# On Considerations of Language in the Diagonal Proof

James R Meyer

31 January 2019

**Abstract**

This paper analyzes the diagonal proof while taking careful account of considerations of language. The analysis shows that the principle of the method of the proof and the essential result of the proof is thereby clarified, since such analysis eliminates certain assumptions that are commonly associated with the proof.

Version History:

Version 2: Introduced the concept that for any enumeration of real numbers there is a corresponding set of real numbers. Clarified that the definition of recursive enumerations is not itself contradictory.

## 1 The Independent Existence of Numbers

The term 'diagonal proof' is commonly used to refer to proofs based on an original proof by Georg Cantor.[2] In this paper, various presentations of the proof will be analyzed while taking careful account of considerations of language. The diagonal proof is commonly cited as proving the 'existence' of real numbers as actual non-physical entities, and which exist completely independently of any finite definition of some language; that is, that there is no finite definition whatsoever that can define such a number (indefinable or inaccessible numbers). [1]

Our first consideration concerns whether the notion of the independent 'existence' of real numbers and sets of real numbers is compatible with the principle of the diagonal proof.

**Theorem 1.1.** *Real numbers and sets of real numbers cannot exist as entities that exist in a non-physical manner completely independently of any finite definition.*

*Proof.* We begin with the initial assumption that real numbers and sets of real numbers have a real existence that is independent of any finite definition, and we apply the method of the diagonal proof.

Given any initial well-defined infinite enumeration $r_1(x)$ of real numbers between 0 and 1, for example the non-integer part of all different binary expansions of the numbers $\sqrt{n}, n \in \mathbb{N}$, we assume that a diagonal number $d_1$ exists for that enumeration, independently of any finite definition that might be made regarding that enumeration.

[1] By a finite definition of a mathematical entity, we mean that there exists a complete definition of the mathematical entity by some sequence of symbols of some language system.

Then there exists a set $A_1$ of numbers (in the binary base) that corresponds precisely to that enumeration, given by:

> Each odd digit of each number in the set $A_1$ corresponds to the ordinal number of the enumeration by the rule that there is as many 1's as the ordinal number of the enumeration. The even digits correspond exactly to the digits of the number that is enumerated. This gives, for example:
>
> > for the 1st enumerated number the corresponding element of the set $A_1$ is $0.1x0x0x0x0x0x0x0x\ldots$
> >
> > for the 2nd enumerated number the corresponding element of the set $A_1$ is $0.1x1x0x0x0x0x0x0x\ldots$
> >
> > for the 3rd enumerated number the corresponding element of the set $A_1$ is $0.1x1x1x0x0x0x0x0x\ldots$
>
> and so on, where the $x$'s are the digits of the enumerated number. This gives a set of numbers that correspond precisely to the enumeration.
>
> Given this first set $A_1$ there exists another set of numbers $A_2$ as follows:
>
> The numbers of $A_2$ are given from each element of $A_1$ by changing the first odd digit that is currently 0 to 1. The set $A_2$ also includes one additional number $d_{A_1}$ which corresponds to the diagonal number $d_1$ of the enumeration $r_1$. We will call such numbers quasi-diagonal numbers. The quasi-diagonal number $d_{A_1}$ has 1 as its initial digit, 0 for all the remaining odd digits, and the even digits are given by the diagonal number of the even digits of the numbers of $A_1$, giving a number in the form of $0.1x0x0x0x0x0x0x0x\ldots$. Hence the even digits of the quasi-diagonal number $d_{A_1}$ are precisely those of the diagonal number $d_1$ of the enumeration $r_1$.
>
> We also note that an enumeration $r_2$ can be defined that corresponds to the set $A_2$, and whose first number is the diagonal number of the enumeration $r_1$, and where the remaining elements of the enumeration are the elements of the enumeration $r_1$, where each ordinal has been shifted up by 1.
>
> Similarly there exists another set $A_3$ which contains a quasi-diagonal number $d_{A_2}$ that corresponds to the diagonal number $d_2$ of $r_2$, and whose other elements are the elements of $A_2$, and an enumeration $r_3$ can be defined that corresponds to the set $A_3$. And so on for a set $A_4$ and a corresponding enumeration $r_4$, and for $A_5$ and $r_5$ etc... .

