The Nested Interval Theorem

How can a blind pediatrician measure the height of a child? Given that the doctor is taller than the child, consider the following suggested procedure:

Step 1: Swing the arm at shoulder level so as to slap the child. This will be a "miss". Call it m_1. Next, swing the arm at knee-level to obtain a "hit" (marked by vocalizations from the child). Call it h_1.
Note that the actual height of the child, c, satisfies $h_1 < c < m_1$.

Step 2: Try to slap the child halfway between the shoulder- and knee-level. If the result is eliciting vocalizations from the child, mark the result as a "hit" (h_2). Otherwise mark it as a "miss" (m_2). If the result is h_2, re-lable $m_1 = m_2$ If the result is m_2, re-lable $h_1 = h_2$.
Notice that $h_2 < c < m_2$ and $m_2 - h_2 = \frac{m_1 - h_1}{2}$

⋮

Step n: Having obtained h_{n-1} and m_{n-1}, swing halfway between h_{n-1} and m_{n-1}. Lable the result of this experiment as either h_n or m_n depending on the presence or absence of vocalization signals. Re-lable either $h_{n-1} = h_n$ or $m_{n-1} = m_n$

accordingly.

Notice that $h_n < c < m_n$ and $m_n - h_n = \frac{m_1 - h_1}{2^n}$.

Continue in this fashion ad infinitum.

Will this procedure be appropriate to determine the exact height of the patient? The answer is "yes" in $\mathbb{R}$. But before we can explain why this is, we will need the following observation:

Thm: A monotone, bounded sequence of real numbers converges in $\mathbb{R}$.

Proof: Let $\{x_n\}_{n=1}^{\infty} \subset \mathbb{R}$ be monotone and bounded.

We first suppose that $\{x_n\}$ is increasing (That is $x_m \leq x_n$ whenever $m < n$). Now, since $\{x_n\}$ is bounded, we may set $x = \sup_{n \in \mathbb{N}} x_n$ (a real number). We will show that

$x = \lim_{n \to \infty} x_n$.

Let $\epsilon > 0$. Since $x - \epsilon < x = \sup_{n\in\mathbb{N}} x_n$, we must have $x_N > x - \epsilon$ for some N. But then, for any $n \geq N$, we have $x - \epsilon < x_N \leq x_n \leq x$. (Why?) That is, $|x - x_n| < \epsilon$ for all $n \geq N$. Consequently, $\{x_n\}$ converges and $x = \sup_{n\in\mathbb{N}} x_n = \lim_{n\to\infty} x_n$.

Finally, if $\{x_n\}$ is decreasing, consider the increasing sequence $\{-x_n\}$. From the first part of the proof, $\{-x_n\}$ converges to $\sup_{n\in\mathbb{N}}(-x_n) = -\inf_{n\in\mathbb{N}} x_n$. It then follows that $\{x_n\}$ converges to $\inf_{n\in\mathbb{N}} x_n$.

The Nested Interval Theorem: If $\{I_n\}_{n=1}^{\infty}$ is a sequence of closed, bounded, nonempty intervals in $\mathbb{R}$ with $I_1 \supset I_2 \supset I_3 \supset \ldots$, then $\bigcap_{n=1}^{\infty} I_n \neq \emptyset$. If, in addition, $\text{length}(I_n) \to 0$, then $\bigcap_{n=1}^{\infty} I_n$ contains precisely one point.

Proof: Write $[a_n, b_n]$. Then $I_n \supset I_{n+1}$ means that $a_n \leq a_{n+1} \leq b_{n+1} \leq b_n$ for all n. Thus, $a = \lim_{n\to\infty} a_n = \sup_{n\in\mathbb{N}} a_n$ and $b = \lim_{n\to\infty} b_n = \inf_{n\in\mathbb{N}} b_n$ both exist (as finite real numbers) and satisfy $a \leq b$ (Why?) Thus we must have $\bigcap_{n=1}^{\infty} I_n = [a, b]$. Indeed, if $x \in I_n$ for all n, then $a_n \leq x \leq b_n$ for all n, and hence $a \leq x \leq b$. Conversely, if $a \leq x \leq b$, then $a_n \leq x \leq b_n$ for all n. That is, $x \in I_n$, for all n. Finally, if $b_n - a_n = \text{length}(I_n) \to 0$, then $a = b$ and so $\bigcap_{n=1}^{\infty} I_n = \{a\}$.

