A Mathematical Excursion: From the Three–door Problem to a Cantor-Type Set

Jaume Paradís, Pelegrí Viader, and Lluís Bibiloni

1. INTRODUCTION. We invite you, reader, on a mathematical trip. Our starting point is a well-known problem, the three-door problem (also known as the Monty Hall problem); our endeavors to solve it take us to a beautiful representation system for the real numbers in $(0, 1]$ which, in turn, provides us with a nice enumeration of the positive rationals; as a bonus we can easily prove the irrationality of e. Our trip ends in the dark region of mysterious sets, where we find a simple description of a Cantor-type perfect set contained in $(0, 1]$.

2. STARTING POINT: THE THREE-DOOR PROBLEM. Mathematics has always been enriched by a diversity of games and intellectual curiosities. These have provided an endless supply of problems that have acquired a life of their own, far removed from the recreational aspect of their origins. For example, the first building blocks of probability owe their existence to the analysis of gambling games carried out by Fermat and Pascal in the beginning of the XVIIth century. Undoubtedly Fermat himself was much attracted to mathematics thanks to Bachet's *Problèmes plaisants et délectables* of 1612 **[1]**, which was an introduction to Bachet's most famous book: the Latin translation of Diophantus' *Arithmetica*, in whose margins Fermat wrote the note that made his major theorem famous. Another important instance, E. Lucas' *Récréations Mathématiques* **[15]**, was a source of interesting problems at the beginning of the present century.

Let us start our excursion by setting a simple problem in the form of a seemingly innocent game.

2.1 The Three-Door Problem. In a TV contest, one of three shut doors hides a wonderful prize while the other two open onto a dismal void. The host proposes that the contestant choose one of the three doors. Then, as the contest rules establish, the host opens one of the other two doors wide showing the absence of any prize and offers the contestant the possibility of changing his/her choice. The contestant has to make a decision: to change or not to change.

Our mathematical challenge is to help the contestant make this decision by finding the probability of both possibilities. A (widely accepted) solution to the problem assigns probability $1/3$ to the option not to change and $2/3$ to the option to change. One way to reach this conclusion is the following reasoning:

The probability of choosing the right door in the first place is unquestionably $1/3$. *The probability that one of the other doors hides the prize is then* $2/3$. *If we choose not to change when we are offered the chance, our door still has the same probability of success*, $1/3$, *while now, the other door has a probability* $2/3$ *of hiding the prize.*

There are numerous references to this problem in the literature: see [**22**], [**23**], [**26**], [**6**] or [**2**].

2.2 A Reformulation of the Problem. Now tackle the same problem with a slightly different setting:

> In a TV contest, the host randomly hides a single prize in one of several boxes. The contestant chooses a box and then the host—who knows where the prize is—picks a box different from the one the contestant chose, opens it, and shows the empty contents to the contestant and the audience. The empty box is then discarded and, at this point, the contestant is permitted to choose a new, different, box, or may stick with the old one. In the first case, the contestant chooses a new box and holds it, while all remaining boxes, together with the one rejected by the contestant, are jumbled randomly. The same process continues until two boxes are left: the one the contestant holds and another one. The contestant is then offered the last possibility of change. After that, a mathematician, having followed the whole process attentively, says: "The contestant's probability of winning is $11/42$."
>
> A second mathematician, who has been fast asleep during the whole contest and does not know the initial number of boxes, but is familiar with the rules of the show, wakes up, hears the last utterance, and says: "From what my colleague says I deduce that initially there were 7 boxes and the contestant changed on two occasions, when there were 4 and 3 boxes to choose from."
>
> How did the two mathematicians reach their conclusions?

We call the preceding reformulation and generalization of the three-door problem the *n-box* problem; boxes are more suitable than doors when it comes to jumbling them randomly.

2.3 A Hint. We suggest our reader try to find the probability of each of the 2^{n-2} possible strategies that our contestant can follow, if n is the initial number of boxes in the game. There are 2^{n-2} because any set of choices can be described as a string of 1 and 0 (1 for changing and 0 for sticking) and there are $n-2$ offers of change.

A strategy can be represented by a strictly decreasing sequence of positive integers $\{n, a_k, a_{k-1}, \dots, a_2, a_1\}$ such that $n > a_k > a_{k-1} > \cdots > a_2 > a_1 \geq 1$, where a_i denotes that a change of boxes was made when there were a_i boxes to choose from (notice that $a_k \neq n-1$).