By the assumption of independent existence, every such set $A_n$ exists, and every element of every such set exists, and every quasi-diagonal number $d_{A_1}, d_{A_2}, d_{A_3}, \ldots, d_{A_n}$ exists, and the corresponding diagonal numbers $d_1, d_2, d_3, \ldots, d_n$ also exist.

Hence there can be no diagonal number that is given by the above process and which is not in some such set $A_n$. Hence there exists an enumerable set that includes the numbers corresponding to the original enumeration $r_1$ and *all* of the subsequent diagonal numbers by the above process.[2]

---

[2] Note that it is completely immaterial to the point in question (which is that numbers that 'exist' independently of any definition leads to a direct contradiction) as to whether that set constitutes all real numbers.

We now have a contradiction, since there exists an enumerable set $A^*$ whose elements are the elements of the original set $A_1$ and every subsequent quasi-diagonal number and which correspond to every number of the original enumeration $r_1$ and *every* subsequent diagonal number, but it is also the case that by the principle of the diagonal proof, for any such set there must always 'exist' a diagonal number that is *not* in that set, and hence there must 'exist' a quasi-diagonal number that is *not* in the set $A^*$.

Hence the initial assumption that real numbers and sets of real numbers have a real existence that is independent of any finite definition is incorrect, and the theorem is proved. □

### 1.1 Contradictions arising from independent existence

It might be thought that the above definition itself is paradoxical without requiring any reference to independent existence. However, discarding the assumption of independent existence immediately dissolves the contradiction of an enumerable set which at the same time includes and does not include its own diagonal number. This is because the defined process is a recursive process that never completes, and where at each recursion, the first item in the enumeration is a new number that was not in the previous enumeration. Hence the definition cannot define any one enumeration where there is a defined singular specific initial number and which includes all diagonal numbers. The definition only defines a series of enumerations, each of which includes only a finite quantity of diagonal numbers. Hence it is the case that provided one does not assume that all sets of numbers actually exist in some sort of independent non-physical manner, a paradoxical contradiction does not arise from the above definition.

We now have the situation where, on the one hand, it is claimed that the diagonal proof proves that real numbers have an actual existence that is independent of any finite definition, while on the other hand, the notion that real numbers have such an independent existence leads to a direct contradiction. If the concept of mathematical proof is to have any logical significance, the cause of this contradiction must be elucidated and eliminated.

Hence it is critical that any presentation of the diagonal proof should proceed in a completely logical manner without making any hidden assumptions. Since in any well-defined proof, all definitions used in the proof are defined in some language system, there must be a full consideration of how language is used in any diagonal proof and whether different levels of language are involved. It is to be noted that in any presentation of the diagonal proof, the diagonal number is defined in some language system by some finite definition, and that definition necessarily references an enumeration function, which is also defined in some language system.

## 2 Analyzing the Diagonal Proof

In the following, presentations of the diagonal proof will be analyzed with respect to a consideration as to which language system the diagonal number and the enumeration function belong to.

For convenience, the proof will be analyzed in terms of the binary system. This simplifies the proof since in the binary system, the only symbols used to define numbers are 0, 1 and the binary point; the proof is not compromised in any way by such simplification

Naturally, it is the case that regardless of whichever base is chosen, there are real numbers that cannot be expressed by a finite sequence of numerical digits. Some rational numbers have an expansion that continues infinitely with a repeating string of numerical digits; no irrational number can be represented by a finite string of the 0s and 1s of binary notation.

By way of a comparison with diagonal proofs that take into account language systems that may be involved, we first give a typical presentation of a diagonal proof [3] where there is no consideration of language systems. [4]

## 2.1 A conventional presentation of the diagonal proof

To be proved:

2.1.a) For every function that enumerates a set of real numbers between 0 and 1, there exists some real number that is between 0 and 1 which is not enumerated by that function.

2.1.b) There is no function that enumerates all real numbers between 0 and 1.

