ZFC - Empirical

Andreas Reif

Allmendingen, Germany, article@andreasreif.de

15th August 2015

Empirical science should have a logically rigid foundation of its basic assumptions. What is the minimum aggregation of rules to make science empirical?
The ZFC[1] axioms[2] applied to experiments provide this foundation. An empirical validation, especially an experiment, is valid, if and only if the axioms of ZFC are applied in the following manner.
At each definition or axiom, the original set theory is presented, and explained. Then the interpretation for empirical science follows.

Basic Definitions 1.: Most basic, there is a realm[3] of things. Classes exist in the realm. A class exists if the realm holds an individual thing of that class. Sheer existence is enough to make a thing part of the realm.
Empirical science is the class of existing phenomenas, actings or descriptions which tell about existing or non existing things. It assumes that there is more than empirical science, this greater area is the realm.
Example: To proof light emission by electrical, resistive heating, a mechanism made publicly available during the 19th century, a circuit with a battery, wires and a light bulb is used. This technology serves as a minimalistic example throughout this article. The technical components, as they lie in front of the experimentor, form a set. These

---

[1] Ernst Zermelo, „Untersuchungen über die Grundlagen der Mengenlehre I“, (Investigations in the foundations of set theory I), 1908, `http://gdz.sub.uni-goettingen.de/dms/load/pdf/?PPN=PPN235181684_0065`, p.261, abbrev.: 1908b. Axioms of Foundation, Replacement and Pairing: Über Grenzzahlen und Mengenbereiche, 1930, `http://matwbn.icm.edu.pl/ksiazki/fm/fm16/fm1615.pdf`, abbrev.: 1930. Axiom of infinity is used despite 1930, p.30

[2] For the axiomatic method see Adolf Fränkel, „Einleitung in die Mengenlehre“, (Introduction to set theory), 3rd edition, Berlin, 1928, `http://gdz.sub.uni-goettingen.de/dms/load/pdf/?PPN=PPN373206852`, p.268

[3] German: 'Bereich' or 'Grundbereich', common translations: 'area' or 'basic domain', for example in: Ebbinghaus et al., „Ernst Zermelo - Collected Works/Gesammelte Werke: Volume I/Band I - Set Theory, Miscellanea/Mengenlehre, Varia“, 2010, p.180, `http://books.google.de/books?id=XB2nd2ovakIC`. The common translations are too strongly associated with control or boundaries which does not match the original notion. Realm as used here has the same meaning like in the phrase 'the realm of possibilities'. Although here it is not the realm of possibilities, but of the existing things.

means and all other things necessary for the experiment are part of the realm. So the air the experimenter needs for breathing is part of the realm, as it has to be there to run the experiment, although it is not mentioned.
Counterexample: Yesterday, the experimenter made a simple working circuit. He thinks he used blue wires, but it were the red ones. So the experiment as he remembers it has never happened and does not belong to the realm of existent things.
Counterexample: Steps to construct a square from a circle with the same area by compass and straight-edge construction is not existent and is therefore not part of the realm.
Example: The attempt to construct a square from a circle is an event, has happened and is therefore part of the realm.

Example: The simple experiment to proof effects of electricity assumes that it is done in the usual environment of humans, that is in the atmosphere. This atmosphere has certain electrical properties not explicitely defined. It works as an isolator. In vacuum the experiment may fail because there is a short circuit of the batteries' poles so the bulb may not glow. The atmosphere as an environment is supporting the experiment although usually not desribed.
Counterexample: Given the assumption, that everything is connected with everything in unexpected ways. If this would hold, nothing can be distinguished from another.

2.: In a realm there are basic relations (German: Grundbeziehungen) between things a, b in the form $a \in b$ (a is element of b). A thing is a set if and only if it can comprise another thing; except the null set, which is also a set and includes nothing. Sets are a part of classes.
In empirical sciences the elements can be analyzed further and so science describes elements containing other elements.
Example: The atoms were considered as not dividable, hence their name. It turned out that they consists of smaller particles. Independent of the question if the set theory considers urelements (elements which do not contain other elements), the elementary particles can be considered as sets. If a researcher assumes there are undividable parts to be found, those would be urelements in his theory.
Example: Set of all electromagnetic waves with a wavelength of smaller than 400 nm. The set contains the wave but a physical container for them is not stated by that set definition.
Example: Set of all experiments not confirming the hypothesis. No physical container for the experiments is stated, nevertheless the set contains them logically.

