular, all open intervals and all closed intervals as well as their countable unions are measurable.

We now define the Lebesgue measure as the function $\lambda : M \to [0, \infty]$ given by

$$\lambda(A) = \inf \left(\sum_{k=1}^{\infty} l(I_k) \right)$$

where the infimum is taken over all collections of closed intervals I_k satisfying $A \subset \bigcup_{k=1}^{\infty} I_k$ and $l(I)$ describes the length of an interval I. An important property of λ is that it is countably additive, that is if $\{A_i\}$ is a countable collection of pairwise disjoint measurable sets then

$$\lambda \left(\bigcup_i A_i \right) = \sum_i \lambda(A_i).$$

In particular we note that if we take $A = \{a\}$ for some $a \in \mathbb{R}$, then in the definition of the measure we can take arbitrarily small intervals I_k and hence $\lambda(A) = 0$. Using the countable additivity we can then conclude that the Lebesgue measure of any countable set is zero. We say that a property holds almost everywhere (abbreviated a.e.) on a set if the measure of the set of points where it does not hold is zero. If a function $f : \mathbb{R} \to \overline{\mathbb{R}}$ (where $\overline{\mathbb{R}}$ is the extended real numbers $\mathbb{R} \cup \{+\infty, -\infty\}$) satisfies the condition that the set $\{x \in \mathbb{R} : f(x) > a\}$ is measurable for every $a \in \mathbb{R}$, then f is said to be a Lebesgue measurable function.

We are now ready to define the Lebesgue integral. For every $A \subseteq \mathbb{R}$ we define the characteristic function of A as the function $1_A : \mathbb{R} \to \mathbb{R}$ given by

$$1_A(x) = \begin{cases} 1, & \text{if } x \in A \\ 0, & \text{otherwise.} \end{cases}$$

For these functions we define the Lebesgue integral as

$$\int_{\mathbb{R}} 1_A d\lambda := \lambda(A)$$

Functions on the form $\sum_{k=0}^{\infty} a_k 1_{A_k}$ where $a_k \in \overline{\mathbb{R}}$ and $A_k \in M$ are known as countably simple functions. If all a_k are non-negative, then the integral is defined as

$$\int_{\mathbb{R}} \left(\sum_{k=0}^{\infty} a_k 1_{A_k} \right) d\lambda := \sum_{k=0}^{\infty} a_k \lambda(A_k).$$

For a function f, define $f^+ = \max(f, 0)$, $f^- = \max(-f, 0)$ so that $f = f^+ - f^-$. For any measurable simple function f we then set

$$\int_{\mathbb{R}} f d\lambda := \int_{\mathbb{R}} f^+ d\lambda - \int_{\mathbb{R}} f^- d\lambda$$

if at least one of the two integrals in the right hand side is finite. Finally, for any function $f : \mathbb{R} \to \overline{\mathbb{R}}$, we define the upper integral

$$\overline{\int_{\mathbb{R}}} f d\lambda := \inf \left\{ \int_{\mathbb{R}} g d\lambda : g \text{ is measurable, countably simple, and } g \geq f \text{ a.e.} \right\}$$

and the lower integral

$$\underline{\int_{\mathbb{R}}} f d\lambda := \sup \left\{ \int_{\mathbb{R}} g d\lambda : g \text{ is measurable, countably simple, and } g \leq f \text{ a.e.} \right\}.$$

We say that the integral of a measurable function $f : \mathbb{R} \to \overline{\mathbb{R}}$ exists if the upper and lower integrals are equal and denote their common value by $\int_{\mathbb{R}} f d\lambda$. If this value is finite we say that f is Lebesgue integrable.

The Lebesgue integral behaves much like the Riemann integral. In particular, any function that is Riemann integrable is also Lebesgue integrable and the values of the integrals are the same. With this in mind we will use the familiar notation $\int_{\mathbb{R}} f(x)dx$ for Lebesgue integrals in the following sections.

Of course we also want to be able to integrate over subsets of the real line so for any measurable subset $A \in M$ and measurable function $f : \mathbb{R} \to \overline{\mathbb{R}}$ we define

$$\int_A f d\lambda := \int_{\mathbb{R}} 1_A f d\lambda.$$

5.3 Bounded variation

Yet another notion of smoothness of functions is if they are of bounded variation.

