Algorithmic Definitions of Singular Functions

by

Daniel Bernstein

A thesis submitted in partial fulfillment of the requirements for Departmental Honors

in Mathematics

Davidson College

May 13, 2013

Advisor: Richard D. Neidinger

Second Reader: Irl C. Bivens

Contents

Abstract

A function f on an interval $[a, b]$ is *singular* if $f(a) < f(b)$, f is increasing (non-decreasing), f is continuous everywhere and $f'(x) = 0$ almost everywhere. In this thesis, we focus on the strictly increasing singular functions of DeRham and Minkowski. We propose algorithmic definitions of each of these functions, and use these new definitions to provide alternative proofs of known properties about the values and derivatives of these functions. Before that, we give some background on sets of discontinuity for general functions, for increasing functions, and some background on the derivatives of increasing functions.

1 Sets of Discontinuity and Increasing Functions

In early calculus courses, most of the functions we encounter are continuous at every point in their domain. We sometimes encounter functions that have at most a finite number of discontinuities on any finite interval but this is usually as discontinuous as it gets. One might wonder if it is possible to define a function on a finite interval that is discontinuous at every point in its domain. More generally, one might wonder about what restrictions exist on the set of points at which a function is discontinuous.

1.1 Dirichlet-Like Functions

Consider the function $D : (0,1) \to (0,1)$ defined as follows:

$$D(x) = \begin{cases} 1 & \text{if } x \in (0,1) \cap \mathbb{Q}, \\ 0 & \text{if } x \in (0,1) \setminus \mathbb{Q}. \end{cases}$$

Dirichlet proposed this function in 1829 and it was groundbreaking because hitherto, people generally thought about defining functions only using analytic expressions. It gives us an example of a function defined on an interval that is discontinuous at every point in its domain. We can extend the main idea behind Dirichlet's function to create other functions that are discontinuous on unusual sets.

Now, we give an example of a function $E : (0,1) \to (0,1)$ that is continuous at exactly one point:

$$E(x) = \begin{cases} x - \frac{1}{2} & \text{if } x \in (0,1) \cap \mathbb{Q} \\ \frac{1}{2} - x & \text{if } x \in (0,1) \setminus \mathbb{Q} \end{cases}$$

Figure 1 shows an approximate graph of E.

This function is continuous at the point $x = \frac{1}{2}$ and discontinuous everywhere else. Thomae used the idea of Dirichlet to define the function $T : (0,1) \to (0,1)$ which is

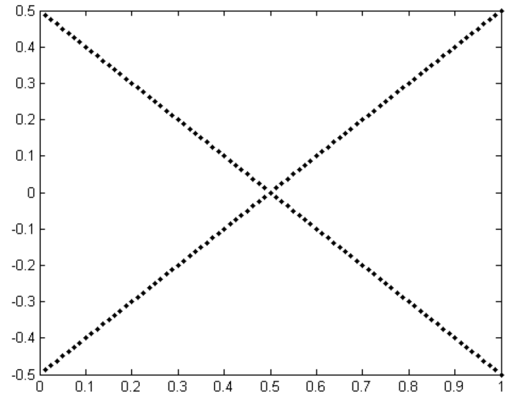
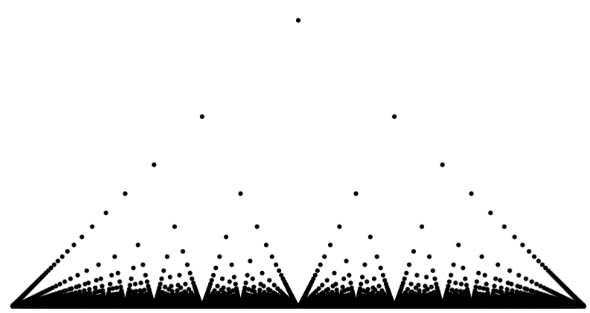


Figure 1: Graph of E

continuous at every irrational in $(0,1)$ and discontinuous at every rational in $(0,1)$. Figure 2 shows the graph of T.

$$T(x) = \begin{cases} \frac{1}{d} & \text{if } x \in \mathbb{Q} \text{ and has denominator } dinreducedform \\ 0 & \text{if } x \text{ is irrational} \end{cases}$$



Graph of T. Source: http://en.wikipedia.org/wiki/File:Thomae_function_(0,1).svg

From looking at its graph, it isn't too surprising that T it is discontinuous at each rational. It is less obvious that T is continuous at each irrational, so we will prove that formally. Let $x \in (0,1) \setminus \mathbb{Q}$. Let $\varepsilon > 0$. Let $q \in \mathbb{N}$ be such that $\frac{1}{q} < \varepsilon$. For each $i \in \{2, 3, ..., q\}$ let

$$\delta_i = \min_{p \in \{1,2,...,i-1\}} \{|x - \frac{p}{i}|\}.$$

Let $\delta = \min\{\delta_i\}$. Let $c \in (x - \delta, x + \delta)$. We will now show that $|f(x) - f(c)| < \varepsilon$, completing the proof that f is continuous at x. Since x is irrational, $f(x) = 0$ and since f is nonnegative, we have $|f(x) - f(c)| = |f(c)| = f(c)$. So now if we show $f(c) < \varepsilon$ then we are done. If c is irrational, $f(c) = 0$ so we would be done. So assume c is rational. Assume $c = \frac{n}{d}$ in reduced form. Note that $d > q$ because we defined δ such that the interval $(x - \delta, x + \delta)$ contains no fractions with denominator less than or equal to q. So $f(c) = \frac{1}{d} < \frac{1}{q} < \varepsilon$.

One might wonder, can a function be discontinuous on any arbitrary set? As we will see, the answer is no.

1.2 Sets of Discontinuity

Definition. *Let f be some function defined on $\mathbb{R}$. We define the* set of discontinuity of f*, D_f, to be the set of all points at which f is discontinuous.*

Definition. *We say that a set S is a F_σ set if S can be expressed as a countable union of closed sets.*

The goal of this section is to show that for every function f, the set D_f is an F_σ set. The structure of this section is taken from some of the exercises in [1]. First we need some more definitions.

Definition. *Let f be a function defined on $\mathbb{R}$ and let $\alpha > 0$. We say that f is α-continuous at x if there exists some $\delta > 0$ such that for all $y, z \in (x - \delta, x + \delta)$, $|f(y) - f(z)| < \alpha$.*

It follows directly from the definition of continuity that a function f is continuous at x if and only if f is α-continuous at x for each $\alpha > 0$.

Definition. *Let f be some function defined on $\mathbb{R}$. Let $\alpha > 0$. We define the* set of α-discontinuity of f*, $D_f(\alpha)$, to be the set of all points at which f is α-discontinuous.*

Lemma 1. *Let f be a function defined on $\mathbb{R}$, let $\alpha > 0$. Then $D_f(\alpha)$ is closed.*

Proof. Let x be a limit point of $D_f(\alpha)$. We will show that f is not α-continuous at x. Let $\{x_n\}$ be a sequence of points in $D_f(\alpha)$ such that $x_n \to x$ as $n \to \infty$. Let $\delta > 0$. We want to find $y, z \in (x-\delta, x+\delta)$ such that $|f(y) - f(z)| \geq \alpha$. Let n be such that $|x - x_n| < \delta$ and let $\hat{\delta} = \delta - |x - x_n|$. Since f is not α-continuous at x_n, then there exist $y, z \in (x - \hat{\delta}, x + \hat{\delta})$ such that $|f(y) - f(z)| \geq \alpha$. Note that $(x_n - \hat{\delta}, x_n + \hat{\delta}) \subset (x - \delta, x + \delta)$ and thus $y, z \in (x - \delta, x + \delta)$, so f is not α-continuous at x. □

Lemma 2. *Let f be a function defined on $\mathbb{R}$. Let $0 < \alpha_1 < \alpha_2$. Then $D_f(\alpha_2) \subseteq D_f(\alpha_1)$.*

Proof. Let $x \in D_f(\alpha_2)$. So f is not α_2-continuous at x. So for all $\delta > 0$, there exist $y, z \in (x - \delta, x + \delta)$ such that $|f(y) - f(z)| \geq \alpha_2 > \alpha_1$. So f is also not α_1-continuous at x. □

Theorem 3. *Let f be a function defined on $\mathbb{R}$. Then D_f is an F_σ set.*

Proof. Since each set $D_f(\alpha)$ is closed, we are done if we show

$$D_f = \bigcup_{q \in \mathbb{Q}^+} D_f(q).$$

Let $x \in D_f$. Then there exists some $\alpha > 0$ such that f is not α-continuous at x, because otherwise f would be continuous at x. So $x \in D_f(\alpha)$. Let $q \in \mathbb{Q}^+$ such that $q < \alpha$. By Lemma 2, $D_f(\alpha) \subseteq D_f(q)$ and thus $x \in D_f(q)$. Now let $x \in D_f(q)$ for some $q \in \mathbb{Q}^+$. Since f is not q-continuous at x, we have that f is not continuous at x and thus $x \in D_f$. □

An example of a set that is not F_σ is the set of all irrational numbers in an interval. So there cannot exist a function that is discontinuous only at each irrational in a given interval. The proof of this is outside the scope of this paper.

1.3 Increasing Functions

We say that a function f is *increasing* if $x < y$ implies $f(x) \leq f(y)$. We say that a function f is *strictly increasing* if $x < y$ implies $f(x) < f(y)$. Note that all strictly increasing functions are increasing. As we will see, the criteria for the set on which an increasing function is discontinuous are more restrictive then that for a general function.

Theorem 4. *Let $f : (0,1) \to \mathbb{R}$ be increasing. Then for each $c \in (0,1)$, $\lim_{x\to c^+} f(x)$ and $\lim_{x\to c^-} f(x)$ exist.*

Proof. Let $c \in (0,1)$. Note that since f is increasing, f is bounded below on $(c,1)$ by $f(c)$. Therefore, by the completeness of the reals, f has a greatest lower bound on $(c,1)$ so we can define

$$L = \inf_{x\in(c,1)} f(x).$$

Now let $\varepsilon > 0$. Since L is a greatest lower bound, there exists some $y \in (c,1)$ such that $f(y) < L + \varepsilon$. Let $\delta = y - c$; note that $y > c$ so $\delta > 0$ and $y = c + \delta$. This gives us

$$f(c+\delta) < L + \varepsilon.$$

Since f is increasing, for any $x \in (c, c+\delta)$, we have

$$f(x) \leq f(c+\delta) < L + \varepsilon$$

and thus

$$f(x) - L < \varepsilon.$$

Since L is a lower bound for f on $(c,1)$, $L \leq f(x)$ and thus

$$|f(x) - L| = f(x) - L < \varepsilon,$$

thus showing that $\lim_{x\to c^+} f(x)$ exists and is equal to L. Showing the existence of

$\lim_{x\to c^-} f(x)$ is similar. $\square$

Definition. *Given a function $f : D \to \mathbb{R}$ where $D \subseteq \mathbb{R}$, we say that f has a* jump discontinuity *at x if* $\lim_{x\to c^+} f(x)$ *and* $\lim_{x\to c^-} f(x)$ *both exist but are not equal.*

Theorem 5. *Let $f : (0,1) \to \mathbb{R}$ be increasing. Then for all $c \in (0,1)$, either f is continuous at c or f has a jump discontinuity at c.*

Proof. Let $c \in (0,1)$. The limit $\lim_{x\to c} f(x)$ either exists or doesn't exist.

Assume that $\lim_{x\to c} f(x)$ exists. We will now show that $f(c) = \lim_{x\to c} f(x)$. Assume for the sake of contradiction that $f(c) \neq \lim_{x\to c} f(x)$. Then $f(c) < \lim_{x\to c} f(x)$ or $f(c) > \lim_{x\to c}$. Assume $f(c) < \lim_{x\to c} f(x)$. Then $f(c) < \lim_{x\to c^-} f(x)$. Then there exists some $x < c$ such that $f(x) > f(c)$ which is a contradiction because f is increasing. Assume $f(c) > \lim_{x\to c} f(x)$. Then $f(c) > \lim_{x\to c^+} f(x)$. Then there exists some $x > c$ such that $f(x) < f(x)$ which is a contradition because f is increasing. Thus if $\lim_{x\to c} f(x)$ exists, it must equal $f(c)$ so f is continuous at c.

Now assume that $\lim_{x\to c} f(x)$ does not exist. By Theorem 4, $\lim_{x\to c^+} f(x)$ and $\lim_{x\to c^-} f(x)$ both exist so we must have $\lim_{x\to c^+} f(x) \neq \lim_{x\to c^-} f(x)$ for $\lim_{x\to c} f(x)$ to not exist. So f has a jump discontinuity at c. $\square$

Theorem 6. *Let $f : (0,1) \to \mathbb{R}$ be increasing. Let D_f denote the set of points at which f is discontinuous. Then D_f is countable.*

Proof. Informally, we prove this by choosing a rational number in the interval defined by each jump discontinuity. Since our function is increasing, no two such intervals will overlap, so each rational number chosen in this way will be unique. More formally, we will construct a function $g : D_f \to \mathbb{Q}$ that is one-to-one.

Let $c \in D_f$. Then $\lim_{x\to c^+} f(x)$ and $\lim_{x\to c^-} f(x)$ both exist with $\lim_{x\to c^-} f(x) < \lim_{x\to c^+} f(x)$. Let $q_c \in \mathbb{Q} \cap (\lim_{x\to c^-} f(x), \lim_{x\to c^+} f(x))$ and define $g(c) = q_c$.