2.4 The Solution to the Problem. If no change whatsoever is made, the probability of winning is obviously $1/n$. If a last-minute change is made (when there is only one box offered besides that initially chosen), the probability of winning is $(n-1)/n$.

If we describe any other strategy by the convention described in the preceding hint, the first change is made when the contestant can choose from a_k boxes. The probability of choosing the right box is the probability of having previously chosen the wrong one times the probability of choosing correctly among a_k boxes, that is:

$$p_k = \left(1 - \frac{1}{n}\right)\frac{1}{a_k}.$$

For the contestant's next change, the same reasoning shows that

$$p_{k-1} = (1-p_k)\frac{1}{a_{k-1}} = \left(1 - \left(1 - \frac{1}{n}\right)\frac{1}{a_k}\right)\frac{1}{a_{k-1}},$$

which can be expressed as

$$p_{k-1} = \frac{1}{a_{k-1}} - \frac{1}{a_{k-1}\cdot a_k} + \frac{1}{a_{k-1}\cdot a_k \cdot n}.$$

Iterating the process, we have for the last change

$$p_1 = (1-p_2)\frac{1}{a_1} = \frac{1}{a_1} - \frac{1}{a_1\cdot a_2} + \cdots + \frac{(-1)^{k-1}}{a_1\cdot a_2 \cdots a_k} + \frac{(-1)^k}{a_1\cdot a_2 \cdots a_k \cdot n}. \quad (2.1)$$

In (2.1) we have $1 \le a_1 < a_2 < \cdots < a_k < n-1$. A strategy is described by a subset of $\{1,2,\ldots,n-2\}$; ($\varnothing$ corresponds to the strategy of making no change at all). For each strategy $\{a_1, a_2, \ldots, a_k\} \subset \{1,2,\ldots,n-2\}$, the probability of winning is given by (2.1).

This accounts for our first mathematician's assertion, as

$$n=7,\ k=2,\ a_2=4,\ a_1=3, \quad \text{and} \quad p_1 = \frac{1}{3} - \frac{1}{3\cdot 4} + \frac{1}{3\cdot 4\cdot 7} = \frac{11}{42}.$$

To help understand our second mathematician's claim, we can play a little with what we have and find a few more probabilities in the case $n = 7$. A patient completion of Table 1 (2^5 entries) would show a most interesting fact: different strategies correspond to different probabilities. If similar tables for other values of n were made, we would notice that in all cases the probabilities obtained were different, not only within one table but also among different tables. This motivates the following result.

Theorem 1. *Any rational number $p/q \in (0,1]$ has a unique representation*

$$\frac{p}{q} = \frac{1}{a_1} - \frac{1}{a_1\cdot a_2} + \cdots + \frac{(-1)^{k-1}}{a_1\cdot a_2\cdots a_k}, \quad (2.2)$$

where a_i are positive integers such that

$$1 \le a_1 < a_2 < \cdots < a_{k-1} < a_k - 1.$$

TABLE 1

Strategy	Probability
$\{1,2,3,4,5\}$	$1 - \frac{1}{2} + \frac{1}{2\cdot 3} - \frac{1}{2\cdot 3\cdot 4} + \frac{1}{2\cdot 3\cdot 4\cdot 5} - \frac{1}{2\cdot 3\cdot 4\cdot 5\cdot 7} = \frac{62}{105}$
$\{1,2,3,4\}$	$1 - \frac{1}{2} + \frac{1}{2\cdot 3} - \frac{1}{2\cdot 3\cdot 4} + \frac{1}{2\cdot 3\cdot 4\cdot 7} = \frac{53}{84}$
$\{1,2,3\}$	$1 - \frac{1}{2} + \frac{1}{2\cdot 3} - \frac{1}{2\cdot 3\cdot 7} = \frac{9}{14}$
$\{2,4,5\}$	$\frac{1}{2} - \frac{1}{2\cdot 4} + \frac{1}{2\cdot 4\cdot 5} - \frac{1}{2\cdot 4\cdot 5\cdot 7} = \frac{111}{280}$

