Exercise 21. *Show that any ternary decimal of the form* $0.a_1a_2\cdots a_n11$ *(base 3) is not an element of* Δ.

Proof. Because $0.a_1a_2\cdots a_n1 < 0.a_1a_2\cdots a_n11 < 0.a_1a_2\cdots a_n12$ and $(0.a_1a_2\cdots a_n1, 0.a_1a_2\cdots a_n12)$ will be chopped off, thus $0.a_1a_2\cdots a_n11$ is not an element of Δ. $\square$

Exercise 22. *Show that* Δ *contains no (nonempty) open intervals.*

Proof. We have the total length of $[0,1] \setminus \Delta = 1$. Thus Δ cannot contain any open interval. $\square$

Exercise 23. *Show that the endpoints of* Δ *other that 0 and 1 can be written as* $0.a_1a_2\cdots a_{n+1}$ *(base 3), where each* a_k *is 0 or 2, except* a_{n+1}*, which is either 1 or 2. That is, the discarded "middle third" intervals are of the form* $(0.a_1a_2\cdots a_n1,\ 0.a_1a_2\cdots a_n2)$*, where both entries are points of* Δ *written in base 3.*

Proof. We will prove the part "that is" using mathematical induction. For $n = 1$, the discarded interval is $(0.1, 0.2)$, and for $n = 1$, the discarded intervals are $(0.01, 0.02)$ and $(0.21, 0.22)$. Thus the hypothesis holds. Assume that the hypothesis holds for $n = i$, that is, the discarded interval has the form $(0.a_1a_2\cdots a_i1,\ 0.a_1a_2\cdots a_i2)$ where both entries are points of Δ, we will prove that it is also hold for $n = i + 1$.

Notice that the discarded interval in the $i{+}1$th step cannot be $(0.a_1a_2\cdots a_{i+1}0,\ 0.a_1a_2\cdots a_{i+1}1)$ because it is the first third of the remaining interval $(0.a_1a_2\cdots a_{i+1},\ 0.a_1a_2\cdots a_{i+1}1)$ in the ith step. It cannot be $(0.a_1a_2\cdots a_{i+1}2,\ (0.a_1a_2\cdots a_{i+1} + \frac{1}{3^{i+2}}))$ because it is the last third of the remaining interval $(0.a_1a_2\cdots a_{i+1},\ (0.a_1a_2\cdots a_{i+1} + \frac{1}{3^{i+2}}))$. Thus it can only have the form $(0.a_1a_2\cdots a_{i+1}1,\ 0.a_1a_2\cdots a_{i+1}2)$. $\square$

Exercise 24. *Show that* Δ *is perfect; that is, every point is* Δ *is the limit of a sequence of distinct points from* Δ*. In fact, show that every point in* Δ *is the limit of a sequence of distinct endpoints.*

Proof. We will split a point in Δ into two cases. Case 1, it contains finitely many digits (base 3). That means, it can be represented as $0.a_1a_2\cdots a_n$ be a point in Δ. If $a_n = 1$, defined a sequence $x_i = 0.a_1a_2\cdots a_n - \frac{1}{3^n} + \frac{2}{3^{n+1}} + \frac{2}{3^{n+2}} + \cdots + \frac{2}{3^{n+i}}$. Clearly, $x_i \to 0.a_1a_2\cdots a_n$ and for any i, x_i contains only 0 and 2. Thus x_i is an end point.

If $a_n = 2$, we defined $x_i = 0.a_1a_2\cdots a_n + \frac{1}{3^n} - \frac{2}{3^{n+1}} - \frac{2}{3^{n+2}} - \cdots - \frac{2}{3^{n+i}}$. Clearly, $x_i \to 0.a_1a_2\cdots a_n$ and x_i contains only 0's and 2's except for the last digit, which is 1. Thus $x_i \in \Delta$.

If $a_n = 0$, then that number is either 1 or 0. If $0.a_1a_2\cdots a_n = 0$, then let $(x_i) = \frac{1}{3^i}$ and if $0.a_1a_2\cdots a_n = 1$ then $(x_i) = 1 - \frac{1}{3^i}$. We can easily check that $x_i \in \Delta$ and $x_i \to 0.a_1a_2\cdots a_n$.