3.: If any element of N also element of M, so N is a subset of M. Is $M \in N$ and $N \in R$ so it is $M \in R$. If no elements are shared between M and N, or if no element of M is also element of N at the same time, they are disjunctive.
In empirical sciences the transivity of enclosure is basic for every analysis envolving more than one level. The enclosure can be pyhiscally (Example: An atom encloses particles.) or logically (Example: All electrons which are in a certain shell.). Counterexample Transitivity: An electromagnetic wave does not contain a particle, nor vice versa. Con-

tainment is not necessary to make a theory scientific.
Example for disjunctive elements: Experiment with carbon fibre in laboratory 1. Same experiment at the same time in another lab.
Counterexample for disjunctive elements: Like example before, but lab 2 lends battery from lab 1 to execute the same experiment later.

4.: A statement is called 'definite' (German:definit) if its validity is decided by the basic relations empowered by the axioms and the general logical laws. A class statement (German: Klassen-Aussage) $\mathfrak{E}(x)$ in which the variable x can become all objects of the class $\mathfrak{K}$ is also definite if it is definite for each individual object of $\mathfrak{K}$.[4]
In empirical science, the generalization to several cases which are not yet individualized (all elements of a class) is a definite statement if it is valid in all cases.
A state or any event in an experiment is an element in a set. A description, also the writing and reading of it are also elements of a set.
Example: The emission of an electron from an atom at a timeframe and in a defined space is an event and therefor an element.
Example: The screen of a computer which measures temperature in an experiment shows "50°F". This is an event and is therefore an element.
Example: The hypothesis that bodies which have weight attract each other by gravitational force is definite for all such bodies.
Example: The hypothesis that hooke's law holds for all springs is definite for all springs.
Counterexample: There is a prediction that the stock price of the share of company XYZ will soon recover. This is not definite for all periods of time, because 'soon' is not clear.
Example: The prediciton itself is an element.
Counterexample: The shaman can heal those who are pure at heart. This is not definite because the criteria of being pure at heart cannot be decided regarding each individual.
Counterexample: 'The electric circuit XYZ works in the range of 5 to 20 Volts in all applicable cases.' As this statements contains the general clause 'in all applicable cases' this statement is not definite if the applicable cases are not defined.

Axiom 1: Definiteness Two sets are equal(=) if they have the same and no other elements.
An event is the same if it is happening at the same time and space. (Identity) An event is equal if it is not happening at the same time but is otherwise not distinguishable from an event at another time. (Equalness)
Example for identity: According to the lab book the researcher A experienced the result r at Oct 10th, 2012. In an article, he referred to this result r. The reference to the result and the personal experience refer to the identical result r.
Example for identity: According to the lab book the researcher A experienced the result r at Oct 6th, 2012. In an article, he referred to this result r. The reference to the result

[4] $\mathfrak{K}$ old german for K, Klasse; $\mathfrak{E}$ old german for E, may stand for Element-Funktion(element function) or Entscheidung(decision).

and the experience refer to the identical result r.
Counterexample for identity: Result r war found at Oct. 10th, 13th and 30th of 2014. The results are not identical, they are equal.
Example for equalness, counterexample for identity: Researcher A sets up a certain experiment and states a hypothesis which he sees as confirmed if the bulb glows. He saw the bulb glowing in his experiment in 2010, like he saw it in 2008 and 2007. The glowing was not distinguishable for him so it is equal. The glowing is a set, the glowings are sets, the elements in each set are the same. But as they happen at another time, they are not identical. As it is the same experiment, the set of the experiments' means are elements. These are the same means and same procedures so the reruns are equal.
Counterexample for equalness: Researcher A saw the bulb glowing in his experiment. In a later repetition the bulb glows much shorter than the expected several seconds. He concludes that the repitition does not show the same result.
Example and counterexample for equalness: Researcher A did not see the bulb glowing in his experiment although his hypothesis predicts a glowing. He assumes an error in the experiment and looks for a fix in his experiment or in his hypothesis. The not glowing is not distinguishable from doing nothing with the bulb. It is an empty set, all empty sets are equal. This can not be used to confirm a hypothesis because the falsifiability is not provided, for more on falsifiability see below at axiom of segregation.