Definition 5.6. The total variation of a function $f : [a, b] \to \mathbb{R}$, denoted $\bigvee_a^b f$ is

$$\bigvee_a^b f = \sup \left\{ \sum_{i=1}^n |f(x_i) - f(x_{i-1})| \right\}$$

where the supremum is taken over all partitions $a = x_0 < x_1 < \ldots < x_n = b$. If this is finite we say that f is of bounded variation.

Clearly, a function of bounded variation must be bounded. Functions of bounded variation need not be continuous on all of their domain however. As an example, consider a piecewise constant function. The total variation will be the sum of the sizes of jumps but those same jumps are points of discontinuity. We also note that if $[x, y] \subseteq [a, b]$, then $\bigvee_x^y f \leq \bigvee_a^b f$. A consequence of this is that compositions of functions that are of bounded variation are also of bounded variation. Indeed, if $g : [c, d] \to [a, b]$ and $f : [a, b] \to \mathbb{R}$ are of bounded variation, then

$$\bigvee_c^d (f \circ g) \leq \bigvee_a^b f$$

since the image of g lies in the interval $[a, b]$. Another simple case is if f is monotone on $[a, b]$. Then the total variation is simply given by

$$\bigvee_a^b f = |f(a) - f(b)|.$$

Now consider two functions $f, g : [a, b] \to \mathbb{R}$ that are of bounded variation and let $a = x_0 < x_1 < \ldots < x_n = b$ be a partition of the interval. Then

$$\sum_{i=1}^{n} |(f+g)(x_i) - (f+g)(x_{i-1})| \leq \sum_{i=1}^{n} |f(x_i) - f(x_{i-1})| + \\ + \sum_{i=1}^{n} |g(x_i) - g(x_{i-1})| \leq \bigvee_a^b f + \bigvee_a^b g.$$

We can conclude that the sum of to functions of bounded variation are also of bounded variation. Similarly for products we get

$$\begin{aligned} &\sum_{i=1}^{n} |(fg)(x_i) - (fg)(x_{i-1})| = \\ &= \sum_{i=1}^{n} |f(x_i)g(x_i) - f(x_i)g(x_{i-1}) + f(x_i)g(x_{i-1}) - f(x_i)g(x_{i-1})| \leq \\ &\leq \sum_{i=1}^{n} |f(x_i)||g(x_i) - g(x_{i-1})| + \sum_{i=1}^{n} |g(x_{i-1})||f(x_i) - f(x_{i-1})| \leq \\ &\leq \sup |f| \bigvee_a^b g + \sup |g| \bigvee_a^b f. \end{aligned}$$

Since the last expression is finite, we find that products of functions of bounded variation are themselves also of bounded variation. If we also assume that $\inf |g| > 0$, then the quotient $\frac{f}{g}$ will also be of bounded variation. To see this, let $h(x) = \frac{1}{x}$ and note that on any closed interval that does not include the origin, h is monotone and hence of bounded variation. Thus $h \circ g = \frac{1}{g}$ is a composition of functions of bounded variation like we discussed earlier so it is also of bounded variation. Finally we note that $\frac{f}{g}$ is a product of two functions of bounded variation, and hence itself of bounded variation.

While functions that are of bounded variation need not be even continuous on their domains, we can use the theory of the Lebesgue measure and integral to get a similar result. For a proof see [11].

Theorem 5.7. *If a function $f : [a, b] \to \mathbb{R}$ is of bounded variation, then f is differentiable almost everywhere on $[a, b]$ and f' is Lebesgue integrable on $[a, b]$.*

Since differentiablity at a point implies continuity at that point, it follows from this theorem that functions of bounded variation are also continuous almost everywhere. The main purpose for introducing the notion of bounded variation in this text is that absolute continuity is stronger than this property.

Theorem 5.8. *If* $f : [a, b] \to \mathbb{R}$ *is absolutely continuous on* $[a, b]$ *then* f *is of bounded variation on* $[a, b]$.