Proof: The interval $(0,1]$ can be expressed as the disjoint union $(0,1] = \bigcup_{n=1}^{\infty}((n+1)^{-1}, n^{-1}]$, so any $\alpha \in (0,1]$ belongs to one of the intervals $((n+1)^{-1}, n^{-1}]$. Consequently,

$$\alpha = \frac{1}{n} - \lambda_1\left(\frac{1}{n} - \frac{1}{n+1}\right) = \frac{1}{n} - \frac{\lambda_1}{n(n+1)},$$

for some $\lambda_1 \in [0,1)$. If we denote $\alpha_1 = \lambda_1/(n+1)$, we have $\alpha = (1-\alpha_1)/n$ and $\alpha_1 \in (0,(n+1)^{-1})$. Applying the same procedure to α_1 we get $\alpha_1 = (1-\alpha_2)/m$ $(m > n)$, and we eventually get an expansion of the form (2.2):

$$\alpha = \frac{1}{n} - \frac{1}{n \cdot m} \cdot \alpha_2 \qquad (m > n).$$

The algorithm that leads to (2.2) can be summarized by iterating the two operations

$$a_i = \left\lfloor \frac{1}{\alpha_{i-1}} \right\rfloor \qquad \text{with} \qquad \alpha_0 = \alpha, \tag{2.3}$$
$$\alpha_i = 1 - \alpha_{i-1} \cdot a_i,$$

where $\lfloor x \rfloor$ denotes the greatest integer less or equal than x.

If α is irrational, all the α_i are irrational and the algorithm never terminates. This proves more than promised in the phrasing of the theorem; it proves the existence of an infinite expansion of the form (2.2) for any irrational in $(0,1]$.

If $\alpha = p/q$ is a rational number in lowest terms, the algorithm becomes a modified Euclidean algorithm. If we divide q by p:

$$q = a_1 \cdot p + r_1 \qquad (0 \le r_1 < p),$$

it is obvious that

$$\left\lfloor \frac{q}{p} \right\rfloor = a_1 \qquad \text{and} \qquad \frac{r_1}{q} = \alpha_1.$$

Next we would perform the division of q by r_1:

$$q = a_2 \cdot r_1 + r_2 \qquad (0 \le r_2 < r_1),$$

and so on. Since the sequence of remainders r_i is strictly decreasing $(p > r_1 > r_2 > \cdots)$, the algorithm eventually terminates with $r_k = 0$. Therefore the expansion (2.2) is finite and the last two divisions are

$$q = a_{k-1} \cdot r_{k-2} + r_{k-1} \qquad (0 \le r_{k-1} < r_{k-2})$$
$$q = a_k \cdot r_{k-1}\cdot .$$

Thus, $a_{k-1} \cdot r_{k-2} = (a_k - 1) \cdot r_{k-1}$, and since $r_{k-1} < r_{k-2}$ we have $a_{k-1} < a_k - 1$.

The uniqueness of the expansions comes from the double inequality

$$\frac{1}{a_1 + 1} < \frac{1}{a_1} - \frac{1}{a_1 \cdot a_2} + \cdots \le \frac{1}{a_1}.$$

The only duplicate expansion is obtained in the finite case, due to the equality

$$\frac{1}{n+1} = \frac{1}{n} - \frac{1}{n(n+1)}.$$

That is the reason for the exclusion of two consecutive integers at the end of the expansion. ■

We denote the expansion (2.2) by

$$\langle a_1, a_2, \dots \rangle.$$

Now, we see how our second mathematician could reconstruct the events of the contest. Starting with the probability 11/42, we compute $42 = 11 \cdot \mathbf{3} + 9 = 9 \cdot \mathbf{4} + 6 = 6 \cdot \mathbf{7}$. Thus

$$\frac{11}{42} = \frac{1}{3} - \frac{1}{3 \cdot 4} + \frac{1}{3 \cdot 4 \cdot 7} = \langle 3, 4, 7 \rangle.$$

The uniqueness of the expansion allows the second mathematician to say:

> *There were* 7 *boxes and the contestant changed on two occasions, when there were* 4 *and* 3 *boxes to choose from.*

3. A REPRESENTATION SYSTEM FOR THE REAL NUMBERS IN (0,1]. We have solved our generalized n-door problem and, at the same time, have discovered a system of representation for the real numbers α in (0, 1]:

$$\alpha = \frac{1}{a_1} - \frac{1}{a_1 a_2} + \frac{1}{a_1 a_2 a_3} - \cdots$$

where $1 \leq a_1 < a_2 < a_3 < \cdots$.

