The Real Numbers and Cantorian Set Theory

Cantor[1] *has created a Paradise for us…* **David Hilbert**[2]

> *In 1964 he*[3] *made the rare decision to serve on a public commission, responsible for choosing mathematics textbooks for California's grade schools… This was the era of the so-called new mathematics in children's education: the much debated effort to modernize the teaching of mathematics by introducing such high-level concepts as set theory and non-decimal number systems… Feynman did not take the side of the modernizers. Instead he poked a blade into the new-math bubble. He argued to his fellow commissioners that sets, as presented in reformers textbooks, were an example of the most insidious pedantry: new definitions for the sake of introducing words without introducing ideas. A proposed primer instructed first-graders: "Find out if the set of lollipops is equal in number to the set of girls."* ***Feynman described this as a disease*** [my emphasis, JC]. *It removed clarity without adding any precision to the normal sentence: "find out if there are just enough lollipops for the girls." Specialized language should wait until it is needed, he said, and the peculiar language of set theory never is needed.* ***He found that the new textbooks did not reach the areas in which set theory does begin to contribute content beyond the definitions: the understanding of different degrees of infinity*** [my emphasis, JC], *for example.* GENIUS (Richard Feynman and modern physics, biography of RF) **James Gleick**, 1992

> *I first had occasion to hear of the theory of sets at a lecture conducted by I. M. Gelfand*[4] *for Moscow school children. He was then just beginning his teaching career… During the course of two hours he told us about what seemed to us to be completely improbable things: that there are just as many natural numbers as there are rational numbers, and that there are just as many points in an interval as there are in a square.* **Stories about Sets** (1968, translated from Russian), **N. Ya. Vilenkin**

These notes (a *skeletal* summary of the real work that took place in lengthy discussions/arguments in my office) are intended primarily for my second year BA students, and should be read only **after** those in-class discussions/arguments. Separate notes, on rational and irrational numbers and irreducible polynomials, and related decimal expansions and continued fraction expansions, are available in **Maple** worksheets.

[1] Georg Cantor (1845-1918), revolutionary mathematician, creator of the theory of transfinite sets. http://www-groups.dcs.st-andrews.ac.uk/~history/Mathematicians/Cantor.html

[2] David Hilbert (1862-1943), one of the greatest mathematicians of all time. http://www-groups.dcs.st-andrews.ac.uk/~history/Mathematicians/Hilbert.html

[3] Richard Feynman (1918-1988), the renowned US physicist, winner of the Nobel Prize in Physics. http://www-groups.dcs.st-andrews.ac.uk/~history/Mathematicians/Feynman.html

[4] Israil Gelfand (1913-), legendary Ukrainian mathematician, now working (in his 90th year) in the US. http://www-groups.dcs.st-andrews.ac.uk/~history/Mathematicians/Gelfand.html

In my own case I learned Cantorian set theory through reading (initially in one of Felix Klein[5]'s books, later Cantor's Dover book on Transfinite Numbers. In 1971 I came upon Vilenkin's little gem of a book, *Stories about Sets*) and reflection. My current[6] students learned about it, sitting in my office, being asked questions (initially about Hilbert's Hotel (Section 1)), and arguing over points…

Section 1 **Hilbert's Hotel** (rooms, buses, rooms occupancy cards, associated binary expansions)

Section 2 **Cantor enters:** his classic proofs concerning the natural and rational numbers, and the natural and real numbers. Some concepts and terminology: **finite**, **infinite**, **countable** (**denumerable**), **uncountable**. Some standard results concerning countable and infinite sets.

Section 3 **Algebraic** and **transcendental** numbers. Cantor's proof that transcendental numbers *existed*. Euler's conjecture concerning some possible transcendental numbers, and Hilbert's 7th problem.

Section 4 "*I see it, but I do not believe it…*" Cantor (in a letter to Dedekind[7]) What did Cantor see that he could not, at first, believe?

Section 5 Cantor's 'nested interval' proof.

Section 6 Higher levels of infinitude: the **power set** of a set (another of Cantor's great creations).

Section 1

Introduction. We begin our journey into 'Cantor's Paradise' by thinking about an *imaginary* hotel (Hilbert's Hotel), one with an infinite[8] number of rooms; actually one with a *very particular kind*[9] of infinite number of rooms: one whose rooms may be *listed* in a rather special way. Let us record an actual working:

Definition. A ***Hilbert hotel***, *H*, is one with 'rooms' $R_1, R_2, R_3, \ldots$, such that for **every** 'room' R_n (for $n \in \mathbf{N}$) of *H*, *H* has **another** room R_{n+1} (it is understood that *H* has no other rooms but these R_i, and thus the rooms of *H* correspond in a 1-1 way to the standard natural numbers: every room corresponds to some natural number, and conversely). (*H*, it must be understood, has no 'last' room, like an ordinary hotel.)

[5] http://www-groups.dcs.st-andrews.ac.uk/~history/Mathematicians/Klein.html

[6] In the 1980's I gave a series of lectures for adults at University College Dublin, as part of its Extra-Mural programme, and one of the topics I included was the *Infinite*. I ran them (the lectures), not as lectures, but as discussions, rather than presenting material in textbook style: definition(s), theorem, proof, more definitions, more theorems, more proofs… I believe that a journey to Hilbert's Hotel is a good starting point for getting into real set theory, not the school-fed set theory junk. Yes, I'm most definitely in the Feynman camp.

[7] http://www-groups.dcs.st-andrews.ac.uk/~history/Mathematicians/Dedekind.html

[8] Whatever that means… In the early stages we proceed with abandon, **as if** we **know** what 'infinite' means.

[9] Which will later be termed '*countable*' or '*denumerable*'.

Let us make the understanding (which is not essential, but which enables us to get started on our Cantorian journey) that the hotel has a **room occupancy rule** dictating that only 1 person (at most) may occupy a room. An associated:

Definition. By a ***Hilbert vector*** we mean one with an infinite number of co-ordinates[10], in which each entry is 0 or 1; thus a Hilbert vector, *v*, is one of the form

$$(\varepsilon_1, \varepsilon_2, \varepsilon_3, \ldots, \varepsilon_n, \ldots)$$

where $\varepsilon_i = 0 \text{ or } 1$ for every $i \in \mathbf{N}$. (We may think of a Hilbert vector as being a sort of room occupancy card in the *obvious sense*: should every room be occupied, then the corresponding Hilbert vector would have $\varepsilon_i = 1$ for all *i*; whereas if every other room from the first onwards was occupied, while every other room from the second onwards was empty, then the related Hilbert vector would be (1, 0, 1, 0, 1, 0, …)).

Note. An apparently obvious remark is that we consider two Hilbert vectors to be the same ('equal') if, and only if, their co-ordinates agree in *every* place; put another way, two Hilbert vectors are considered to be different (not 'equal') if, and only if, their co-ordinates differ in *at least one* place. So, for example:

- v_1 = (1, 1, 0, **1**, 0, 1, 0, …) and v'_1 = (1, 1, 0, **0**, 0, 1, 1, …) are **not** equal (differ) since they don't agree in (at least) the 4th coordinate. (The fact that they happen also to differ – in the example shown – in the 7th coordinate, is of no consequence.)

Important observation, with a fundamental, later import[11]. With each Hilbert vector we may associate a real number – with value somewhere between 0 and 1 – as follows: let *v* be a Hilbert vector with $v = (\varepsilon_1, \varepsilon_2, \varepsilon_3, \ldots, \varepsilon_n, \ldots)$, then, from *v*, we may create a real number $r \in [0, 1]$ by setting

$$r = \frac{\varepsilon_1}{2^1} + \frac{\varepsilon_2}{2^2} + \frac{\varepsilon_3}{2^3} + \ldots + \frac{\varepsilon_n}{2^n} + \ldots$$

In fact, $(0\,\mathbf{.}\varepsilon_1\,\varepsilon_2\,\varepsilon_3 \ldots \varepsilon_n \ldots)_2$ is the binary expansion of the real number *r*. (Just as every real number *r* in [0, 1] has a decimal expansion, every such number has a binary expansion. There's no mystery about this.)

But now, an important point:

- whereas two **different** real numbers will have **different** binary expansions (if they didn't, the two numbers would be the same)

- two **different** (looking) binary expansions **may** give rise to the **same** real number. Thus, for example, $\frac{1}{2}$ has two different looking binary expansions: $(0\mathbf{.1}0000\ldots\textit{ad infinitum})_2$, and $(0\mathbf{.01111}\ldots\text{ad infinitum})_2$.

[10] Again, of a very particular type (the *same* type as already considered): one co-ordinate per natural number.

[11] In connection with Cantor's famous 'diagonal decimal' proof…

Also, for example, $\frac{7}{8}$ has two different looking binary expansions: $(0.\mathbf{1110000}\ldots\textit{ad infinitum})_2$, and $(0.\mathbf{1101111}\ldots\text{ad infinitum})_2$.

This should not shock one; rather it should be viewed merely as *being in the nature of things* (of course it all rather takes for granted that one knows about – and ideally understands – subtleties in connection with the non-trivial concept of the sum of an infinite series). From our point of view, this will have later import in connection with this sort of phenomena: a real number **may** have two different (looking) decimal expansions:

$$\tfrac{1}{20} = (0.0500000\ldots)_{10} = (0.0499999\ldots)_{10}$$

Of course it is only for *some* real numbers that this phenomenon occurs; in fact it **only** occurs for those rational numbers (in reduced form) whose denominators are divisible only by 2 or 5 (related, obviously, to the '10' base): 2, 4, 5, 8, 10, 16, 20, etc. Rational numbers with such denominators have decimal expansions which are eventually all 0's or, equivalently, all 9's.

Definition. By a ***Hilbert bus***, *B*, we understand one with 'passengers' $P_1, P_2, P_3, \ldots$, (a passenger P_n for every $n \in \mathbf{N}$).

Our journey began with a single (thought-provoking!)

Question. Can an 'extra' passenger be accommodated in a filled Hilbert hotel (with the **room occupancy rule** in force) without someone already accommodated having to give up a room? Or perhaps it is impossible?

Answer. For a novice there was a surprising (at first) answer… (The great Danish physicist Neils Bohr[12] once said of anyone claiming to understand Quantum Mechanics that they had failed to understand the problem; here one might say of someone who is not surprised that they, too, have failed to understand the point…) One then reflected that the 'surprise' should be seen merely as an initial reflection of the nature of the 'infinite', which is radically different from the 'finite'.