If that point contains infinitely many digits: $0.a_1a_2\cdots$, then a_i can be either 0 or 2 and there are infinitely many 2's. Let $(x_i) = 0.a_1a_2\cdots a_{t_i}$ such that x_i contains i 2's. Thus $x_i \in \Delta$ and $x_i \to 0.a_1a_2\cdots$.

Thus there is always exist a sequence of points in Δ that converges to a point in Δ. $\square$

Exercise 26. *Let $f : \Delta \to [0,1]$ be the Cantor function and let $x, y \in \Delta$ with $x < y$. Show that $f(x) \le f(y)$. If $f(x) = f(y)$, show that x has two distinct binary decimal expansions. Finally, show that $f(x) = f(y)$ if and only if x and y are consecutive endpoints of the form $x = 0.a_1a_2\cdots a_n1$ and $y = 0.a_1a_2\cdots a_n2$ (base 3).*

Proof. Any endpoints t of Δ has the form

$$\sum_{n=1}^{\infty} \frac{2b_{t,n}}{3^n}$$

for $b_n = 0,\ 1$. Thus if $x < y$, then there exists an integer k such that $b_{x,1} = b_{y,1}, b_{x,2} = b_{y,2}, \cdots, b_{x,k-1} = b_{y,k-1}$ and $b_{x,k} < b_{y,k}$. Thus

$$\sum_{n=1}^{\infty} \frac{2b_{x,n}}{2^n} \le \sum_{n=1}^{\infty} \frac{2b_{y,n}}{2^n}$$

or $f(x) \le f(y)$. If $f(x) = f(y)$ then x and y must be of the forms $0.b_{x,1}\cdots b_{x,k-1}0111\cdots$ and $0.b_{y,1}\cdots b_{y,k-1}1$ (base 2). Thus $x = 0.b_{x,1}\cdots b_{x,k-1}1\cdots$ and $y = 0.b_{y,1}\cdots b_{y,k-1}2$ (base 3). ☐

Exercise 27. *Fix $n \ge 1$, and let $I_{n,k}, k = 1, \cdots, 2^{n-1}$ be the component subintervals of the nth level Cantor set I_n. If $x, y \in \Delta$ with $|x - y| < 3^{-n}$, show that x and y are in the same component $I_{n,k}$. For this same pair of points show that $|f(x) - f(y)| \le 2^{-n}$.*

Proof. To show that x and y are in the same component $I_{n,k}$, we will prove that after the nth step, any removed interval has the length of 3^{-n} or above. Indeed, by exercise 23, the middle third intervals have the form $(0.a_1a_2\cdots a_t1, 0.a_1a_2\cdots a_t2)$ where $t = 0,\ 1, \cdots, n-1$. Thus all the intervals are longer or equal to 3^{-n}. Thus if $|x - y| < 3^{-n}$, x, y must be in the same component $I_{n,k}$.

If x, y are in the same component $I_{n,k}$, then either $I_{n,k} = (0.a_1a_2\cdots a_{n-1}0, 0.a_1a_2\cdots a_{n-1}1)$ or $I_{n,k} = (0.a_1a_2\cdots a_{n-1}2, 0.a_1a_2\cdots a_{n-1}2 + 3^{-n})$. In the first case, we have

$$|f(x) - f(y)| \le |0.a_1a_2\cdots a_{n-1}0 - 0.a_1a_2\cdots a_{n-1}1| \text{ (base 2)} = 2^{-n}.$$

Similar with the second case, we have $|f(x) - f(y)| \le 2^{-n}$. ☐

Exercise 28. *Let $f : \Delta \to [0,1]$ be then Cantor function (as originally defined). Check that $f(x) = \sup\{f(y) : y \in \Delta, y \le x\}$ for any $x \in \Delta$.*

Proof. In exercise 26, we already proved that if $x, y \in \Delta$ and $x < y$ then $f(x) \le f(y)$. Now, for any $f(v) \in \{f(y) : y \in \Delta, y \le x\}$, if $v < x$ then $f(v) \le f(x)$. if $v = x$ then $f(v) = f(x)$. Thus $f(x)$ is an upper bound for $\{f(y) : y \in \Delta, y \le x\}$. Moreover, $f(x) \in \{f(y) : y \in \Delta, y \le x\}$, thus $\sup\{f(y) : y \in \Delta, y \le x\} = f(x)$ for any $x \in \Delta$. ☐