The first mathematicians who paid any attention to these expansions were Lambert (1770) and Lagrange (1798); see **[13]** and **[12]**. Later, Ostrogradsky (†1860), and Sierpiński (1911) were the first to develop a few of their numerical properties; see **[19]** and **[27]**. Pierce (1929) used the model in an algorithm to find algebraic roots of polynomials, **[18]**. Some authors have attached Pierce's name to the expansion that had been previously referred to as "Lambert fractions" or "ascending fractions." In a 1986 presentation, Shallit **[24]** studied the metric theory of the model following the methods used for the non-alternated expansions (Engel's series) developed in 1947 by Borel **[3]** and Lévy **[14]**, and later by Erdős, Rényi, and Szüsz **[4]**, improved by Rényi in 1962 **[20]**. There is also a 1987 paper by A. Knopfmacher and J. Knopfmacher **[10]**, who use the model to construct the real numbers. Some interesting new results related to Pierce expansions can be found in **[25]** and **[11]**.

The infinite Pierce expansion $\langle 1, 2, 3, 4, \dots \rangle$ is the Taylor expansion of $1 - e^x$ for $x = -1$:

$$1 - \frac{1}{e} = \frac{1}{1!} - \frac{1}{2!} + \frac{1}{3!} - \cdots = \langle 1, 2, 3, \dots \rangle.$$

Incidentally, this proves the irrationality of e. Other examples are not so obvious:

$$\langle 1, 3, 5, 7, \dots \rangle = \frac{1}{\sqrt{e}} \sum_{n=0}^{\infty} \frac{1}{2^n \cdot n! \cdot (2n+1)}.$$

As a system of representation, Pierce expansions are not bad: Truncating the expansion of $\alpha = \langle a_1, a_2, \dots, a_n \rangle$ at level n, provides quite a good approximation to α:

$$|\alpha - \langle a_1, a_2, \dots, a_n \rangle| < \frac{1}{a_1 \cdots a_n \cdot a_{n+1}}, \tag{3.1}$$

which, in the worst case ($a_i = i$, $i = 1, 2, \dots$), is of the order $1/(n+1)!$.

4. A NEW ENUMERATION FOR THE POSITIVE RATIONALS. The most famous enumeration of the positive rationals is the diagonal ordering

$$\frac{1}{1}, \frac{2}{1}, \frac{1}{2}, \frac{3}{1}, \frac{2}{2}, \frac{1}{3}, \frac{4}{1}, \frac{3}{2}, \frac{2}{3}, \frac{1}{4}, \frac{5}{1}, \frac{4}{2}, \frac{3}{3}, \frac{2}{4}, \frac{1}{5}, \dots$$

All fractions appear in this scheme, repeated infinitely many times; p/q appears in position $(1/2)(p+q-1)(p+q-2)+q$. After suppressing repetitions, to determine the position of the irreducible ones is, as Prof. Hardy says in [7, p. 1], *more complicated*: the computational complexity of the diagonal ordering algorithm, if one suppresses all repetitions, is exponential. This problem is intimately related to the representation of rational numbers [**17**].

The basic idea is to use the binary representation of a positive integer n as a string of 0 and 1. Some of these strings can be considered strategies in our generalized n-box problem. We use the corresponding Pierce expansion to assign a rational to our n.

To any strictly increasing finite sequence of positive integers $\{a_1, a_2, \dots, a_k\}$ with $1 \le a_1 < a_2 < \cdots < a_k$, we associate the positive integer $n = 2^{a_1-1} + 2^{a_2-1} + \cdots + 2^{a_k-1}$, or, what amounts to the same, the number n that in the binary system is written, from right to left, as 1 in positions $a_1, a_2, \dots, a_k$ and 0 elsewhere. For example,

$$\{1,3,5,8\} \longrightarrow 2^0 + 2^2 + 2^4 + 2^7 = 10010101.$$

Now, to any rational number $p/q \in (0,1]$ we associate its Pierce expansion $\langle a_1, a_2, \dots, a_k \rangle$, which may be regarded as the strictly increasing finite sequence of positive integers $\{a_1, a_2, \dots, a_k\}$, where $a_k > 1 + a_{k-1}$. Its corresponding positive integer has the binary form $10\dots$, with a 0 in the next-to-last position as we go from right to left. To any rational number $q/p > 1$, we associate the Pierce expansion corresponding to its inverse $p/q = \langle a_1, a_2, \dots, a_k \rangle$ and we then consider the strictly increasing finite sequence of positive integers $\{a_1, a_2, \dots, a_{k-1}, a_k - 1, a_k\}$. Its corresponding positive integer has the binary form $11\dots$, with a 1 in the next-to-last position as we go from right to left.