To see that g is one-to-one, let $c, d \in D_f$ such that $g(c) = g(d)$. Assume for the sake of contradiction that $c \neq d$ and without loss of generality, assume $c < d$. Since

$g(c) \in (\lim_{x\to c^-} f(x), \lim_{x\to c^+} f(x))$ and $g(d) \in (\lim_{x\to d^-} f(x), \lim_{x\to d^+} f(x))$, we must have

$$\lim_{x\to c^+} f(x) > g(x) = g(d) > \lim_{x\to d^-} f(x). \tag{1}$$

Let $\{c_n\}$ be a decreasing sequence such that $c_n \to c$. Let $\{d_n\}$ be an increasing sequence such that $d_n \to d$. Since $c < d$ and $\{c_n\}$ is decreasing and $\{d_n\}$ is increasing, there must exist some N such that for all $n \geq N, c_n < d_n$. Since f is monotone, then for each $n \geq N$, $f(c_n) < g(d_n)$ implying that $\lim_{x\to c^+} f(x) \leq \lim_{x\to d^-} f(x)$ which contradicts (1). □

Can any countable set be the set of discontinuity for a monotone function? As it turns out, yes. We address this in the theorem below.

Theorem 7. *Let $D = \{x_1, x_2, ...\} \subset (0,1)$. Then there exists a monotone function $f : (0,1) \to (0,1)$ such that f is discontinuous precisely at the points in D.*

Proof. We define f as follows:

$$f(x) = \sum_{n:x_n<x} \frac{1}{2^n}.$$

We will now show that f is monotone. Let $x, y \in (0,1)$ such that $x < y$. Then note that

$$f(y) = \sum_{n:x_n<y} \frac{1}{2^n} = \sum_{n:x_n<x} \frac{1}{2^n} + \sum_{n:x\leq x_n<y} \frac{1}{2^n} = f(x) + \sum_{n:x\leq x_n<y} \frac{1}{2^n} \geq f(x)$$

since each term in the sum is positive. So f is monotone.

Now we show that f is discontinuous at each x_i by showing that the right and left hand limits of f at x_i are unequal for any i.

$$\lim_{x\to x_i^+} f(x) = \lim_{x\to x_i^+} \sum_{n:x_n<x} \frac{1}{2^n} \geq \sum_{n:x_n\leq x_i} \frac{1}{2^n} > \sum_{n:x_n<x_i} \frac{1}{2^n} = f(x_i).$$

$$\lim_{x \to x_i^-} f(x) = \lim_{x \to x_i^-} \sum_{n: x_n < x} \frac{1}{2^n} \leq \sum_{n: x_n < x_i} \frac{1}{2^n} = f(x_i).$$

So $\lim_{x \to x_i^+} f(x)$ and $\lim_{x \to x_i^-} f(x)$ are unequal and thus f is discontinuous at each x_i.

□

We now give a theorem about *strictly* increasing functions which will be useful later.

Theorem 8. *Let $f : (0,1) \to (0,1)$ be surjective and strictly increasing. Then f is continuous.*

Proof. Let $x \in (0,1)$. Let $\varepsilon > 0$. Let $z_1 \in (f(x) - \varepsilon, f(x)) \cap (0,1)$ and let $z_2 \in (f(x), f(x) + \varepsilon)) \cap (0,1)$. Since f is surjective, there exist $y_1, y_2 \in (0,1)$ such that $f(y_1) = z_1$ and $f(y_2) = z_2$. Since f is strictly increasing, $y_1 < x < y_2$. Let $\delta = \min\{x - y_1, y_2 - x\}$, note $\delta > 0$. Let $c \in (x - \delta, x + \delta)$. Assume $c < x$. Since f is increasing and since $y_1 \leq x - \delta$, we have $z_1 \leq f(x - \delta) \leq f(c) \leq f(x)$ and thus $f(x) - \varepsilon < c < f(x) + \varepsilon$. The case where $c > x$ is similar. □

We now consider properties of the derivatives of increasing functions.

2 Derivative of an Increasing Function

In this section we will prove a major theorem about existence properties of the derivative of an increasing function. Then, we will briefly introduce The Cantor Function (for a more complete discussion of The Cantor Function, see [5]).

2.1 Existence

Earlier we saw that it is possible to construct a function that is discontinuous everywhere, but if we want the function to be monotone, it can be discontinuous only at a countable set. Since differentiability implies continuity, it follows that we can construct a function that is not differentiable anywhere. If we restrict our attention to continuous functions, this is still possible (see [1, p.146] for an example of a continuous, nowhere differentiable function). What if we restrict our attention to monotone functions? Can we construct a monotone function that is not differentiable anywhere? As we will see, such a function cannot exist - we will prove that the derivative of a monotone function must exist almost everywhere. But first, we need some preliminary results.

Theorem 9 (Fatou's Lemma). *If $\{f_n\}_{n=1}^{\infty}$ is a sequence of nonnegative measurable functions and $f_n(x) \to f(x)$ for almost all x in some set E, then*

$$\int_E f \leq \liminf \int_E f_n$$

For a proof of Fatou's Lemma, see [9].

Definition. *Let I be an interval and let E be a set. Then $l(I)$ denotes the length of I and $m(E)$ denotes the Lebesgue measure of E.*

Definition. *Let $E \subseteq \mathbb{R}$. Let $\mathscr{I}$ be a collection of intervals of positive length that covers E. Then $\mathscr{I}$* ***covers E in the sense of Vitali*** *if for each $\varepsilon > 0$ and for each $x \in E$, there exists an interval $I \in \mathscr{I}$ such that $x \in I$ and $l(I) < \varepsilon$. We may also say that $\mathscr{I}$ is a* ***Vitali covering*** *of E.*

Lemma 10. *Let E be a set of finite Lebesgue measure and let $\mathscr{I}$ be a set of closed intervals of positive length that covers E in the sense of Vitali. Then, given $\varepsilon > 0$, there exists a finite disjoint collection of intervals $\{I_1, \ldots, I_N\}$ such that*

$$m\left(E \setminus \bigcup_{n=1}^{N} I_n\right) < \varepsilon.$$

Proof. Let O be an open set of finite measure that contains E. We first define the set $\mathscr{I}^* \subseteq \mathscr{I}$ where each $I \in \mathscr{I}^*$ is contained in O, then we note that $\mathscr{I}^*$ is also a Vitali covering of E. To see this, we let $\varepsilon > 0$ and let $x \in E$. Since O is open, there exists a

$\delta > 0$ such that $(x-\delta, x+\delta) \subset O$. Let $\hat{\varepsilon} = \min\{\varepsilon, \delta\}$. Since $\mathscr{I}$ is a Vitali covering of E, there exists some $I \in \mathscr{I}$ such that $x \in I$ and $l(I) < \hat{\varepsilon}$. Then, I is contained entirely within O so $I \in \mathscr{I}^*$. Since $l(I) < \varepsilon$, $\mathscr{I}^*$ is a Vitali covering of E. So, now we will assume without loss of generality that each $I \in \mathscr{I}$ is contained entirely within O.

Now, we inductively create a sequence $\{I_n\}$ of disjoint intervals in $\mathscr{I}$. Let I_1 be any interval in $\mathscr{I}$. Supposing that $I_1, \dots I_n$ have all been chosen. If $E \subseteq \bigcup_{k=1}^{n} I_k$, then for any $\varepsilon > 0$, we have $m\left(E \setminus \bigcup_{n=1}^{N} I_n\right) < \varepsilon$ so we would be done. So, assume this is not the case. Now we show how to choose I_{n+1}. Let k_n be the supremum of the lengths of intervals in $\mathscr{I}$ that do not intersect any of the intervals $I_1, \dots, I_n$. Since each element of $\mathscr{I}$ is contained in O, we have $k_n \leq m(O) < \infty$. Then, we choose an interval $I_{n+1} \in \mathscr{I}$ such that I_{n+1} is disjoint from all of $I_1, \dots, I_n$ and $l(I_{n+1}) > \frac{1}{2}k_n$. We can choose I_{n+1} to be disjoint from the finite collection $I_1, \dots, I_n$ because $\mathscr{I}$ is a Vitali covering, so I_{n+1} can be arbitrarily small. We can choose I_{n+1} to have length greater than $\frac{1}{2}k_n$ because k_n is the supremum of the lengths of intervals that are disjoint from $I_1, \dots, I_n$.

Thus, we have an infinite sequence $\{I_n\}$ of disjoint intervals of $\mathscr{I}$. Since $\bigcup_{n=1}^{\infty} I_n \subseteq O$, we have $\sum_{n=1}^{\infty} l(I_n) \leq m(O) < \infty$. This implies two things. First, it implies $\lim_{n\to\infty} l(I_n) = 0$. Otherwise there would exist some $B > 0$ such that for some positive integer k, for each $n \geq k$, we would have $l(I_n) \geq B$ which would give $\sum_{n=1}^{\infty} I_n = \infty$. Second, this implies that for any $\varepsilon > 0$, we can choose a positive integer N such that

$$\sum_{n=N+1}^{\infty} l(I_n) < \varepsilon/5.$$

Now, let

$$R = E \setminus \bigcup_{n=1}^{N} I_n.$$

If we show that $m(R) < \varepsilon$, then we will be done.

Let $x \in R$. Note that R does not intersect $\bigcup_{n=1}^{N} I_n$. Then, since $\bigcup_{n=1}^{N} I_n$ is the union of finitely many closed intervals and thus itself closed, we can find a $\delta > 0$ such that

$(x-\delta, x+\delta)$ does not intersect $\bigcup_{n=1}^{N} I_n$. Then, since $\mathscr{I}$ is a Vitali covering, there exists an $I \in \mathscr{I}$ such that $x \in I \subset (x-\delta, x+\delta)$. So I also does not intersect $\bigcup_{n=1}^{N} I_n$.

Now let n be a positive integer such that for all $i \leq n$, I does not intersect I_i. Then, since k_n is the supremum of the lengths of the intervals that do not intersect any of $I_1, \ldots, I_n$, $l(I) \leq k_n$. By construction, $l(I_{n+1}) > \frac{1}{2}k_n$ which gives $l(I) \leq k_n < 2l(I_{n+1})$. Now, I_{n+1} may intersect I, and for n large enough, it will because $\lim_{i\to\infty} l(I_i) = 0$ and each interval in $\mathscr{I}$ has positive length. So fix n to be the smallest integer such that I intersects I_{n+1}. Note that we must have $n \geq N$ and $l(I) < k_n < 2l(I_{n+1})$. Since $x \in I$ and since I has a point in common with I_{n+1}, the distance from x to the midpoint of I_{n+1} is at most $l(I) + \frac{1}{2}l(I_{n+1}) \leq \frac{5}{2}l(I_n)$. For each i, define J_i to be the interval that shares a midpoint with I_i that is five times the length. Then x must be contained in the interval J_{n+1}. Since x is an arbitrary point in R, we have that

$$R \subseteq \bigcup_{i=N+1}^{\infty} J_i.$$

Hence

$$m(R) \leq \sum_{i=N+1}^{\infty} l(J_i) = 5 \sum_{i=N+1}^{\infty} l(I_i) < \varepsilon$$

□

Definition. *Let f be a function on the real numbers. The following four quantities are called the* ***derivates*** *of f at x:*

$$D^+f(x) = \overline{\lim_{h\to 0^+}} \frac{f(x+h)-f(x)}{h} \qquad , D^-f(x) = \overline{\lim_{h\to 0^-}} \frac{f(x+h)-f(x)}{h},$$

$$D_+f(x) = \underline{\lim_{h\to 0^+}} \frac{f(x+h)-f(x)}{h} \qquad , D_-f(x) = \underline{\lim_{h\to 0^-}} \frac{f(x+h)-f(x)}{h}.$$

By definition of limit superior and limit inferior, we have $D^+f(x) \geq D_+f(x)$ and $D^-f(x) \geq D_-f(x)$. The function f is differentiable at x if and only if $D^+f(x) =$

$D_+f(x) = D^-f(x) = D_-f(x) \neq \infty$, and the derivative of f at x is the common value.

Theorem 11 (Lebesgue's Theorem)**.** *Let $f : [a, b] \to \mathbb{R}$ be an increasing function. Then f is differentiable almost everywhere.*

Proof. We prove this theorem by showing that the subset of $[a, b]$ whose elements have unequal derivates is of measure zero. We consider only the set E where for all $x \in E$, $D^+f(x) > D_-f(x)$. The cases where other derivates are unequal is similarly handled. Now, let u, v be rational numbers with $u > v$ and define

$$E_{u,v} = \{x : D^+f(x) > u > v > D_-f(x)\}.$$

Then we have the following identity:

$$E = \bigcup_{u,v \in \mathbb{Q}, u>v} E_{u,v}.$$

This union is countable so if we show that an arbitrary $E_{u,v}$ is of measure zero, then it follows that E is of measure zero and we are done.