Axiom 2: Elementary Sets There is an improper (an exceptional) set which does not contain elements (null set). There is a set which comprises any single element and no other element. If there are any elements a, b in the realm, there is a set which comprises those 2 and no other element.
Interpreting the null set for empirical science done in two steps:
1.: Most obvious the Null-Hypothesis H_0 correspondends to the Null-Set: If a hypothesis derive a H_0, the H_1 is not valid.
2.: More general, as any property can be seen as an hypothesis: At any time and space, any property can be defined as zero at least as a point of reference. Each property can only be defined as a difference to a virtual case in which it is zero.
This holds even if a property an only have 2 states. In this cases it is the difference to a state (or a sequence of it) which can be zero.
Example: The fact that zero can not be achieved in reality due to inevitable quantum fluctuations complies with the axiom. These fluctuations can only be defined if an absolute zero value is defined. So a zero value is most basic for any description, even and especially for near null phenomena.
Example: The Casimir experiment assumes that if the plates would not experience a force towards each other there would be zero force.
Counterexample: If it is not possible to assume the possibility of zero force, the casimir experiment would not be possible.
Counterexample: The gnostic idea that empty space is filled with power which can manifest forces. This as a basic idea excludes the possibility of null and makes any existence of any phenomena and its non-existence explainable. So it must be excluded by empiri-

cal science because it does allow to reject any forecasts.

Interpreting elementary sets with 1 element for empirical science: The axiom states that each single mean of an experiment may be uncovered & identified. If this is not possible, the experiment would introduce magic as the functionality and properties of each mean could not be defined.
Example: The researcher R examines the battery of a measurement device. He wants to be sure that it has enough voltage to make the device work properly. So the battery can be identifiable (as one individual thing) and forms a set of technical means containing a single element.

For elementary sets with 2 elements, see Axiom of Pairing below.

Axiom 3: Segregation The class statement (German: Klassenaussage) of M is definite to all elements of a set M, so a subset of M exists which contains only elements for which the class statement is true.
Explanation within set theory: An element's property which decides if this element belongs to a certain set is either true or false. Otherwise it is not definite - having the property is not defined. Therefore it does not belong to the set nor its complementary set. There is subset of each set for which the property is true. Segregation limits the sets to a size so they are free of contradictions. Sets are not to be built independently (abstractly), but are built up from the bottom of basic relations (German: Grundbeziehungen des Bereichs). The criteria must be proofed being definite for each element. If the relation 'is element of' cannot be proofed, it is not a set.(Removes Richard's and Russell's paradox.)

For empirical science, this axiom offers several aspects, which are considered as most basic in the scientific community. These aspects are the steps in the researcher's working cycle: First, a general statement is created which may be a progress in theory, in empirical science a hypothesis, in set theory a set definition. It is meant to be applied by deduction. Second, the possibility of failing in one or more cases: The priority of the cases over the general statement what validity is concerned. (Induction) Third, the solution by segregating the cases which do not fit.

1.) Deduction is prior for making progress in theories. (While confirmation is pending.) After a phase of creativity a researches asks: Is a hypothesis H_1 valid for a lot of cases? The claim, the outlook for a possibly successful deduction - mindful created as a probe for real, repeated doing in future. This is a possible progress in science. In set theory, the axiom of segregation assumes that the mathematician expresses the definition of a set. After that there is a step of validation. So a new definition of a set are the first steps to make the world of known[5] logical statements grow. So creating a hypothesis is a way to say: There is a set in which elements have $\mathfrak{E}(element) = true$.

[5]„Known "to avoid the discussion if mathematical statements are found or invented. The amount of known statements is independent of this distinction.