Conversely, to any positive integer n written in the binary system as

$$2^{a_1} + 2^{a_2} + \cdots + 2^{a_k} \quad \text{with} \quad 0 \le a_1 < a_2 < \dots < a_k,$$

we assign the rational number

$$\langle 1+a_1, 1+a_2, \dots, 1+a_k \rangle \in (0,1] \quad \text{if } a_k > 1 + a_{k-1},$$

$$\frac{1}{\langle 1+a_1, 1+a_2, \dots, 1+a_{k-2}, 1+a_k \rangle} > 1 \quad \text{if } a_k = 1 + a_{k-1}.$$

The uniqueness of the Pierce expansion of any rational number in $(0,1]$ ensures the bijectivity of the map just defined between the positive integers and the positive rationals.

An example may help us understand the map we have just defined. Let us find what rational occupies place 10^{12} in our enumeration. First we write 10^{12} in binary:

$$10^{12} = 2^{39} + 2^{38} + 2^{37} + 2^{35} + 2^{31} + 2^{30} + 2^{28} + 2^{26} + 2^{23} + 2^{21} + 2^{18} + 2^{16} + 2^{12},$$

which corresponds to the fraction (in this case a rational greater than 1):

$$\frac{1}{\langle 13,17,19,22,24,27,29,31,32,36,38,40\rangle} = \frac{94232197736202240}{6843703050416119}.$$

The algorithm to find the fraction occupying a given place n has a polynomial computational complexity. The inverse algorithm would also have polynomial complexity, assuming the correctness of a conjecture formulated by Erdős and Shallit in [5] concerning the upper bound of the length of the finite Pierce expansion of p/q:

$$\text{length of the Pierce expansion of } \frac{p}{q} = O\big((\log q)^2\big).$$

5. A CLOSER LOOK AT PIERCE EXPANSIONS. Let us contemplate what we have accomplished and examine Pierce expansions more closely. For each real number α in (0, 1], define its i-th projection $\omega_i(\alpha)$ to be the map that assigns to α its i-th partial quotient: if $\alpha = \langle a_1, a_2, a_3, \dots \rangle$, then $\omega_i(\alpha) = a_i$.

A *cylinder* of order k is the set of numbers such that the first k partial quotients are fixed:

$$C(a_1, a_2, \dots, a_k) = \{\alpha \in (0,1] : \omega_1(\alpha) = a_1, \omega_2(\alpha) = a_2, \dots, \omega_k(\alpha) = a_k\}.$$

A cylinder of any order is an interval of length

$$|C(a_1, a_2, \dots, a_k)| = \frac{1}{a_1 \cdot a_2 \cdots a_k \cdot (1 + a_k)}.$$

Moreover, a cylinder of order k is the disjoint union of all the cylinders of order $k+1$ contained in it:

$$C(a_1, a_2, \dots, a_k) = \bigcup_{j=1+a_k}^{\infty} C(a_1, a_2, \dots, a_k, j).$$

We can also consider generalized cylinders, in which the fixed partial quotients are not the first k; they are not intervals, though they are still unions of intervals. The simplest is

$$H[\omega_k(\alpha) = n] = \{\alpha \in (0,1] : \omega_k(\alpha) = n\},$$

which is a union of cylinders:

$$H[\omega_k(\alpha) = n] = \bigcup_{1 \le a_1 < a_2 < \cdots < a_{k-1} \le n-1} C(a_1, a_2, \dots, a_{k-1}, n).$$

Consequently, its Lebesgue measure is:

$$\lambda(H[\omega_k(\alpha) = n]) = \sum_{1 \le a_1 < a_2 < \cdots < a_{k-1} \le n-1} \frac{1}{a_1 a_2 \cdots a_{k-1} n(n+1)}$$

$$= \frac{1}{n(n+1)} \sum_{1 \le a_1 < a_2 < \cdots < a_{k-1} \le n-1} \frac{1}{a_1 a_2 \cdots a_{k-1}}. \quad (5.1)$$

One way to evaluate the last sum in (5.1) is to multiply inside by $(n-1)!$ and divide outside by the same quantity:

$$\sum_{1 \le a_1 < \cdots < a_{k-1} \le n-1} \frac{1}{a_1 a_2 \cdots a_{k-1}} = \frac{1}{(n-1)!} \sum_{1 \le a_1 < \cdots < a_{n-k} \le n-1} a_1 a_2 \cdots a_{n-k}. \tag{5.2}$$

