Naive Introduction To Diagonal Constructions

In the proceeding an attempt to present an informative and non-formal introduction to the method of diagonalization is given. Diagonalization is a process of constructing from a set of things a thing which is distinct from each member of the set. The use of the term naive in the title, is to signify that where we are not formal in this introduction, we will be relying on an appeal to the readers experience in order to attempt to convey understanding. In the following our informal definitions are presented.

Naive Definitions:

Definition 1.

1. When S is said to be a set, we mean S is a collection of things, where the word thing will be allowed to be any concept or object. To motivate a set we think of a basket which has things placed in it, where the set is not the basket but rather the uniqueness of the totality of the things collected in the basket. So when the basket contains only a banana and an orange, this is a set which is different from when the basket contains a banana an orange and an apple. And the empty basket would constitute a set of nothing. Again the basket is meant to motivate. When we write $x \in S$ we are saying that x is a thing in the set S. One of the defining properties of a set is as follows:

   For all points x, given we have both

   $$x \in S \text{ yields } x \in X$$

   and

   $$x \in X \text{ yields } x \in S,$$

   then we will have $X = S$

   Which in essences states that sets are completely determined by the things they are a collection of. Taking x is an element of S to be $x \in S$ the last statement could be restated as: Sets are completely determined by their elements. We will make it clear later why formally our sets are not only restricted to being collections of mathematical things, but also why we can not consider any arbitrary collection of mathematical things a set.
2. Common notation for sets uses left and right brackets $\{, \}$ to signify a basket and then what sits between the brackets is the content of the basket defining the set. Note distinct elements of a set will be separated by a ,. So $\{a, b, c\}$ is the set of the things a, b, c where $a \in \{a, b, c\}$, $b \in \{a, b, c\}$, $c \in \{a, b, c\}$.

We note from the above that

$$\{a, b, c\} = \{a, c, b\} = \{b, a, c\} =$$

$$\{b, c, a\} = \{c, a, b\} = \{c, b, a\}$$

and that there is no point in listing an element of a set twice. When dealing with sets that are defined by a formula $\phi(x)$ and which are to hard to write in basic bracket notation, we may use the notation:

$$\{z : \phi(z) \wedge z \in U\}$$

where this notation is saying that $a \in \{z : \phi(z) \wedge z \in U\}$ when and exactly when $\phi(a)$ is true and $a \in U$. Here we are taking $\phi(z)$ to be making a statement about z, and then $\phi(a)$ is the statement made about a in place of z, where then the statement $\phi(a)$ is true. Note that the $\wedge$ symbol is the formal symbol for (and). Also it should be noted that the set U is appearing because later, as we shale see, not having the set there in the general form, will lead to contradictions. As a last moment note we should point out that the empty set is denoted as $\emptyset$ and it can be attained from any set U as $\emptyset = \{z : z \in U \wedge z \neq z\}$.

3. A function will be a very specific kind of mapping, where a mapping is a process of taking things to things. Indeed the function is meant to model the action of taking a thing and putting it with another thing. Here the key property is that we can not take a thing and put it whole with many different things. Again this is motivational and only to be taken to the point that one has an image of what is being talked about. So if x is a thing, then a function can take x to a thing y but if $z \neq y$, then the function should not take x to both y and z. Note $\{y, z\}$ is a set and sets are things, and thus a function is free to take x to $\{y, z\}$ and still have only take x to one thing. So functions are very specific maps and what are the points of discourse matters, but in any context they will always take a thing to only one point of discourse no matter how many things they act on. Here the set of things a function acts on will be called its domain, while the set of things that elements from its domain are mapped into will be called its target or codomain. We write $f : A \to B$ to mean that f is a function with domain A and codomain B. Given $f : A \to B$, for $a \in A$ we will write $f(a)$ for the unique thing in B that a is sent to under f. It should be noted that it can happen that $f : A \to B$ and for $x \in A$ and $y \in A$ with $x \neq y$ we have $f(x) = f(y)$, which highlights the fact that saying that $f(x)$ is a unique thing that x is mapped to in B is only stating that f takes x to only one thing in B and this allows that it is possible for other things in A to be taken to $f(x)$ also (naturally if y is taken to $f(x)$ we have the above $f(x) = f(y)$).
4. Given $f : A \to B$, the range of f will be $\{f(z) : z \in A\}$. Note here that $\{f(z) : z \in A\}$ is being written in place of the more formal:

$$\{b : b \in B \wedge (\text{There exists } z \in A \text{ where } f(z) = b)\}$$

And this practice of omitting understood components of the formal definition is often used when the formal form is too long.