If $E_{u,v}$ is empty, then we are done so assume that $E_{u,v}$ is nonempty. Let $s = m(E_{u,v})$ and let $\varepsilon > 0$. Let O be an open set that contains $E_{u,v}$ such that $m(O) < s + \varepsilon$. For an $x \in E_{u,v}$, we can choose an arbitrarily small $h > 0$ such that $[x - h, x]$ is contained in O and so that

$$\frac{f(x) - f(x - h)}{h} < v. \tag{2}$$

For each $x \in E_{u,v}$, define $\mathscr{I}_x$ to be the collection of all intervals $[x - h, x]$ that are contained in O and satisfy (2) and let $\mathscr{I} = \bigcup_{x \in E_{u,v}} \mathscr{I}_x$. Then $\mathscr{I}$ is a Vitali covering of $E_{u,v}$ so by Lemma 10, we can choose a finite, disjoint collection $\{I_1, \ldots, I_N\}$ of intervals in $\mathscr{I}$ such that

$$m(E_{u,v} \setminus \bigcup_{n=1}^{N} I_n) < \varepsilon.$$

For each $n = 1, \ldots, N$ let x_n and h_n be the x and h that were used to create I_n. Then for each $n = 1, \ldots, N$, (2) gives

$$f(x_n) - f(x_n - h_n) < vh_n.$$

We can sum all of these to get

$$\sum_{n=1}^{N} f(x_n) - f(x_n - h_n) < v \sum_{n=1}^{N} h_n. \tag{3}$$

Note that h_n is the width of the interval I_n and that I_n's are disjoint and contained within O which gives us $\sum_{n=1}^{N} h_n \leq m(O) < s + \varepsilon$. Plugging this into (3) gives

$$\sum_{n=1}^{N} f(x_n) - f(x_n - h_n) < v(s + \varepsilon). \tag{4}$$

Let $A_n = \mathrm{Int}(I_n) \cap E_{u,v}$ and let $A = \bigcup_{n=1}^{N} A_n$. Then, $m(A) > s - \varepsilon$. Since each point in A is in the interior of some I_n, then for each $y \in A$, we can choose an arbitrarily small k such that $[y, y + k]$ is contained in some I_n; and since each y in A is also in $E_{u,v}$, we have $D^+ f(x) > u$, so we can choose an arbitrarily small k such that $\frac{f(y+k)-f(y)}{k} > u$. Now, for each $y \in A$, let $\mathscr{J}_y$ be the collection of intervals $[y, y + k]$ with k small enough to meet both of the above conditions. Let $\mathscr{J} = \bigcup_{y \in A} \mathscr{J}_y$. Then $\mathscr{J}$ is a Vitali covering of A, so by Lemma 10, we can choose a finite disjoint collection $\{J_1, \ldots, J_M\}$ of intervals in $\mathscr{J}$ where $J_i = [y_i, y_i + k_i]$ such that

$$m(A \setminus \bigcup_{i=1}^{M} J_i) < \varepsilon.$$

Then we also have

$$m\left(\bigcup_{i=1}^{M} J_i\right) = \sum_{i=1}^{M} k_i > m(A) - \varepsilon = s - 2\varepsilon. \tag{5}$$

Recall that by construction of the ks, we have $\frac{f(y_i+k_i)-f(y_i)}{k_i} > u$ and thus

$$\sum_{i=1}^{M} f(y_i + k_i) - f(y_i) > u \sum_{i=1}^{M} k_i. \tag{6}$$

Then (5) and (6) give

$$\sum_{i=1}^{M} f(y_i + k_i) - f(y_i) > (s - 2\varepsilon)u. \tag{7}$$

By construction, each interval J_i is contained within some I_n. Then for each i, let n be such that $J_i \subset I_n$. Since f is increasing, the following inequality holds:

$$f(x_n - h_n) \leq f(y_i) \leq f(y_i + k_i) \leq f(x_n).$$

We can use this to get

$$f(x_n) - f(x_n - h_n) \geq f(y_i + k_i) - f(y_i).$$

Note that each J_i is contained inside exactly one I_n. Now fix an n and consider all i such that $J_i \in I_n$. Since f is increasing and since the J_is are disjoint, we have

$$f(x_n) - f(x_n - h_n) \geq \sum_{J_i \subset I_n} f(y_i + k_i) - f(y_i).$$

Thus

$$\sum_{n=1}^{N} f(x_n) - f(x_n - h_n) \geq \sum_{i=1}^{M} f(y_i + k_i) - f(y_i).$$

Applying (4) and (7) gives

$$v(s + \varepsilon) > u(s - 2\varepsilon).$$

Since this is true for each $\varepsilon > 0$, we must have $vs \geq us$. But since $u > v$, then $vs \geq us$ can only be true if $s = 0$.

This shows that the function

$$g(x) = \lim_{h \to 0} \frac{f(x+h) - f(x)}{h}$$

is defined almost everywhere and that f is differentiable when g is finite. We now show that g must be finite almost everywhere.

For each $n \in \mathbb{N}$ and each $x \in [a, b]$, we define

$$g_n(x) = n \left[f\left(x + \frac{1}{n} \right) - f(x) \right]$$

and for each $x > b$, $g(x) = f(b)$.

Then, $g_n(x) \to g(x)$ as $n \to \infty$ for almost all x. So g is measurable. Since f is increasing, $g_n(x) \geq 0$ wherever g_n exists. Then, by Theorem 9,

$$\begin{aligned}
\int_{[a,b]} g \leq \liminf \int_{[a,b]} g_n &= \liminf \left[n \int_a^b \left(f\left(x + \frac{1}{n} \right) - f(x) \right) dx \right] \\
&= \liminf \left[n \int_{\left[b, b+\frac{1}{n}\right]} f - n \int_{\left[a, a+\frac{1}{n}\right]} f \right] \\
&= \liminf \left[f(b) - n \int_{\left[a, a+\frac{1}{n}\right]} f \right] \\
&\leq f(b) - f(a).
\end{aligned}$$

So g is integrable and thus finite almost everywhere. So f is differentiable almost everywhere.

$\square$

2.2 The Cantor Function

One of the first facts we learn about the derivative of a real valued function f is that for any point x in the domain of f, if $f'(x)$ exists, then the value of $f'(x)$ is the slope of the line that is tangent to the graph of f at the point x. Thus it seems natural to assume that if a function $f : [a, b] \to \mathbb{R}$ is increasing and $f(a) < f(b)$, then $f'(x)$ must be positive somewhere between a and b. However, this is not true. The Cantor Function, whose graph is shown in Figure 2, is a counterexample to this intuition.

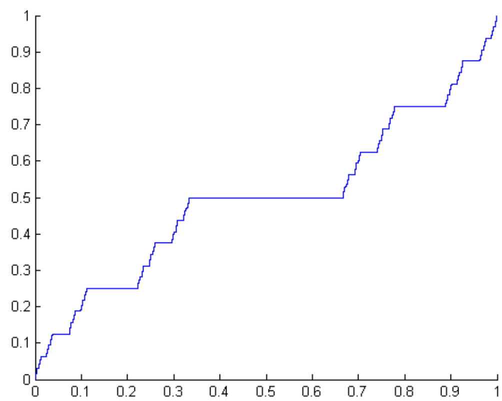


Figure 2: The Cantor Function

This function is interesting because it is continuous and its derivative is zero almost everywhere, yet its graph climbs from 0 to 1. It satisfies the following definition:

Definition 12. *A continuous, increasing function $f : [a, b] \to \mathbb{R}$ that satisfies $f(a) < f(b)$ and $f'(x) = 0$ for almost all $x \in [0, 1]$ is said to be* ***singular****.*

From looking at the graph of The Cantor Function, it isn't that surprising that it is singular; the graph is horizontal in almost all places. This function is almost always constant. One might wonder, are there any singular functions that are *never* constant? Equivalently, are there any strictly increasing singular functions? The answer to this question is yes. We will see two such examples.

3 DeRham's Function

Our first example of a strictly increasing singular function is DeRham's function. There are multiple ways that this function is usually defined. We discuss two of them and then

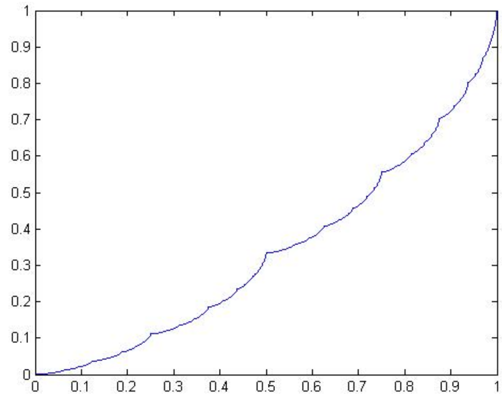


Figure 3: Graph of $\hat{f}_{\frac{1}{3}}(x)$

provide our own definition. We then use our new definition to provide alternative proofs of some known results.

3.1 A Natural Probability Question

Say we have a fair coin that we flip an infinite number of times. We will write down a 0 each time the coin lands on tails and a 1 each time it lands on heads. After generating this infinite string of zeros and ones, we can append a decimal point to the front and interpret the result as a number between 0 and 1 expressed in binary. Call this number R. What does the cumulative distribution function for R look like? We will denote this $f(x)$. So $f(x)$ is the probability that $R < x$. If $R < \frac{1}{2^k}$ for some positive integer k, then our first k flips must be tails which occurs with probability $\frac{1}{2^k}$. Thus, $f(\frac{1}{2^k}) = \frac{1}{2^k}$. In fact, we can show that for any $x \in [0, 1]$, we have $f(x) = x$. So in this case, the cumulative distribution function of R is not particularly interesting. However, if we assume that our coin is unfair, landing on tails with probability $a \neq \frac{1}{2}$, and heads with probability $(1-a)$, then the cumulative distribution function of R, which we will denote $\hat{f}_a(x)$, takes on some more interesting characteristics. Figures 3 and 4 show the graphs of $\hat{f}_{\frac{1}{3}}(x)$ and $\hat{f}_{\frac{2}{3}}(x)$. Note that these functions are not inverses of one another, although they may appear to be at first glance.

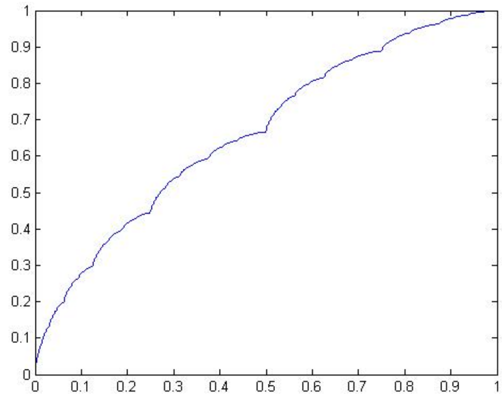


Figure 4: Graph of $\hat{f}_{\frac{2}{3}}(x)$

3.2 Fractal Properties

Note that each part of the curve on either side of the line $x = \frac{1}{2}$ is a shrunken version of the whole curve. More specifically, if we shrink the whole curve to half its size in the x direction and a times its size in the y direction, we get the left half of the curve and if we shrink the whole curve to half its size in the x direction and $1-a$ times its size in the y direction, we get the right half of the curve. We state this property more rigorously with the following theorem.

Theorem 13. *DeRham's function $\hat{f}_a(x)$ satisfies the following:*

$$\hat{f}_a(x) = \begin{cases} a\hat{f}_a(2x) & \text{if } 0 \leq x \leq \frac{1}{2} \\ (1-a)\hat{f}_a(2x-1)+a & \text{if } \frac{1}{2} < x \leq 1 \end{cases}.$$

Proof. Let $x = 0.b_1b_2b_3\ldots$ be a binary representation of x that does not terminate. Recall that $\hat{f}_a(x) = P(R < x)$ and let $R = 0.r_1r_2r_3\ldots$ be the binary representation of R.

Now assume $0 \leq x \leq \frac{1}{2}$. Then $b_1 = 0$ (assuming that if $x = \frac{1}{2}$, then it is expressed as $0.0111\ldots$). Then, $R < x$ if and only if $r_1 = 0$ and $0.r_2r_3r_4\ldots < 0.b_2b_3b_4\ldots$. We have $r_1 = 0$ with probability a. Since $x_1 = 0$, $2x = 0.b_2b_3b_4\ldots$. Thus $0.r_2r_3r_4\ldots < 0.b_2b_3b_4\ldots$

with probability $\hat{f}_a(2x)$. So if $0 \leq x \leq \frac{1}{2}$, then $\hat{f}_a(x) = a\hat{f}_a(2x)$.

Now assume $\frac{1}{2} < x \leq 1$. Then $b_1 = 1$. So if $r_1 = 1$, then we must have $0.r_2r_3r_4\ldots < 0.b_2b_3b_4\ldots$. Also, $2x - 1 = 0.b_2b_3b_4\ldots$. So $0.r_2r_3r_4\ldots < 0.b_2b_3b_4\ldots$ is true with probability $\hat{f}_a(2x-1)$. If $r_1 = 0$, then the values of $r_2, r_3, r_4, \ldots$ are unrestricted. Since $r_1 = 1$ with probability $1-a$ and $r_1 = 0$ with probability a, if $\frac{1}{2} < x \leq 1$, then $\hat{f}_a(x) = (1-a)\hat{f}_a(2x-1) + a$. □

Not only does DeRham's function satisfy the formula given in Theorem 13, it is the unique continuous solution to that functional equation. See [3] for more details.

DeRham uses this definition to prove that $\hat{f}_a$ is singular for all $a \neq \frac{1}{2}$. In [6], Kawamura provides an alternate definition and uses it to explicitly identify sets of real numbers on which $\hat{f}_a'(x) = 0$ and the sets on which $\hat{f}_a'(x)$ does not exist.

3.3 Algorithmic Definition

We now define a function $f_a(x)$ with an algorithm. We will later show that $f_a(x) = \hat{f}_a(x)$ for all x. Then we use this new definition to get alternative proofs for the aforementioned results of DeRham and Kawamura.