The right-hand sum in (5.2) can be viewed as the coefficient of x^k in the polynomial $x(x+1)(x+2)\cdots(x+n-1)$, which is a *Stirling number of the second kind*. The properties of Stirling numbers (both of the first and the second kind) can be found in [**8**, pp. 243–253], whose notation we follow: $\begin{bmatrix} n \\ k \end{bmatrix}$. Thus,

$$x(x+1)(x+2)\cdots(x+n-1) = \begin{bmatrix} n \\ 1 \end{bmatrix} x + \begin{bmatrix} n \\ 2 \end{bmatrix} x^2 + \cdots + \begin{bmatrix} n \\ n \end{bmatrix} x^n. \tag{5.3}$$

Using this notation, the sums in (5.2) are equal to $\begin{bmatrix} n \\ k \end{bmatrix}/(n-1)!$. Finally, we have

$$\lambda(H[\omega_k(\alpha) = n]) = \frac{1}{n(n+1)} \cdot \frac{\begin{bmatrix} n \\ k \end{bmatrix}}{(n-1)!} = \frac{\begin{bmatrix} n \\ k \end{bmatrix}}{(n+1)!}. \tag{5.4}$$

With (5.4) and a very simple property of Stirling numbers:

$$\begin{bmatrix} n \\ 1 \end{bmatrix} + \begin{bmatrix} n \\ 2 \end{bmatrix} + \cdots + \begin{bmatrix} n \\ n \end{bmatrix} = n!$$

which follows from (5.3) by considering $x = 1$, it is easy to prove the following:

Theorem 2. *The set of real numbers in* $(0,1]$ *whose Pierce expansion contains the integer n has Lebesgue measure* $1/(n+1)$.

Theorem 2 has an immediate corollary:

Theorem 3. *The set of real numbers in* $(0,1]$ *whose Pierce expansion does not contain the integer n has Lebesgue measure* $n/(n+1)$.

It is not difficult to generalize Theorem 3:

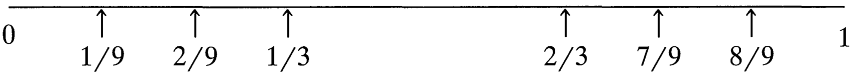
Ttheorem 4. *The set of real numbers in* $(0,1]$ *whose Pierce expansion does not contain the distinct integers m and n has Lebesgue measure* $nm/(n+1)(m+1)$.

All these results can be found in [**24**].

6. A CANTOR-TYPE PERFECT SET. A set in **R** that is closed and has no isolated points is said to be *perfect*. Such a set coincides with the set of its limit-points (its *derived* set). The easiest example of a perfect set in **R** is a closed interval, but there are perfect sets that not only are not intervals, they do not even contain any interval. A classic example of this behavior is Cantor's ternary set

$$\mathscr{C} = \left\{ \alpha \in [0,1] : \alpha = \sum_{i=1}^{\infty} \frac{a_i}{3^i}, \text{ with } a_i \in \{0,2\} \right\},$$

that is to say, the set of all real numbers in [0, 1] that can be written in the ternary system without the digit 1. Geometrically, Cantor's set can be described by iterating indefinitely the following procedure: From $[0,1]$ we suppress the central open interval $(1/3, 2/3)$ and from the remaining two intervals, we suppress the corresponding central open intervals $(1/9, 2/9)$ and $(7/9, 8/9)$, and so on. The points that remain after the suppression of all these open intervals constitute Cantor's set.



Cantor's set is a perfect set; it is uncountable (a consequence of being perfect) and has measure zero (its complement in $[0,1]$ is the union of countably many disjoint open intervals with total length 1). The interior of $\mathscr{C}$ is obviously empty. See **[21]** for details.

Now, let us consider the set C of real numbers in $(0,1]$ whose Pierce expansion contains no odd integers. According to Theorem 4, the Lebesgue measure of C is

$$\lambda(C) = \prod_{n=1}^{\infty}\left(1 - \frac{1}{2n}\right) = 0.$$

The set C is uncountable since we can establish a one-to-one correspondence between its elements and $(0,1]$:

$$\langle a_1, a_2, \ldots, a_n, \ldots \rangle \leftrightarrow \langle \frac{a_1}{2}, \frac{a_2}{2}, \ldots, \frac{a_n}{2}, \ldots \rangle.$$

It is also easy to prove that C, like Cantor's set, is perfect and its interior is empty.