2.) Induction is prior to deduction for validation. If a single element on which a statement applies, fails, this is a contradiction which has to be handled.
Possibility of failing is the price of generalization. But failing is a very anthropomorphic notion, what corresponds to this in set theory: If a definition of a set leads to an empty set or the definition leads to a logical object which is not a set, but only a class, because the element function is not definite for all elements. The result 'empty set' can be a desired outcome in a proof, so it may be a success depending on the context. It still comes close to the notion of failing, because the definition does not yield a result, it 'fails'. In empirical sciences, an hypothesis can fail by finding a counterexample. So the possibility of failing is the core of empirical and mathematical scientific statements.[6] So both approaches consider the possibility of failing and consider a fail as a contradiction which has to be handled in a further step.
3.) Solution by segregation. In set theory, the contradiction is resolved by diminishing the elements from the set the original statement creates. So the set has a smaller size then. The contradicting cases are segregated. In empirical science, the hypothesis must be diminished, because one contradicting case is enough to make it invalid. Alternatively, the hypothesis is refined so the previous statement can be kept after adding a complement. ('All swans are white, exceptions are: ... ' saves the 'all swans are white' theory which then becomes a more sloppy 'swans in general are white'-theory.)

Example from the technical world for keeping a hypothesis and diminishing another: A 19th century experimentator who wants to create a light bulb uses a fiber too thick and long to get it properly heated. The idea of light emitting by electrical heating was defined by: There is a set of things which allow light emitting by electrical heating which is durable enough to everyday appliance. This is the hypothesis, in the terms of set theory. Refined by a fail: If it works at all, it must be done differently. Which is just holding up the belief and refine it by segregating the contradictive cases. These cases are not included in the original, too big, not reasonable sized set.

Example (change of implicit definition): The hypothesis that particles cannot travel by speed of light contradicts the fact that photons travel by speed of light and are also particles. So the hypothesis may be read as concerning 'particles except photons'. (Hypothesis contradicts existence of other elementary particles traveling by speed of light,

[6] By setting induction prior for validation, hypotheses which do not contradict any theoretical outcome of an experiment are segregated. Falsifiability means to define, which outcome would contradict the hypothesis. Falsifiability lacks if no experiment(experience) can be mentioned which would contradict the hypothesis if experienced. Such an hypothesis is not 'always true', but a sure fault. As no virtual experiment can falsify it, the hypothesis does not belong to the class of things on which empiricism rules. More details how this ZFC-Empirical maps to falsificationism will be part of future publications. The main logical difference is that falsificationism states hypothesis as 'all x are/lead to y' while ZFC states hypthesis as 'there is a set M of elements m_n which have the property P'. This primary focus on the single element excludes contradictions: If a statement does not hold in one single case, for falsificationism, it is not valid, for set theory, the element just does not belong to the set (reasonable sized sets). Set theory is the more general theory and will provide very practical insights for the creation and basic structure of a hypothesis.

too.)

Counterexample (too general hypothesis): Hypothesis given: 'A carbon fiber emits light when heated.' This is wrong in this generality. For example, a temperature above about 480°C is required. It can be read - in everyday speech - as 'can emit light' when some conditions are met. In scientific speech, it is good practice to mention the most important condition, like an important threshold. It can also be read as statement of existence: Somewhere, somehow light is emitted by heating a carbon fiber. Such an statement is not falsifiable and therefore not part of an empirical science.

Category of counterexamples for missing segregation: A hypothesis is held dispite the contradicting facts. The known counterexamples are only accepted by the word of mouth if at all. (Emperor's clothes syndrom, spiral of silence)

Counterexample - contrast case for validating vs. finding a hypothesis: Experimentator does an experiment and fails. Later he succeeds by defining a new hypothesis with creativity techniques and executes a confirming experiment. The finding of a new hypothesis in this case is (as usually) creativity in a wider sense, not confirmation itself.

Example: The unproven hypothesis that particles can be accelerated above the speed of light contradicts the fact that acceleration devices do not manage to do so. So either the hypothesis cannot be held up, or the runs cannot be generalized. (Statements of existence cannot be diminished, but made implausible.)