Ex. (a) Please note that it is essential that the intervals used in the nested interval theorem be both closed and bounded. Indeed, $\bigcap_{n=1}^{\infty}[n,\infty)=\emptyset$ and $\bigcap_{n=1}^{\infty}(0,1/n]=\emptyset$.

(b) Suppose that $\{I_n\}$ is a sequence of closed intervals with $I_n \supset I_{n+1}$, for all n and with length$(I_n)\to 0$ as $n\to\infty$. If $\bigcap_{n=1}^{\infty} I_n=\{x\}$, then any sequence of points $\{x_n\}_{n=1}^{\infty}$, with $x_n \in I_n$ for all n, must converge to x. (Why?)

The Nested Interval Theorem will prove to be a very useful tool in the next few topics.

Perfect Sets

Def: A set P is called perfect if it is empty or if it is a closed set and every point of P is a limit point of P.

Ex. (a) $\mathbb{R}$, $(-\infty,a]$, and $[a,\infty)$ as well as any closed and bounded interval $[a,b]$ $(a<b)$ are perfect sets. (Why?)

(b) The sets (a,b), $[a,b]\cup\{c\}$ $(b<c)$, $\mathbb{Q}$, and $\mathbb{R}\setminus\mathbb{Q}$ are not perfect sets. The sets (a,b), $\mathbb{Q}$, and $\mathbb{R}\setminus\mathbb{Q}$ fail to be closed, even though every point in each of these sets is a limit point of the sets. The set $[a,b]\cup\{c\}$ fails to be a perfect set, because c is not a limit point of $[a,b]\cup\{c\}$.

(c) Let $\{x_n\}_{n=1}^{\infty}$ be convergent in (M,d). That is $x_n \xrightarrow{d} x$ in M. Then the set $\{x_n : n\geq 1\}\cup\{x\}$ is not perfect. Although the set is closed, only x is a limit point.

Notice that in all of the above listed examples of perfect subsets of $\mathbb{R}$, the perfect sets turned up to be uncountable sets. Must this always be the case? According to the next theorem, the answer is "yes"!

Thm: Let P be a perfect subset of $\mathbb{R}$. Then P is uncountable.

Proof: Suppose to the contrary that $P = \{x_1, x_2, \dots, x_n, \dots\}$ is countable. We shall try to obtain a contradiction.
To do so, let I_1 be any closed interval centered at x_1 of length$(I_1) \leq 1$. Then, since $x_1 \in P$ and P is perfect, it follows that x_1 is a limit point of P. In particular, $(I_1 \setminus \{x_1\}) \cap P \neq \emptyset$ (Why?). Let n_2 be the smallest integer for which $x_{n_2} \in (I_1 \setminus \{x_1\}) \cap P$ and let I_2 be any closed interval centered at x_{n_2} of length$(I_2) \leq 1/2$ such that $I_2 \subset I_1$ and $x_1 \notin I_2$. Observe that by the minimality of n_2, $x_k \notin I_2$ for any $k < n_2$ (Why?).
Since $x_{n_2} \in P$, it is a limit point of P and therefore $(I_2 \setminus \{x_{n_2}\}) \cap P \neq \emptyset$.
Let n_3 be the smallest integer for which $x_{n_3} \in (I_2 \setminus \{x_{n_2}\}) \cap P$.
Set I_3 to be any closed interval centered at x_{n_3} of length$(I_3) \leq 1/3$ such that $I_3 \subset I_2$ and $x_{n_2} \notin I_3$.
Continuing in this fashion, we obtain a nested sequence of closed intervals $I_1 \supset I_2 \supset I_3 \supset \dots$ such that length$(I_n) \mapsto 0$ as $n \to \infty$ and $x_k \in I_m$ for all $k < n_m$. By the Nested Interval Theorem,