1. Clearly, there are functions that enumerate at least some real numbers between 0 and 1. For such a function $r(x)$, the function generates a correspondence between the number 1 and a unique real number. We designate by $r(1)$ the real number corresponding to the number 1, and by $r(2)$ the real number corresponding to the number 2, and so on.
2. For any such function enumerating a set of real numbers, we can define a real number $\boldsymbol{d}$ as follows:

   The initial symbols of the number are zero followed by a point, viz: '0.'.
   Then if the first digit of the first number in the enumeration is 0 the first digit of the new number $\boldsymbol{d}$ is 1; if it is 1 the first digit of the new number $\boldsymbol{d}$ is 0. Similarly, if the second digit of the second number in the enumeration function is 0 the second digit of the new number $\boldsymbol{d}$ is 1; if it is 1 the second digit of the new number $\boldsymbol{d}$ is 0. And so on, so that in general:

   > if the $n^{\text{th}}$ digit of $r(n)$ is 0, then the $n^{\text{th}}$ digit of the new number $\boldsymbol{d}$ is 1, otherwise it is 0.

3. This real number $\boldsymbol{d}$ differs from every other real number enumerated by the function since it is different from every number enumerated by the function by at least one digit. For any finite enumeration function, the number $\boldsymbol{d}$ is a rational number, since the sequence of digits is finite. But if the function enumerates infinitely many real numbers, then the definition of $\boldsymbol{d}$ defines a real number that has an infinite expansion, and for which there may not be a corresponding finite sequence of digits 0 and 1 that has the same real number value.

---

[3] Note that in the binary system some real numbers can be defined in two different ways, for example, 0.1 can also be defined as the infinitely repeating 0.01111111.... This is sometimes used as a criticism of the diagonal argument, since it is conceivable that, while the number generated by the diagonal argument is a different symbol sequence to every number in the enumeration, it is possible that it is the *same value* as one of the numbers in the enumeration. However this criticism is easily obviated, see 4 *Appendix: Dual Definition*.

[4] See, for example, Hodges[4] for a typical presentation.

4. So, given any function that enumerates a set of real numbers, it is always possible to define a real number that is not given by that enumeration - the diagonal number. This proves 2.1.a.

5. Now suppose that there can be a function that enumerates every real number between 0 and 1.

6. This results in a contradiction, because the diagonal number would be on the one hand defined as a number that is included in the enumeration, but on the other hand it cannot be included in the enumeration because it differs from every number in the enumeration, since it is always different to the $n^{\text{th}}$ number in the enumeration at the $n^{\text{th}}$ digit.

7. That means that the supposition that there can be a function that enumerates every real number between 0 and 1 is incorrect.

8. Therefore there cannot be a function that enumerates every real number between 0 and 1. This proves 2.1.b.

We note that the above presentation makes no mention of the language systems involved.

## 2.2 The diagonal proof for a given language system

We now analyze the proof by including a consideration of the language systems involved. We first analyze the diagonal proof for a given language system, where the enumeration function, the real numbers enumerated by that function, and the diagonal number all belong to the same language system.

To be proved:

2.2.a) For every function of a given language system $\boldsymbol{L}$ that enumerates some set of real numbers of that language system $\boldsymbol{L}$ between 0 and 1, there exists some symbol sequence of that language system that represents a real number between 0 and 1 in that language system $\boldsymbol{L}$ and which is not enumerated by that function.

2.2.b) For any given language system $\boldsymbol{L}$, there is no function of that language system $\boldsymbol{L}$ that enumerates all sequences of that language system $\boldsymbol{L}$ that represent real numbers between 0 and 1 in that language system $\boldsymbol{L}$.

1. It is evident that there are functions of language systems that include at least some real numbers of that language system. For such a function $r(x)$ of the given language system $\boldsymbol{L}$, the function generates a correspondence between each natural number and a unique real number. We designate by $r(n)$ the real number corresponding to the natural number $n$.

2. For any such function enumerating a set of real numbers, we can define a real number of the a given language system $\boldsymbol{L}$, and which we designate as $\boldsymbol{d}$, by the following rule:

   The initial symbols of the number are zero followed by a point, viz: '0.'.
   For the subsequent digits, if the $n^{\text{th}}$ digit of $r(n)$ is 0, then the $n^{\text{th}}$ digit of the new number $\boldsymbol{d}$ is 1, otherwise it is 0.