5. When $f : A \to B$ and $\{f(z) : z \in A\} = B$, we say that f is onto.
6. When $f : A \to B$ and for all $x \in A$, $y \in A$ we have that if $x \neq y$, then $f(x) \neq f(y)$, then we say that f is (1 to 1).
7. If $f : A \to B$ where f is (1 to 1) and onto, then we will say that f is a bijection or a complete pairing. Here the facts that :

    i. f takes each thing in A to a thing in B

    ii. No two things in A go to the same thing in B

    iii. Everything in B has something in A which is taken to it

    Yields that our map completely pairs off the things in A with the things in B. Since they are paired off we take them to have the same number of things. So the number of things in A is identical to the number of things in B. However perhaps number imports to much because we might have both A and B be infinite and then we don't consider their pairing to say that they each have a number like 5 assigned to them. To be clear when there is a bijection from A to B, then we will say that A and B have the same cardinality. In a complete development of set theory we have a standard for cardinalities and each degree of the standard is called a cardinal number, so that A and B having the same cardinality will mean they are in bijection with a set called a cardinal number.

Diagonal Constructions

Definition 2. An ordered quadruple $\langle I, X, f, \phi \rangle$ gives a template for a diagonal construction if and only if

1. $f : I \to X$ is onto.
2. There is a y so that $\forall a \in I[\phi(a, f(a), y) \wedge \neg\phi(a, f(a), f(a))]$

In the above $\forall$ is short hand for (for all). It should also be noted that in condition 2 of the above definition we are asserting that the formula is true for some y. So the above definition says, that f is onto and that the formula ϕ with y put in it is true for each $a, f(a)$, and the formula is false if y is replaced with $f(a)$ for each a. Also note that $\neg$ is the logical (not) symbol. We also note that because f is onto and we have in condition 2 that for all $a \in I$, $\neg\phi(a, f(a), f(a))$ is true while for all $a \in I$,

$$\phi(a, f(a), y)$$

is true, it must be the case that $y \neq f(a)$ for any $a \in I$, and thus $y \notin X$.

Examples

Example 3. From Cantors Work

* We take $\mathbb{N}$ to be the counting numbers with zero, $\mathbb{Q}$ to be the rational numbers, $\mathbb{R}$ to be the real numbers, and $\pi_n : \mathbb{R} \to \mathbb{N}$ to be the function where $\pi_n(r)$ is the non-negative number from 0 to 9 that is the nth term after the decimal of the decimal expansion of r, where whole numbers will be taken to have all zero's after the decimal. We give $\phi_1(z, x, y)$ as the formula $\pi_z(x) \neq \pi_z(y) \wedge y \in \mathbb{R}$.

** Taking $f : \mathbb{N} \to \mathbb{R}$ to be any function, and letting $X = \{f(n) : n \in \mathbb{N}\}$, we will have our quadruple be:

$$\langle \mathbb{N}, X, f, \phi_1 \rangle$$

.

We show that $\langle \mathbb{N}, X, f, \phi_1 \rangle$ is a triple for a diagonalization by showing that for some y, ϕ_1 meets condition 2 of our definition. To the aforementioned ends we give $k \in \mathbb{R}$ so that k is between zero and one, and for all $n \in \mathbb{N}$ we have:

$$\pi_n(k) = \begin{cases} \pi_n f(n) + 1 & \text{if } \pi_n f(n) = 0 \\ \pi_n f(n) - 1 & \text{if } \pi_n f(n) > 0 \end{cases}$$

We now note that for all $n \in \mathbb{N}$ we have by construction that

$$\pi_n k \neq \pi_n f(n) \wedge k \in \mathbb{R}$$

so that $\phi_1(n, f(n), k)$ is true for k and all n. To see that condition 2 is met it only remains for us to show that for all $n \in \mathbb{N}$, $\neg\phi_1(n, f(n), f(n))$ is true. To see this we note that $\pi_n f(n) \neq \pi_n f(n)$ is always false, and therefore $\neg(\pi_n f(n) \neq \pi_n f(n))$ is always true and thus

$$\neg(\pi_n f(n) \neq \pi_n f(n)) \vee \neg(f(n) \in \mathbb{R})$$

is always true where then by de-Morgan's law's we then have

$$\neg(\pi_n f(n) \neq \pi_n f(n) \wedge f(n) \in \mathbb{R})$$

is true for all $n \in \mathbb{N}$ and thus

$$\neg\phi_1(n, f(n), f(n))$$

is true for all $n \in \mathbb{N}$. From the preceding we have

$$\forall n \in \mathbb{N}[\phi_1(n, f(n), k) \wedge \neg\phi_1(n, f(n), f(n))]$$

. Note that $\vee$ means (or) and that A or B is true exactly when atleast one of A or B is true, and note that de-Morgan's Law's give that

$$\neg(A \wedge B) \Leftrightarrow (\neg A) \vee (\neg B)$$

which means $\neg(A \wedge B)$ and $(\neg A) \vee (\neg B)$ always have the same truth value.