$\bigcap_{n=1}^{\infty} I_n = \{x\}$ for some $x \in \mathbb{R}$. Notice, however, that x is a limit point of P because it is the limit point of the center points of the intervals I_n. Thus, as P is closed, we must have $x \in P$. And here is the contradiction; x cannot be any of the x_m, since $m \le n_m$ and $x_m \notin I_{m+1}$.
Thus P must be uncountable as desired.

Ex. (a) Although we are still lacking the means to prove it, it can be shown that if P is a nonempty perfect subset of (M,d) in which every Cauchy sequence converges, P must be an uncountable set.

(b) Let (M,d) be a discrete metric space. Suppose $P \subset M$ is not empty. Then P is not perfect; $B_1(x) = \{x\}$ for any $x \in P$. Hence x cannot be a limit point of P. What went wrong? Every Cauchy sequence in (M,d) must converge in (M,d). Why don't we have perfect sets other than $\emptyset$?

(c) Let $M = \mathbb{Q}$ under the usual metric of $\mathbb{R}$. Then $P = [0,1] \cap \mathbb{Q}$ is perfect in $\mathbb{Q}$. Notice, however, that $P \subset \mathbb{Q}$ and must therefore be countable. Does this contradict (a)? Absolutely not! Lots of Cauchy sequences of elements of $\mathbb{Q}$ fail to converge in $\mathbb{Q}$.

It appears that nonempty perfect subsets are rather large. One would expect these sets to occupy some space on the real number line, for instance. However, reality is stranger than intuition.

It seems that perfect subsets of $\mathbb{R}$ can be so constructed as to include almost all of $\mathbb{R}$ and yet be so thin that they fail to contain a single interval, no matter how small this interval may be.

But before we can present our argument, a few definitions will prove handy.

Def: Let A be a subset of a metric space (M,d). If $x \in A$ and x is not a limit point of A, then x is called an isolated point of A.

Remark: With this definition, we see that a set is perfect if it is closed and has no isolated points.

Def: A set A is said to be dense in (M,d) (or, as some authors say, everywhere dense) if $\overline{A} = M$. That is, every point of M is either in A or is a limit point of A if A is dense in M.

Def: A set A is said to be nowhere dense in (M,d) if $(\overline{A})^{\circ} = \emptyset$.

Ex. (a) The sets $\mathbb{Q}$ and $\mathbb{R} \setminus \mathbb{Q}$ are dense subsets of $\mathbb{R}$.

(b) Let $A = \{1/n : n \geq 1\}$, then A is nowhere dense in $\mathbb{R}$, because $\overline{A} = A \cup \{0\}$ and $(\overline{A})^{\circ} = \emptyset$.

(c) Let (M,d) be discrete. If $A \subset M$, then A is both a closed and an open subset of M (Why?). Thus $A = \overline{A}$ and $A = A^{\circ}$. Thus, if A is not empty $A = \overline{A} = (\overline{A})^{\circ} \neq \emptyset$. Thus A is not nowhere dense in M. Notice, however, that the statement "not nowhere dense" is not equivalent to the phrase "dense". In fact, the only dense subsets of M is M itself. Every point of a discrete space is an isolated point.

(d) In general, "not nowhere dense" is not the same as dense. In $\mathbb{R}$, the set $A=(0,1)$ is not dense because $\bar{A}=[0,1]\neq\mathbb{R}$. However, A is not nowhere dense as $(\bar{A})^{\circ}=A=(0,1)\neq\emptyset$. The term "nowhere dense" is an unfortunate choice of language that is sadly too common in the literature to avoid. You should think of nowhere dense sets as "thin" sets or sets that are far from having even a single neighborhood.