3. In this way, a real number $\boldsymbol{d}$ of the given language system $\boldsymbol{L}$ is defined (the 'diagonal' number); this number $\boldsymbol{d}$ differs from every other real number enumerated by the function since it is different from every number enumerated by the function by at least one digit. For any finite enumeration function, the number $\boldsymbol{d}$ is a rational number, since the sequence of digits is finite. But if the function enumerates infinitely many real numbers, then the definition of $\boldsymbol{d}$ defines a real number in terms of an infinite expansion, and for which there may not be a finite sequence of digits 0 and 1 that has the same real number value ( i.e., the number may not have a trailing sequence of all 0s or of all 1s).

4. So, given any function of the given language system $\boldsymbol{L}$ that enumerates a set of real numbers of that given language system $\boldsymbol{L}$, it is always possible to define a real number of that given language system $\boldsymbol{L}$ that is not given by that enumeration (the diagonal number). This proves 2.2.a.

5. If it is now supposed that there can be a function the given language system $\boldsymbol{L}$ that enumerates every real number of that given language system $\boldsymbol{L}$ between 0 and 1, that results in a contradiction, because the number $\boldsymbol{d}$ would be on the one hand defined as a number that is given by the enumeration, but on the other hand it cannot be given by the enumeration because it differs from every number in the enumeration, since it is always different to the $n^{\text{th}}$ number in the enumeration at the $n^{\text{th}}$ digit.

6. That means that the supposition that there can be a function of the given language system $\boldsymbol{L}$ that enumerates every real number of that given language system $\boldsymbol{L}$ between 0 and 1 is incorrect, and there cannot be a function of a given language system $\boldsymbol{L}$ that enumerates every real number of that language system $\boldsymbol{L}$ between 0 and 1. This proves 2.2.b.

## 2.3 The diagonal proof for two different language systems

Now we consider the method of the proof where the enumeration function belongs to a different language to the language of the real numbers that are enumerated. In such a case, the enumeration function is a meta-language, and the language of the real numbers that are to be enumerated is an object language; the sequences of symbols of this object language are objects of the meta-language, and do not represent syntax of the meta-language.

In this case the propositions for which a proof is to be attempted are:

2.3.a) For any function of a given language system $\boldsymbol{L_1}$ that enumerates sequences of another language system $\boldsymbol{L_2}$ that represent real numbers between 0 and 1 in that language system $\boldsymbol{L_2}$, there exists a real number between 0 and 1 which is not enumerated by that function.

2.3.b) For any given language system $\boldsymbol{L_2}$, there is no function of *any* language system that can enumerate *all* sequences of that language system $\boldsymbol{L_2}$ that represent real numbers between 0 and 1 in that language system $\boldsymbol{L_2}$

As in the previous case, we designate by $r(n)$ the real number corresponding to the natural number $n$.

It is required that the enumeration function belongs to the language system $\boldsymbol{L_1}$. But, in order that a diagonal number is definable, this function (and hence the language system $\boldsymbol{L_1}$) must be able to determine any digit of any binary expansion of any sequence of symbols of the language system $\boldsymbol{L_2}$ that represents a real number in the language system $\boldsymbol{L_2}$. It must be borne in mind that every representation of an irrational number in the language system $\boldsymbol{L_2}$ cannot be by a sequence of digits 0 and 1, but by some finite sequence of other symbols which define that number.

But there is no logical reason to make the assumption that the language system $\boldsymbol{L_1}$ can make a complete determination of the binary expansion from any such finite sequence of symbols that belongs to the language system $\boldsymbol{L_2}$, and which has a numerical value in that language system $\boldsymbol{L_2}$, according to the rules of syntax of that system $\boldsymbol{L_2}$.[5]

It is of course imperative that any mathematical proof must provide a logical basis for every step of the proof, without having any reliance on any non-axiomatic assumptions. We note that the conventional presentation of the diagonal proof provides no basis for an assumption that a diagonal number is in fact definable from such an enumeration as indicated above. And, as shown in Theorem 1.1, the assumption that the diagonal number simply 'exists' without requiring any actual finite definition leads to a contradiction.

Hence there is no logically valid proof that there is a diagonal number for an enumeration of real numbers where the enumeration is in a different language to the numbers that are enumerated, and the proposition 2.3.a is not proved.