Exercise 29. *Prove that the extended Cantor function $f : [0,1] \to [0,1]$ is increasing.*

Proof. For any $x, y \in [0,1]$ and $x < y$, we have $\{f(t) : t \in \Delta, t \le x\} \subset \{f(t) : t \in \Delta, t \le y\}$. Thus $f(x) = \sup\{f(t) : t \in \Delta, t \le x\} \le \sup\{f(t) : t \in \Delta, t \le y\} = f(y)$. Thus f is increasing. ☐

Exercise 30. *Check that the construction of the generalized Cantor set with parameter α leads to a set of measure $1-\alpha$; that is, check that the discarded intervals now have total length α.*

Proof. At the nth stage, we discard 2^{n-1} intervals, each has length $\alpha 3^{-n}$, thus the total length of the discarded interval is

$$\begin{aligned}\sum_{n=1}^{\infty} \alpha 2^{n-1} 3^{-n} &= \alpha \sum_{n=1}^{\infty} 2^{n-1} 3^{-n} \\ &= \frac{1}{2}\alpha \sum_{n=1}^{\infty} \alpha 2^{n} 3^{-n} \\ &= \frac{1}{2}\alpha \left(\frac{\frac{2}{3}^{\infty} - 1}{\frac{2}{3} - 1} - 1 \right) \\ &= \alpha.\end{aligned}$$

Thus the discarded interval has length α. $\square$

Exercise 32. *Deduce from Theorem 2.17 that a monotone function $f : \mathbb{R} \to \mathbb{R}$ has points of continuity in every open interval.*

Proof. First, remind that theorem 2.17 states as follow, if $f : (a, b) \to \mathbb{R}$ is monotone, then f has at most countably many points of discontinuity in (a, b), all of which are jump discontinuities.

Back to the problem, for a monotone function $f : \mathbb{R} \to \mathbb{R}$, assume that there exists an open interval (a, b) such that f discontinuous at every point in (a, b). Then the set of discontinuity points consists of $\{x : a < x < b\}$. And since this set is uncountable, the set of discontinuity points of f is also uncountable, which is contradict to theorem 2.17. Thus f must have points of continuity in every open interval. $\square$

Exercise 34. *Let $D = \{x_1, x_2, \cdots\}$, and let $\epsilon_n > 0$ with $\sum_{n=1}^{\infty} \epsilon_n < \infty$. Define $f(x) = \sum_{x_n \le x} \epsilon_n$. Check the following:*

(i) f is discontinuous at the points of D.

Proof. For any x_i, we have $f(x_i-) = \sum_{x_n < x_i} \epsilon_n$ and $f(x_i) = \sum_{x_n \le x_i} \epsilon_n = \sum_{x_n < x_i} \epsilon_n + \epsilon_i > f(x_i-)$. Thus f is discontinuous at the points of D. $\square$

(ii) f is right-continuous everywhere.

Proof. For any t, if $t \notin D$ then f continuous at t, thus f is right-continuous at t. If $t \in D$, assume that $t = x_i$. We will prove that $f(x_i)$ is right-continuous using the definition.

By the definition, f is right-continuous at x_i if and only if for any $\epsilon > 0$, there exists $\delta > 0$ such that, if $x_i < y < x_i + \delta$, then $|f(y) - f(x)| < \epsilon$, or

$$\begin{aligned} &\Big|\sum_{x_n \le y} \epsilon_n - \sum_{x_n \le x_i} \epsilon_n\Big| < \epsilon \\ \Leftrightarrow &\sum_{x_i < x_n \le y} \epsilon_n < \epsilon.\end{aligned}$$