(e) While $\mathbb{R}$ is everywhere dense in itself, $\mathbb{R}$ is nowhere dense when it is considered as a subset of $\mathbb{R}^2$ (Why?)

Ex. Suppose that into a zombie infested city a special forces unit has been sent to search for survivors. The progress of the operation is monitored in a remote HQ. If the red dots represent the positions of the special forces soldiers and the black dots are positions of detected zombies, what can you say about the outcome of this operation?

Solution: Every point on the screen is either a zombie or a limit point of zombies. Thus, topologically speaking, the set of all zombie points is dense in the set of points.
Notice also that the set of all red points is nowhere dense; All neighborhoods of the red points have been breached, with no chace for their recovery.
Sadly, the operation is a fiasco. All of our brave soldiers will soon join the ranks of the living-dead. But look at the bright side: We have learned a few important concepts.
you should think of nowhere dense sets as extreemely dangerous for prey; nowhere dense sets are the oposites of thick sets.

We are finally ready for the promised argument.
Lemma 1: Let E be a closed subset of (M,d) and $E^{(i)}$ be the set of all isolated points of E. Then $E \setminus E^{(i)}$ is a closed subset of (M,d).
Proof: Let $\{x_n\}_{n=1}^{\infty} \subset E \setminus E^{(i)}$ be a sequence that converges to $x \in M$. To show that $E \setminus E^{(i)}$ is closed, we must prove that $x \in E \setminus E^{(i)}$. Notice, however, that since E is closed $x_n \to x \Rightarrow x \in E$. Furthermore, this means that x is a limit point of E so x is not an isolated point. Hence $x \notin E^{(i)}$. It follows that $x \in E \setminus E^{(i)}$ as desired.

Lemma 2: Let E be a closed subset of $\mathbb{R}$. Then the set of all isolated points of E, $E^{(i)}$, is at most countable.

Proof: Every isolated point of E is contained in an open interval that has no other points of E. In other words, if $x \in E^{(i)}$, $x \in I_x$ such that $I_x \cap E = \{x\}$ and I_x is open.

Thus $E^{(i)} \subset \bigcup_{x: x \in E^{(i)}} I_x$ where $I_x \cap I_y = \emptyset$ if $x \neq y$.

Since each interval contains a rational, this union is countable. This implies that $E^{(i)}$ is also countable as each I_x contains only one point of $E^{(i)}$.

Thm: There exist perfect subsets of $\mathbb{R}$ that contain nearly all of $\mathbb{R}$ and yet fail to have even a single rational number. Such sets must be nowhere dense.

Proof: Let $\epsilon > 0$. List all the rational numbers in a sequence $\{r_n\}_{n=1}^{\infty}$ and put each r_n at the center of an open interval I_n of $\text{length}(I_n) = \frac{\epsilon}{2^n}$. If $G = \bigcup_{n=1}^{\infty} I_n$ then G is open and $\text{Length}(G) \leq \sum_{n=1}^{\infty} \text{length}(I_n) = \sum_{n=1}^{\infty} \frac{\epsilon}{2^n} = \epsilon$. Setting $E = G^c$, we see that E is a closed set whose size must be infinite. That is $\text{length}(E) = \infty$ (Why?).

Thus E must be uncountable. By lemma 1, $E \setminus E^{(i)}$ is also a closed set and by lemma 2, $E \setminus E^{(i)}$ must be uncountable (because we delete at most countably many points). Finally, observe that $P = E \setminus E^{(i)}$ is a perfect subset of $\mathbb{R}$.