Fundamental points

1. The passengers from an infinite[13] number of Hilbert buses may be accommodated in a **single** Hilbert hotel (later that may be *viewed* as a way of *understanding* that the rational numbers and the natural numbers have the same cardinality)

2. The **entire** collection/set of **all** Hilbert vectors **cannot** be allocated in the rooms of a **single** Hilbert hotel in a 1-1 way (later that may be viewed as a way of understanding that the real numbers and the natural numbers do **not** have the same cardinality[14])

[12] http://www-groups.dcs.st-andrews.ac.uk/~history/Mathematicians/Bohr_Niels.html

[13] 'Infinite' of a *very particular kind* (countable/denumerable).

[14] Strictly, that the set of real numbers in the interval [0, 1] and the natural numbers do not have the same cardinality. However, since there is a 1-1 correspondence between the points on **any** two line

A[15] proof of 1. The passengers[16] from the Hilbert buses are as follows:

In bus 1: $P_{1,1}, P_{1,2}, P_{1,3}, \dots, P_{1,n}, \dots$
In bus 2: $P_{2,1}, P_{2,2}, P_{2,3}, \dots, P_{2,n}, \dots$
In bus 3: $P_{3,1}, P_{3,2}, P_{3,3}, \dots, P_{3,n}, \dots$
… …
In bus *m*: $P_{m,1}, P_{m,2}, P_{m,3}, \dots, P_{m,n}, \dots$
… …

(The key idea is to consider passengers by their location along the '**short diagonals'**: those passengers with **fixed** suffix sums (2, 3, 4, 5, …) Consider the 'suffix sums', the values of the '$m+n$', where *m* gives the number of the bus, and *n* gives the number of the passenger within that bus; then, for each natural number (from 2 onwards) there are only a finite number of *m*'s and *n*'s which sum to that number:

- there is only one passenger with suffix sum 2, namely $P_{1,1}$
- there are two passengers with suffix sum 3, namely $P_{1,2}$ and $P_{2,1}$
- there are three passengers with suffix sum 4, namely $P_{1,3}$, $P_{2,2}$, and $P_{3,1}$

  … … and, in general:
- there are $(n-1)$ passengers with suffix sum *n*: $P_{1,n-1}, P_{2,n-2}, \dots, P_{n-1,1}$

Thus the entire 'doubly-infinite' collection/set of all passengers may be accommodated in a single Hilbert hotel by placing

- $P_{1,1}$ in the first room
- $P_{1,2}$ and $P_{2,1}$ in the next two rooms
- $P_{1,3}$, $P_{2,2}$, and $P_{3,1}$ in the next three rooms

  … … and, in general (for all $n \in \mathbf{N}, n \geq 2)$:
- $P_{1,n-1}, P_{2,n-2}, \dots, P_{n-1,1}$ in the next $(n-1)$ rooms

Comment. That result – expressed as it is – may be easily translated into the **first surprise** of Cantorian set theory: there is a 1-1 correspondence between the set of all rational numbers[17], and the set of all natural numbers.

Proof of 2. Suppose the entire collection/set of Hilbert vectors can be distributed in a 1-1 way in the rooms of a single Hilbert hotel (we now proceed to use the famous

segments (be one a finite line, and the other a finite, or semi-infinite, or doubly infinite line segment (i.e. the 'reals'), then the 'strictly' doesn't carry any weight.

[15] There are many quite different ways of proving this result, and I will give another one later.

[16] And so each 'passenger' may be *thought of* as being related to a rational number in the obvious way: passenger $P_{m,n}$ is related to the rational number $\frac{m}{n}$.

[17] As formulated here it would only set up the 1-1 correspondence for rationals (not in reduced form) with positive numerators and denominators, but it is a trivial exercise to extend it to all rationals (it's merely two Hilbert buses, and one extra passenger. Do you see why?)

‘diagonal’ argument that Cantor introduced into Mathematics; the reason for the nomenclature ‘diagonal’ will be immediately clear)

In room 1 is the Hilbert vector $v_1 = (\boldsymbol{\varepsilon}_{\mathbf{1,1}}, \varepsilon_{1,2}, \varepsilon_{1,3}, \ldots, \varepsilon_{1,n}, \ldots)$
In room 2 is the Hilbert vector $v_2 = (\varepsilon_{2,1}, \boldsymbol{\varepsilon}_{\mathbf{2,2}}, \varepsilon_{2,3}, \ldots, \varepsilon_{2,n}, \ldots)$
In room 3 is the Hilbert vector $v_3 = (\varepsilon_{3,1}, \varepsilon_{3,2}, \boldsymbol{\varepsilon}_{\mathbf{3,3}}, \ldots, \varepsilon_{3,n}, \ldots)$
… …
In room n is the Hilbert vector $v_n = (\varepsilon_{n,1}, \varepsilon_{n,2}, \varepsilon_{n,3}, \ldots, \boldsymbol{\varepsilon}_{\boldsymbol{n,n}}, \ldots)$
… …

Define the Hilbert vector $\mathbf{v}$ by[18] $\mathbf{v} = (\boldsymbol{\varepsilon}_{\mathbf{1}}, \boldsymbol{\varepsilon}_{\mathbf{2}}, \boldsymbol{\varepsilon}_{\mathbf{3}}, \ldots, \boldsymbol{\varepsilon}_{\boldsymbol{n}}, \ldots)$, where $\boldsymbol{\varepsilon}_{\boldsymbol{n}} \neq \boldsymbol{\varepsilon}_{\boldsymbol{n,\,n}}$, for all $n \in \mathbf{N}$; then that $\mathbf{v}$ is **not** in **any** room. **Why**? It’s obvious: $\mathbf{v}$ cannot be in any room since its n^{th} coordinate differs (by definition/construction) from the n^{th} coordinate of the Hilbert vector supposedly allocated to the n^{th} room of the Hilbert hotel.

Section 2

Cantor enters

Introduction. Our main interest in this course is to understand some of the revolutionary ideas introduced into Mathematics by the truly great genius mathematician, Georg Cantor. Cantor opened up great vistas of mathematical thought with his (gradual) introduction of his theory of transfinite numbers, or ‘set theory’.

Getting started. ‘Set theory’ begins with *counting*: how many of this are there, how many of that are there? Are there the same number of this as that? Are there more of this than that? Indeed, what *exactly* do we mean by ‘*same’* and ‘*more*’? In our everyday lives these are terms that most people take so much for granted, they would hardly believe that anyone would be **troubled** by their meanings:

- If there are 98 pupils in school A, and 89 pupils in school B, then there are **more** pupils in school A than in school B (alternatively: the **number** of pupils in school A is **greater** than the number of pupils in school B, or there are **fewer** pupils in school B than in school A, or the **number** of pupils in school B is **less** than the number of pupils in school A)

- If there are 98 pupils in school C, then there are the **same number** of pupils in school A as in school C (alternatively: the number of pupils in school C is the **same** as the number of pupils in school A, or the number of pupils in school C is the **same** as the number of pupils in school A

It could be said that one only begins to be troubled when one considers infinite collections/sets of things/elements. Of course there is a history to all of this, going back at least to Aristotle, and later Galileo and others.

[18] This definition entails the classic Cantor diagonal construction, one of Cantor’s great creations.

How many primes are there between (say) 1 and 10? Of course there are four (the 'set[19]' $S_1 = \{2, 3, 5, 7\}$ whose 'elements' are 2, 3, 5, and 7). And how many squares are there between (say) 1 and 20? Of course there are four (the 'set' $S_2 = \{1, 4, 9, 16\}$ whose 'elements' are 1, 4, 9, and 16). Comparing 'finite' collections *seems* fairly straightforward, but what about comparing two infinite collections/sets ('infinite' in the sense that both contain and infinite number of things/elements). Indeed, what exactly is meant by an 'infinite' number of things/elements[20].

Some (to be modified) definitions and intuitive results. At one time the meanings of 'equal/more/less in number' were so taken for granted that precise meanings might not have been recorded. And the consequences of those definitions, too, were so taken for granted. For later purposes it's quite important to record them, and see those consequences.

Definition 1. Let A and B be two collections/sets[21] of things/elements, then A and B are said to have the ***same*** numbers of things/elements (or to have the ***same cardinality***) if there is a 1-1 correspondence between their elements. **Notation**. $|A| = |B|$.

Examples 1.

- Let A be {p, q, r} and B = {2, 6, 23}, then A and B have the same cardinality since there is the 1-1 correspondence between the elements of A and those of B: p $\leftrightarrow$ 2, q $\leftrightarrow$ 6, r $\leftrightarrow$ 23.

- Let[22] A be {2, 4, 6, 8, 10, **...** } and B = {3, 6, 19, 12, 15, **...**}, then A and B have the same cardinality since there is the 1-1 correspondence the elements of A and those of B: $2 \leftrightarrow 3$, $4 \leftrightarrow 6$, $6 \leftrightarrow 29$, $8 \leftrightarrow 12$, $10 \leftrightarrow 15$, **...** , (in general:) $2n \leftrightarrow 3n$ (for all $n \in \mathbf{N}$).

Definition 2. Let A and B be two collections/sets[23] of things/elements, then A is said to have a ***greater*** numbers of things/elements (or to have ***greater cardinality***) if there is a 1-1 correspondence between the elements of B and a **proper** subset of A.

Convention. If A has a greater number of elements that B, we may also say that B has *fewer* elements than A (or that the cardinality of B is *less* than the cardinality of A).

Notations. $|A| > |B|$ (alternatively $|B| < |A|$).

[19] I hope I can be forgiven for straightaway using 'set theory' language before we have really entered the world of Set Theory. It would be a little artificial of me to not use the language (since almost every school child has encountered it), and instead use words like – say – 'thing(s)' (for element(s)') or 'collection(s)' (for 'set(s)')

[20] So far we have used terms like 'finite' (number of) and 'infinite' (number of) as though we knew what they meant. Of course it is patently nonsensical to say something *like*: a set is 'finite' if it contains a 'finite' number of elements (one would hardly say that a 'good' person is a person who does 'good' things…)

[21] 'Finite' or 'infinite'.

[22] It is intended that A consists of all multiples of 2, while B consists of all multiples of 3.

[23] 'Finite' or 'infinite'.

Examples 2.

- Let A be {p, q, r, s} and $B = \{2, 6, 23\}$, then A has a **greater** number of elements than B, since there is the 1-1 correspondence between the elements of B and a **proper** subset of A: $\mathrm{p} \leftrightarrow 2$, $\mathrm{q} \leftrightarrow 6$, $\mathrm{r} \leftrightarrow 23$.