*** Here we note the important property of the diagonalization template, which is $k \notin X$. We take k in the above to be the diagonalization of X with respect to f given by the template, and we will refer to k as the diagonalization when non-ambiguous.

**** From the above we have that for every $f : \mathbb{N} \to \mathbb{R}$, given

$$X = \{f(n) : n \in \mathbb{N}\}$$

we have a diagonalization k for f and thus $k \notin X$. We now note that there is an $f : \mathbb{N} \to \mathbb{R}$ so that $\mathbb{Q} = \{f(n) : n \in \mathbb{N}\}$ and then we have that the k defined above for this f would yield $k \in \mathbb{R} - \mathbb{Q}$ and thus k is irrational.

***** We can see that there can not be an $f : \mathbb{N} \to \mathbb{R}$ where $\mathbb{R} = \{f(n) : n \in \mathbb{N}\}$, since then the k defined above which is the diagonalization for this given f yields $k \in \mathbb{R} - \mathbb{R} = \emptyset$ and since $\emptyset$ is the empty set and contains no elements we have a contradiction. This last result is a proof that the cardinality of $\mathbb{R}$ is strictly greater than the cardinality of $\mathbb{N}$.

Example 4. From Cantor's Work

* We will take $P(I)$ to be the set of subsets of I where S is a subset of I if and only if it is the case that if $x \in S$, then $x \in I$. An example of a subset is: $\{1, 2\}$ is a subset of $\{1, 2, 3\}$ and then

$$P(\{1, 2, 3\}) = \{\emptyset, \{1\}, \{2\}, \{3\}, \{1, 2\}, \{1, 3\}, \{2, 3\}, \{1, 2, 3\}\}$$

. Note that $P(\{1, 2, 3\})$ is a set whose elements are sets and that the empty set $\emptyset$ and all of $\{1, 2, 3\}$ are elements of $P(\{1, 2, 3\})$. Also note that $1 \notin P(\{1, 2, 3\})$, $2 \notin P(\{1, 2, 3\})$ and $3 \notin P(\{1, 2, 3\})$.

** For $f : I \to P(I)$ we take $X = \{f(a) : a \in I\}$. We will have $\phi_2(z, x, y)$ be the formula $y \in P(I) \wedge (z \in y \leftrightarrow z \notin x)$

*** We now show that $\langle I, X, f, \phi_2\rangle$ is a template for a diagonalization. To these ends we take $k = \{a : a \in I \wedge a \notin f(a)\}$. We now have that $\phi_2(a, f(a), k)$ holds by the construction of k. We note that $f(a) \in P(I)$ for all I and that $a \in f(a) \leftrightarrow a \notin f(a)$ says that a is in $f(a)$ when and exactly when a is not in $f(a)$ and clearly this is never going to be true. We now have that

$$f(a) \notin P(I) \vee \neg(a \in f(a) \leftrightarrow a \notin f(a))$$

is true since $\neg(a \in f(a) \leftrightarrow a \notin f(a))$ is true. It then follows from de-Morgan's laws that $\neg\phi_2(a, f(a), f(a))$ is true for all $a \in I$. We now have

$$\forall a \in I[\phi_2(a, f(a), k) \wedge \neg\phi_2(a, f(a), f(a))]$$

is true and we have that condition 2 is satisfied for $\langle I, X, f, \phi_2\rangle$.

**** From the above we have for all I for all $f : I \to P(I)$ the k constructed above as the diagonalization for f yields that $k \in P(I) - \{f(a) : a \in I\}$.

***** It follows from the above that there can not be a function $f : I \to P(I)$ where the range of f is $P(I)$, since then we have $P(I) = \{f(a) : a \in I\}$ and consequently the k we have where k is a diagonalization for f yields $k \in P(I) - P(I) = \emptyset$ and a contradiction is reached.

****** So from the above we have that $P(\mathbb{N})$ has cardinality strictly greater than the cardinality of $\mathbb{N}$. It turns out the cardinality of $\mathbb{R}$ is equal to the cardinality of $P(\mathbb{N})$.

Example 5. From Russel's Work

* For a set X, let $id : X \to X$ so that $id(z) = z$ for all $z \in X$. id is called the identity function because it takes things to themselves. Our $\phi_3(z, x, y)$ will be the formula $(z \notin x \leftrightarrow z \in y)$.