7. A CANTOR-TYPE PERFECT SET OF TRANSCENDENTAL NUMBERS. In 1851 J. Liouville established a very important result that permitted him to exhibit, for the first time in mathematics, a transcendental real number (a real number that is not the root of any polynomial equation with rational coefficients). A real number α is said to be algebraic of degree n if there is a polynomial of degree n (but not lower) with rational coefficients that has α as a root.

Liouville's Theorem. [**16**, pp. 87–93] *If α is algebraic of degree n, $(n > 1)$, there exists a constant M (depending on α) such that*

$$\left|\alpha - \frac{a}{b}\right| \geq \frac{M}{b^n}.$$

for all rational numbers a/b.

Consider the following Pierce expansion

$$l_p = \langle p^{2!}, p^{3!-2!}, \ldots, p^{n!-(n-1)!}, \ldots \rangle = \frac{1}{p^{2!}} - \frac{1}{p^{3!}} + \frac{1}{p^{4!}} - \cdots + \frac{(-1)^n}{p^{n!}} + \cdots,$$

where p is any positive integer.

The number l_p is transcendental because (3.1) tells us that

$$\left|l_p - \left(\frac{1}{p^{2!}} - \frac{1}{p^{3!}} + \cdots + \frac{(-1)^k}{p^{k!}}\right)\right| = \left|l_p - \frac{a}{p^{k!}}\right| < \frac{1}{p^{(k+1)!}} < \frac{1}{(p^{k!})^k}, \quad (7.1)$$

which would contradict Liouville's theorem if l_p were algebraic of degree k.

We may now consider the set L_p of all real numbers in $(0,1]$ whose Pierce expansion contains only integers extracted from the Pierce expansion of l_p. It is at once seen that L_p has measure zero, is uncountable, and (since all its elements satisfy inequalities like (7.1)) consists entirely of transcendental numbers.

8. LOOKING BACK. We have started with an interesting and controversial problem, the Monty Hall dilemma (which has a totally probabilistic set-up), and have reached some very peculiar subsets of $[0, 1]$: uncountable closed sets with an empty interior, without isolated points and of measure zero—the same structure as Cantor's ternary set, with the added feature of being formed exclusively by transcendental numbers. The connection between such apparently distant concepts is a beautiful system for real number representation, Pierce expansions, which *exactly* describe the probability of each one of the possible strategies that can be followed by the contestant in a generalization of the Monty Hall problem: the n-box problem. We have also encountered a nice (and new) enumeration of the positive rationals that is based on both the strategies in the n-box problem and Pierce expansions.

Undoubtedly we have overlooked many unexplored places, but we hope you have enjoyed the few we have been lucky enough to find.

REFERENCES

1. C. Bachet de Méziriac, *Problèmes plaisants et délectables*, 1st ed., 1612. 3rd ed., Gauthier–Villars, Paris, 1874.
2. E. Barbeau, Fallacies, flaws, and flimflam: The problem of the car and the goats, *College Math. J.* **24** (1993) 149–154.
3. E. Borel, Sur les développements unitaires normaux, *Comptes Rendus* **225** (1947) 74–79.
4. P. Erdős, A. Rényi, and P. Süsz, On Engel's and Sylvester's series, *Annales Univ. Sci. Budapest, Sectio Math.* **1** (1958) 7–32.
5. P. Erdős and J. O. Shallit, New bounds on the Length of Finite Pierce and Engel Series, *Séminaire de Théorie des nombres de Bordeaux* **3** (1991) 43–53.
6. L. Gillman, The car and the goats, *Amer. Math. Monthly* 99 (1992) 3–7.
7. G. H. Hardy, *A Course of Pure Mathematics*, Cambridge Univ. Press, 8th ed., London, 1941.
8. R. L. Graham, D. E. Knuth, and O. Patashnik, *Concrete Mathematics*, Addison-Wesley, 6th ed., Reading, Mass., 1990.
9. A. Ya. Khintchine, *Continued Fractions*. Translated by Peter Wynn. P. Noordhoff, Ltd., Groningen, 1963.
10. A. Knopfmacher and J. Knopfmacher, Two constructions of the real numbers via alternating series, *Internat. J. Math. and Math. Sci.* **12** (1989) 603–613.
11. A. Knopfmacher and M. E. Mays, Pierce expansions of ratios of Fibonacci and Lucas numbers and polynomials, *Fibonacci Quarterly* **33** (1995) 153–163.
12. J. L. Lagrange, Essai d'analyse numérique sur la transformation des fractions, *Journal de l'École Polytechnique*, Vème cahier, t. II, prairial, an VI. Also in *Oeuvres* ed. J. A. Serret, Tome VII, Section quatrième, 291–313.
13. J. H. Lambert, *Verwandlung der Brüche*: *Beiträge zum Gebrauche der Mathematik und deren Anwendung*, Part 2. Berlin 1770.
14. P. Lévy, Remarques sur un théoreme de M. Émile Borel, *Comptes Rendus* **225** (1947) 918–919.
15. E. Lucas, *Récréations mathématiques*, Gauthier–Villars, Paris, 1891–1894. Reprinted by A. Blanchard, Paris, 1960.
16. I. Niven, *Irrational Numbers*, Carus Math. Monographs, **11**. Wiley, NY, 1956.
17. J. Paradís, Ll. Bibiloni, and P. Viader, On Actually Computable Bijections between **N** and $\mathbb{Q}^+$, *Order* **13** (1996) 369–377.
18. T. A. Pierce, On an Algorithm and its Use in approximating Roots of Algebraic Equations, *Amer. Math. Monthly* **36** (1929) 532–525.
19. E. Ya. Remez, On series with alternating signs, which may be related to two algorithms of M. V. Ostrogradski for the approximation of irrationals numbers, *Matem. Nauk.*, (N.S.) 6,5 **45** (1951) 33–42.
20. A. Rényi, A new Approach to the theory of Engel's series, *Annales Univ. Sci. Budapest, Sectio Math.* **5** (1962) 25–32.