- Let[24] A be {2, 4, 6, 8, 10, **...** } and B = {4, 8, 12, 16, 20, **...**}, then A – **according to the above Definition 2** – has a **greater** number of elements than B since there is the 1-1 correspondence the elements of B and a **proper** subset of A: $4 \leftrightarrow 4$, $8 \leftrightarrow 8$, $12 \leftrightarrow 12$, $16 \leftrightarrow 16$, $20 \leftrightarrow 20$, **...** , (in general:) $2(2n) \leftrightarrow 4n$ (for all $n \in \mathbf{N}$). In fact, B is a proper subset of A.

The historical rejection. A complete study of the pre-history of Cantor's work is not possible in a short course, but it is not an over-simplification to record that the **essential source of disquiet** before Cantor's time (and his supportive fellow-colleague Dirichlet played a great part in clarification some of Cantor's omissions) was that the above definitions **appeared only to make sense in the case of finite collections/sets**.

Meaning? Consider, briefly, some of the (perhaps unwritten) truths one tends **to take for granted** with respect to the meanings of 'same/equal', 'more/greater', 'less/fewer'; I will write telegraphically (and do remember this merely records discussions held at greater length in active classroom exchanges):

1. If $|A| = |B|$ then $|B| = |A.|$ (that would hardly raise an eyebrow: if A has the same cardinality as B, then B has the same cardinality as A).

2. If $|A| = |B|$ and $|B| = |C|$ then $|A| = |C|$ (that, too, would hardly surprise: if A has the same cardinality as B, and B has the same cardinality as C, then A has the same cardinality as C).

3. If $|A| > |B|$ and $|B| > |C|$ then $|A| > |C|$ (that, too, would hardly surprise: if A has greater cardinality than B, and B has greater cardinality than C, then A greater cardinality than C). That is known as the 'Trichotomy Law'.

4. If $|A| > |B|$ then (surely?)$|A| \neq |B|$ and (surely?) $|A| \not< |B|$ (that would **normally** be **laughed at**: if A has greater cardinality than B, then (surely!) A could **not** have the **same** cardinality as B, **nor** could A have a **smaller** cardinality as B).

#1, #2 and #3 cause no difficulties (with respect to the definitions), #4 is **deeply troubling with respect to 'infinite' sets**, which is one of the principal reasons for the earlier (historical) 'rejection of infinite' because of the 'paradoxes' associated with Galileo. For example

- there **appears** to be the **same** number of **natural numbers** as **even natural numbers**: $1 \leftrightarrow 2$, $2 \leftrightarrow 4$, $3 \leftrightarrow 6$, **...** , (in general:) $n \leftrightarrow 2n$ (for $n \geq 1$).

[24] Here it is intended that A consists of all multiples of 2, while B consists of all multiples of 4.

- at the same time there **appears** to be **more** natural numbers than even ones: **don't** pair '1' from **N** with **any** even natural number, and then establish the following 1-1 correspondence (between the **rest** of **N** and all the even natural numbers): $2 \leftrightarrow 2$, $3 \leftrightarrow 4$, $4 \leftrightarrow 6$, **...** , (in general:) $n \leftrightarrow 2n-2$ (for $n \geq 2$).
- and at the same time there **appears** to be **more** even natural numbers than natural ones: **don't** pair '2' from the even natural numbers with any natural number, and then establish the following 1-1 correspondence (between the **rest** of the even natural numbers and all of **N**): $1 \leftrightarrow 4$, $2 \leftrightarrow 6$, $3 \leftrightarrow 8$, **...** , (in general:) $n \leftrightarrow 2n+2$ (for $n \geq 1$).

In short. Adopting 2**N** as the (reasonable) notation for the even natural numbers, we have the **apparently** nonsensical: $|\mathbf{N}| = |2\mathbf{N}|, |\mathbf{N}| > |2\mathbf{N}|, \text{and} \, |\mathbf{N}| < |2\mathbf{N}|.$ One should try to empathise with earlier thinkers; it was entirely reasonable that they rejected the infinite: **it appeared truly bizarre**. The 'Trichotomy Law' that generations took for granted for 'finite' collections *appeared* to be violated for 'infinite' collections.

We looked at some variations of the Galilean observation (you can make up your own), to mention just two:

- If A and B are two circles, then there is a 1-1 correspondence between their points.
- If L and l are any two line segments, the same is true, even if one is of finite length and the other is not.

Thinking aloud; what could have happened, but didn't. It *could* have happened that the rational numbers, **Q**, and the natural numbers, **N**, could not have been paired in a 1-1 way, **since** the former are '**dense**' while the latter are '**discrete**'. **But** in December 1873 cantor proved that the elements of **N** and **Q** may be put in 1-1 correspondence. That was a **surprise**! Next (as a natural question to have pondered), the real numbers **R** are the rationals and the 'irrationals' – **I** (say) – together, and the latter are also 'dense'. **If** the elements of **I** could be put in 1-1 correspondence with those of **N**, then **R** (two Hilbert bus loads, as it were. You see that?) and **N** would have the same number of elements!! So, the question that Cantor asked himself in December 1973 (he wrote to Dedekind about it) was:

Can (the elements of) **N** and **R** can a be paired in a 1-1 way?

A historic moment. It could be said that true 'set theory' began with Cantor's remarkable discovery (some three weeks after he began to think about it) of:

The first great Cantor theorem. There is no 1-1 correspondence between the natural and the real numbers (so giving the **first ever example** of two infinite collections/sets that are **genuinely different** in 'cardinality'). In fact (since there is no difference – in terms of cardinality – between the number of points on any two line segments) there is no 1-1 correspondence between the natural numbers and the real numbers on the unit line segment.

Cantor's 'diagonal decimal' proof (with its slight, rectifiable flaw[25]). Suppose that $r_1, r_2, r_3, r_4, \ldots$ is a **complete** enumeration of **all** real numbers in [0, 1].

Now, every real number in $[0, 1]$ is expressible as a *decimal*, and we have

$$r_1 = (0\boldsymbol{.}\mathbf{r_{1,1}}r_{1,2}r_{1,3}\ \cdots\ r_{1,n}\ \cdots)_{10}$$
$$r_2 = (0\boldsymbol{.}r_{2,1}\mathbf{r_{2,2}}r_{2,3}\ \cdots\ r_{2,n}\ \cdots)_{10}$$
$$r_3 = (0\boldsymbol{.}r_{3,1}r_{3,2}\mathbf{r_{3,3}}\ \cdots\ r_{3,n}\ \cdots)_{10}$$

… … In general

$$r_n = (0\boldsymbol{.}r_{n,1}r_{n,2}r_{n,3}\ \cdots\ \mathbf{r_{n,n}}\ \cdots)_{10}$$

where $0 \le r_{i,\,j} \le 9$ for all $i, j \in \mathbf{N}$.

Now, define the real number $\boldsymbol{r}$ by[26] $\boldsymbol{r} = (0\boldsymbol{.r_1\, r_2\, r_3} \ldots \boldsymbol{r_n} \ldots)_{10}$ where the $\{\boldsymbol{r_n}\}$ are chosen so that $\boldsymbol{r_1} \ne \boldsymbol{r_{1,1}}$, $\boldsymbol{r_2} \ne \boldsymbol{r_{2,2}}$, $\boldsymbol{r_3} \ne \boldsymbol{r_{3,3}}, \ldots$, $\boldsymbol{r_n} \ne \boldsymbol{r_{n,n}}$, for all $n \in \mathbf{N}$.

[**Aside.** Recall the earlier

> Define the Hilbert vector $\boldsymbol{v}$ by $\boldsymbol{v} = (\boldsymbol{\varepsilon_1}, \boldsymbol{\varepsilon_2}, \boldsymbol{\varepsilon_3}, \ldots, \boldsymbol{\varepsilon_n}, \ldots)$, where $\boldsymbol{\varepsilon_n} \ne \boldsymbol{\varepsilon_{n,\,n}}$, for all $n \in \mathbf{N}$; then that $\boldsymbol{v}$ is **not** in **any** room. **Why**? It's obvious: $\boldsymbol{v}$ cannot be in any room since its n^{th} coordinate differs (by definition/construction) from the n^{th} coordinate of the Hilbert vector supposedly allocated to the n^{th} room of the Hilbert hotel.

for we would now like to say:]

Then $\boldsymbol{r}$ **cannot** be paired with *any* natural number since its n^{th} decimal digit differs (by definition/construction) from the n^{th} decimal digit of the real number supposedly paired with the n^{th} natural number. Thus there is no 1-1 correspondence between the real numbers in [0, 1] and the natural numbers. [End of flawed proof.]

And what *is* the flaw? It's quite simple (and easily rectified), and is entirely to do with the fact that *some* real numbers have two *different looking* decimal expansions.

Suppose, e.g., that there had been a 1-1 correspondence commencing like this:

[25] It's not uncommon for some low-level texts to present the proof with the flaw (not much of a 'proof'!) without comment. Of course it depends on the *intended* readership. Even Kac and Ulam (two outstanding Polish mathematicians) present the flawed proof in their (popular) *Mathematics and Logic*. Of course they probably made a judgement to protect – as it were – their readers from too many technicalities. In a book, where space is limited, that is acceptable, but it would not be right – in my view – to do so in an *active* teaching situation. Perhaps let it ride at first exposure (to see if anyone objects), and then draw attention to the *rectifiable* flaw.

[26] It is, of course, the 'diagonal' decimal construction.

$$r_1 = (0.\mathbf{1}0000000\ \ldots)_{10}$$
$$r_2 = (0.7\mathbf{8}652109\ldots)_{10}$$
$$r_3 = (0.02\mathbf{5}87602\ldots)_{10}$$
$$r_4 = (0.371\mathbf{3}77661\ldots)_{10}$$
$$\ldots\ \ldots \text{ etc,}$$

the *suggestion* being that the decimal value of r_1 continues all 0's after that initial '1', and that **none** of the diagonal digits is a '9'. Suppose, then, that in forming '$\boldsymbol{r}$' one choose $\boldsymbol{r_1} = \mathbf{0}$ (so $\boldsymbol{r_1} \neq \boldsymbol{r_{1,1}}$) and $\boldsymbol{r_n} = \mathbf{0}$ for **all** $\boldsymbol{n} \in \mathbf{N}$ ($\boldsymbol{n} \geq 2$) (and thus $\boldsymbol{r_n} \neq \boldsymbol{r_{n,n}}$).
But then the created '$\mathbf{r}$' has value $(0.\mathbf{099999999}\ldots)_{10}$ which is, of course, the **same** number as $r_1 = (0.\mathbf{10000000}\ \ldots)_{10}$. One would then not have established the *claimed* impossibility of a 1-1 correspondence.