Algorithm 1 Given input $x \in [0,1]$, generate $f_a(x)$

```
Suppose x has binary representation 0.d_1 d_2 d_3 ...
L_1 ← 0
U_1 ← 1
M_0 ← 1
for n = 1, 2, 3, ... do
  M_n ← (1 − a)L_n + aU_n
  if d_n = 1 then
    L_{n+1} ← M_n
    U_{n+1} ← U_n
  else
    U_{n+1} ← M_n
    L_{n+1} ← L_n
  end if
end for
All three of the sequences defined converge to a common value. This value is f_a(x).
```

We can think of this algorithm as defining a sequence of closed intervals $\{[L_n, U_n]\}_{n=1}^{\infty}$. Note that if $d_k = 0$, then the length of $[L_k, U_k]$ is a times the size of $[L_{k-1}, U_{k-1}]$, and if $d_k = 1$, then the length of $[L_k, U_k]$ is $1 - a$ times the size of $[L_{k-1}, U_{k-1}]$. From this, we can see that the lengths of the intervals tends to 0 as $n \to \infty$, and since the intervals are closed, then there is exactly one point contained in all of them. This justifies our claim at the end of the algorithm. It is also worth noting that after k iterations of the algorithm, the largest value that $f_a(x)$ could take is U_k, which occurs when all the digits after the kth are 1. Similarly, the smallest value that $f_a(x)$ could take is L_k, which occurs when all the digits after the kth are 0.

Each dyadic rational has two distinct binary representations. One such representation ends in an infinite string of zeros, while the other ends in an infinite string of ones. For any dyadic x, we can assume either binary representation and our algorithm will generate the same value of $f_a(x)$. In the case of infinite zeros, L_k will eventually be fixed at $f_a(x)$, while U_k converges down to $f_a(x)$. In the case of infinite ones, U_k will become fixed at $f_a(x)$, while L_k converges up to $f_a(x)$. In the next section, we prove that $f_a(x) = \hat{f}_a(x)$ - that our algorithmic definition indeed defines DeRham's function. Before we get to that, we prove some other results about $f_a(x)$.

3.4 Analytic Properties

Now we will use our algorithmic definition of $f_a(x)$ to show that it is strictly increasing and singular. First we show that it is strictly increasing.

Theorem 14. *For $a \in (0, 1)$, the function $f_a(x)$ is strictly increasing.*

Proof. Let $x, y \in [0, 1]$ such that $x < y$. Consider the binary expansions of x and y and if either is dyadic, assume it is expressed in the terminating way. Since $x < y$, then there exists some k such that x has a 0 in the k^{th} digit and y has a 1. Let i be the smallest such k. Then, if L_{ix}, U_{ix} and L_{iy}, U_{iy} are the i^{th} endpoints for x and y

respectively, then $U_{ix} = L_{iy}$. We already know that $f_a(y) \geq L_{iy}$. Also, $f_a(x) \leq U_{ix}$. However, if $f_a(x) = U_{ix}$, then that would require x to have only ones from the i^{th} point on. This would make x a dyadic rational, but expressed in the non-terminating way, and we assumed that x would be terminating if dyadic. So $f_a(x) < U_{ix} = L_{iy} \leq f_a(y)$. □

Now we show that $f_a(x)$ is singular. We start by showing continuity.

Theorem 15. *For $a \in (0, 1)$, the function $f_a(x)$ is continuous.*

Proof. We claim that $f_a(x)$ is surjective. Let $y \in [0, 1]$. We now generate an infinite string $b_1b_2b_3\ldots$ of zeros and ones. We now define sequences $\{w_n\}_{n=1}^{\infty}$, $\{u_n\}_{n=1}^{\infty}$ and $\{m_n\}_{n=1}^{\infty}$ as follows. Let $w_1 = 0$ and $u_1 = 1$. For each $n \in \mathbb{N}$, let $m_n = aw_n + (1-a)u_n$. If $m_n \geq y$, then let $u_{n+1} = w_n$, let $w_{n+1} = w_n$ and let $b_n = 1$ Otherwise, if $m_n < y$, then let $u_{n+1} = u_n$, let $w_{n+1} = m_n$ and let $b_n = 0$ Now we can interpret $0.b_1b_2b_3\ldots$ as a number in $[0, 1]$ expressed in binary. Let x take this value. Then $f(x) = y$. Since $f_a(x)$ is surjective and strictly increasing, Theorem 8 implies that $f_a(x)$ is continuous. □

Before we prove that $f_a'(x) = 0$ when it exists, we need to show that for any function f, we can approximate $f'(x)$ with secant lines. We do this with two lemmas.

Lemma 16. *Let $\{x_n\}_{n=1}^{\infty}$ and $\{y_n\}_{n=1}^{\infty}$ and x be such that $x_n \to x$ and $y_n \to x$ as $n \to \infty$. Let $\{r_n\}_{n=1}^{\infty}$ be a sequence such that $0 \leq r_n \leq 1$ for all n. Then $r_nx_n + (1-r_n)x_n \to x$ as $n \to \infty$.*

Proof. Let $\varepsilon > 0$. Let N be such that for all $n \geq N$, $|x_n - x| < \varepsilon$ and $|y_n - x| < \varepsilon$.

Then,

$$\begin{aligned}
|r_n x_n + (1-r_n)y_n - x| &= \\
|r_n(x_n - x) + (1-r_n)(y_n - x)| &\leq \\
r_n|x_n - x| + (1-r_n)|y_n - x| &< \\
r_n\varepsilon + (1-r_n)\varepsilon &= \\
\varepsilon. &
\end{aligned}$$

□

Lemma 17. *Let f be a function that is differentiable at x. Let $\{a_n\}_{n=1}^{\infty}$ and $\{b_n\}_{n=1}^{\infty}$ be such that $a_n \to x$ and $b_n \to x$ as $n \to \infty$ where $a_n \leq x \leq b_n$ with $a_n \neq x$ or $b_n \neq x$ for all n. Then*

$$\frac{f(b_n) - f(a_n)}{b_n - a_n} \to f'(x)$$

as $n \to \infty$.

Proof. We start with the following algebraic manipulation:

$$\begin{aligned}
\frac{f(b_n) - f(a_n)}{b_n - a_n} &= \\
\frac{f(b_n) - f(x) + f(x) - f(a_n)}{b_n - a_n} &= \\
\frac{b_n - x}{b_n - a_n}\left(\frac{f(b_n) - f(x)}{b_n - x}\right) + \frac{x - a_n}{b_n - a_n}\left(\frac{f(x) - f(a_n)}{x - a_n}\right). &
\end{aligned}$$

Then note that $\frac{f(b_n)-f(x)}{b_n-x} \to f'(x)$ and $\frac{f(x)-f(a_n)}{x-a_n} \to f'(x)$ as $n \to \infty$. Also note that since $a_n \leq x \leq b_n$ for all n, $0 \leq \frac{b_n-x}{b_n-a_n} \leq 1$ for all n. Since $\frac{x-a_n}{b_n-a_n} = 1 - \frac{b_n-x}{b_n-a_n}$, we can apply Lemma 16 to get

$$\frac{b_n - x}{b_n - a_n}\left(\frac{f(b_n) - f(x)}{b_n - x}\right) + \frac{x - a_n}{b_n - a_n}\left(\frac{f(x) - f(a_n)}{x - a_n}\right) \to f'(x)$$

as $n \to \infty$. □

Theorem 18. *For $a \neq \frac{1}{2}$ and $0 \leq x \leq 1$, we have $f_a'(x) = 0$ if it exists.*

Proof. Let $x \in [0, 1]$ and let $0.b_1b_2b_3\ldots$ be the binary expansion of x. Define

$$x_n^- = 0.b_1b_2\ldots b_n000\ldots\,,$$
$$x_n^+ = 0.b_1b_2\ldots b_n111\ldots\,.$$

Now, note that for any n we have the following:

$$L_n = f_a(x_n^-) \tag{8}$$
$$U_n = f_a(x_n^+) \tag{9}$$
$$x_n^- \leq x \leq x_n^+ \tag{10}$$
$$x_n^+ - x_n^- = \left(\frac{1}{2}\right)^n. \tag{11}$$

We can use the above to approximate the derivative of f_a at x with a secant line. We define

$$d_n(x) = \frac{f_a(x_n^+) - f_a(x_n^-)}{x_n^+ - x_n^-}.$$

The value $d_n(x)$ is the slope of a secant line through the graph of f_a near x. So by Lemma 17, if $f_a'(x)$ exists, then $d_n(x) \to f_a'(x)$ as $n \to \infty$. Applying (8),(9) and (11) to our definition of $d_n(x)$ gives

$$d_n(x) = \frac{U_n - L_n}{\left(\frac{1}{2}\right)^n}.$$

Now, for each n, define

$$k_n(x) = k_n = \text{ is the number of 1s in the first } n \text{ binary digits of } x.$$

As we noted earlier, if $d_k = 0$, then the length of $[L_k, U_k]$ is a times the size of $[L_{k-1}, U_{k-1}]$ and if $d_k = 1$, then the length of $[L_k, U_k]$ is $1-a$ times the size of $[L_{k-1}, U_{k-1}]$.

Therefore, since the length of $[L_1, U_1] = [0, 1]$ is 1, the length of $[L_n, U_n]$ must be $(1 - a)^{k_n} a^{n-k_n}$. So

$$f_a(x_n^+) - f_a(x_n^-) = U_n - L_n = (1-a)^{k_n} a^{n-k_n}, \tag{12}$$

which gives

$$d_n(x) = \frac{(1-a)^{k_n} a^{n-k_n}}{\left(\frac{1}{2}\right)^n} = \frac{(1-a)^{k_n} a^{n-k_n}}{\left(\frac{1}{2}\right)^{k_n} \left(\frac{1}{2}\right)^{n-k_n}} = (2-2a)^{k_n} (2a)^{n-k_n}. \tag{13}$$

Now that we have a concrete expression for the value of each $d_n(x)$, we can start to figure out the limit of the sequence $\{d_n(x)\}_{n=1}^{\infty}$. Recall that this is of interest because if $f_a'(x)$ exists, then its value must equal the limit of the sequence in (13).

Now assume that $d_n(x) \to c$ as $n \to \infty$ for some real constant c. If $c \neq 0$, then we must have $\frac{d_{n+1}(x)}{d_n(x)} \to \frac{c}{c} = 1$. We will now show that the sequence $\{\frac{d_{n+1}(x)}{d_n(x)}\}_{n=1}^{\infty}$ cannot converge to 1. From this, it will follow that the sequence $\{d_n(x)\}_{n=1}^{\infty}$ either does not converge, or that it converges to 0. Thus the derivative of $f_a(x)$ must equal 0 wherever it exists. So now we will show that $\frac{d_{n+1}(x)}{d_n(x)}$ does not tend to 1. Note that

$$\frac{d_{n+1}(x)}{d_n(x)} = \frac{(2-2a)^{k_{n+1}} (2a)^{n+1-k_{n+1}}}{(2-2a)^{k_n} (2a)^{n-k_n}} = (2-2a)^{k_{n+1}-k_n} (2a)^{1+k_n-k_{n+1}}.$$

Since k_i denotes the number of 1s in the first i digits of x, we must have either $k_{n+1} = k_n$ or $k_{n+1} = k_n + 1$. The first case gives $\frac{d_{n+1}(x)}{d_n(x)} = 2a$ and the second case gives $\frac{d_{n+1}(x)}{d_n(x)} = 2(1-a)$. So each term in our sequence is either $2a$ or $2(a-1)$. So if our sequence is to converge, it must converge to $2a$ or to $2(1-a)$. Neither one of these values is 1 since $a \neq \frac{1}{2}$. So if the derivative of f_a exists at a point x, it must be 0. $\square$

Corollary 19. *For $a \neq \frac{1}{2}$, $f_a'(x) = 0$ almost everywhere.*

Proof. We get this from the previous theorem and Lebesgue's theorem (Theorems 18 and 11). $\square$

Now we show that our definition of DeRham's function is equivalent to the probability definition.

Theorem 20. *For all x, $f_a(x) = P(R < x)$.*

Proof. First, we show by induction that for all x and for all n, $f_a(x_n^+) = P(R < x_n^+)$ and $f_a(x_n^-) = P(R < x_n^+)$. For the base case, note that for all x, we have $x_0^+ = 1$ and $x_0^- = 0$ and $P(R < 1) = 1 = f_a(1)$ and $P(R < 0) = 0 = f_a(0)$. For the inductive step, assume that $f_a(x_n^+) = P(R < x_n^+)$ and $f_a(x_n^-) = P(R < x_n^-)$. If $b_{n+1} = 0$ (the $(n+1)^{\text{th}}$ digit of x), then $x_n^- = x_{n+1}^-$ so we're done by the induction hypothesis. If $b_{n+1} = 1$, then

$$f_a(x_{n+1}^-) = (1-a)L_n + aU_n = L_n + a(U_n - L_n) =$$
$$f_a(x_n^-) + a(f_a(x_n^+) - f_a(x_n^-)) = P(R < x_n^-) + aP(x_n^- \leq R < x_n^+).$$

This represents the probability that $R < x_n^-$ or $x_n^- < R < x_n^+$ and the $(n+1)^{\text{th}}$ digit of R is 0. Then we are done because $R < x_{n+1}^-$ if and only if these conditions are met. Showing the inductive step for $f(x_{n+1}^+)$ is similar.