Proof. Since f is absolutely continuous there is $\delta > 0$ such that

$$\sum |f(y_i) - f(x_i)| < 1$$

whenever $\{[x_i, y_i] : i = 1, \ldots, n\}$ is a finite collection of pairwise disjoint subintervals of $[a, b]$ satisfying

$$\sum_{i=1}^{n} |y_i - x_i| < \delta.$$

Now let N be the smallest integer such that $N > \frac{b-a}{\delta}$ and let $a_j = a + \frac{j(b-a)}{N}$ for $j = 0, 1, \ldots, N$. We find that $a_j - a_{j-1} = \frac{b-a}{N} < \delta$ and so

$$\bigvee_{a_{j-1}}^{a_j} f < 1$$

for $j = 1, \ldots, N$. Hence

$$\bigvee_a^b f = \sum_{j=1}^{N} \bigvee_{a_{j-1}}^{a_j} f < N$$

so f is of bounded variation. □

As a consequence of this theorem, properties of functions of bounded variation such as boundedness and differentiability almost everywhere also hold for absolutely continuous functions. In particular, we note that if an absolutely continuous function were nowhere differentiable then it could not also be differentiable almost everywhere. Thus functions such as the Weierstrass functions discussed earlier cannot be absolutely continuous.

5.4 The fundamental theorem of calculus

In this section we shall look at indefinite integrals, that is a function $F : [a, b] \to \mathbb{R}$ that is of the form

$$F(x) = \int_a^x f(t)dt + C$$

where $C \in \mathbb{R}$. In the case where f is Riemann integrable, the Fundamental Theorem of calculus tells us that

(i) F is continuous on $[a, b]$. If f is continuous at a point $x_0 \in [a, b]$, then F is differentiable at x_0 and $F'(x_0) = f(x_0)$.

(ii) If there is a differentiable function $G(x)$ on $[a, b]$ such that $G'(x) = f(x)$, then

$$\int_a^x f(t)dt = G(x) - G(a).$$

If we instead want to look at the case when f is Lebesgue integrable, then the statement must be changed somewhat. The result in (i) is generalised by the following lemma.

Lemma 5.9.

(i) If $f : [a, b] \to \mathbb{R}$ is absolutely continuous and $f' = 0$ almost everywhere, then f is constant.

(ii) If f is Lebesgue integrable on $[a, b]$ and $F : [a, b] \to \mathbb{R}$ is given by

$$F(x) = \int_a^x f(t)dt + C$$

then $F' = f$ almost everywhere.

A proof of this can be found in [10]. Another result we will need is the following, a proof of which is found in [2, p. 131].

Lemma 5.10. *If $|f|$ is integrable on $[a, b]$, then for every $\epsilon > 0$ there is $\delta > 0$ such that if $E \subseteq [a, b]$ is measurable and $\lambda(E) < \delta$, then*

$$\left| \int_E f(t)dt \right| < \epsilon. \tag{4}$$

We are now ready to establish the connection between absolutely continuous functions and the fundamental theorem of calculus.

Theorem 5.11. *A function $F : [a, b] \to \mathbb{R}$ is absolutely continuous if and only if*

$$F(x) = F(a) + \int_a^x f(t)dt$$

for some function f that is Lebesgue integrable on $[a, b]$.

Proof. First, assume that F is the indefinite integral of f as given above and let $\epsilon > 0$. By Lemma 5.10 we can find δ such that the inequality in (4) holds. Now let $\{[x_i, y_i] : i = 1, \dots, n\}$ be a finite collection of mutually disjoint subintervals of $[a, b]$ satisfying

$$\lambda\left(\bigcup_{i=1}^n [x_i, y_i]\right) = \sum_{i=1}^n |y_i - x_i| < \delta.$$

Then

$$\sum_{i=1}^n |F(y_i) - F(x_i)| = \sum_{i=1}^n \int_{[x_i,y_i]} f(t)dt = \int_{\bigcup [x_i,y_i]} f(t)dt < \epsilon$$

by Lemma 5.10. Thus F is absolutely continuous on $[a, b]$.