All's well though, since a rectification can be made. **Instead** of the earlier

> Now, define the real number $\boldsymbol{r}$ by $\boldsymbol{r} = (0.\boldsymbol{r_1}\,\boldsymbol{r_2}\,\boldsymbol{r_3}\ \ldots\ \boldsymbol{r_n}\ \ldots)_{10}$ where the $\{\boldsymbol{r_n}\}$ are chosen so that $\boldsymbol{r_1} \neq \boldsymbol{r_{1,1}}, \boldsymbol{r_2} \neq \boldsymbol{r_{2,2}}, \boldsymbol{r_3} \neq \boldsymbol{r_{3,3}}, \ldots$, $\boldsymbol{r_n} \neq \boldsymbol{r_{n,n}}$, for all $n \in \mathbf{N}$.

proceed as follows

> define the real number $\boldsymbol{r}$ by $\boldsymbol{r} = (0.\boldsymbol{r_1}\,\boldsymbol{r_2}\,\boldsymbol{r_3}\ \ldots\ \boldsymbol{r_n}\ \ldots)_{10}$ where the $\{\mathbf{r_n}\}$ are chosen so that **none** of them are **0** or **9** and $\boldsymbol{r_1} \neq \boldsymbol{r_{1,1}}$, $\boldsymbol{r_2} \neq \boldsymbol{r_{2,2}}, \boldsymbol{r_3} \neq \boldsymbol{r_{3,3}}, \ldots$, $\boldsymbol{r_n} \neq \boldsymbol{r_{n,n}}$, for all $n \in \mathbf{N}$.

Now '$\boldsymbol{r}$'is **none** of the r_1 , r_2 , r_3 , r_4 , ... , and so there is no such enumeration of the real numbers in [0, 1]. (In short, using terminology about to be introduced, the set of real numbers in [0, 1] is **not countable** (or is '**non-denumerable**', or is '**uncountable**').

Comment. You should see the connection with the room allocation cards in the Hilbert hotel: the cards are automatically different if they **merely differ** in a **single** coordinate, but the real numbers that may be associated with the cards are not necessarily different. One could alter the room occupancy rule – which in out earlier telling allowed at most 1 person per room – and instead allow up to 9 people per room. The new Hilbert vectors would be $(\varepsilon_1, \varepsilon_2, \varepsilon_3, \ldots, \varepsilon_n, \ldots)$, with $0 \leq \varepsilon_i \leq 9$ for every $i \in \mathbf{N}$. The occupancy cards (1, 0, 0, 0, 0, 0, *ad infinitum*) and (0, 9, 9, 9, 9, 9, *ad infinitum*) would be different, but the two real numbers associated with them are the same, since both equal $\frac{1}{10}$.

Anyone who is familiar with the mathematics of 'continued fractions' (like you, my 2nd year students) has no difficulty in reframing Cantor's diagonal argument to prove that there is no 1-1 correspondence between the natural numbers and (e.g.) the set of all irrational numbers[27] in [0, 1].

Proof. Suppose $i_1, i_2, i_3, i_4, \ldots$ is a **complete** enumeration of **all** irrational numbers in [0, 1]. Now, every irrational number in $[0, 1]$ is expressible as a **unique** *continued fraction*, and we have

$$i_1 = [0, \boldsymbol{a_{1,1}}, a_{1,2}, a_{1,3}, \ \ldots \ , a_{1,n} \cdots]$$
$$i_2 = [0, a_{2,1}, \boldsymbol{a_{2,2}}, a_{2,3}, \ \ldots \ , a_{2,n} \cdots]$$
$$i_3 = [0, a_{3,1}, a_{3,2}, \boldsymbol{a_{3,3}}, \ \ldots \ , a_{3,n} \cdots]$$

… … In general

$$i_n = [0, a_{n,1}, a_{n,2}, a_{n,3}, \ \ldots \ , \boldsymbol{a_{n,n}} \cdots]$$

where $a_{j,k} \in \mathbf{N}$ for all $j, k \in \mathbf{N}$.

Now, define the irrational number ***i*** by $\boldsymbol{i} = [0, \boldsymbol{a_1}, \boldsymbol{a_2}, \boldsymbol{a_3}, \ \ldots \ , \boldsymbol{a_n} \cdots]$ where the $\{\boldsymbol{a_n}\}$ are chosen so that $\boldsymbol{a_1} \neq \boldsymbol{a_{1,1}}, \boldsymbol{a_2} \neq \boldsymbol{a_{2,2}}, \boldsymbol{a_3} \neq \boldsymbol{a_{3,3}}, \ldots, \boldsymbol{a_n} \neq \boldsymbol{a_{n,n}}$, for all $n \in \mathbf{N}$. Then ***i*** is **not** any one of the $\{i_n\}$. **Why?** It's obvious: ***i*** cannot be any of the $\{i_n\}$ because its n^{th} partial quotient differs (by definition/construction) from the n^{th} partial quotient of the irrational number supposedly paired with the n^{th} natural number. (I trust you recognise that I've simply transferred – with appropriate changes – the argument establishing the impossibility of allocating the Hilbert hotel room allocation cards in the rooms of a single Hilbert hotel) [End of proof]

After all that excitement, we come back to earth to record some standard definitions and elementary results.

Definition 1. A set *S* is said to be ***finite*** if there is no 1-1 correspondence between the elements of *S* and those of $S \cup \{a\}$, '*a*' some element adjoined to *S*.

Comment. I really only record this definition for the sake of it. At one time there was a temptation to say that a set is 'finite' (in the sense – as above – that it has only a 'finite' number of elements) if it has elements $\{a_1, a_2, \ldots, a_n\}$ for some $n \in \mathbf{N}$. However it was (rightly) felt that definition was unsatisfactory, because it is somehow saying that a set is 'finite' if it is, well, 'finite' (what is that '*n*' supposed to be? A 'finite' natural number?). Even Cantor himself struggled (like a poet might over the spelling of a word) to give a satisfactory definition, and it was, in fact, Dedekind who provided the first satisfactory one.

So, now you know!!

[27] One needs to know that every irrational number α has a *unique* expansion of the form:

$$[a_0, a_1, a_2, \ldots, a_n, \ldots \textit{ ad infinitum}]$$

where a_0 is the 'integral part' of α, and the other a_i are natural numbers, the so-called 'partial quotients' of α.

Definition 2. A set S is said to be ***infinite***[28] if there **is** a 1-1 correspondence between the elements of S and those of $S \cup \{a\}$, 'a' some element adjoined to S.

Definition 3. A set S is said to be ***countable*** (another term that is used is *denumerable*) if its elements can be put in 1-1 correspondence with **N**, the (infinite!) set of natural numbers[29].

Definition 2a. A set is said to be ***infinite***[30] if it contains a countable subset.
Comment. So a set is 'infinite' if it has elements $a_1, a_2, a_3, \ldots$, one for **every** natural number. (Tongue-in-cheek: *ad infinitum*!)

Definition 4. A set is said to be ***uncountable*** if it is **infinite** but **not countable**.
Comment. So, the set of real numbers (all of them, or any non-zero interval of them) is uncountable. The reals provided the first historical example of such a set.

Some elementary standard results.

Simple result 1. Let A be any countable set, then the set obtained from A by adjoining any ('extra') finite number of elements is also countable (i.e. let $A' = A \cup \{e_1, \ldots, e_r\}$, then A' is also countable).

Comment. You may think of this as being about accommodating an extra finite number of passengers turning up at an already filled Hilbert hotel.

Proof. Since A is countable, then its elements are some $a_1, a_2, a_3, \ldots a_n, \ldots$. Now, define a function f on $A \cup \{e_1, \ldots, e_r\}$ as follows:

- $f(e_1) = e_1,\ f(e_2) = e_2,\ \ldots,\ f(e_r) = e_r,$
- $f(a_1) = a_{r+1},\ f(a_2) = a_{r+2},\ \ldots,\ f(a_n) = a_{r+n},\ \ldots$ (all $n \in \mathbf{N}$)

In short, $e_1, e_2, \ldots, e_r, a_1, a_2, a_3, \ldots a_n, \ldots$, is a new countable listing of the elements of $A \cup \{e_1, \ldots, e_r\}$.

Simple result 2. Let A be any infinite set[31], then the set obtained from it by adjoining any finite number of 'extra' elements has the same number of elements as A itself (i.e. there is a 1-1 correspondence between the elements of A and $A \cup \{e_1, \ldots, e_r\}$, where $\{e_1, \ldots, e_r\}$ are the 'extra' elements).

[28] This is one possible definition; there's another shortly.
[29] Of course that definition rather takes for granted that one (somehow) knows what that set '**N**' **is**. If one wanted to, then one could study exactly that question: what **is** '**N**'? Such a study would involve considering (say) the Peano axioms. A web reference for Peano: http://www-groups.dcs.st-andrews.ac.uk/~history/Mathematicians/Peano.html
[30] It might be correct to say that this is most people's innate understanding of what an infinite set **is**.
[31] Not necessarily countable.

Proof. Since A is infinite then it contains a countable subset A' (say). Let $\{e_1, \ldots, e_r\}$ be the 'extra' elements (not in A), and let $A' = \{a_1, a_2, a_3, \ldots a_n, \ldots\}$ be the countable subset of A. Now, define a function f on $A \cup \{e_1, \ldots, e_r\}$ as follows

- If $a \in A, a \notin A'$, then let $f(a) = a$, otherwise
- Let $f(e_1) = e_1,\ f(e_2) = e_2,\ \ldots,\ f(e_r) = e_r,$
  $f(a_1) = a_{r+1},\ f(a_2) = a_{r+2},\ \ldots,\ f(a_n) = a_{r+n},\ \ldots$ (all $n \in \mathbf{N}$)

Simple result 3. 'A countable union of countable sets is countable' (an old reliable; of course it's merely the accommodating of an infinite[32] number of Hilbert buses in a single Hilbert hotel).

Proof. (Aside. Of course one could say that we've already 'proved' this result, or rather say that it would just be a matter of translating the 'passengers' proof into this new language. So what I am going to do here – merely to show a variation – is to give another different, standard proof. This one will use a particular way of factoring natural numbers:)

A simple result from Number Theory. Every natural number n may be expressed in the form $2^a(2b+1)$ for some non-negative integers a and b; furthermore, that representation is unique.

Comment. The first part is straightforward:

- if n is odd, then $n = 2^0(2b+1)$, for some non-negative integer b,
- if n is even, then let 2^a be the largest power of 2 that divides n, and so $n = 2^a n',\ n' \in \mathbf{N}$, where n' must be odd (otherwise $n' = 2n'',\ n'' \in \mathbf{N}$, which would give $n = 2^a n' = 2^a(2n'') = 2^{a+1}n''$, which would mean that n was divisible by 2^{a+1}, conflicting with 2^a being the largest power of 2 dividing n. So, that n' must be odd.)