If x is a dyadic rational, then $x = x_n^-$ for some n so we are done. Otherwise, if x is not a dyadic rational, then for all n, we have $x_n^- < x < x_n^+$. Since $f_a(x)$ is increasing, it follows that $f_a(x_n^-) < f_a(x) < f_a(x_n^+)$. Applying the above gives $P(R < x_n^-) < f_a(x) < P(R < x_n^+)$. Since $x_n^- < x < x_n^+$, we also have $P(R < x_n^-) < P(R < x) < P(R < x_n^+)$. As stated earlier, the sequences defined by $P(R < x_n^-)$ and $P(R < x_n^+)$ both converge to the same real number. Since the above inequalities are true for all n, this means that $P(R < x) = f_a(x)$. □

It is natural to wonder where $f_a'(x)$ exists and where it does not exist. We start off by showing that this function is not differentiable at the dyadic rationals.

Theorem 21. *Let $x \in [0,1]$ be a dyadic rational. Then $f_a'(x)$ does not exist.*

Proof. Since x is a dyadic rational, it has a binary representation that ends with an infinite string of 0s. Let $0.b_1b_2\ldots b_i000\ldots$ be this representation. Recall the definition of $d_n(x)$ from the proof of Theorem 18 and recall that if $f_a'(x)$ exists, then $d_n(x) \to f_a'(x)$. Also

$$d_n(x) = (2-2a)^{k_n}(2a)^{n-k_n}.$$

Now, assume that $n \geq i$. Recall that we are working with the binary representation of x such that each binary digit of x beyond the i^{th} digit is a 0. This means that $k_n = k_i$ which gives

$$d_n(x) = (2-2a)^{k_n}(2a)^{n-k_n} = (2-2a)^{k_i}(2a)^{n-k_i} =$$
$$(2-2a)^{k_i}(2a)^{i-k_i}(2a)^{n-i} = d_i(x)(2a)^{n-i}.$$

So if $a > \frac{1}{2}$, we have $d_n(x) \to \infty$ as $n \to \infty$, implying that the derivative doesn't exist. Note that if $a < \frac{1}{2}$ then $d_n(x) \to 0$ as $n \to \infty$. However, we can define a sequence analogous to $d_n(x)$ that is based on the other binary representation of x (the representation that ends in an infinite string of 1s) that will tend to infinity when $a < \frac{1}{2}$.

Let $0.e_1e_2\ldots e_j111\ldots$ be the binary representation of x such that each binary digit of x beyond the j^{th} digit is a 1. We now define $d_n(x)$ exactly as before. Assume $n \geq j$. This means that $j - k_j = n - k_n$ which gives

$$d_n(x) = (2-2a)^{k_n}(2a)^{n-k_n} =$$
$$(2-2a)^{k_n-k_j}(2-2a)^{k_j}(2a)^{j-k_j} = (2-2a)^{n-j}d_j(x).$$

So if $2-2a > 1$, or equivalently, if $a < \frac{1}{2}$, then $d_n(x) \to \infty$ as $n \to \infty$. It is also interesting to note that for this way of defining $d_n(x)$, if $a > \frac{1}{2}$ then $d_n(x) \to 0$ as $n \to \infty$. $\square$

3.5 Points at Which $f_a'(x)$ Exists

We will now identify a subset of $[0,1]$ where f_a is differentiable. This will have to do with what we will call the *density* of zeros and ones in x [6]. Heuristically, the density of zeros in x is the probability that a digit chosen at random from the binary representation of x is a zero. The density of ones in x is similarly defined. Since we already know that f_a is not differentiable at any dyadic rational, we will not bother with the dyadic rationals. This means that in defining the density of zeros and ones in x, we will not need to worry about non-unique binary representations. We introduce the following notation

Definition 22. *Given an $x \in [0,1]$ such that x is not a dyadic rational, we define p_n to be the position of the n^{th} zero in the binary representation of x. In addition, we define the two functions*

$$D_0(x) = \lim_{n\to\infty} \frac{n}{p_n},$$

$$D_1(x) = 1 - D_0(x).$$

There do exist numbers for which the densities of 0s and 1s do not exist.

Theorem 23. *Let $a \in \left(0, \frac{1}{2}\right) \cup \left(\frac{1}{2}, 1\right)$. Let $x \in [0,1]$ such that x is not a dyadic rational and $D_0(x)$ and $D_1(x)$ both exist with $0 < D_0(x), D_1(x) < 1$ and $a^{D_0(x)}(1-a)^{D_1(x)} < \frac{1}{2}$. Then $f_a'(x) = 0$.*

Before we prove this, we will explore what it means for some $x \in [0,1]$ to satisfy $a^{D_0(x)}(1-a)^{D_1(x)} < \frac{1}{2}$. Almost all numbers have equal density of zeros and ones in their binary representation (see [7]), so $D_0(x) = D_1(x) = \frac{1}{2}$ is true for almost all $x \in [0,1]$. Note that if $D_0(x) = D_1(x) = \frac{1}{2}$, then $a^{D_0(x)}(1-a)^{D_1(x)} = \sqrt{a}\sqrt{1-a}$ which is always less than $\frac{1}{2}$. To see this, note that each of $\sqrt{a}$ and $\sqrt{1-a}$ is less than 1 and at least one must be less than $\frac{1}{2}$. So, from this it follows that almost all $x \in [0,1]$ satisfy $a^{D_0(x)}(1-a)^{D_1(x)} < \frac{1}{2}$. So by Theorem 23, $f_a'(x)$ exists almost everywhere. Note that

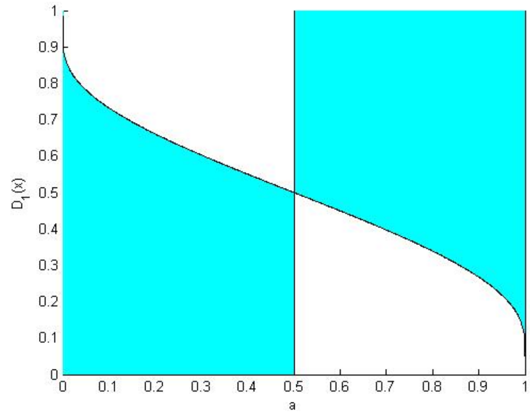


Figure 5: Values of a and $D_1(x)$ that satisfy the hypothesis of Theorem 23

we already proved this result a different way (see Corollary 19).

Now, we look at how a given value of a affects the density of 0s and 1s of xs where the derivative of $f_a(x)$ exists. By replacing $D_0(x)$ with $1 - D_1(x)$ in $a^{D_0(x)}(1-a)^{D_1(x)} < \frac{1}{2}$, we get $a^{1-D_1(x)}(1-a)^{D_1(x)} < \frac{1}{2}$. Using exponent rules and taking the natural logarithm of both sides gives

$$D_1(x) \ln \frac{1-a}{a} < -\ln 2a.$$

Now, if $a < \frac{1}{2}$, then $\ln \frac{1-a}{a} > 0$ which gives

$$D_1(x) < \frac{-\ln 2a}{\ln \frac{1-a}{a}}.$$

Otherwise, if $a > \frac{1}{2}$, then $\ln \frac{1-a}{a} < 0$ which gives

$$D_1(x) > \frac{-\ln 2a}{\ln \frac{1-a}{a}}.$$

Figure 5 shows the graph of this function of a, bounding densities D_1 where the hypothesis of Theorem 23 holds. Points in the shaded area represent combinations of a and $D_1(x)$ such that $f'_a(x)$ exists and equals zero. No points along the borders are shaded. All points with $a \neq \frac{1}{2}$ and $D_1(x) = \frac{1}{2}$ are shaded, as expected by the earlier discussion.

We could also show that if $a^{D_0(x)}(1-a)^{D_1(x)} > \frac{1}{2}$, then $f_a'(x)$ does not exist since the slope approaches ∞. This and some borderline cases where $a^{D_0(x)}(1-a)^{D_1(x)} = \frac{1}{2}$ are treated in [6]. We now present an original proof of Theorem 23.

Proof. We only show that $\lim_{h\to 0^+} \frac{f_a(x+h)-f_a(x)}{h} = 0$. The proof for the limit coming from the negative side is similar. Let $\varepsilon > 0$. We will now show that we can choose an h such that $\frac{f_a(x+h)-f_a(x)}{h} < \varepsilon$. Without loss of generality, assume that $\varepsilon < \frac{1}{2a^{D_0(x)}(1-a)^{D_1(x)}} - 1$. Now, note that the sequence $\frac{p_{n+1}}{p_n}$ converges to 1. This is true because $\frac{p_{n+1}}{p_n} = \frac{p_{n+1}}{n+1} \cdot \frac{n+1}{p_n} = \frac{p_{n+1}}{n+1} \cdot \left(\frac{n}{p_n} + \frac{1}{p_n}\right)$ which converges to $\frac{1}{D_0(x)} \cdot D_0(x) = 1$.

Note that for any n, $p_{n+1} > p_n$ which means that $\frac{p_{n+1}}{p_n} > 1$. This means that we can pick a positive integer N_1 such that for all $n \geq N_1$, we have $\frac{p_{n+1}}{p_n} < 1 + \log_2\left(\sqrt[3]{\varepsilon+1}\right)$. Note that the sequence $\{1 - \frac{n}{p_n}\}_{n=1}^{\infty}$ converges to D_1 which means that we can pick a positive integer N_2 such that for all $n \geq N_2$, we have $1 - \frac{n}{p_n} > D_1(x) + \log_{1-a}\left(\sqrt[3]{\varepsilon+1}\right)$ (since $0 < 1-a < 1$, the value of $\log_{1-a}\left(\sqrt[3]{\varepsilon+1}\right)$ is negative). Note that the sequence $\{\frac{n-1}{p_n}\}_{n=1}^{\infty}$ converges to D_0 which means that we can pick a positive integer N_3 such that for all $n \geq N_3$, $\frac{n-1}{p_n} > D_0(x) + \log_a\left(\sqrt[3]{\varepsilon+1}\right)$ (as before, since $0 < a < 1$, the value of $\log_a\left(\sqrt[3]{\varepsilon+1}\right)$ is negative). Since x is not a dyadic rational, the binary representation of x must have an infinite number of zeros. Thus, p_n is unbounded. Also, since $\varepsilon < \frac{1}{2a^{D_0(x)}(1-a)^{D_1(x)}} - 1$, we have $2a^{D_0(x)}(1-a)^{D_1(x)}(1+\varepsilon) < 1$. This means that we can choose a positive integer N_4 such that $\left(2a^{D_0(x)}(1-a)^{D_1(x)}(1+\varepsilon)\right)^{p_n} < \varepsilon$ for all $n \geq N_4$.

Now, let $N = \max\{N_1, N_2, N_3, N_4\}$ and pick an h satisfying

$$0 < h < 2^{-p_N}. \tag{14}$$

Fix an n such that $x_{p_n}^+ < x + h \leq x_{p_n-1}^+$. Then, we have

$$0 < x_{p_n}^+ - x < h. \tag{15}$$

Note that x and $x_{p_n}^+$ agree out to $p_{n+1} - 1$ decimal places, and that when they differ at the p_{n+1}^{th} place, x has digit 0 and $x_{p_n}^+$ has digit 1. So $x_{p_n}^+ - x > 2^{-(p_{n+1})}$. Then, by (15), $h > 2^{-(p_{n+1})}$. Putting this together with (14) gives $p_{n+1} > p_N$. This inequality says that the $(n+1)^{\text{th}}$ 0 lies further out then the N^{th} 0 which implies $n \geq N$. Since f_a is strictly increasing, (10) gives $f_a(x+h) - f_a(x) < f_a(x_{p_n-1}^+) - f_a(x_{p_n-1}^-)$. Using this and applying (12), we have

$$\begin{aligned}\frac{f_a(x+h) - f_a(x)}{h} < \frac{f_a(x_{p_n-1}^+) - f_a(x_{p_n-1}^-)}{2^{-p_{n+1}}} = \frac{(1-a)^{p_n-n}a^{n-1}}{2^{-p_{n+1}}} &= \\ \left(\left(\frac{(1-a)^{p_n-n}a^{n-1}}{2^{-p_{n+1}}}\right)^{\frac{1}{p_n}}\right)^{p_n} &= \\ \left(2^{\frac{p_{n+1}}{p_n}}\,(1-a)^{1-\frac{n}{p_n}}\,a^{\frac{n-1}{p_n}}\right)^{p_n}&.\end{aligned}$$

Now, since $n \geq N$, we can apply the inequalities from the definitions of N_1, N_2, N_3 and N_4:

$$\begin{aligned}\left(2^{\frac{p_{n+1}}{p_n}}\,(1-a)^{1-\frac{n}{p_n}}\,a^{\frac{n-1}{p_n}}\right)^{p_n} &< \\ \left(2^{1+\log_2\left(\sqrt[3]{\varepsilon+1}\right)} \cdot (1-a)^{D_1(x)+\log_{1-a}\left(\sqrt[3]{\varepsilon+1}\right)} \cdot a^{D_0(x)+\log_a\left(\sqrt[3]{\varepsilon+1}\right)}\right)^{p_n} &= \\ \left(2(1-a)^{D_1(x)}a^{D_0(x)}(\varepsilon+1)\right)^{p_n} &< \varepsilon. \qquad \square\end{aligned}$$

4 Minkowski's Question-Mark Function

We now discuss Minkowski's Question-Mark function, denoted $?(x)$. Like DeRham's function, this is strictly increasing yet it has derivative zero almost everywhere. There are two common ways to define $?(x)$ (both produce the same result) and they are detailed in [2]. We propose an alternative way to define $?(x)$.

4.1 Algorithmic Definition of $?(x)$

For any rational number q, we define $\text{num}(q)$ and $\text{denom}(q)$ to be the numerator and denominator of q when it is expressed in lowest terms. We define $\text{denom}(0)$ to be 1.