We know that

$$\sum_{x_i<x_n\leq y} \epsilon_n \leq \sum_{x_i<x_n\leq x_i+\delta} \epsilon_n,$$

thus it's sufficient to prove that

$$\sum_{x_i<x_n\leq x_i+\delta} \epsilon_n < \epsilon$$

for some $\delta > 0$. Notice that $\sum_{n=1}^{\infty} \epsilon_n < \infty$, thus

$$\lim_{N\to\infty} \sum_{n=N}^{\infty} \epsilon_n = 0,$$

which means, by the definition, there exists $\delta' \in \mathbb{N}$ such that

$$\sum_{n=\delta'}^{\infty} \epsilon_n < \epsilon.$$

Now, let $\delta = \min\{x_t | t \leq \delta' \text{ and } x_t > x_i\} - x_i$, then for any $x \in \{x_1, x_2, \cdots, x_{\delta'}\}$, either $x \leq x_i$ or $\delta < x - x_i$, which means $x_i + \delta < x$. Thus $(x_i, x_i + \delta) \cap \{x_1, x_2, \cdots, x_{\delta'}\} = \varnothing$. Thus

$$\sum_{x_i<x_n\leq x_i+\delta} \epsilon_n < \sum_{n=\delta'}^{\infty} \epsilon_n < \epsilon.$$

Since such δ exists, f is always right-continuous. $\square$

(iii) f is continuous at each point $x \in \mathbb{R} \setminus D$. How might this construction be modified so as to yield a strictly increasing function with these same properties?

Proof. Since f is already right-continuous everywhere, it's sufficient to prove that $f(x)$ is left-continuous for any $x \in \mathbb{R} \setminus D$. Indeed, we have

$$f(x) = \sum_{x_n\leq x} \epsilon_n = \sum_{x_n<x} \epsilon_n$$

since $x_n \neq x$ for all n. Moreover,

$$f(x-) = \lim_{t\to x^-} \sum_{x_n\leq t} \epsilon_n = \sum_{x_n<x} \epsilon_n.$$

Clearly, $f(x) = f(x-)$, we get $f(x)$ is also left-continuous. Thus $f(x)$ is continuous for all $x \in \mathbb{R} \setminus D$.

Let $g(x) = x + f(x)$, since $g'(x) = 1$, g is strictly increase, and g has the same continuous properties as f. $\square$

Exercise 35. *Let $f : [a, b] \to \mathbb{R}$ be increasing, and let (x_n) be an enumeration of the discontinuities of f. For each n, let $a_n = f(x_n) - f(x_n-)$ and $b_n = f(x_n+) - f(x_n)$ be the left and right "jumps" in the graph of f, where $a_n = 0$ if $x_n = a$ and $b_n = 0$ if $x_n = b$. Show that $\sum_{n=1}^{\infty} a_n \leq f(b) - f(a)$ and $\sum_{n=1}^{\infty} b_n \leq f(b) - f(a)$.*

Proof. First, notice that

$$\begin{aligned}\sum_{n=1}^{\infty} a_n + b_n &= f(x_1) - f(x_1-) + f(x_1+) - f(x_1) + \cdots \\ &= f(x_1+) - f(x_1-) + f(x_2+) - f(x_2-) + \cdots .\end{aligned}$$

Now, we will prove that for the set $X = \{x_n | x_n \text{ is a point of discontinuity of } f\}$, we always have

$$\sum_{n=1}^{|X|} a_n + b_n \leq f(\sup(X)+) - f(\inf(X)-).$$

Indeed, if $|N| = 1$, then the result becomes $f(x_1+) - f(x_1-) \leq f(x_1+) - f(x_1-)$, which is trivial. Assume that the result holds for $|X| = k$, we will prove that for $|X| = k+1$, this result also holds. Indeed, for $X = \{x_1, x_2, \cdots, x_{k+1}\}$, without loss of generality, assume that $\inf(X) = x_1$. And since f is increasing, we have $f(x_1+) \leq f(\inf(X \setminus \{x_1\})$. Thus

$$\begin{aligned}\sum_{n=1}^{k+1} a_n + b_n &= f(x_1+) - f(x_1-) + \sum_{n=2}^{k+1} f(x_n+) - f(x_n-) \\ &\leq f(x_1+) - f(x_1-) + f(\sup(X \setminus \{x_1\}) - f(\inf(X \setminus \{x_1\}) \\ &\leq f(\sup(X \setminus \{x_1\}) - f(x_1-) \\ &= f(\sup(X)+) - f(\inf(X)-).\end{aligned}$$