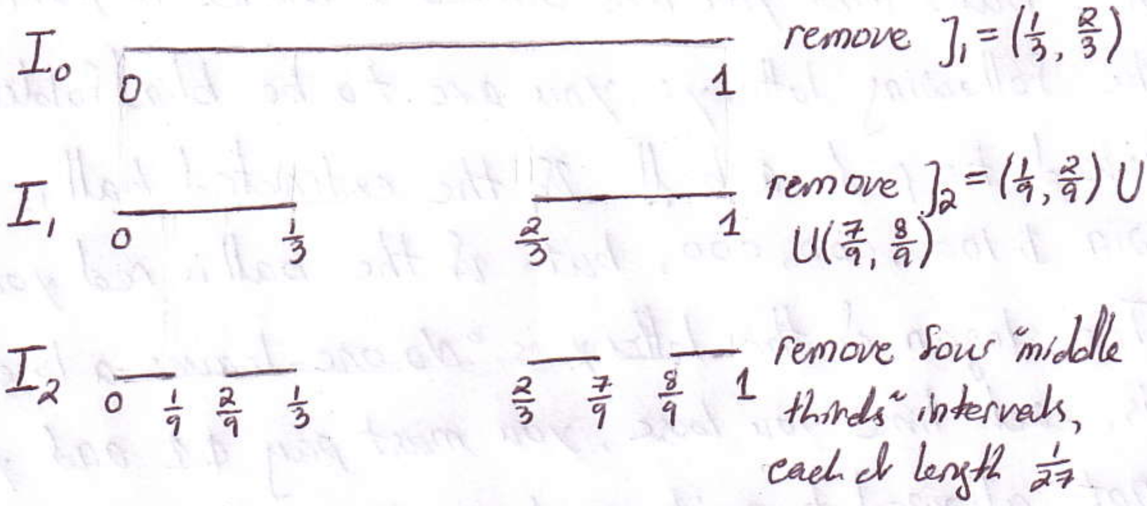
The Cantor Set

Suppose that an urn contains equally many red and blue balls and you are offered a chance to participate in the following lottery: you are to be blindfolded and asked to pick a ball. If the extracted ball is blue, you win $1000,000,000, but if the ball is red you loose. The slogan of this lottery is "No one leaves a loser." That is, each time you loose, you must pay ¢1 and you are not allowed to quit until you have drawn a blue ball and won the prize. So what do you say? Do you want to play?

Suppose that each ball in the urn is marked with a unique number from the interval $[0,1]$. The blue balls may therefore be regarded as a subset of $[0,1]$. This means that the lottery game can be simulated by a procedure that extracts a number from $[0,1]$ at random.

Consider the process of successively removing "middle thirds" from the interval $[0,1]$.

We continue this process inductively. At the nth stage we construct I_n from I_{n-1} by removing 2^{n-1} disjoint, open, "middle thirds" intervals from I_{n-1}, each of length 3^{-n}; we will call this discarded set J_n. Thus, I_n is the union of 2^n closed subintervals of I_{n-1}, and the complement of I_n in $[0,1]$ is $J_1 \cup \ldots \cup J_n$.

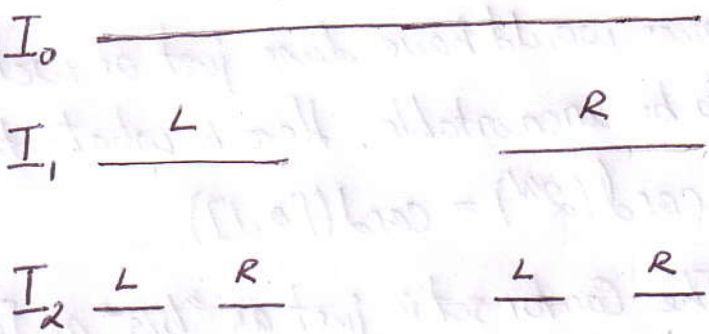


The Cantor set Δ is defined as the set of points that still remain at the end of this process, in other words, the "limit" of the sets I_n. More precisely, $\Delta = \bigcap_{n=1}^{\infty} I_n$. It follows from the nested interval theorem that $\Delta \neq \emptyset$, but notice that Δ is at least countably infinite. The endpoints of each I_n are in Δ:

$$0, 1, \tfrac{1}{3}, \tfrac{2}{3}, \tfrac{1}{9}, \tfrac{2}{9}, \ldots \in \Delta.$$

We will refer to these points as the endpoints of Δ, that is, all of the points in Δ of the form $a/3^n$ for some integers a and n.