Algorithm 2 Given input $x \in [0,1]$, generate the base 2 representation of $?(x)$

```
Binary Representation ← 0.
w ← 0
u ← 1
for n = 1, 2, 3, ... do
    m ← (num(w)+num(u)) / (denom(w)+denom(u))
    if x ≥ m then
        append a 1 to the right of Binary Representation
        w ← m
    else
        append a 0 to the right of Binary Representation
        u ← m
    end if
end for
return Binary Representation
```

We now show the steps this algorithm would perform in order to perform the computation of $?(\frac{3}{8})$.

Table 1: Computation of $?(\frac{2}{5}) = \frac{3}{8}$

$[w, u]$	m	test	ans
$[\frac{0}{1}, \frac{1}{1}]$	$\frac{1}{2}$	$x < \frac{1}{2}$	0.0
$[\frac{0}{1}, \frac{1}{2}]$	$\frac{1}{3}$	$x \geq \frac{1}{3}$	0.01
$[\frac{1}{3}, \frac{1}{2}]$	$\frac{2}{5}$	$x \geq \frac{2}{5}$	0.011
$[\frac{2}{5}, \frac{1}{2}]$	$\frac{3}{7}$	$x < \frac{3}{7}$	0.0110
$[\frac{2}{5}, \frac{3}{7}]$	$\frac{5}{12}$	$x < \frac{5}{12}$	0.01100
$[\frac{2}{5}, \frac{5}{12}]$	$\frac{7}{17}$	$x < \frac{7}{17}$	0.011000...

As you can see, the value in the rightmost column will continue to be $0.011000\ldots0_2 = \frac{3}{8}$ as $x < m$ will always be true after the fourth query, as x becomes the lower endpoint.

Figure 6 shows the graph of $?(x)$.

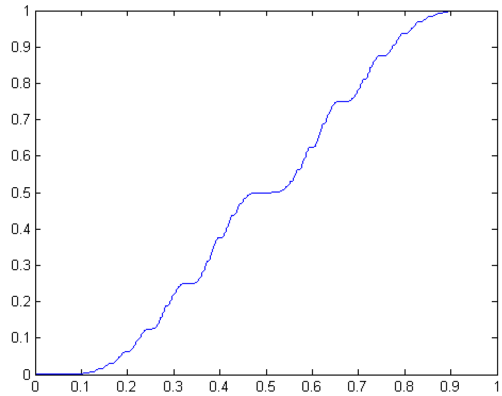


Figure 6: Graph of $?(x)$

As stated earlier, $?(x)$ is a strictly increasing singular function. While we will not prove this in this thesis, we will show how this algorithmic definition is equivalent to the well-established definition. In order to do this, we rephrase some of the details of the algorithm.

Algorithm 3 For some x, generate $?(x)$

ans $\leftarrow 0$
$w_1 \leftarrow 0$
$u_1 \leftarrow 1$
$m_0 \leftarrow 1$
for $n = 1, 2, 3, \ldots$ **do**
  $m_n \leftarrow \frac{\text{num}(w_n)+\text{num}(u_n)}{\text{denom}(w_n)+\text{denom}(u_n)}$
  if $x \geq m_n$ **then**
    ans $\leftarrow$ ans $+\frac{1}{2^n}$
    $w_{n+1} \leftarrow m_n$
    $u_{n+1} \leftarrow u_n$
  else
    $u_{n+1} \leftarrow m_n$
    $w_{n+1} \leftarrow w_n$
  end if
end for
return ans

4.2 Notes and Remarks

An m_n created in this algorithm will be referred to as a *mediant* and any w_n, u_n will be called *endpoints*. The subscripts have no algorithmic importance, but they enable us to analyze the algorithm more easily. Note that in all cases, $m_0 = 1$. In Corollary 29 which follows, we show that all fractions that define m_n (and hence w_n and u_n) are in reduced form so numerators and denominators simply add as shown and cancellation is not relevant.

Remark 24. *At every iteration of the loop, w_n, u_n and m_n are rational. Furthermore, the values of* $\mathrm{denom}(w_n)$, $\mathrm{denom}(u_n)$ *are monotone increasing and* $\mathrm{denom}(m_n)$ *is strictly increasing with each iteration.*

Remark 25. *For every $n \geq 1$,* $\mathrm{denom}(m_n) > \mathrm{denom}(w_n)$ *and* $\mathrm{denom}(m_n) > \mathrm{denom}(u_n)$ *and $w_n < m_n < u_n$.*

Remark 26. *The inequality $w_n \leq x \leq u_n$ is always true.*

4.3 Mediants, Endpoints and an Algebraic Property

We show that ? maps every rational number to a dyadic rational number. To do this, we show that when using the algorithm to calculate $?(x)$ for some rational x, then for some n, $m_n = x$. In order to get to this point, we need to prove a few lemmas. In this context, we will assume that a and a' are nonnegative integers and b and b' are positive integers.

Lemma 27. *If $\frac{a}{b} < \frac{a'}{b'}$ then $\frac{a}{b} < \frac{a+a'}{b+b'} < \frac{a'}{b'}$.*

Proof. Showing that $\frac{a}{b} < \frac{a+a'}{b+b'}$:

$$\frac{a}{b} = \frac{a\frac{b+b'}{b}}{b+b'} = \frac{a + \frac{a}{b}b'}{b+b'} < \frac{a + \frac{a'}{b'}b'}{b+b'} = \frac{a+a'}{b+b'}.$$

Showing that $\frac{a'}{b'} > \frac{a+a'}{b+b'}$:

$$\frac{a'}{b'} = \frac{a'\frac{b'+b}{b'}}{b+b'} = \frac{a' + \frac{a'}{b'}b}{b+b'} > \frac{a + \frac{a}{b}b}{b+b'} = \frac{a+a'}{b+b'}. \qquad \square$$

Lemma 28. *If $\frac{a}{b}$ and $\frac{a'}{b'}$ are the endpoints of an interval in an iteration of the algorithm to find $?(x)$, then $a'b - b'a = 1$.*

Proof. We induct on the iteration through the loop. For the base case we look at the first iteration through the loop. Here, our endpoints are $\frac{a}{b} = \frac{0}{1}$ and $\frac{a'}{b'} = \frac{1}{1}$. Then $a'b - b'a = 1 \cdot 1 - 1 \cdot 0 = 1$. For the inductive step, let $\frac{a}{b}, \frac{a'}{b'}$ be the endpoints at iteration $n-1$ through the loop and assume $a'b - b'a = 1$ holds. Then, at iteration n, the endpoints are either $\frac{a}{b}, \frac{a+a'}{b+b'}$ or $\frac{a+a'}{b+b'}, \frac{a'}{b'}$. In the first case, we have

$$(a'+a)b - a(b'+b) = a'b + ab - ab' - ab = a'b - b'a = 1$$

and in the second we find

$$a'(b+b') - b'(a+a') = a'b + a'b' - b'a - b'a' = a'b - b'a = 1. \qquad \square$$

Corollary 29. *As generated by the algorithm, each endpoint is in reduced form.*

Proof. Let $\frac{a}{b}$ and $\frac{a'}{b'}$ be endpoints. Let d, m, n be nonnegative integers such that $a = nd$ and $b = md$. So d is a divisor of a and b. By Lemma 28 we have

$$a'b - b'a = 1,$$
$$a'md - b'nd = 1,$$
$$a'm - b'n = \frac{1}{d}.$$

If $d > 1$, then $\frac{1}{d}$ would not be integer. Then, since a', m, b', n are all integers, $d = 1$ so the greatest common divisor of a and b is 1. The proof that $\frac{a'}{b'}$ is in reduced form is similar.

$\square$

Lemma 30. *If $\frac{a}{b}$ and $\frac{a'}{b'}$ are the endpoints of an interval in an iteration of the algorithm to calculate $?(x)$ for some x, then $\frac{a+a'}{b+b'}$ is the unique fraction between $\frac{a}{b}$ and $\frac{a'}{b'}$ with smallest denominator.*

Proof. From Lemma 27 we have $\frac{a}{b} < \frac{a+a'}{b+b'} < \frac{a'}{b'}$. Consider any fraction $\frac{x}{y}$ (x, y integers) such that $\frac{a}{b} < \frac{x}{y} < \frac{a'}{b'}$. Then $a'y - b'x > 0$, and since a', b', x, y are integers, $a'y - b'x \geq 1$. Then

$$\frac{a'}{b'} - \frac{a}{b} = \frac{a'}{b'} - \frac{x}{y} + \frac{x}{y} - \frac{a}{b} = \frac{a'y - b'x}{b'y} + \frac{bx - ay}{by} \geq \frac{1}{b'y} + \frac{1}{by} = \frac{b+b'}{bb'y}. \tag{16}$$

From this and from Lemma 28 it follows that $\frac{b+b'}{bb'y} \leq \frac{a'b-b'a}{bb'} = \frac{1}{bb'}$, and thus $y \geq b + b'$. If $y > b + b'$ then $\frac{x}{y}$ does not have the smallest denominator among fractions between $\frac{a}{b}$ and $\frac{a'}{b'}$, so $y = b + b'$. We now show that $\frac{a+a'}{b+b'}$ is the unique value for $\frac{x}{y}$. Then

$$\frac{a'}{b'} - \frac{a}{b} = \frac{a'b - b'a}{bb'} = \frac{1}{bb'} = \frac{y}{bb'y} = \frac{b+b'}{bb'y},$$

and thus the inequality in (16) becomes an equality. This gives us $a'y - b'x = 1$ and $bx - ay = 1$. Solving for x and y gives $x = a + a'$ and $y = b + b'$, and these solutions are unique. $\square$

Lemma 31. *Let $x \in [0, 1]$. If $\{w_n\}$ and $\{u_n\}$ are respectively the sequence of lower and upper endpoints generated in using the algorithm to calculate $?(x)$, then $u_n - w_n \to 0$ as $n \to \infty$.*

Proof. Letting $u_n = \frac{a'}{b'}$ and $w_n = \frac{a}{b}$, we have $u_n - w_n = \frac{a'b-b'a}{bb'}$. By Lemma 28, this equals $\frac{1}{bb'} = \frac{1}{\mathrm{denom}(w_n)\mathrm{denom}(u_n)}$ which converges to 0 by Remark 24. $\square$

Theorem 32. *Let $x \in [0, 1]$. If $\{m_n\}$ is the sequence of mediants generated when we use the algorithm to calculate $?(x)$, then $m_n \to x$ as $n \to \infty$.*

Proof. Let $\{u_n\}$ and $\{w_n\}$ be the sequences of lower and upper endpoints respectively. Let $\varepsilon > 0$. By Lemma 31 we can choose N such that for all $n \geq N$, $u_n - w_n < \varepsilon$. By Remark 26 and Lemma 27, we have $w_n \leq x, m_n \leq u_n$, so $|x - m_n| \leq u_n - w_n < \varepsilon$. $\square$

Theorem 33. *Let $\frac{x}{y} \in (0, 1)$ be a fraction in lowest terms. Then $\frac{x}{y}$ becomes a mediant at some point in the calculation of $?(\frac{x}{y})$.*

Proof. Let $m_1, m_2, \ldots$ be the set of mediants used to calculate $?(x)$. Let $m_1, m_2, \ldots, m_n$ be the finite subsequence of mediants with denominator strictly less than y. We know such a subsequence exists by Remark 24. Consider the iteration of the loop where m_n becomes one of the endpoints; without loss of generality, assume m_n is the greater endpoint and denote the smaller endpoint w. Then m_{n+1} is the mediant of w and m_n and $\operatorname{denom}(m_{n+1}) \geq y$. By Lemma 30, m_{n+1} is also the unique fraction between w and m_n with smallest denominator, and by Remark 26, $w \leq \frac{x}{y} \leq m_n$ but equality is impossible because all mediants up to m_n (including w) have smaller denominator than y. So $\frac{x}{y}$ must be the mediant of w and m_n. $\square$

Corollary 34. *If $x \in [0, 1]$ is rational then $?(x)$ is a dyadic rational.*

4.4 Analytic Properties

Theorem 35. *? is strictly increasing.*

Proof. Let $a, b \in [0, 1]$ such that $a < b$. I will show that this implies $?(a) <?(b)$. Remember that we calculate $?(a)$ through an iterative process where for each $n \in \mathbb{N}$ we have some a_n, and then we check if $a \geq a_n$, then generate a_{n+1} accordingly and repeat. We do the same process to calculate $?(b)$. This gives us two sequences $\{a_n\}, \{b_n\}$. Note that $a_n \to a$ and $b_n \to b$. This means that if $a_n = b_n$ for every n, then $a = b$ which contradicts our assumption that $a < b$. So we know that there exists some point at which the sequences $\{a_n\}$ and $\{b_n\}$ begin to differ. Let N be the smallest natural number such that $a_{N+1} \neq b_{N+1}$. So for all n such that $1 \leq n \leq N$, $a_n = b_n$. Since the values of a_{n+1}

and b_{n+1} are completely determined by whether $a \geq a_n$ and whether $b \geq b_n$ respectively, we can conclude that for any n such that $1 \leq n \leq N-1$, we have $a_n = b_n \geq a$ if and only if $a_n = b_n \geq b$. Since we generate the binary representations for $?(a)$ and $?(b)$ based only on whether or not $a_n \geq a$ and $b_n \geq b$ respectively, we know that the binary representations for $?(a)$ and $?(b)$ are the same for the first $N-1$ terms right of the decimal point. Since $a_{N+1} \neq b_{N+1}$ we know that exactly one of $a \geq a_N = b_N$ and $b \geq a_N = b_N$ is false. Since $b > a$, $a \geq a_N = b_N$ must be false and thus the N^{th} binary digit of $?(a)$ is 0 and the N^{th} binary digit of $?(b)$ is 1. Since all the preceding digits were the same, this implies $?(b) > ?(a)$. $\square$

In [10], Salem proves that $?'(x)$ is 0 almost everywhere. He defines $?(x)$ using the continued fraction representation of x. The remainder of this thesis will show the relationship between Salem's continued fraction approach and our algorithmic approach.