Some numerical examples.

$$50 = 2^1 \times 25 = 2^{\mathbf{1}}.(2 \times \mathbf{12} + 1)$$
$$53 = 2^0 \times 53 = 2^{\mathbf{0}}.(2 \times \mathbf{26} + 1)$$
$$32 = 2^5 \times 1 = 2^{\mathbf{5}}.(2 \times \mathbf{0} + 1)$$

Proof of 'uniqueness'. Suppose $n = 2^{a_1}(2b_1+1) = 2^{a_2}(2b_2+1)$, for some non-negative integers a_1, b_1, a_2, b_2. Then $a_1 = a_2$, for if not, and, say $a_1 < a_2$, then $(2b_1+1) = 2^{a_2-a_1}(2b_2+1)$, is divisible by 2, whereas $(2b_1+1)$ is odd. Thus $a_1 \not< a_2$, and similarly $a_2 \not< a_1$, and thus $a_1 = a_2$. Hence $2b_1+1 = 2b_2+1$, giving $b_1 = b_2$. Thus the representation $n = 2^a(2b+1)$ is unique.

[32] *Countable* infinite!

Return to proof of countable union result. Let $S_1, S_2, \ldots, S_n, \ldots$ be a countable set of sets, each of which is countable. Then their elements may be listed:

(In S_1 :) $s_{1,1}, s_{1,2}, s_{1,3}, \ldots$

(In S_2 :) $s_{2,1}, s_{2,2}, s_{2,3}, \ldots$

(In S_3 :) $s_{3,1}, s_{3,2}, s_{3,3}, \ldots$

etc

Now, let $n \in \mathbf{N}$, and let $n\,(\text{uniquely}) = 2^{a-1}(2b-1)$ for $a, b \in \mathbf{N}$, and define the function f by $f(n) = s_{a,\,b}$, then that sets up a 1-1 correspondence between the elements of **N** and those of $S_1 \cup S_2 \cup S_3 \ldots \cup S_n \ldots$.

Section 3

Algebraic and transcendental numbers[33]

Introduction to a new way of looking at numbers. As you know, an early classification of (real) numbers distinguished two types: rational and irrational. Simple examples of the latter are $\sqrt{2}, \sqrt[3]{7}, \frac{\sqrt{7}+3}{4}, \frac{3}{11}\sqrt[4]{21} - \frac{2}{31}, \log_{10} 2, \ldots$, less simple[34] are examples like $\pi, \pi^2, \ldots$, extremely difficult ones are $e^{\pi}, 2^{\sqrt{2}}, \ldots$, and – for example – some of unknown status are ones like $(e+\pi), e\pi, \frac{e}{\pi}, \ldots$.

A way of looking at numbers (introduced by the great Euler[35]) was to regard them as being solutions of equations of *a very particular type*. For example, a rational number may be regarded as being a solution of the special equation

$$ax + b = 0 \qquad \ldots \text{(i)}$$

where a and b are **integers** (with $a \neq 0$). Thus, for example,

- $\frac{3}{2}$ is a (in fact, the) solution of the equation $2x - 3 = 0$.

However, a simple number like (e.g.) $\sqrt{2}$, is not a solution of an equation of type (i) above (it is, of course, a solution of the completely trivial equation $0x + 0 = 0$, as is indeed *every* number). But, with a slight change, we see that there is some equation – not much different from (i) (in fact: $ax^2 + bx + c = 0$, a, b and c are **integers** (with $a \neq 0$) – for which $\sqrt{2}$ is a solution, namely the one with $a = 1, b = 0, c = -2$.

All this is leading to:

[33] See also my separate, but related notes on 'Irreducible polynomials' (of the second degree).

[34] To prove, that is.

[35] http://www-groups.dcs.st-andrews.ac.uk/~history/Mathematicians/Euler.html

Historical Note: When Cantor first showed (in a famous paper of early 1874) that the real numbers could not be paired in a 1-1 way with the natural numbers, he did so to establish that there existed 'transcendental' (meaning *non-algebraic*) numbers. So, what is an 'algebraic' number?

Definition: α (real or complex) is said to be ***algebraic over the integers*** **(Z)** if $x = \alpha$ is a solution of *some* polynomial equation of the form

$$a_0x^n + a_1x^{n-1} + \ldots + a_{n-1}x + a_n = 0$$

where[36] $a_0, a_1, \ldots, a_{n-1}, a_n \in \mathbf{Z}$, and $a_0 \neq 0$.

Examples of algebraic[37] numbers.

- All rational numbers are, of course, algebraic. $\frac{3}{2}$ is a (in fact, the) solution of the equation $2x - 3 = 0$.

- $\sqrt{2}$ is a solution of the equation $x^2 - 2 = 0$. Of course one could also say that $\sqrt{2}$ is a solution of the equation $x^4 - 4 = 0$ (just as one could say that $\frac{3}{2}$ is a solution of the equation $4x^2 - 9 = 0$), but – in a sense that can (and will) be made precise – the simplest equation (of the very precise kind in the above definition) for which $\sqrt{2}$ is a solution is $x^2 - 2 = 0$.

- $\frac{3}{7}\sqrt{3} - \frac{1}{2}$ is a solution of the equation $ax^2 + bx + c = 0$, where the integers *a*, *b* and *c* are… (I don't need to tell you what they are; we can find them by setting:

  $$x = \tfrac{3}{7}\sqrt{3} - \tfrac{1}{2}, \text{ then } 14x = 6\sqrt{3} - 7,\ 14x + 7 = 6\sqrt{3},\ (14x+7)^2 = (6\sqrt{3})^2,$$
  $$\text{giving } 196x^2 + 196x + 49 = 108, \text{ and finally } 196x^2 + 196x - 59 = 0.$$

  Thus $x = \frac{3}{7}\sqrt{3} - \frac{1}{2}$ is a solution of the equation $196x^2 + 196x - 59 = 0$.

- $\sqrt{2} + \sqrt{3}$ is an algebraic number, and the coefficients of an appropriate equation are easily found by setting:

  $$x = \sqrt{2} + \sqrt{3}, \text{ then } x^2 = (\sqrt{2} + \sqrt{3})^2 = 2 + 2\sqrt{2} \times \sqrt{3} + 3 = 5 + 2\sqrt{6},$$
  $$\text{giving } x^2 - 5 = 2\sqrt{6},\ (x^2 - 5)^2 = (2\sqrt{6})^2 = 24, \text{ which tidies up to}$$
  $$x^4 - 10x^2 + 25 = 24, \text{ namely } x^4 - 10x^2 + 1 = 0.$$

[36] The 'coefficients.'

[37] It is standard practice to use the simple 'algebraic' for 'algebraic over **Z**'. (So why introduce the earlier terminology? It's simply that in more general settings, the coefficients can be other sorts of exotic mathematical objects, besides '**Z**'.)

- All the above examples are of real algebraic numbers, but there are **complex** algebraic numbers, e.g. $\sqrt{2}+\sqrt{-3}$ $(=\sqrt{2}+i\sqrt{3},$ where '*i*' is, of course, the square root of (−1), is an algebraic number, and the coefficients of an appropriate equation are easily found by setting: $x=\sqrt{2}+\sqrt{-3},$ then $x^2=(\sqrt{2}+\sqrt{-3})^2=2+2\sqrt{2}\times\sqrt{-3}-3=-1+2\sqrt{-6},$ giving $x^2+1=2\sqrt{-6},$ $(x^2+1)^2=(2\sqrt{-6})^2=-24,$ which tidies up to $x^4+2x^2+1=-24,$ namely $x^4+2x^2+25=0.$

  Thus $x=\sqrt{2}+\sqrt{-3}$ is a solution of the equation $x^4+2x^2+25=0,$ which has integral coefficients, and non-zero leading coefficient.

Comment. In introducing the notion of an algebraic number, Euler (or anyone else for that matter) **might well have believed** that every number (real or complex) would be algebraic (there are simply so many polynomials – with integral coefficients, and non-zero leading coefficient – (an infinite number of course!) that one might feel: surely one of those will 'do'?), especially if one bears in mind that the following (of which the last one is especially significant) may be proved:

- the sum, product, difference, ratio (non-zero denominator, of course) of any two algebraic numbers is also an algebraic number (of course the same is true of rational numbers, for example)

- but, much more strikingly, **every** solution of an equation of the form

  $$b_0x^m+b_1x^{m-1}+\ldots+b_{m-1}x+b_m=0$$

  where the $b_0,b_1,\ldots,b_{m-1},b_m\in$ **Algebraic numbers**, and $b_0\neq 0,$ is **itself** an algebraic number

That last result (which I won't be proving; it's not that it's difficult, but it does take so setting up of other ideas and techniques) is really striking!! In my view *it's precisely that*, that **could** well have led one to believe that there **might be no** numbers that aren't algebraic. Let's take a brief look at what it *means*: suppose one took the equation (I am choosing an *especially simple example*, with a point though)

$$\sqrt{\tfrac{2}{3}}x^3+\sqrt{\tfrac{-1}{11}}=0 \qquad \textbf{...} \text{ (i)}$$

in which the coefficients $\sqrt{\tfrac{2}{3}}$ and $\sqrt{\tfrac{-1}{11}}$ are **not** all integers (in fact none are, and the second one is a complex number). **Some** numbers are the solutions of (i), and they are all algebraic numbers!! Why? Simply rearrange (i) by obvious moves:

$$\sqrt{\tfrac{2}{3}}x^3=-\sqrt{\tfrac{-1}{11}},\ (\sqrt{\tfrac{2}{3}}x^3)^2=(-\sqrt{\tfrac{-1}{11}})^2,\ \tfrac{2}{3}x^6=\tfrac{-1}{11},\ 22x^6=-3,\text{ giving}$$
$$22x^6+3=0 \qquad \textbf{...} \text{ (ii)}$$

Now every solution of (i) is a solution of (ii) (but not vice-versa; there will be 3 ‘extra’ solutions), and thus every solution of (ii) is an algebraic number, then so too is every solution of (i).