Proposition 2.3.b is clearly false, since a meta-language can always enumerate *all* symbol sequences of an object language. This is dealt with in detail in the following section.

## 3 Indefinable Real Numbers

It is shown above by straightforward analysis that the conventional simplistic approach to the diagonal proof fails to establish certain rather pertinent facts. Conventionally, any consideration of language systems in relation to the proof is deemed to be unimportant. But certain claims that are made concerning the diagonal proof actually introduce the notion of different levels of language as meta-language and object language. Such arguments are commonly presented in a form such as:

*Proposition:*

*There cannot be any language system that can represent all real numbers.*

*Initial assumption:*[6] *It is possible that every real number can be expressed by some finite sequence of symbols from a finite set of symbols.*

---

[5] If the language system $\boldsymbol{L_1}$ can make a complete determination of the binary expansion from a finite sequence of $\boldsymbol{L_2}$, then it would appear that the language system $\boldsymbol{L_1}$ has the complete determination of the numerical value of that sequence of symbols of the language system $\boldsymbol{L_2}$. If that is the case, then it would appear that such a sequence of symbols $\boldsymbol{L_2}$ in fact represents a real number in that language system $\boldsymbol{L_1}$, which contradicts the initial stipulation that the real numbers in the enumeration belong to the language system $\boldsymbol{L_2}$ rather than the language system $\boldsymbol{L_1}$.

[6] The assumption to be contradicted.

*Then there exists a method of generating an enumeration of these sequences, such that there is a unique correspondence of natural numbers to every one of these combinations of symbols, as follows:*[7]

> *Since there is a finite quantity of such symbols, the symbols can be enumerated; we designate this enumeration as S(m). Given an enumeration of the symbols, then all possible sequences of these symbols can be enumerated, as follows:*
>
> *All sequences consisting of a single symbol are enumerated, according to the above enumeration S(m).*
>
> *All sequences consisting of two symbols follow on the above enumeration, by first taking all of those sequences whose first digit is S(1), and enumerating them according to the enumeration of the second digit of the sequence. Then the same procedure is applied to the sequences whose first digit is S(2), and enumerating them according to the enumeration of the second digit of the sequence. And similarly for the sequences whose first digit is S(3). And so on for the remaining first digits. Then the same procedure is applied for sequences that consist of three symbols. And so on.*
>
> *Since the same procedure can be carried out similarly for all such symbol sequences, there can be an enumeration of all the sequences.*

*But the diagonal proof shows that if there were such an enumeration, then a new real number can be defined from that enumeration, and which is not in that enumeration.*

*Therefore the initial supposition that there is a language system that can express every real number by some finite sequence of symbols from a finite set of symbols is incorrect, and there is no language system that can represent all real numbers.*

The claim has its origins in a paper by Gyula König.[5] It might be noted that Cantor initially refused to accept König's argument, stating, '*Infinite definitions (which are not possible in finite time) are absurdities. ... Am I wrong or am I right?*'[3]

The fallacy in the claim is easily observed, since on the one hand the claim invokes the principle of the real numbers of an object language being referenced by a meta-language, but on the other hand, at the same time, the claim uses a version of the diagonal proof that ignores any consideration whatsoever of meta-language and object language. While the claim forces the enumeration of the real numbers to be in a different language system (a meta-language) to the real numbers that are in the enumeration, at the same time the claim provides no logical step within the diagonal proof that proves that a real number is actually definable from that enumeration (as seen above in Section 2.3).[8]

The claim hence requires the following assumption:
Given any definition of an enumeration of a set of sequences of symbols of a language system, a diagonal number, which is associated with that definition of an enumeration, 'exists' independently of any definition, and there is no requirement that a method of actually defining such a diagonal number be proven.

---

[7] This is, of course, a dictionary style enumeration.

[8] Note that a similar claim is made regarding the power set proof and to Cantor's earlier 1874 proof, see Section 3.2 for details.

It has already been demonstrated in Theorem 1.1 that such an assumption leads to a direct contradiction, and is therefore untenable. The only remaining valid conclusion is that, for any given language system, the diagonal proof proves that there cannot be a function that enumerates all real numbers in that language system *and* which is in the same language system as those numbers.