4.5 Salem's Definition

Salem provides another definition of ? which can be found in [10]. We will prove that our definition is equivalent to Salem's, but first we introduce continued fractions.

4.5.1 Continued Fractions

This section is adapted from [8]. Consider the fraction $\frac{9}{7}$. We can manipulate it in the following way:

$$\frac{9}{7} = 1 + \frac{2}{7} = 1 + \frac{1}{\frac{7}{2}} = 1 + \frac{1}{3 + \frac{1}{2}}.$$

This is the simple continued fraction representation of $\frac{9}{7}$ and we can denote it $[1; 3, 2]$. More generally, the notation $[a_0; a_1, a_2, \ldots, a_n]$ denotes the following continued fraction:

$$a_0 + \cfrac{1}{a_1 + \cfrac{1}{a_2 + \cfrac{1}{\ddots_{a_{n-1} + \frac{1}{a_n}}}}}.$$

We may also have continued fractions that do not terminate. The notation $[a_0; a_1, a_2, \ldots]$ denotes the following infinite continued fraction:

$$a_0 + \cfrac{1}{a_1 + \cfrac{1}{a_2 + \cfrac{1}{\ddots}}}.$$

Every terminating continued fraction represents a unique rational number. However, every rational number can be expressed as a continued fraction in one of exactly two distinct ways. We can see this by looking at the first example:

$$\frac{9}{7} = 1 + \frac{1}{3 + \frac{1}{2}} = 1 + \frac{1}{3 + \frac{1}{1+1}} = 1 + \frac{1}{3 + \frac{1}{1 + \frac{1}{1}}}$$

which means $\frac{9}{7} = [1; 3, 2] = [1; 3, 1, 1]$. More generally, if some rational number x has the continued fraction representation $[a_0; a_1, a_2, \ldots, a_n]$ where $a_n \neq 1$ then x can also be expressed as $[a_0; a_1, a_2, \ldots, a_n - 1, 1]$.

Throughout the rest of this paper we will assume that any continued fraction representation of a rational number x will not end with a 1 unless $x = 1$. We will always assume that a_0 will be an integer and that all of $a_1, a_2, \ldots, a_n, \ldots$ are positive unless we state otherwise.

Every infinite continued fraction represents a unique irrational number. Every irrational number can be expressed as a unique infinite continued fraction.

If $x = [a_0; a_1, a_2, \ldots, a_n]$ or $[a_0; a_1, a_2, \ldots]$, we let x_i denote $[a_0; a_1, a_2, \ldots, a_i]$. The x_i's are called the *convergents* of x. Note that if x is rational, then x equals its last convergent. As we will see in the next few lemmas, the convergents are useful because they serve as good approximations. Before exploring this we must develop the framework necessary to relate convergents to each other.

If we let $x = [a_0, a_1, \ldots]$ (we could also allow this to terminate), we define two sequences which will wind up being the numerators and denominators of the convergents of x:

$$h_{-2} = 0, h_{-1} = 1, h_i = a_i h_{i-1} + h_{i-2} \text{ for } i \geq 0,$$

$$d_{-2} = 1, d{-}1 = 0, d_i = a_i d_{i-1} + d_{i-2} \text{ for } i \geq 0.$$

Theorem 36. *Let $x = [a_0; a_1, a_2, \ldots]$ or $[a_0; a_1, a_2, \ldots, a_n]$. Then for any real number $y \neq 0$, $i \leq n$,*

$$[a_0, a_1, \ldots, a_{i-1}, y] = \frac{yh_{i-1} + h_{i-2}}{yd_{i-1} + d_{i-2}}.$$

Proof. We proceed by induction on i. If $i = 0$, then we have

$$y = \frac{1 \cdot y + 0}{0 \cdot y + 1} = \frac{yh_{-1} + h_{-2}}{yd_{-1} + d_{-2}}.$$

Assume that the result holds for $i - 1$. We then note the following consequence of our notation:

$$[a_0; a_1, a_2, \ldots, a_i, y] = [a_0; a_1, a_2, \ldots, a_{i-1}, a_i + \frac{1}{y}].$$

By the induction hypothesis, we then have

$$[a_0; a_1, a_2, \ldots, a_{i-1}, a_i + \frac{1}{y}] = \frac{(a_i + \frac{1}{y})h_{i-1} + h_{i-2}}{(a_i + \frac{1}{y})d_{i-1} + d_{i-2}}$$

$$= \frac{y(a_i h_{i-1} + h_{i-2}) + h_{i-1}}{y(a_i d_{i-1} + d_{i-2}) + d_{i-1}} = \frac{yh_i + h_{i-1}}{yd_i + d_{i-1}}$$

thus completing the induction. $\square$

Corollary 37. *Let $x = [a_0; a_1, a_2, \ldots]$ or $[a_0; a_1, a_2, \ldots, a_n]$. Then $x_i = \frac{h_i}{d_i}$.*

Remark 38. *Note that if $a_0 \geq 0$, then $\{h_i\}$ and $\{d_i\}$ are strictly increasing.*

Throughout the rest of this paper, all continued fractions mentioned will have $a_0 = 0$ so we will abbreviate $[a_0; a_1, a_2, \ldots, a_n]$ by $[a_1, a_2, \ldots, a_n]$ with the understanding that $a_0 = 0$.

Lemma 39. *If $x = [a_1, a_2, \ldots]$ or $x = [a_1, a_2, \ldots, a_n]$, then each convergent x_i is the rational number closest to x with denominator at most* $\mathrm{denom}(x_i)$. *More rigorously, if n/d is a rational number such that $|x - n/d| < |x - x_i|$ for some $i \geq 1$, then $d >$* $\mathrm{denom}(x_i)$.

Lemma 40. *For any irrational x, the following infinite inequality holds:*

$$x_2 < x_4 < x_6 < \ldots < x < \ldots < x_5 < x_3 < x_1.$$

This inequality also holds for rational x if we include only the first $i - 1$ convergents if $x_i = x$.

Proofs of Lemmas 39 and 40 can be found in [8].

As we will see, Salem's definition of ? turns out to be the same as the algorithmic definition of ?, but for the purposes of proving this, we will denote the function as defined by Salem as $\hat{?}$ For some $x \in [0, 1]$ with continued fraction representation $[a_1, a_2, \ldots, a_n]$ or $[a_1, a_2, \ldots]$, Salem defines $\hat{?}(x)$ by:

$$\hat{?}(x) = \frac{1}{2^{a_1 - 1}} - \frac{1}{2^{a_1 + a_2 - 1}} + \ldots + \frac{(-1)^{n+1}}{2^{a_1 + \ldots + a_n - 1}} + \ldots,$$

where the above sum terminates if and only if the continued fraction terminates.

4.5.2 Connection Between Convergents and Mediants

For a given $x \in [0, 1]$, the convergents of x have significance in the algorithmic calculation of $?(x)$. As the following theorem states, each convergent of x will become a mediant at some point in the calculation of $?(x)$.

Theorem 41. *Let $x \in (0,1)$. If $\{m_n\}$ is the sequence of mediants generated when we use Algorithm 3 to calculate $?(x)$ then $\{x_i\}$, the sequence of convergents to x, is a subsequence of $\{m_n\}$.*

Proof. We first consider the case $x_1 = 1$ and note that since $m_0 = 1$ we are done. Now we consider the cases where $x_1 < 1$. Let $\{a_n\}$ and $\{b_n\}$ be the sequences of lower and upper endpoints respectively. We claim that for each convergent x_i there exists some N such that $a_N < x_i < b_N$, and $a_{N+1} \geq x_i$ or $b_{N+1} \leq x_i$. To see this we first let x_i be some convergent of x and we note that $0 < x_i < 1$, $a_1 = 0$, and $b_1 = 1$. Then we note that if for all n, $a_n < x_i < b_n$, then x_i must never be the fraction with smallest denominator between a_n and b_n because otherwise it would be chosen to be an endpoint and thus equal to some a_n or some b_n. But there are only finitely many fractions with denominator less than $\text{denom}(x_i)$ and the algorithm exhausts them by turning them into endpoints or excluding them from the interval a_n, b_n, so this is a contradiction and thus the claim is proven. Without loss of generality assume $a_{N+1} \geq x_i$. Then $a_{N+1} = m_N$ so $m_N \geq x_i$. By Lemma 30, $\text{denom}(m_N) \leq \text{denom}(x_i)$ so by the contrapositive of Lemma 39, $|x - m_N| \geq |x - x_i|$. Since $m_N = a_{N+1}$, then by Remark 26, $m_N \leq x$. Since $x \geq m_N$, then $|x - m_N| \geq |x - x_i|$ implies $x_i \geq m_N$ and since we are assuming $m_N \geq x_i$, we have $m_N = x_i$. □

Now we will investigate how the sequence of convergents is distributed throughout the sequence of mediants. To explain the pattern that emerges, we will use an example. Table 2 keeps track of the values used in the calculation of $?(\frac{26}{59})$. As a continued fraction, $\frac{26}{59} = [2, 3, 1, 2, 2]$ and so its convergents are $[2] = \frac{1}{2}$, $[2, 3] = \frac{3}{7}$, $[2, 3, 1] = \frac{4}{9}$, $[2, 3, 1, 2] = \frac{11}{25}$ and $[2, 3, 1, 2, 2] = \frac{26}{59}$. For each value of n from 1 to 12, we report the w_n, m_n and u_n. We also write $\frac{26}{59}$ between w_n and m_n if $w_n < \frac{26}{59} < m_n$, and we write $\frac{26}{59}$ between m_n and u_n if $m_n < \frac{26}{59} < u_n$. When the convergents appear as mediants, they are typed in boldface. Note that after every convergent, $\frac{26}{59}$ switches to the other side of the mediant. After each convergent, an additional number is introduced into the

Table 2: Computation of $?(\frac{26}{59})$

n	w_n	x	m_n	x	u_n
1	$\frac{0}{1}$	$\frac{26}{59}$	$\mathbf{\frac{1}{2} = [2]}$		$\frac{1}{1}$
2	$\frac{0}{1}$		$\frac{1}{3} = [2, 1]$	$\frac{26}{59}$	$\frac{1}{2}$
3	$\frac{1}{3}$		$\frac{2}{5} = [2, 2]$	$\frac{26}{59}$	$\frac{1}{2}$
4	$\frac{2}{5}$		$\mathbf{\frac{3}{7} = [2, 3]}$	$\frac{26}{59}$	$\frac{1}{2}$
5	$\frac{3}{7}$	$\frac{26}{59}$	$\mathbf{\frac{4}{9} = [2, 3, 1]}$		$\frac{1}{2}$
6	$\frac{3}{7}$		$\frac{7}{16} = [2, 3, 1, 1]$	$\frac{26}{59}$	$\frac{4}{9}$
7	$\frac{7}{16}$		$\mathbf{\frac{11}{25} = [2, 3, 1, 2]}$	$\frac{26}{59}$	$\frac{4}{9}$
8	$\frac{11}{25}$	$\frac{26}{59}$	$\frac{15}{34} = [2, 3, 1, 2, 1]$		$\frac{4}{9}$
9	$\frac{11}{25}$		$\mathbf{\frac{26}{59} = [2, 3, 1, 2, 2]}$		$\frac{15}{34}$
10	$\frac{26}{59}$		$\frac{41}{93} = [2, 3, 1, 2, 2, 1]$		$\frac{15}{34}$
11	$\frac{26}{59}$		$\frac{67}{152} = [2, 3, 1, 2, 2, 2]$		$\frac{41}{93}$
12	$\frac{26}{59}$		$\frac{93}{211} = [2, 3, 1, 2, 2, 3]$		$\frac{67}{152}$
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$

continued fraction representation of the mediant. This number starts and 1, and at each mediant, it increases by 1 until the mediant reaches the next convergent. Then another number is introduced and the process continues.

Now we will state and prove the above patterns in their general form. We first prove that for any x, the mediants in the computation of $?(x)$ swap between being greater than and less than x after a convergent appears as a mediant.

Lemma 42. *Let $x = [a_1, a_2, \ldots] \in [0, 1]$ and assume that the convergents x_i and x_{i+1} exist.*

Then $x \neq x_i$.

Let n, k be such that $m_n = x_i$ and $m_{n+k} = x_{i+1}$.

Let $j \in \{1, 2, \ldots, k\}$.

If $x_i > x$ then $m_{n+j} \leq x$ with equality possible only if $j = k$.

If $x_i < x$ then $m_{n+j} \geq x$ with equality possible only if $j = k$.

Proof. Throughout the proof, we make implicit use of Lemma 40 which states that $x_l \leq x$ for all even l and $x_l \geq x$ for all odd l. We proceed by induction on i.

For the base case $i = 1$, we have $x_2 \leq x < x_1 = m_n$, and we will show $m_{n+j} \leq x$ for $j \in \{1, 2, \ldots, k\}$. Specifically,

$$x_1 = \frac{1}{a_1} \qquad \text{and} \qquad x_2 = \frac{1}{a_1 + \frac{1}{a_2}}.$$

Since 1 is odd, $x_1 > x$ by Lemma 40. Let n, k be such that $m_n = x_1$ and $m_{n+k} = x_2$ (n and k exist by Lemma 41). Let $j \in \{1, 2, \ldots, k\}$. Note that $w_n = 0$ and $u_{n+1} = \frac{1}{a_1}$ because for each positive integer $l \leq a_1$, we have $\frac{1}{l} \geq \frac{1}{a_1} > x$ and $m_l = \frac{1}{l+1}$.