As we shall see presently, Δ is actually uncountable! This is more than a little surprizing. Just try to imagine how terribly sparse the next few levels of the "middle thirds" diagram would look on the page. Adding even a few more levels defies the limits of typesetting! For good measure we will give two proofs that Δ is uncountable, the first being somewhat combinatorial. Notice that each subinterval of I_{n-1} results in two subintervals of I_n (after discarding a middle third). We lable these two new intervals L and R (for left and right).



As we progress down through the levels of the diagram toward the Cantor set (somewhere far below), imagine that we "step down" from one level to the next by repeatedly choosing either a step to the left (landing on L-interval in the next level below) or a step to the right (landing on R-interval). At each stage we are only allowed to step down to a subinterval of the interval we are presently on - jumping across "gaps" is not allowed! Thus, each string of choices, $LRLRRLLRLLLR\ldots$, describes a unique "path" from the top level I_0 down to the bottom level Δ. The Cantor set, then, is quite literally the "dust" at the end of the trail. Said another way, each such "path" determines a unique sequence of nested subintervals, one from each level,

whose intersection is a single point of Δ.
Conversely, each point $x \in \Delta$ lies at the end of exactly one such path, because at any given level there is only one possible subinterval of I_n on our diagram, call it $\tilde{I}_n$, that contains x. The resulting sequence of intervals $(\tilde{I}_n)$ is clearly nested. (Why?). Thus, the Cantor set Δ is in one-to-one correspondence with the set of all path, that is, the set of all sequences of Ls and Rs. Of course, any two choices would have done just as well, so we might already know to be uncountable. Here is what this means:

$$\operatorname{card}(\Delta) = \operatorname{card}(2^{\mathbb{N}}) = \operatorname{card}([0,1]).$$

Absolutely amazing! The Cantor set is just as "big" as $[0,1]$ and yet it strains the imagination to picture such a sparse set of points. Before we give our second proof that Δ is uncountable, let's see why Δ is "small" in at least one sense. We will show that Δ has "measure zero"; that is, the "measure" or "total length" of all of the intervals in its complement $[0,1] \setminus \Delta$ is 1. Here's why: By induction, the total length of the 2^{n-1} disjoint intervals comprising J_n (the set we discard at the n^{th} stage) is $2^{n-1}/3^n$, and so the total length of $[0,1] \setminus \Delta$ must be

$$\sum_{n=1}^{\infty} \frac{2^{n-1}}{3^n} = \frac{1}{3} \sum_{n=1}^{\infty} \left(\frac{2}{3}\right)^{n-1} = \frac{1}{3} \, \frac{1}{1-\frac{2}{3}} = 1.$$

We have discarded everything!? And left uncountably many points behind!? How bizarre! This simultaneous "bigness" and "smallness" is precisely what makes the Cantor set so intriguing. If you decide to play the lottery, you would run the risk that all the blue balls are labled with numbers from the Cantor set or something analogous; you have a zero probability of picking out a winning ball if that is the case. And since the prize is too good to be true, it is.

Our second proof that Δ is uncountable is based on an equivalent characterization of Δ in terms of ternary (base 3) decimals. Recall that each x in $[0,1]$ can be written, in possibly more than one way, as: $x = 0.a_1a_2a_3\ldots$ (base 3), where each $a_n = 0, 1,$ or 2. This three-way choice for decimal digits (base 3) corresponds to the three-way splitting of intervals that we saw earlier. To see this, let us consider a few specific examples. For instance, the three cases $a_1 = 0, 1,$ or 2 correspond to the three intervals $[0, 1/3]$, $(1/3, 2/3)$, and $[2/3, 1]$

$$I_1 \quad \underset{0 \qquad\qquad 1/3}{\overline{\quad a_1 = 0 \quad}} \qquad a_1 = 1 \qquad \underset{2/3 \qquad\qquad 1}{\overline{\quad a_1 = 2 \quad}} \qquad (\text{why?})$$