Comment. Just absorb the point, and take it that there is a way of proving the general assertion (the so-called ‘closure’ theorem concerning algebraic numbers)

Algebraic numbers before (and just after) Cantor’s 1873 work. In the middle 1700’s Euler wondered if there were any transcendental numbers[38], but could not prove any *suspects* (π, for example) to be so. The first example of a transcendental was proved by the French mathematician Liouville[39] in 1844; the example he gave was the (infinite, of course) decimal:

$$0.\mathbf{11}000\mathbf{1}00000000000000000\mathbf{1}000\ldots$$

namely, $\sum_{n=1}^{\infty} \frac{1}{10^{n!}}$.

Hermite[40] gave a *difficult* proof that the number e $(= \sum_{n=1}^{\infty} \frac{1}{n!})$ is transcendental, and in 1882 Lindemann[41] gave a *very* difficult proof that the classic number, π, is transcendental.

Cantor proved that there **existed** a transcendental number, in fact, he did more: he proved that there existed an **uncountable** number of **real transcendental** numbers. To see how he did it, we just need to introduce three simple concepts in connection with an algebraic number, its **degree**, its **minimal polynomial**, and its **height**.

Definition. An algebraic number is said to be of ***degree*** r if it is a solution of some polynomial equation of degree r, with integer coefficients (leading one non-zero), and is **not** a solution of such an equation of smaller degree.

Examples.

- All rational numbers are automatically of degree 1

- $\sqrt{7}$ (e.g.) is algebraic, of degree 2, since $\sqrt{7}$ is a solution of the quadratic equation $x^2 - 7 = 0$, and cannot be a solution of a 1st degree polynomial equation $ax + b = 0$, integers a and b, $a \neq 0$, since otherwise $\sqrt{7}$ would be rational (which it isn’t)
  Note. $\frac{2}{3}$, is, of course, a solution of the 2nd degree equation $9x^2 - 4 = 0$, and so it is algebraic. But $\frac{2}{3}$ is not of degree ‘2’ since it is a solution of the smaller degree equation $3x - 2 = 0$ (which has integer coefficients)

[38] He introduced the term ‘transcendental’ (‘transcending the power of the algebraic’).
[39] http://www-groups.dcs.st-andrews.ac.uk/~history/Mathematicians/Liouville.html
[40] http://www-groups.dcs.st-andrews.ac.uk/~history/Mathematicians/Hermite.html
[41] http://www-groups.dcs.st-andrews.ac.uk/~history/Mathematicians/Lindemann.html

Definition. Let α be an algebraic number of degree *n*, and suppose α is a solution of the equation $a_0x^n + a_1x^{n-1} + \ldots + a_{n-1}x + a_n = 0$, $a_0, a_1, \ldots, a_{n-1}, a_n \in \mathbf{Z}$, $a_0 \neq 0$; then $a_0x^n + a_1x^{n-1} + \ldots + a_{n-1}x + a_n$ is said to be the ***minimal polynomial*** of α if $a_0 > 0$ and $\gcd(a_0, a_1, \ldots, a_{n-1}, a_n) = 1$.

The point of that definition. Let's look at a particular algebraic number, e.g., $\frac{5+\sqrt{73}}{6}$, which happens to be a solution of the quadratic equation

$$3x^2 - 5x - 4 = 0 \qquad \textbf{...}\ \text{(i)}$$

Of course, being a solution of (i), it is also a solution of

$$-3x^2 + 5x + 4 = 0 \qquad \textbf{...}\ \text{(ii)}$$

And one could also say that $\frac{5+\sqrt{73}}{6}$ is a solution of the equation

$$9x^2 - 15x - 12 = 0 \qquad \textbf{...}\ \text{(iii)}$$

In short, (i) is – in a sense – the simplest polynomial equation (integer coefficients, etc) for which $\frac{5+\sqrt{73}}{6}$ is a solution (being irrational it is not of degree 1), and that '$3x^2 - 5x - 4$' is referred to as its 'minimal polynomial'.

Definition. Let $b(x) = b_0x^n + b_1x^{n-1} + \ldots + b_{n-1}x + b_n$ be a polynomial of (genuine, i.e., $b_0 \neq 0$) degree *n* with integer coefficients, then the ***height*** of *b*(*x*) – denoted by *h*(*b*(*x*)) – is the sum of the absolute values of the coefficients of *b*(*x*).

Examples.

1. $h(2x^2 - 4x + 3) = |2| + |-4| + |3| = 2 + 4 + 3 = 9$
2. $h(-10x^3) = |-10| = 10$
3. $h(-50x^{12} - 40x + 30) = |-50| + |-40| + |30| = 50 + 40 + 30 = 120$

Important comment. For **fixed** degree, there are only a **finite** number of polynomials of **given** height.

Example. If one took the degree to be '5' (say), there would be only a finite number of 5th degree polynomials of height 8 (say). Why? Well, the 6 possible coefficients $b_0, b_1, \ldots, b_5$ would be required to satisfy

$$|b_0| + |b_1| + \ldots + |b_5| = 8$$

and since that **trivially** forces **every** coefficient to lie between -8 and 8 then there are only a finite number of possible *b*'s satisfying that restriction.

Definition. Let $a(x) = a_0x^n + a_1x^{n-1} + \ldots + a_{n-1}x + a_n$ be the minimal polynomial of α; then the ***height*** of α is the height of $a(x)$.

Examples.

1. The height of $\sqrt{2}$ is 3.
2. The height of $\frac{\sqrt{5}+1}{2}$ is 3.
3. The height of $\sqrt{-2}$ is 3.
4. The height of $\sqrt{-1}$ is 2.
5. The height of $\frac{3}{2}$ is 5.
6. The height of $\frac{6}{4}$ is (**also**) 5.

The point of all of these definitions.

- **Cantor argued that there is *only* a *countable* number of algebraic numbers**. Why? There are only a **finite** number of algebraic numbers of given degree *r*, and height *h*. Then, fixing the degree, and **varying** the height, we find that there are only a **countable** number of algebraic numbers of given degree. Finally, we vary the degree, and since there is only a countable number of algebraic numbers for each degree, and only a countable number of degrees (degree 1, degree 2, degree 3, **…**), then there is only a countable number of algebraic numbers altogether.

- Cantor then argued that there must exist transcendental numbers, in fact an uncountable number of them. How? Well, first of all the real numbers are uncountable, but the real algebraic numbers are countable. Thus there must exist some (either a finite number, a countable number, or an uncountable number) real transcendental numbers. There could not be just a finite number of them (since the union of a countable set and a finite set is countable), nor could there be only a countable number (since the union of a countable set and a finite set is countable), and thus there must **exist** an **uncountable number of real transcendental numbers**.

Euler's conjecture re some possible transcendental numbers (leading to 'Hilbert's seventh problem'). I start with some simple, familiar numerical facts:

$$\left(\frac{8}{27}\right)^{\frac{1}{3}} = \frac{2}{3},\ \left(\frac{1}{7}\right)^{-1} = 7,\ \left(\frac{25}{36}\right)^{-\frac{1}{2}} = \frac{6}{5},\ 4^2 = 16,\ \left(\frac{25}{36}\right)^{-1} = \frac{36}{25},\ \textbf{… (make up your own)}$$

Those are deliberately *intended* to be examples of the following: a rational number r_1, raised to a rational power *p*, turning out to be another rational number r_2. Simple, yes? But now consider this: what is the *nature* of the number *p* (is it rational? is it some algebraic number of (perhaps) the 30th degree?, is it transcendental?) in

$$\left(\frac{7}{27}\right)^p = \frac{2}{3}$$

where, as you see, I have merely *changed* the ‘8’ to a ‘7’ in $\left(\frac{8}{27}\right)^{\frac{1}{3}} = \frac{2}{3}$, and put a ‘$p$’ in place of that ‘$\frac{1}{3}$.’ Well, as you will know from our class discussion (with easily given proof), that ‘p’ is irrational. And a similar sort of thing happened with other examples… Then we noted that of course one could make up *trivial* examples like $0^{\sqrt{2}} = 0,\ 0^{\frac{3}{4}} = 0,\ 1^{\pi} = 1,\ 1^{-\sqrt[3]{2}} = 1, \ldots$, and out of all of that Euler posed this question: suppose one has two rational numbers r_1 and r_2 (with $r_1 \neq 0, 1$) related by

$$r_1^{\,p} = r_2$$

then, what is the nature of the number ‘p’. A great intuition of Euler’s **suggested**:

if p is **not rational**, then it is **transcendental**[42]

Euler couldn’t prove it (even he!)…

Hilbert’s (simple) extension of Euler’s conjecture (the seventh problem of his famous list). Let a_1 and a_2 be algebraic numbers (with $a_1 \neq 0, 1$) such that $a_1^{\,p} = a_2$, then p is either rational or transcendental. (Alternatively, the ratio of the logarithms of two algebraic numbers is either rational or transcendental. Or, the normal version of this problem: let α and β be algebraic numbers ($\alpha \neq 0, 1,$ and β irrational) then α^{β} is transcendental.)

(Standard) Examples. $2^{\sqrt{2}}$ is transcendental. e^{π} (being one of *the* values of the *complex* number $(-1)^{-\sqrt{-1}}$) is transcendental.

Comment. Hilbert once opined[43] (in the 1920’s) that the last problem would not be settled in his lifetime… , but it was settled… (a **long** story…)

Section 4

I see it, but I do not believe it…

A ‘higher infinity’ than the reals? When he proved (in December 1873) that there is no 1-1 correspondence between the natural and real numbers, Cantor naturally turned his attention to an obvious next question: is there some collection/set of things that is somehow ‘larger’ (in a then not explicitly defined sense) than the real numbers? An **immediate**, **obvious** candidate was the set of all points in the plane, and, since that set is easily seen to be of the same cardinality as any finite square (with non-zero side, of course!), and there is a 1-1 correspondence between the set of all real numbers and the set of all real numbers on any finite line segment (not a single point, of course!), then the **problem to be settled became**: could one prove that there

[42] An (obvious) alternative formulation is this: the ratio of the logarithms of two rational numbers is either *rational* or *transcendental*.

[43] There is a C.L. Siegel anecdote relating to this in Constance Reid’s *Hilbert* biography.

is **no** 1-1 correspondence between the **points of the unit square** (say) and the set of **points on a unit line** (say).

Why might that have appeared to be a reasonable candidate? Well, if one thinks of a square as being a (vertical, say) layering of lines, each defined by an end point which varies along the entire length of a vertical line segment, and there are uncountably many such end points, then certainly those lined can't be dismantled (as it were) and laid end-to-end to form the doubly infinite line segment of the real numbers.

(Four years later, however, the sensational) Theorem. (Famously he wrote to Dedekind: "*I see it, but I do not believe it*".) There **is** a 1-1 correspondence between the points in the unit square $S_2 = [0, 1] \otimes [0, 1]$, and the real numbers on the unit interval $S_1 = [0, 1]$.