Next we assume that F is absolutely continuous on $[a, b]$ and wish to show that it can be written as an indefinite integral. Let

$$G(x) = F(a) + \int_a^x F'(t)dt$$

where the integrand exists almost everywhere and is integrable by Lemma 5.9. The lemma also tells us that $G'(x) = F'(x)$ for almost every $x \in [a, b]$. Hence $(F - G)'(x) = 0$ for almost every $x \in [a, b]$ and so $F - G$ must be a constant function. Since $F(a) = G(a)$ it then follows that $F = G$ and so F is given by an indefinite integral. □

6 Rectifiable curves

In this section we will study the arc length of curves in $\mathbb{R}^2$, in particular the curves given by the graph of a function.

To start with, we define a curve in $\mathbb{R}^2$ as a continuous function $\gamma : [a, b] \to \mathbb{R}^2$. In particular the graph of a continuous function $f : [a, b] \to \mathbb{R}$ is the curve $\gamma(x) = (x, f(x)), x \in [a, b]$. What we are interested in is the arc length of these curves. To calculate it we approximate by polgonal paths. Let $P = \{x_0, x_1, \ldots, x_n\}$ be a partition of $[a, b]$, that is the elements of P satisfy $a = x_0 < x_1 < \ldots < x_n = b$. Then

$$L(\gamma, P) = \sum_{i=1}^{n} \|\gamma(x_i) - \gamma(x_{i-1})\|$$

describes the length of the polygonal path with vertices given by the elements of P. Now define

$$L(\gamma) = \sup \{L(\gamma, P) : P \text{ is a partition of } [a, b]\}.$$

as the arc length of the curve γ. If $L(\gamma)$ is finite then we say that the curve γ is rectifiable.

This connects to functions of bounded variation introduced in the previous section by the following statement, the proof of which can be found in [11].

Theorem 6.1. *A curve $\gamma = (\gamma_1, \gamma_2) : [a, b] \to \mathbb{R}$ is rectifiable if and only if the functions γ_1 and γ_2 are of bounded variation on $[a, b]$.*

Proof. First we show that γ_1 and γ_2 being of bounded variation implies that γ is rectifiable. To do this we will use that if α and β are positive numbers, then $\sqrt{\alpha + \beta} \le \sqrt{\alpha} + \sqrt{\beta}$. This follows from the observation that

$$\alpha + \beta \le \alpha + 2\sqrt{\alpha\beta} + \beta = (\sqrt{\alpha} + \sqrt{\beta})^2$$

holds for all positive α and β. Let $P = \{x_0, x_1 \ldots, x_n\}$ be a partition of $[a, b]$. We get

$$\begin{aligned} L(\gamma, P) &= \sum_{i=1}^{n} \|\gamma(x_i) - \gamma(x_{i-1})\| = \\ &= \sum_{i=1}^{n} \sqrt{(\gamma_1(x_i) - \gamma_1(x_{i-1}))^2 + (\gamma_2(x_i) - \gamma_2(x_{i-1}))^2} \\ &\leq \sum_{i=1}^{n} |\gamma_1(x_i) - \gamma_1(x_{i-1})| + \sum_{i=1}^{n} |\gamma_2(x_i) - \gamma_2(x_{i-1})| \leq \bigvee_a^b \gamma_1 + \bigvee_a^b \gamma_2 \end{aligned}$$

which is finite by assumption. Thus γ is rectifiable.

On the other hand, if γ is rectifiable, then for any partition $P = \{x_0, x_1 \ldots, x_n\}$ of $[a, b]$ we have

$$\sum_{i=1}^{n} |\gamma_k(x_i) - \gamma_k(x_{i-1})| \leq \sum_{i=1}^{n} \|\gamma(x_i) - \gamma(x_{i-1})\|$$

and so

$$\bigvee_a^b \gamma_k \leq L(\gamma)$$

for $k = 1, 2$ so both γ_1 and γ_2 are of bounded variation. □

This theorem allows us to conclude that, for example, the curve given by the graph of the function

$$g(x) = \begin{cases} x^2 \sin^2\left(\frac{1}{x}\right), x \in (0, 1] \\ 0, x = 0 \end{cases}$$

(see figure 2) that was studied earlier is rectifiable since this function is absolutely continuous and hence of bounded variation. We can of course also use the theorem to show that a curve is not rectifiable.