21. W. Rudin, *Principles of Mathematical Analysis*, McGraw–Hill, New York, 3d ed. 1976.
22. S. Selvin, A problem in probability, *Amer. Statist.* **29** (1975) 67.
23. S. Selvin, On the Monty Hall problem, *Amer. Statist.* **29** (1975) 134.
24. J. O. Shallit, Metric Theory of Pierce Expansions, *Fibonacci Quarterly* **24** (1986) 22–40.
25. J. O. Shallit, Pierce Expansions and Rules for the Determination of Leap Years, *Fibonacci Quarterly* **32** (1994) 416–423.
26. M. Shaughnessy, T. Dick, Monty's dilemma: should you stick or switch?, *Math. Teacher* **84** (1993) 252–256.
27. W. Sierpiński, Sur quelques algorithmes pour développer les nombres réels en séries, (1911). In *Oeuvres choisies*, Tome I, Warszawa, 1974, 236–254.

JAUME PARADÍS graduated in Mathematics at the University of Barcelona and received his Ph.D. in Mathematics at the Universitat Politècnica de Catalunya, under the supervision of Professor J. Egozcue. He now has tenure at the Universitat Pompeu Fabra, Barcelona. His early research includes several books and papers on the History of Algebra, mainly on Viète's and Fermat's contributions. His present research is focused on Pierce expansions and other systems of real number representation and on singular functions defined with their help.
Applied Mathematics, Univ. Pompeu Fabra, Ramon Trias Fargas 25–27, 08005 Barcelona, Spain
jaume.paradis@econ.upf.es

PELEGRÍ VIADER is Associate Professor of Mathematics at the Universitat Pompeu Fabra, Barcelona. He received his Ph.D. in mathematics at the Universitat Politècnica de Catalunya. His main research interests are Pierce expansions and other algorithms of real number representation, singular functions and History of Mathematics.
Applied Mathematics, Univ. Pompeu Fabra, Ramon Trias Fargas 25–27, 08005 Barcelona, Spain
pelegri.viader@econ.upf.es

LLUÍS BIBILONI studied Mathematics at Universitat de Barcelona. He became involved in teacher training, in the Departament de Didàctica de les Matemàtiques i de les Ciències Experimentals, at the Universitat Autònoma de Barcelona where he is professor of Mathematics Education. He received his Ph.D. in mathematics at the Universitat Politècnica de Catalunya. His interests include the Philosophy and History of Science and of Mathematics, and has been deeply influenced by the work of George Pólya on problem solving whose ideas is trying to pursue at a practical teaching level and at research level in the field of Mathematics Education. He is also interested in computer–aided teaching. His favorite mathematical subjects are Number Theory, Analysis, Mathematical Logic, Nonstandard Analysis, Mathematical Physics, and Probability Theory.
Facultat de Ciències de l'Educació, Univ. Autònoma Barcelona, 08193, Bellaterra, Barcelona, Spain.
l.bibiloni@uab.es