Proof[44]. Let $(x, y) \in S_2$, and let x and y have *decimal* expansions given by:

$$x = 0.x_1x_2x_3 \ldots, \; 0 \le x_n \le 9$$
$$y = 0.y_1y_2y_3 \ldots, \; 0 \le y_n \le 9$$

Then, pair the (single) point (x, y) with the (single) real number X (in S_1) defined by

$$X = 0.x_1y_1x_2y_2x_3y_3 \ldots$$

Conversely, let $X \in S_1$, and let X have decimal expansion given by:

$$X = 0.X_1X_2X_3X_4X_5X_6 \ldots, \; 0 \le X_n \le 9$$

Then, pair (the single) X with the (single) point (x, y) (in S_2) where x and y are defined by:

$$x = 0.X_1X_3X_5 \ldots$$
$$y = 0.X_2X_4X_6 \ldots$$

(**Aside**. To get a *feeling* for what is going on, one should consider some **simple**, doable examples. Take $(\frac{1}{2}, \frac{1}{3}) \in S_2$; then $\frac{1}{2} = 0.5000\ldots$ and $\frac{1}{3} = 0.3333\ldots$, and thus $(\frac{1}{2}, \frac{1}{3})$ is paired with $X = 0.53030303\ldots$ (which is $\frac{1}{2} + \frac{3}{99} = \frac{1}{2} + \frac{1}{33} = \frac{35}{66}$). Similarly, take $\frac{4}{7} \in S_1$; then $\frac{4}{7} = 0.571428\,571428\ldots$, and so $\frac{4}{7}$ is paired with $(X, Y) \in S_2$, where $X = 0.512\,512\,512\ldots$ (which is $\frac{512}{999}$), and $Y = 0.748\,748\,748\ldots$ (which is $\frac{748}{999}$).

Of course, those were nice, easy examples (because one knew the exact decimal representations), whereas if one considered an example like (say): take $(\sqrt{2} - 1, \frac{1}{3}) \in S_2$; then $\sqrt{2} - 1$ does have a decimal expansion, but, being an irrational number, it does not have a periodic decimal expansion, and so one can't do the sort of

[44] I present the 'proof' in its standard (flawed) form. Like the earlier 'diagonal-decimal' proof, it suffers from a *rectifiable* flaw.

thing we did above. However, the point is though that $(\sqrt{2}-1, \frac{1}{3}) \in S_2$ **is** paired with a unique element of S_1 (whose decimal expansion will **commence**: 0.43134323133353…, since $\sqrt{2}'s$ decimal expansion commences 1.4142135…) [End of **Aside**] That, then, pairs every element of S_1 with an element of S_2, and vice versa. Thus 'There is a 1-1 correspondence between the real numbers in the unit square $S_2 = [0, 1] \otimes [0, 1]$, and the real numbers on the unit interval $S_1 = [0, 1]$.' [End of 'proof']

The flaw. It is simply that the correspondence is **not** 1-1 (it's as if in comparing the sets A = {p, q, r, s} and B = {2, 6, 23}, that we set up the correspondences: p$\leftrightarrow$2, q$\leftrightarrow$6, r$\leftrightarrow$23, and finally s$\leftrightarrow$6. **Both** 'q' and 's' have been paired with '6'), as we will now see (it all revolves around the phenomenon of non-unique decimal expansions of certain numbers).

Consider *any* α number in [0, 1] which has a non-unique decimal expansion, e.g.:

$$\alpha = 0.74329999\ \ldots = 0.7433000\ \ldots$$

Then note that two **different** elements of $S_2 = [0, 1] \otimes [0, 1]$

$$(x_1, y_1) = (0.73999\ \ldots, 0.42999\ \ldots) = (0.74000\ \ldots, 0.43000\ \ldots) = (\tfrac{74}{100}, \tfrac{43}{100})$$

and

$$(x_1, y_1) = (0.73000\ \ldots, 0.43000\ \ldots) = (\tfrac{73}{100}, \tfrac{43}{100})$$

are both paired with α. (**Advice**. Create your own such example.)

Yes, the flaw can be rectified. But it's just a bit messy… However, here is a slight (correct!) variation of the above:

Theorem. There **is** a 1-1 correspondence between the points in the unit square $S_2 = [0, 1] \otimes [0, 1]$, both of whose co-ordinates are irrational, and the set of irrational numbers in the unit interval $S_1 = [0, 1]$.

Proof. Every irrational number in [0, 1] has a unique continued fraction expansion $[0, a_1, a_2, a_3, \ \ldots\ , a_n \ldots]$ where $a_n \in \mathbf{N}$ for all $n \in \mathbf{N}$. Then, with $(x, y) \in S_2$ and x, y both irrational, we have:

$$x = [0, x_1, x_2, x_3, \ \ldots\ , x_n \cdots]$$
$$y = [0, y_1, y_2, y_3, \ \ldots\ , y_n \cdots]$$

and a function $F(x, y)$ defined by

$$F(x, y) = [0, x_1, y_1, x_2, y_2, x_3, y_3, \ldots\ , x_n, y_n, \ldots]$$

has irrational values in the unit interval, and there is now no flaw with regard to being 1-1. **[End of proof]**

Section 5

Cantor's 'nested interval' proof

The most commonly seen (in text books) proof of Cantor's that the real numbers are uncountable is his classic diagonal-decimal proof, but another really elegant proof of his is the 'nested interval' one (originally it was a bit more cumbersome, but it was tidied up by Dedekind). It requires at one critical point an appeal to the so-called:

Fundamental property of the real numbers. Let $\{a_n\}$ be any monotonic **increasing** sequence of *real* numbers (i.e. $a_1 \le a_2 \le a_3 \le \ldots \le a_n \le \ldots$) that is bounded above (i.e. there is some A such that $a_n \le A$ for all $n \in \mathbf{N}$) then $\{a_n\}$ converges (i.e. $\lim\limits_{n\to\infty} a_n$ exists).

Comment. This seemingly innocuous property is extraordinarily subtle, and completely encapsulates the **critical** difference there is between the *rational* and the *real* numbers: whereas there are monotonic increasing sequences of *rational* numbers that are bounded above, the sequence $\{a_n\}$ does **not necessarily** converge to a rational number (can you think of an example?)

(Of course there are monotonic increasing sequences of *rational* numbers that are bounded above and $\{a_n\}$ converges; can you think of examples?) There is also the:

Fundamental property of the real numbers (alternative version). Let $\{b_n\}$ be any monotonic **decreasing** sequence of *real* numbers (i.e. $b_1 \ge b_2 \ge b_3 \ge \ldots \ge b_n \ge \ldots$) that is bounded **below** (i.e. there is some B such that $b_n \ge B$ for all $n \in \mathbf{N}$) then $\{b_n\}$ converges (i.e. $\lim\limits_{n\to\infty} b_n$ exists).

A marriage of these two results may be made, and be called:

The fundamental nested interval property of real numbers. Let $\{a_n\}$ and $\{b_n\}$ be monotonic increasing and monotonic decreasing sequences of *real* numbers such that $a_n \le b_n$ for all $n \in \mathbf{N}$, then there is some real number that lies in all the nested closed intervals $\{I_n\}$, where $I_n = [a_n, b_n]$.

Comment. This result is an immediate consequence of the two earlier fundamental results. Being 'closed' is an important element in its validity (the nested *open* intervals $(0,1), (0,\tfrac{1}{2}), (0,\tfrac{1}{3}), (0,\tfrac{1}{4}), (0,\tfrac{1}{5}), \ldots$ do **not** share a common point).

Now we are ready for:

The nested interval proof that the real numbers are uncountable. Suppose that r_1 , r_2 , r_3 , r_4 , ... is a countable enumeration of all real numbers[45].

First choose *any* a_1, b_1 such that $a_1 < b_1$ and $r_1 \notin I_1 = [a_1, b_1]$, and successively define intervals $I_2, I_3, I_4, \ldots$, as follows:

- $a_1 \le a_2, b_2 \le b_1$, and $r_2 \notin I_2 = [a_2, b_2]$
- $a_2 \le a_3, b_3 \le b_2$, and $r_3 \notin I_3 = [a_3, b_3]$

  … … In general:
- $a_{n-1} \le a_n, b_n \le b_{n-1}$, and $r_n \notin I_n = [a_n, b_n]$

Then – by the fundamental nested interval property – there exists some $r \in \mathbf{R}$ with $r \in I_n$ for all $n \in \mathbf{N}$, and $r = r_i$ for some $i \in \mathbf{N}$. But then $r_i \notin I_i = [a_i, b_i]$ is impossible, and so an countable enumeration of the reals is impossible.

Comment. Of course what makes that proof 'work' is the **fundamental** property…
If one considered just the rational numbers, **Q**, one may construct a nested sequence of closed intervals of rational intervals (i.e., the end points and **all** interior points are rational) with **no** rational number common to all of them: for example, consider the (*real*) irrational number $\sqrt{2}$ and its '*L* and *R*-approximations' (or, if you prefer, the alternating 'partial convergents' obtained from its non-terminating continued fraction expansion). Then, arrange its *L*-apprs in ascending order (of size) $\frac{1}{1}, \frac{7}{5}, \frac{41}{29}, \ldots$, and its *R*-apprs in descending order (of size): $\frac{3}{2}, \frac{17}{12}, \frac{99}{70}, \ldots$; then the closed intervals (all with rational end points) $\left[\frac{1}{1}, \frac{3}{2}\right], \left[\frac{7}{5}, \frac{17}{12}\right], \left[\frac{41}{29}, \frac{99}{70}\right], \ldots$ have **no** rational point in common (because, of course, the only point that they actually have in common – from a higher vantage point, as it were – is the *real irrational* number $\sqrt{2}$).

Section 6

The 'power set' of a set

Question. Given *any* set *S*, does there exist a set S' with a *higher* finite level of cardinality than *S*? In other words, is there a set S'–*ideally* (though not necessarily) obtained, *somehow*, from *S*– such that the elements of *S* and S' *cannot* be put in 1-1 correspondence, *but* the elements of *S can* be put in 1-1 correspondence with some subset of S'.

A response. The answer is *completely trivial*, of course, if the set *S* has a *finite* number of elements: simply form the set $S' = S \cup \{a\}$, where '*a*' is some 'extra' element (not in *S*). In short, simply add an extra element to *S*, so *increasing* its cardinality by '1'. Of course, that is not the only way to form a larger set from a finite set, but at least it does what we require. However it would appear to be a much *more*

[45] In what follows it should be understood that all the defined *a*'s and *b*'s are real numbers.

difficult thing to achieve a similar construction in the infinite case. I very much doubt that anyone–who didn't come to hear what Cantor actually came up with–would find a way unless aided. Reflect on Cantor's own initial belief–which he held onto for some three years, before he eventually demonstrated its invalidity–that the set of points in the plane (which is, of course, equivalent to the set of points in any square of non-zero side) provided a higher level of infinitude than that of the real numbers.