Example 6.2. We now consider the curve given by the graph of the function

$$f(x) = \begin{cases} x \sin(\frac{1}{x}), x \in (0, 1] \\ 0, x = 0 \end{cases}$$

seen in figure 3.

We have seen that this function is not absolutely continuous, but it is in fact not even of bounded variation and hence the curve is not rectifiable. To see this we consider the partition

$$\begin{aligned} x_0 &= 0 \\ x_j &= \frac{2}{\pi(n - j)}, \text{ for } j = 1, \ldots, n - 1 \\ x_n &= 1. \end{aligned}$$

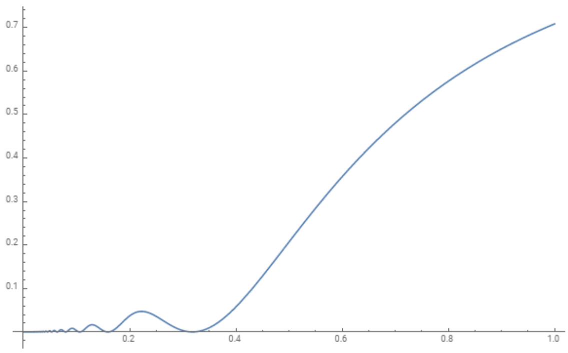

Figure 2: The graph of $x^2 \sin^2(\frac{1}{x})$

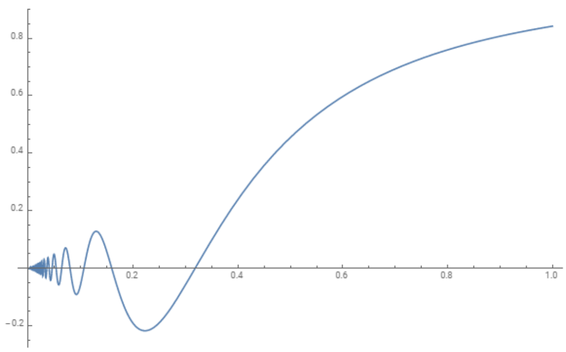

Figure 3: The graph of $x \sin(\frac{1}{x})$

If we choose n to be even, then we find that for j that is even and satisfying $0 < j < n$, it holds that

$$f(x_j) = \frac{2}{\pi(n-j)} \sin\left(\frac{\pi}{2}(n-j)\right) = 0.$$

This gives us the following estimate

$$\sum_{j=1}^{n} |f(x_j) - f(x_{j-1})| \geq \sum_{j=1}^{\frac{n}{2}} |f(x_{2j-1})| =$$
$$= \sum_{j=1}^{\frac{n}{2}} \left| \frac{2}{\pi(n-2j+1)} \sin\left(\frac{\pi}{2}(n-2j+1)\right) \right| = \frac{2}{\pi} \sum_{j=1}^{\frac{n}{2}} \frac{1}{n-2j+1}.$$

However the last expression can be made arbitrarily big by choosing n sufficiently big. We can thus conclude that the total variation of f is infinite and so f is not of bounded variation.

Unfortunately this does not help us actually calculate the length of curves. If we make some extra assumptions on γ, then we can use the following integral formula.

Theorem 6.3. *Let $\gamma = (\gamma_1, \gamma_2) : [a,b] \to \mathbb{R}^2$ be a curve such that $\gamma_1(t)$ and $\gamma_2(t)$ are continuously differentiable on $[a,b]$. Then γ is rectifiable and*

$$L(\gamma) = \int_a^b \sqrt{\left(\frac{d\gamma_1}{dt}\right)^2 + \left(\frac{d\gamma_2}{dt}\right)^2} dt.$$

Proof. Since γ_1 and γ_2 are continuously differentiable on a compact set, their respective derivatives are bounded. As a consequence γ_1 and γ_2 are Lipschitz continuous and hence absolutely continuous and in particular of bounded variation. We have thus established that γ is rectifiable.