## 3.1 Fully formal proofs of the diagonal proof

It might be noted that there are formal proofs of the diagonal argument; for example A Cantor Trio: Denumerability, the Reals, and the Real Algebraic Numbers by Ruben Gamboa & John Cowles. [9] There are also formal proofs of arguments similar to the diagonal proof and which also prove that there is no function that defines a one-to-one correspondence between the natural numbers and the real numbers, for example see The Real Numbers are Uncountable at the Metamath Proof Explorer. [10]

While there are claims that these proofs prove the general non-denumerability of the real numbers, in fact, such proofs only prove that there cannot be an enumeration function of all real numbers of a language, and which is in the same language as the language for those real numbers. Naturally, in a correctly formulated fully formal proof, it must be the case that any function that is to map the natural numbers to real numbers must necessarily be in the same language as the real numbers of the proof. Hence these formal proofs do not prove that there cannot be a language with a finite number of symbols and which can express all real numbers.

## 3.2 Applicability to all similar results

The conclusion above also applies to other proofs that give similar proofs to the naïve diagonal proof. This includes Cantor's 1874 proof[1] that there can be no function that gives a one-to-one correspondence of real numbers and natural numbers. It also applies to the power set proof, which is essentially Cantor's original diagonal proof in another guise; the power set proof states that, given a set A with an infinite number of elements, there cannot be a function that provides a one-to-one correspondence of every element of each element of the Power Set of A to each element of the set A. As for the diagonal proof, the power set proof is only valid for a function that is in the same language as the sets referred to by the function.

# 4 Appendix: Dual Definition

A common criticism of the diagonal argument is based on the fact that if a number has a finite binary definition, it will also have a definition which results in a sequence that, after some finite digit the sequence continues as an infinitely continuing series of 1s, or as an infinitely continuing series of 0s. For example, 0.1011101, which is $^{93}/_{128}$ in standard base 10 fractional notation, can also be defined by 0.1011100111... where the expansion continues with infinitely many 1s. The criticism is that there could be an enumeration which results in a diagonal number that continues after a certain digit with infinitely many 1s and which has the same numerical value as a number in the enumeration and which has a finite binary

[9] http://www.cs.uwyo.edu/ ruben/static/pdf/cantor-trio-final.pdf

[10] http://us.metamath.org/mpegif/mmcomplex.html#uncountable

definition; or an enumeration which results in a diagonal number that continues after a certain digit with infinitely many 0s and which has the same numerical value as a number that continues after a certain digit with infinitely many 1s.

However, there is an elementary method for obviating this problem: instead of operating on single digits, define the diagonal number in terms of consecutive pairs of digits of the $n^{\text{th}}$ number in the enumeration, according to the following rule:

If the $(2n-1)^{\text{th}}$ and $(2n)^{\text{th}}$ digits of the $n^{\text{th}}$ number in the enumeration are 1 and 0 respectively then the $(2n-1)^{\text{th}}$ and $(2n)^{\text{th}}$ digits of the diagonal number are 0 and 1 respectively, otherwise they are 1 and 0.

In this way, the diagonal number cannot be a number that continues with infinitely many 1s after some digit. Furthermore, since the diagonal number does not terminate, it cannot have a dual representation, so that it cannot be a number that continues with infinitely many 1s after some digit and for which there is a finite representation.

# References


[1] Georg Cantor. Ueber eine eigenschaft des inbegriffs aller reellen algebraischen zahlen. *Journal für die reine und angewandte Mathematik*, 77:258–262, 1874.

[2] Georg Cantor. *Über eine elementare Frage der Mannigfaltigkeitslehre*. Druck und Verlag von Georg Reimer, 1892.

[3] Georg Cantor. Briefe, ed. Herbert Meschkowski, Winfried Nilson, 1991.

[4] Wilfrid Hodges. An editor recalls some hopeless papers. *Bulletin of Symbolic Logic*, 4(1):1–16, 1998.

[5] Julius König. Über die Grundlagen der Mengenlehre und das Kontinuumproblem. *Mathematische Annalen*, 61(1):156–160, 1905.
English translation online (On the foundations of set theory and the continuum problem): "https://www.jamesrmeyer.com/infinite/konig-on-foundations-english.html".