Given the background–natural numbers, rationals, irrationals, reals, the difficulty of *seeing/creating/discovering/guessing* a larger infinity than the reals–it should come as a shock to see Cantor's stunningly simple creation of a *general* method for producing a greater level of infinity from any give level. This is something to savour and appreciate, before which anyone, who doesn't already know what Cantor did, should think about it.

Definition: Let S be a set (finite or infinite); by $P(S)$, the '***power set***' of S is meant the set of all subsets of S (i.e. $P(S) = \{S' | S' \subseteq S\}$).[46] (It is a standard convention to denote the power set of S by 2^S – and indeed to refer to the cardinality of $P(S)$ as being $2^{|S|}$ – where $|S|$ denotes the cardinality of S (i.e. how many elements there are in S). A reason for that nomenclature will be clear after considering a few examples.)

Examples.

- Let $S_3 = \{a, b, c\}$ be a set with three elements, then the power set of S_3 has 8 elements: $\{a\}, \{b\}, \{c\}, \{a,b\}, \{a, c\}, \{b, c\}, \{a, b, c\}$, and ϕ (the 'empty set', representing *no choice*).
- Let $S_4 = \{a, b, c, d\}$ be a set with four elements, then the power set of S_4 has 16 elements $\{a\}, \{b\}, \{c\}, \{d\}, \{a, b\}, \{a, c\}, \{a, d\}, \{b, c\}, \{b, d\}, \{c, d\}$, $\{a, b, c\}$, $\{a, b, d\}$, $\{a, c, d\}$, $\{b, c, d\}$, $\{a, b, c, d\}$, and ϕ.
- Let S be the set of **all** natural numbers; then the power set of S has (a lot of!) elements, *some* of which are: {1}, {2}, {3}, … {1, 2}, {1, 3}, {1, 4}, … {2, 3}, {2, 4}, {2, 5}, … … , {120, 121}, {120, 122}, … , {1, 2, 3}, {1, 2, 4}, … , … … … , {1, 2, 3, 4, 5, 6, … }, … {2, 3, 5, 7, 11, … } (the primes!), … {2, 4, 6, 8, 10, 12, … } (all even nat. nos.), … … … (you must try to **imagine** many different types of examples of such elements), and, of course, ϕ.

A reason for the 'power set' terminology. All, in fact, should be clear after considering just the first example above. $S_3 = \{a, b, c\}$ has 3 elements, and laboriously we listed all elements of $P(S_3)$. However, without listing all (8) elements of $P(S_3)$ we may see that it has $2^3 (= 8)$ elements by *thinking of* subsets of $S_3 = \{a, b, c\}$ as being of the form {X, Y, Z} where

X is either '*a*' or blank (and so X has – as it were – 2 'values')
Y is either '*b*' or blank (and so Y has – as it were – 2 'values')

[46] There is a good reason for the name–power set–and the associated notation, which will be clear in a moment.

Z is either '*c*' or blank (and so Z has – as it were – 2 'values')

(So, e.g., the subset {*a*, *c*} is simply {*a*, blank, *c*}, or, e.g., ϕ (the empty subset) is simply {blank, blank, blank}

By simple combinatorics, then, the number of subsets of S_3 is 2^3. Indeed it should be seen that for any finite *n*, the number of subsets of $S_n = (a_1, a_2, \ldots, a_n)$ is 2^n.

Comment. Another way of seeing it (again when *n* is finite) is via a familiar use of the binomial theorem. Start with

$$(1+x)^n = {}^nC_0 + {}^nC_1x + {}^nC_2x^2 + \ldots + {}^nC_{n-1}x^{n-1} + {}^nC_nx^n \quad \textbf{...} \text{ (i)}$$

and set $x = 1$ to give:

- The LHS of (i) is simply 2^n
- The RHS of (i) is ${}^nC_{\mathbf{0}} + {}^nC_{\mathbf{1}} + {}^nC_{\mathbf{2}} + \ldots + {}^nC_{\boldsymbol{n-1}} + {}^nC_{\boldsymbol{n}}$, and the connection should now be **obvious** since the individual binomial coefficients in turn count how many subsets there are formed from *n* objects by successively choosing **none** of them, **1** of them, **2** of them, **...** , ***n*** of them.

So much, then, for nomenclature.

It was another of Cantor's remarkable discoveries that the power set of *any* set, produces a higher level of cardinality (be it a finite set–in which case it is, of course, trivial–or an infinite set (definitely non-trivial!)). It may come as a surprise that a proof of that remarkable fact is extraordinarily simple (with the usual hindsight!)

Theorem (Cantor). There is **no** 1-1 correspondence between the elements of *S* and those of *P*(*S*).

Comment. I could give the proof immediately (it is very short, but it could well leave you with a feeling of bewilderment), and I believe that you will find it easier to follow the proof if I first of all motivate it by considering a special case of it–the simplest case–where the set *S* is the 'smallest' infinite set, the set **N** of natural numbers.

A thought experiment. Imagine for a moment that the elements of **N**, and its power set *P*(**N**), *could* be put in 1-1 correspondence, and (just to get us going) suppose that the early correspondences went *something like* this (elements from N given first, with their corresponding paired elements from *P*(**N**) coming second):

'1' paired with $s_1 = \{4, 8, 790\}$
'2' paired with $s_2 = \{2, 4, 6, 8, 10, 9876\}$
'3' paired with $s_3 = \{1, 2, 4, 5, 6, \ldots$ *ad infinitum*$\}$
'4' paired with $s_4 = \{4\}$

‘5’ paired with $s_5 = \{1, 2, 3, 5, 10, 15, 20, 25, 30, 35, \ldots$ *ad infinitum*$\}$
‘6’ paired with $s_6 = \phi$ (the empty subset of **N**)
‘7’ paired with $s_7 = \{7, 1000000000000765\}$ etc

I wish to draw your attention to a very important notion, which takes its **meaning** *directly* from the *supposed* 1-1 correspondence; look at the following reproduction of the above initial terms, but with certain elements from **N** being **highlighted**:

‘**1**’ paired with $s_1 = \{4, 8, 790\}$
‘2’ paired with $s_2 = \{2, 4, 6, 8, 10, 9876\}$
‘**3**’ paired with $s_3 = \{1, 2, 4, 5, 6, \ldots$ *ad infinitum*$\}$
‘4’ paired with $s_4 = \{4\}$
‘5’ paired with $s_5 = \{1, 2, 3, 5, 10, 15, 20, 25, 30, 35, \ldots$ *ad infinitum*$\}$
‘**6**’ paired with $s_6 = \phi$ (the empty subset of **N**)
‘7’ paired with $s_7 = \{7, 1000000000000765\}$ etc

What I am drawing your attention to is that

- *some* of the natural numbers–in the above *supposed* 1-1 pairing they happen to be 2, 4, 5, 7, (and possibly some others, further along)–*are* members of the subsets with which they are paired

- whereas others–in the above they are **1**, **3**, **6**, (and possibly some others, further along)–are **not** members of the subsets with which they are paired.

I would like to call each of the former natural numbers *internal* elements *with respect to the* (supposed) *pairing*, and call each call each of the latter natural numbers ***external*** elements *with respect to the* (supposed) *pairing*.

Next, form the subset of *P*(**N**) which consists of **all** external *n*’s in the above supposed 1-1 pairing; from the above illustration that subset $(s_r, \text{some } r \in \mathbf{N})$ would look like this:

$s_r = \{$**1**, **3**, **6**, with some, or possibly, no other *n*’s$\}$

(Thus ‘*r*’ is the natural number with which s_r is paired.) Now (the crux!) we see that the supposed 1-1 correspondence is impossible! Why? Simply think about the nature of that ‘*r*’. Is it ‘internal’ or ‘external’? (Of course, it must be one or the other!)

- It **can’t** be internal, because, if it were, then it would be one of the numbers **1**, **3**, **6**, … , all of which are external.
- It **can’t** be external, because, if it were, then it would **not** be one of the numbers **1**, **3**, **6**, … , whereas they are precisely **all** the external *n*’s.

Thus the *s*, consisting of all external *n*’s is not paired with any element of **N**, proving that a 1-1 correspondence between the elements of **N** and *P*(**N**) is impossible.

Comment. A proof of the general version of Cantor's theorem just proceeds along similar lines (whether *S* be finite or infinite). Also, since the elements of *S* can *clearly* be paired in 1-1 way with *some* of the elements of *P*(*S*) (simply take all the singletons in *P*(*S*); i.e., all those elements of *P*(*S*) of the form $\{a\}$, where '*a*' varies over all the elements of *S*), it means that *P*(*S*) has 'more' elements than *S*).

Proof. Suppose there is a 1-1 map from *S* to *P*(*S*), *f*, say. Define

- '*e*' is an *internal* element of *S* (with respect to '*f*') if $e \in f(e)$ (in *P*(*S*))
- '*e*' is an *external* element of *S* (with respect to '*f*') if $e \notin f(e)$ (in *P*(*S*))

For non-empty *S*, external elements **do** exist[47], and – as above – considering the subset of all external elements leads easily to a conflict. [End of proof]

Comment. It is obvious, then, that the power set *P*(*S*) of any set *S* has 'more' elements than *S* since the above theorem establishes that they do not have the same number of elements, and it is obvious that *S* has the same number of elements as some subset of *P*(*S*): simply make each element '*e*' of *S* be paired with the {e} subset of *P*(*S*).

Final comment. Concerning the work of Cantor, we have merely scratched the surface…

[47] In the introductory S = **N** case of this theorem we rather *took for granted* that there were some external elements (in the event of there being a 1-1 correspondence…), and one should raise this **objection**: did the 'proof' not rely upon assuming that there were some external elements? Of course it did. To get the proof-idea introduced I did lead you to the notion of 'internal' and 'external' elements (with respect to a supposed 1-1 correspondence). However, it is easy to see that such elements **must** exist **if** there is a 1-1 correspondence: for **if** there is a 1-1 correspondence **and** there are **no** 'external' elements then **every** element would be internal. But then, taking every subset of *P*(*S*) consisting of a singleton – $\{a\}$ – the element of *S* with which $\{a\}$ would have to be paired would be '*a*' itself (from *S*)