For the case $|X| = \infty$, let $X_k = \{x_1, x_2, \cdots, x_k\}$, we have

$$\sum_{n=1}^{k} a_n + b_n \leq f(\sup(X_k)+) - f(\inf(X_k)-).$$

Take the limit both sides, we get

$$\lim_{k\to\infty} \sum_{n=1}^{k} a_n + b_n \leq \lim_{k\to\infty} f(\sup(X_k)+) - \lim_{k\to\infty} f(\inf(X_k)-).$$

Notice that since $\sup(X_k)$ is increasing and $\inf(X_k)$ is decreasing as $k \to \infty$, the sequence $f(\sup(X_k)+) - f(\inf(X_k)-)$ is increasing and be bounded by $f(\sup[a, b]+) - f(\inf[a, b]-) = f(a) - f(b)$. Thus $\lim_{k\to\infty} f(\sup(X_k)+) - f(\inf(X_k)-)$ is converge. Thus the sequence $\sum_{n=1}^{k} a_n + b_n$ is bounded by $f(a) - f(b)$ too. Moreover, because $a_n, b_n \geq 0$, $\sum_{n=1}^{k} a_n + b_n$ is increasing, thus it is also converge. Thus

$$\sum_{n=1}^{\infty} a_n + b_n \leq f(\sup(X)+) - f(\inf(X)-).$$

Now, since f is increasing and $\sup(X), \inf(X) \in [a, b]$, we always have $f(\sup(X)+) - f(\inf(X)-) \leq f(b) - f(a)$. Thus

$$\sum_{n=1}^{|X|} a_n + b_n \leq f(b) - f(a),$$

for X finite and

$$\sum_{n=1}^{\infty} a_n + b_n \leq f(b) - f(a),$$

for X infinite, which is a stronger result, compared to $\sum_{n=1}^{\infty} a_n \leq f(b) - f(a)$ and $\sum_{n=1}^{\infty} b_n \leq f(b) - f(a)$. □

Exercise 36. *In the notation of Exercise 35, define* $h(x) = \sum_{x_n \leq x} a_n + \sum_{x_n < x} b_n$. *Show that h is increasing and that $g = f - h$ is both continuous and increasing. Thus, each increasing function f can be written as the sum of a continuous increasing function g and a "pure jump" function h.*

Proof. Since both a_n and b_n is non-negative,

$$h(x) = \sum_{x_n \leq x} a_n + \sum_{x_n < x} b_n$$

is increasing. For any $x > y$, we have

$$g(x) - g(y) = (f(x) - f(y)) - (h(x) - h(y)).$$

But by exercise 35, we know that

$$f(x) - f(y) \geq \sum_{y \leq x_n \leq x} a_n + b_n,$$

thus $g(x) - g(y) > 0$, which means g is increasing. Now, if $x \notin X$, then $f(x)$ and $h(x)$ is continuous. Thus $g(x) = f(x) - h(x)$ is also continuous. If $x \in X$, let $x = x_i$, we will prove that $g(x-) = g(x) = g(x+)$. Indeed,

$$\begin{aligned}
g(x_i) - g(x_i-) &= (f(x_i) - f(x_i-)) - (h(x_i) - h(x_i-)) \\
&= a_i - (\sum_{x_n \leq x_i} a_n + \sum_{x_n < x_i} b_n - \sum_{x_n < x_i} a_n - \sum_{x_n < x_i} b_n) \\
&= a_i - a_i \\
&= 0.
\end{aligned}$$

Thus $g(x_i) = g(x_i-)$. Moreover, we have

$$\begin{aligned}
g(x_i+) - g(x_i) &= (f(x_i+) - f(x_i)) - (h(x_i+) - h(x_i)) \\
&= b_i - (\sum_{x_n \leq x_i} a_n + \sum_{x_n \leq x_i} b_n - \sum_{x_n \leq x_i} a_n - \sum_{x_n < x_i} b_n) \\
&= b_i - b_i \\
&= 0,
\end{aligned}$$

which means $g(x_i+) = g(x)$. Thus g is continuous. □