We now show by induction on j that $u_{n+j} = \frac{1}{a_1}$ and $m_{n+j} = \frac{j}{ja_1+1}$. Note that $w_{n+1} = 0$ and $u_{n+1} = \frac{1}{a_1}$. To see this, note that for each $l < a_1$, we have $\frac{1}{l} > \frac{1}{a_1}$ and $m_l = \frac{1}{l+1}$. Also, $m_n = \frac{1}{a_1} > x$ so $u_{n+1} = m_n$. For the base case $j = 1$, we get our mediant

$$m_{n+1} = \frac{1+0}{a_1+1} = \frac{1}{a_1+1} \leq \frac{1}{a_1 + \frac{1}{a_2}} \leq x.$$

Now let $2 \leq j \leq k$. Our induction hypothesis is that $u_{n+j-1} = \frac{1}{a}$ and $m_{n+j-1} = \frac{j-1}{(j-1)a_1+1}$. First note that the sequence defined by $\frac{l}{a_1 l+1} = \frac{1}{a_1+\frac{1}{l}}$ is strictly increasing. Also, $l = a_2$ is the unique value such that $\frac{1}{a_1+\frac{1}{l}} = x_2$. So $k = a_2$. Therefore,

$$m_{n+j-1} = \frac{j-1}{(j-1)a_1+1} \leq \frac{1}{a_1+\frac{1}{k}} = \frac{1}{a_1+\frac{1}{a_2}} = x_2 \leq x.$$

So $w_{n+j} = m_{n+j-1} = \frac{1}{a_1+\frac{1}{j-1}}$ and $u_{n+j} = u_{n+j-1} = \frac{1}{a_1}$. Both these fractions are expressed in reduced terms, so our mediant is

$$m_{n+j} = \frac{j-1+1}{(j-1)a_1+1+a_1} = \frac{j}{ja_1+1} = \frac{1}{a_1+\frac{1}{j}}.$$

So for all $1 \leq j \leq k$, $m_{n+j} = \frac{1}{a_1+\frac{1}{j}} \leq \frac{1}{a_1+\frac{1}{a_2}} = x_2 \leq x$.

We now do the inductive step for $i \geq 2$. Assume the lemma is true for $i-1$. We assume i is odd and thus $x_i > x$. The proof for the even case is similar. Let n, k be such that $x_i = m_n$ and $x_{i+1} = m_{n+k}$. We now show by induction on j that $m_{n+j} \leq x$ for all $j \in \{1, 2, \ldots, k\}$.

We first show that $w_{n+1} = w_n = x_{i-1}$. By Theorem 41, there exists an l such that $x_{i-1} = m_l$. Then since $i-1$ is even, $x_{i-1} < x$ so $w_{l_{1+1}} = x_{i-1}$. Now we apply our induction hypothesis that the main lemma is true for $i-1$. This gives us that for all p such that $l + p \leq n$, we have $m_{l+p} \geq x$ so $w_{l+p+1} = w_{l+p} = x_{i-1}$. For $l + p = n$, $w_{n+1} = w_n = x_{i-1}$.

Now we do the base case $j = 1$. Since $i+1$ is odd, $x_{i+1} \leq x$. So $m_{n+1} \leq x$ follows from $m_{n+1} \leq x_{i+1}$. We will show $m_{n+1} \leq x_{i+1}$ by showing $m_{n+1} - w_{n+1} \leq x_{i+1} - w_{n+1}$. Let a, a', b, b' be positive integers such that $x_{i-1} = w_{n+1} = \frac{a}{b}$ and $x_i = m_n = \frac{a'}{b'}$ where these fractions are in reduced form. Since $m_n = x_i > x$, we have $u_{n+1} = m_n = \frac{a'}{b'}$, and so $m_{n+1} = \frac{a+a'}{b+b'}$. By Theorem 36, $x_{i+1} = \frac{a_{i+1}a'+a}{a_{i+1}b'+b}$, so

$$x_{i+1} - w_{n+1} = \frac{a_{i+1}a' + a}{a_{i+1}b' + b} - \frac{a}{b} = \frac{a_{i+1}(a'b - b'a)}{a_{i+1}bb' - b^2}.$$

Also, $m_{n+1} = \frac{a+a'}{b+b'}$ so,

$$m_{n+1} - w_{n+1} = \frac{a + a'}{b + b'} - \frac{a}{b} = \frac{a'b - b'a}{bb' - b^2} \leq \frac{a_{i+1}(a'b - b'a)}{a_{i+1}bb' - b^2} = x_{i+1} - w_{n+1}.$$

Now we do the inductive step for $2 \leq j \leq k$. We use strong induction. Our induction hypothesis will be that $m_{n+l} < x_{i+1}$ for all $l < j$. Then $w_{n+l+1} = m_{n+l}$ for all $l < j$, and $u_{n+l+1} = u_{n+1} = m_n = x_i$. In particular, $w_{n+j} = m_{n+j-1}$ and $u_{n+j} = x_i$. Just as in the $j = 1$ case, let a, a', b, b' be positive integers such that $x_{i-1} = w_{n+1} = \frac{a}{b}$ and

$x_i = m_n = \frac{a'}{b'}$ where these fractions are in reduced form. Then $w_{n+l} = \frac{a+(l-1)a'}{b+(l-1)b'}$. So

$$m_{n+j} = \frac{a + (j-1)a' + a'}{b + (j-1)b' + b'} = \frac{a + ja'}{b + jb'} = \frac{\frac{1}{j}a + a'}{\frac{1}{j}b + b'}.$$

Note that this sequence is strictly increasing and that $j = a_{i+1}$ is the unique value such that $\frac{\frac{1}{j}a+a'}{\frac{1}{j}b+b'} = x_{i+1}$. So $k = a_{i+1}$ and thus

$$m_{n+j} = \frac{a + ja'}{b + jb'} \leq \frac{a_{i+1}a' + a}{a_{i+1}b' + b} = x_{i+1} \leq x.$$

□

Corollary 43. *Let $x \in [0,1]$. Fix $i \geq 2$ such that $x \neq x_i$. Let n be such that $m_n = x_i$ (such an n exists by Theorem 41).*

If i is odd, then $w_n = x_{i-1}$.

If i is even, then $u_n = x_{i-1}$.

Lemma 44. *If $m_n = x_{i-1}$ then $m_{n+a_i} = x_i$.*

Proof. Let n be such that $m_n = x_{i-1}$. We will prove the case for when $i-1$ is odd. The even case is similar. Since $i-1$ is odd, $w_n = x_{i-2}$ by Corollary 43. Let k be such that $m_{n+k} = x_i$. We will show that $k = a_i$ and then we will be done. By Lemma 42, for all $j \in \{1, 2, \ldots, k\}$, $m_{n+j} \leq x$ and thus $u_{n+j} = x_{i-1} = \frac{h_{i-1}}{d_{i-1}}$. Also note that for $j < k$, $m_{n+j} = w_{n+j+1}$. Since $w_{n+1} = x_{i-2}$,

$$m_{n+1} = \frac{h_{i-1} + h_{i-2}}{d_{i-1} + d_{i-2}},$$
$$m_{n+2} = \frac{(h_{i-1} + h_{i-2}) + h_{i-1}}{(d_{i-1} + d_{i-2}) + d_{i-1}} = \frac{2h_{i-1} + h_{i-2}}{2d_{i-1} + di - 2},$$
$$m_{n+k} = \frac{kh_{i-1} + h_{i-2}}{kd_{i-1} + d_{i-2}}.$$

Since k is defined such that $m_{n+k} = x_i$, we have

$$\frac{kh_{i-1} + h_{i-2}}{kd_{i-1} + d_{i-2}} = x_i = \frac{h_i}{d_i} = \frac{a_i h_{i-1} + h_{i-2}}{a_i d_{i-1} + d_{i-2}},$$

and thus $k = a_i$ as desired.

□

Lemma 45. *If* $x = [a_1, a_2, \ldots]$ *or* $[a_1, a_2, \ldots, a_k]$ *then* $m_n = x_i$ *if and only if* $n = a_1 + a_2 + \ldots + a_i - 1$

Proof. We proceed by induction on i. We already have the base case $i = 1$ in the proof of Lemma 42. For the inductive step, let $k = a_1 + a_2 + \ldots + a_{i-1} - 1$ and assume that $m_k = x_{i-1}$. Then, by Lemma 44, $m_{k+a_i} = x_i$. So letting $n = k + a_i$ we have $m_n = x_i$ where $n = a_1 + a_2 + \ldots + a_i - 1$. □

4.5.3 Agreement with Salem's Extension

Lemma 46. *Let* a, k *be nonnegative integers. Then* $\frac{1}{2^a} - \frac{1}{2^{a+k}} = \frac{1}{2^{a+1}} + \frac{1}{2^{a+2}} + \ldots + \frac{1}{2^{a+k}}$.

Proof.

$$\begin{aligned} \frac{1}{2^a} - \frac{1}{2^{a+k}} &= \\ \frac{1}{2^{a+1}} + \frac{1}{2^{a+1}} - \frac{1}{2^{a+k}} &= \\ \frac{1}{2^{a+1}} + \frac{1}{2^{a+2}} + \frac{1}{2^{a+2}} - \frac{1}{2^{a+k}} &= \ldots = \\ \frac{1}{2^{a+1}} + \frac{1}{2^{a+2}} + \ldots + \frac{1}{2^{a+k}} + \frac{1}{2^{a+k}} - \frac{1}{2^{a+k}} &= \\ \frac{1}{2^{a+1}} + \frac{1}{2^{a+2}} + \ldots + \frac{1}{2^{a+k}}. \end{aligned}$$

□

Theorem 47. *The function ? defined by Algorithm 3 is the same as the function* $\hat{?}$ *defined by Salem.*

Proof. Assume that x is irrational, so the continued fraction does not terminate. Using Lemma 46 we can rewrite Salem's extension $\hat{?}$ as follows

$$\hat{?}(x) = \sum_{k=1}^{\infty} \sum_{j=a_{2k-1}}^{a_{2k-1}+a_{2k}-1} \frac{1}{2^j}.$$

We can see that this is exactly what the algorithm does. For each n, Algorithm 3 adds $\frac{1}{2^n}$ to the running sum whenever $x \geq m_n$. By Lemma 42, this happens every iteration that follows the iteration where the mediant was an odd convergent of x until, but not including, the first iteration after the next even convergent is a midpoint. By Lemma 45, these are the iterations that are numbered $a_k, a_k+1, \ldots, a_k+a_{k+1}-1$ for all odd k.

For rational x, the proof is similar. □

5 Conclusion and Ideas For Further Study

We were able to use our algorithm to find a reasonably elegant proof for the singularity of $f_a(x)$ and explicit sets on which $f_a'(x)$ did and did not exist.h. Maybe we can find something similar for $?(x)$. In [4], it is shown that $?'(x)$ does not exist for any x with continued fraction representation where the terms do not exceed 4. The proof of this is very messy. As of now, we are able to use the algorithmic definition to find a much neater proof of the much weaker claim that that $?'(x)$ does not exist when x has continued fraction expansion that is eventually periodic of period length at most two, where the numbers in the period do not exceed 4. Hopefully, we will be able to extend this and investigate more of the known properties about $?(x)$ and $f_a(x)$ using our algorithms.

Another idea is to use these algorithmic definitions to explore the connections between Minkowski's Question-Mark Function, DeRham's Function, their inverses and other variations. We already know of one connection. At each step of the algorithm for $?(x)$, we have an interval $[w, u]$ and we choose the mediant by adding the numerators and denominators of u and v. If instead, for some $a \in (0, 1)$, we choose the point that is

a part of the way between w and u, our algorithm will generate $f_a^{-1}(x)$. Perhaps there is an algorithmic template that can describe all strictly increasing singular functions.

References

[1] Stephen Abbott. *Understanding Analysis.* Springer, New York, 2001.

[2] Randolph M. Conley. A survey of the minkowski ?(x) function. Master's thesis, West Virginia University, Morgantown, 2003.

[3] Georges DeRham. On some curves defined by functional equations. *Rendiconti del Seminario Matematico dell'Universita e del Politecnio di Torino*, pages 101–113, 1957. Translated by Ilan Vardi.

[4] Anna A. Dushistova and Nikolai G. Moshchevitin. On the derivative of the minkowski question mark function $?(x)$. *Fundementalnaya i Prikladnaya Matematika*, 2010.

[5] Bernard R. Gelbaum and John M. H. Olmsted. *Counterexamples in Analysis.* Holden-Day, 1964.

[6] Kiko Kawamura. On the set of points where lebesgue's singular function has the derivative zero. *Proceedings of the Japan Academy*, 87(A), October 2011.

[7] Davar Khoshenvisan. Normal numbers are normal. *CMI Annual Report*, 2006.

[8] I. Niven, H. Zuckerman, and H. Montgomery. *An Introduction to the Theory of Numbers.* Wiley, 5th edition edition, 1991.

[9] H. L. Royden. *Real Analysis.* Macmillan, New York, 1963.

[10] R. Salem. On some singular monotonic functions which are strictly increasing. *Transactions of the American Mathematical Society*, 53(3):427–439, May 1943.