The standard Cantor set $C$, named after mathematician George Cantor, is defined as follows. Start with the interval $C_0 = [0, 1]$. Cut out the central third, which leaves two intervals, $C_1 = [0, \frac{1}{3}] \cup [\frac{2}{3}, 1]$. Now cut out the central third of each of those intervals, which leaves four intervals, $C_2 = [0, \frac{1}{3}] \cup [\frac{2}{3}, \frac{3}{8}] \cup [\frac{5}{8}, \frac{7}{9}] \cup [\frac{8}{9}, 1]$. And keep going forever. The points that are left form the Cantor set $C = \bigcap_k C_k$.

The total length of $C$ is zero. To see this, note that at step $k$, $C_k$ consists of $2^k$ intervals of length $(\frac{1}{3})^k$. So the total length of $C_k$ must be $2^k (\frac{1}{3})^k$. But that shrinks to zero as $k$ increases. The total length of $C$ can be no more than that of any of the sets $C_k$, so it must be zero!

There’s a straightforward way to tell if a number $x$ is in $C$, and it has to do with how $x$ can be expressed in base 3 as a series of digits. Writing a number $x \in [0, 1]$ as a base 3 series of digits means you express $x$ in the form

$$x = \frac{d_1}{3} + \frac{d_2}{3^2} + \frac{d_3}{3^3} + \cdots = \langle 0.d_1 d_2 d_3 \ldots \rangle_3$$

where each $d_j \in \{0, 1, 2\}$. For example,

$$\frac{1}{3} = \langle 0.1 \rangle_3 \quad \frac{2}{3} = \langle 0.2 \rangle_3 \quad \frac{1}{9} = \langle 0.01 \rangle_3 \quad \frac{2}{9} = \langle 0.02 \rangle_3 \quad \frac{5}{9} = \langle 0.12 \rangle_3$$

Here’s how to convert a number $x$ to its base 3 digits. Multiplying $x = \langle 0.d_1 d_2 d_3 \ldots \rangle_3$ by 3 yields $3x = \langle d_1 . d_2 d_3 \ldots \rangle_3$ which means that the integer part of $3x$ is the first digit, $d_1 = \lfloor 3x \rfloor$, and the fractional part is $3x - d_1 = \langle 0.d_2 d_3 d_4 \ldots \rangle_3$. So let $x_2 = 3x - d_1$. Now $d_2 = \lfloor 3x_2 \rfloor$, and we can let $x_3 = 3x_2 - d_2$ and keep going.

As with decimal notation, base 3 representations can repeat. This happens with fractions whose denominator is not a power of 3. Here are some examples. The line over a group of digits indicates that they repeat.

$$\frac{1}{2} = \langle 0.\overline{1} \rangle_3 \quad \frac{4}{5} = \langle 0.2\overline{10} \rangle_3 \quad \frac{1}{10} = \langle 0.002\overline{20} \rangle_3 \quad \frac{1}{7} = \langle 0.0\overline{102120} \rangle_3$$

Recall that in base 10, we can represent a terminating decimal with a series of repeating 9s: $\frac{1}{2} = 0.5 = 0.49 \overline{9}$ The same happens in base 3: $\frac{1}{3} = \langle 0.1 \rangle_3 = \langle 0.0\overline{2} \rangle_3$.

A number is in $C$ if it can be represented in base 3 using only digits 0 and 2. Here’s why: At the first step, all numbers of the form $\langle 0.1 \ldots \rangle_3$ are removed to form $C_1$. At the second step, all remaining numbers of the form $\langle 0.01 \ldots \rangle_3$ or $\langle 0.21 \ldots \rangle_3$ are removed to form $C_2$. And so on. You have to use the $\langle \ldots \overline{2} \rangle_3$ representation for numbers on the right end of an interval of $C_k$. For example, the representation $\frac{1}{3} = \langle 0.0\overline{2} \rangle_3$ means it’s $\frac{1}{3} \in C$. 
Solve these problems:

1. Convert the number $\frac{7}{9}$ to base 3. \( \langle 0.21 \rangle_3 \)

2. Convert the number $\frac{4}{11}$ to base 3. \( \langle 0.10021 \rangle_3 \)

3. Determine whether these numbers are in $C$. Answer “yes” or “no”.
   
   \[
   \begin{array}{cccc}
   \frac{9}{10} & = & \langle 0.2200 \rangle_3 & \text{YES} \\
   \frac{4}{15} & = & \langle 0.201 \rangle_3 & \text{NO} \\
   \frac{1}{25} & = & \langle 0.001 \ldots \rangle_3 & \text{NO} \\
   \frac{12}{13} & = & \langle 0.220 \rangle_3 & \text{YES}
   \end{array}
   \]

There are still a lot of points in $C$. In fact, there are as many points in $C$ as there are in the whole interval $[0, 1]$! Here’s why: Given $x \in C$, consider the digits of $x = \langle 0.d_1d_2d_3 \ldots \rangle_3$, and every $d_j$ is either 0 or 2. Now map that to the number $f(x) = \langle 0.b_1b_2b_3 \ldots \rangle_2$ where

\[
    b_j = \begin{cases} 
    0 & \text{if } d_j = 0 \\
    1 & \text{if } d_j = 2
\end{cases}
\]

and yes, that’s a base 2 series for $f(x)$. Note that $f$ gives a perfect one-to-one correspondence between numbers in $C$ and numbers in $[0, 1]$, so in that sense, there are as many numbers in $C$ as in the whole unit interval, despite the fact that $C$ has a total length of 0.

More problems:

1. Find $f\left( \frac{1}{10} \right)$ and express it as an exact fraction in ordinary decimal notation.

   \[
   f\left( \frac{1}{10} \right) = \frac{1}{5}
   \]

2. The number $t$ is in $C$:

   \[
   t = \langle 0.20200222 \rangle_3
   \]

   Express $t$ and $f(t)$ as exact fractions in ordinary decimal notation.

\[
   t = \frac{2443}{3280} \quad f(t) = \frac{167}{255}
\]