** We now take our quadruple to be $\langle X, X, id, \phi_3 \rangle$ where X is a set of sets. To see that the given quadruple is a template for a diagonalization we will take $k = \{z : z \in X \wedge z \notin z\}$. Provided everything is a set, we now have that by definition for all $z \in X$

$$z \notin id(z) \leftrightarrow z \in k$$

is true, noting that $id(z) = z$. Since it is not possible to have a z where $z \notin id(z) \leftrightarrow z \in z$ we have

$$\neg(z \notin id(z) \leftrightarrow z \in id(z))$$

is true for all $z \in X$. The preceding yields that for all $z \in X$ we have:

$$(z \notin id(z) \leftrightarrow z \in k) \wedge \neg(z \notin id(z) \leftrightarrow z \in id(z))$$

The preceding yields that

$$\forall z \in X[\phi_3(z, id(z), k) \wedge \neg\phi(z, id(z), id(z))$$

is true and therefore the second condition for $\langle X, X, id, \phi_3 \rangle$ being a template for a diagonalization is met. Noting that the first is trivially met, we have that k is the diagonalization for id. We note that similar to our other examples we have $k \notin X$.

*** As was done in the preceding examples we will take the range of our function to be too large to be the range of a function with the specified domain. Since id will always be a function and since our domain and range are the same in this example, we will be forced to conclude that the range can't be a set. To this end we take X to be the set of all sets and we note if k is the diagonalization of $id : X \to X$ given above, then $k \notin X$ and a contradiction with k being a set is met. It follows to avoid a contradiction we must assume that there can not formally be a set of all sets.

**** This example is very important. It yields that we can not take any collection of things to be sets in a formal set up for set theory. The resolution that is given for our set theory comes in two parts.

1. (Naive Axiom of Comprehension(see next page for formal)) In order to keep the useful set builder notation we restrict it to allow only for new sets to be given as formulated subsets of an existing set. And thus we take as an axiom: For all sets S and all formulas $\phi(z)$ with only the free variable z we have
$$\{z : z \in S \wedge \phi(z)\}$$
is a set. This is how I gave the definition of the set builder notation at the beginning of this, so it wouldn't present itself as a modification. But originally mathematicians had tried to take $\{z : \phi(z)\}$ as a set for all formulas $\phi(z)$ with z free. In essence this type of formulation allows for any collection of things with a formula to be a set, and the above example just showed that that leads to contradictions. Note here free means that z is a variable we can freely substitute things for. So $z + 1 = 2$ is a formula with z free and then $\{z : z \in \mathbb{N} \wedge (z + 1 = 2)\} = \{1\}$. Also note that our formulation is naive in that we will allow for our formula to import things like 1, 2 and + where in set theory these are in fact defined sets and not fixed constants. So we in formal practice may have our ϕ have more than one free variable and then in the formal axiom schema of comprehension all variables are quantified. See the below note for the formal comprehension schema.
2. (Axiom of Foundation) If X is a set and there exists a $y \in X$, then there exists a $z \in X$ where for all $y \in X$ we have $y \notin z$.

   The Axiom of foundation states that all non-empty sets must have an element that does not contain any other element of the set as an element. From here we get that there is no set X which is the set of all sets, because then $X \in X$ and we have that the set $\{X\}$ contradicts the axiom of foundation. Which is good since such an X leads to contradiction. We also note that there are more axioms in set theory than the axiom of comprehension and foundation, and these axioms allow for things like the formulation of $\{S\}$ for all sets S.

   For the sake of a degree of completeness, there is the power set axiom which gives: For all sets Y there exists a set $P(Y)$ where $S \in P(Y)$ if and only if S is a subset of Y.

   We then have from the power set axiom and comprehension that for all sets S,
$$\{z : z \in P(S) \wedge z = S\}$$
is a set. And then since
$$\{z : z \in P(S) \wedge z = S\} = \{S\}$$
we have for all sets S, $\{S\}$ is a set.

Formal Axiom of Comprehension schema: For all n and for all ϕ with free variables $z, w_1, ...w_n$ we will have

$$\forall Y \forall w_1, ..., \forall w_n \exists X \forall z[z \in X \leftrightarrow (z \in Y \wedge \phi(z))]$$

Here it should be noted that $\exists$ is the symbol for there exists. It also needs to be noted that in mathematics we are very particular about there being a difference in meaning between saying $\forall z \exists X$ and $\exists X \forall z$. When we say $\forall z \exists X$ we mean that for any z given an x can then be found for the given z. When we say $\exists X \forall z$ we mean that single X can be given for all choices of z.