To see that the arc length is given by the formula above, let $P = \{x_0, x_1, \ldots, x_n\}$ be a partition of $[a,b]$. By the mean value theorem we can find points $s_i \in (x_{i-1}, x_i)$ and $t_i \in (x_{i-1}, x_i)$ satisfying

$$\gamma_1(x_i) - \gamma_1(x_{i-1}) = \gamma_1'(s_i)(x_i - x_{i-1})$$

and

$$\gamma_2(x_i) - \gamma_2(x_{i-1}) = \gamma_2'(t_i)(x_i - x_{i-1})$$

for $i = 1, \ldots, n$. Then

$$L(\gamma, P) = \sum_{i=1}^{n} \|\gamma(x_i) - \gamma(x_{i-1})\| =$$
$$= \sum_{i=1}^{n} \sqrt{(\gamma_1(x_i) - \gamma_1(x_{i-1}))^2 + (\gamma_2(x_i) - \gamma_2(x_{i-1}))^2} =$$
$$= \sum_{i=1}^{n} \sqrt{(\gamma_1'(s_i))^2 + (\gamma_2'(t_i))^2}(x_i - x_{i-1}).$$

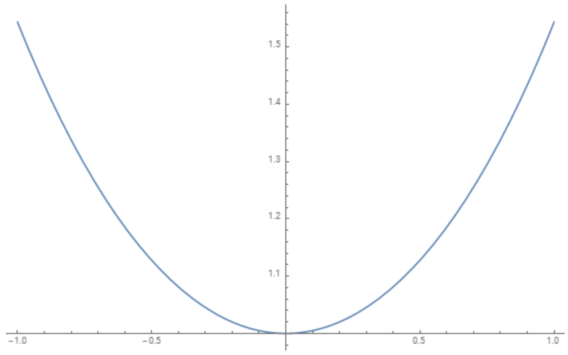

Figure 4: The graph of $\cosh(x)$

This is the Riemann sum of the desired integral associated with the partition P. Taking finer refinements, i.e. letting $\max_{i=1,\dots n}(x_i - x_{i-1}) \to 0$, thus makes this sum approach the desired integral. $\square$

Due to the square root in the integrand in this formula, it is often difficult to use this to actually calculate these values analytically and numerical methods may be needed. Nonetheless, there are functions for which this formula proves useful.

Example 6.4. Consider the curve γ given by the graph of $\cosh(x)$ defined on $[-1, 1]$, seen in figure 4. Using the formula from the last theorem and the identity $1 = \cosh^2(x) - \sinh^2(x)$ we find that

$$L(\gamma) = \int_{-1}^{1} \sqrt{1 + \sinh^2(x)}dx = \int_{-1}^{1} \sqrt{\cosh^2(x)}dx = \int_{-1}^{1} \cosh(x)dx =$$
$$= [\sinh(x)]_{-1}^{1} = e + \frac{1}{e}.$$

References

[1] Cristopher Apelian and Steve Surace. *Real and Complex Analysis*. Chapman & Hall/CRC, 2010.

[2] Soo Bong Chea. *Lebesgue Integration*. Springer, 1995.

[3] Renato Fiorenza. *Hölder and locally Hölder Continuous Functions and Open Sets of Class $C^k, C^{k,\alpha}$*. Springer, 2016.

[4] Godfrey Hardy. Weierstrass's non-differentiable function. *Transactions of the American Mathematical Society*, 17, 1916.

[5] Jian-Ke Lu. *Boundary Value Problems for Analytic Functions*. World Scientific Publishing, 1993.

[6] Walter Rudin. *Principles of Mathematical Analysis*. McGraw Hill, 1976.

[7] Gerald Teschl. *Ordinary differential equations and Dynamical Systems*. American Mathematical Society, 2012.

[8] Mathematics Department Coffee Room Wayne State University. Every convex function is locally lipschitz. *The American Mathematical Monthly*, 79, 1972.

[9] William Ziemer. *Modern Real analysis*. Springer, 2017.

[10] Lecture notes: Absolutely continuous functions. `https://sites.ualberta.ca/~rjia/Math418/Notes/chap3.pdf`. Accessed 2019-05-20.

[11] Lecture notes: Functions of bounded variation. `https://sites.ualberta.ca/~rjia/Math418/Notes/chap2.pdf`. Accessed 2019-05-22.