There is some ambiguity at the endpoints:

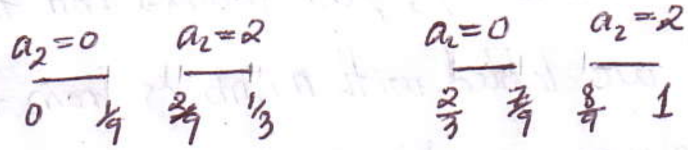
$$1/3 = 0.1 \text{ (base 3)} = 0.0222\ldots \text{ (base 3)}$$
$$2/3 = 0.2 \text{ (base 3)} = 0.1222\ldots \text{ (base 3)}$$
$$1 = 1.0 \text{ (base 3)} = 0.2222\ldots \text{ (base 3)}$$

but each of these ambiguous cases has at least one representation with a_1 in the proper range.

Next, the figure below shows the situation for I_2

$a_1 = 0$ and $\qquad$ $a_1 = 2$ and

I_2



(why?)

Again, some confusion is possible at the endpoints:

$$1/9 = 0.01 \text{ (base 3)} = 0.00222\ldots \text{ (base 3)}$$

$$8/9 = 0.22 \text{ (base 3)} = 0.21222\ldots \text{ (base 3)}$$

We will take these few examples as proof of the following

Thm: $x \in \Delta$ if and only if x can be written as $\sum_{n=1}^{\infty} \frac{a_n}{3^n}$, where each a_n is either 0 or 2.

Thus the Cantor set consists of those points in $[0,1]$ having some base 3 decimal representation that excludes the digit 1. Knowing this we can list all sorts of elements of Δ. For example, $1/4 \in \Delta$ because $1/4 = 0.020202\ldots$ (base 3). The above theorem also leads to another proof that Δ is uncountable; or, rather, it gives us a new way of writing the old proof. The proof used sequences of 0s and 1s, and now we find ourselves with sequences of 0s and 2s; the connection isn't hard to guess.

Corollary: Δ is uncountable; in fact, Δ is equivalent to $[0,1]$.

Proof: By altering our notation we can easily display a correspondence between Δ and $[0,1]$. Each $x \in \Delta$ may be written

$x = \sum_{n=1}^{\infty} \frac{2b_n}{3^n}$, where $b_n = 0$ or 1, and now we define the

Cantor function $f : \Delta \to [0,1]$ by

$$f\left(\sum_{n=1}^{\infty} \frac{2b_n}{3^n}\right) = \sum_{n=1}^{\infty} \frac{b_n}{2^n} \quad (b_n = 0, 1)$$

That is,

$$f(0.a_1 a_2 a_3 \ldots \text{(base 3)}) = 0.\frac{a_1}{2}\frac{a_2}{2}\frac{a_3}{2}\ldots \text{(base 2)} \quad (a_n = 0, 2).$$

Now f is clearly onto, and hence we have a second proof that Δ is uncountable. (Why?) But f isn't one-to-one; here's why:

$$f(1/3) = f(0.0222\ldots \text{(base 3)}) = 0.0111\ldots \text{(base 2)}$$
$$= 0.1 \text{ (base 2)} = f(0.2 \text{ (base 3)}) = f(2/3).$$

The same phenomenon occurs at each pair of endpoints of any discarded "middle third" interval (i.e, a subinterval of J_n):

$$f(1/9) = f(0.00222\ldots \text{(base 3)}) = 0.00111\ldots \text{(base 2)}$$
$$= 0.01 \text{ (base 2)} = f(0.02 \text{ (base 3)}) = f(2/9).$$

It is easy to see that f is increasing; that is, if $x, y \in \Delta$ with $x < y$, then $f(x) \leq f(y)$. We leave it as an exercise to check that $f(x) = f(y)$ if and only if x and y are endpoints of a discarded "middle third" interval. Thus, f is one-to-one except at the endpoints of Δ (a countable set), where it's two-to-one. It follows that Δ is equivalent to $[0,1]$. (How?)