Axiom 4: Powerset of subsets There is a set which contains all and only the subsets of a set. Example: $A = \{a, b, c\}; \mathcal{P}(A) = \{\{\}, \{a\}, \{b\}, \{c\}, \{a, b\}, \{a, c\}, \{b, c\}, \{a, b, c\}\}$
Interpretation for empirical science: The means of an experiment can be grouped and examined arbitrarely. The outcome of the examination can vary with each grouping. The axiom adresses control. The means of an experiment can be rearranged, changed - this must not be limited by methodology. It defines the control the experimenter has over his experiment. There are practical limits like limited resources, but it must be possible that the experiment's means can be rearranged, they build a set, the possible changes are the powerset of this set.
Example: The simple circuit can be rearranged. The wires can be connected differently, so the wire 1 does connect plus pole, then connects negative pole with the bulb und wire 2 vice versa, the bulb will still glow. The experimenter can disconnect one wire, the bulb remains dark.
Counterexample: The simple ciruit is casted within a beautiful design lamp, given it is built in a way so that the light bulb gets destroyed if a modification is tried, for example to avoid reverse engineering to protect intellectual property of the design. In this case the axiom is violated and the pocket lamp does not proof light emmission by resistive heating by electricity.

Axiom 5: Set Union There is a set S which contains all and only the elements of the elements of a set A. The order of mentioning or adding does not matter. The commutative, associative, distributive laws hold, interpreted for the general case of sets.
If an experiment only works in one setting without tolerance, it is not clear if the output yields from the stated rule or from a factor which is hidden by the overly restrictive conditions. Tolerance is also required for the definition of the confirming output.[7] Otherwise the passing corridor would be infinitely sharp and no confirmation would be possible. Tolerance requires the definition of a union set to define a range of values. The axiom also states that they are independent from their sequence of their appearence. So the results can vary randomly within the tolerance, so they are a confirmation.

An experiment must be able to change within tolerances and the output must be equal within tolerances, as long as the interesting mechanism is stable. The axiom also enables tolerance about time. So an experiment stays the same if executed at another time. Isomorphism of time results from containment relationship: An experiment is valid for all times, so it is valid for each defined time in any defined interval.
Example: The light bulb glows, then the wires are disconnected, it glows for 50 seconds. Therefore effects of electricity are proofed.
The timely order of the setup is not necessary per se for the necessary conditions of an experiment. Order is only a matter of human efficiency.
Example: To achieve a salted solution, first fill pure water into a bowl, then pour the salt in the vessel. A different sequence yields the same result.
Example: To close the circuit, connect wire one with the plus pole of the battery to the bulb. Then connect the minus pole to the other connector of the bulb. The other sequence yields the same result.
Of course there are many, expecially technically and methodological timely orders which must be followed. But these must be part of the description. From the axiomatic point of view, the timely order does not matter without special mentioning.
For practical reasons, timely order matters a lot in everyday life. So the researcher first drinks coffee before he goes to work. He may search for the battery first, because he knows that it may take longer to find them. A chemist always puts water first in a test tube then the acid - to avoid foam. If these sequences are a matter of the subject he researches for, they must be stated directly or be very common to the community so they are implicitly defined.

It also provides tolerance of what is considered as a valid confirmation. It makes axiom 1 more concrete: It makes the definition of what is the same experiment applicable

[7] Popper, Logic, p.108f. A theory with higher precision is logically easier to falsify: „Thus the rule that theories should have the highest attainable degree of testability (and thus allow only the narrowest range) entails the demand that the degree of precision in measurement should be raised as much as possible.
It is often said that all measurement consists in the determination of coincidences of points. But any such determination can only be correct within limits. There are no coincidences of points in a strict sense. Two physical 'points' - a mark, say, on the measuring-rod, and another on a body to be measured - can at best be brought into close proximity; they cannot coincide, that is, coalesce into one point. ... Thus an interval, a range, always remains.“

to non-trivial cases. So it can be that A=B, not only A=A.

Axiom 6: Choice From every set T of non-zero sets (M, N, O, ...) there is a set S which is defined by exactly one element of each of the sets (M, N, O, ...).
We assume certain, important properties (hypothesis) and claim that we can get as close to the desired properties as our resources allow. The scientific procedure is getting closer to the truth by choosing predictions and finding their confirmations. A prediction (hypothesis) is chosen among the infinite set of possible descriptions of an experiment. And the means to reproduce the outcome which acts as confirmation are chosen from an infinite set of possible means.

A valid hypothesis is the same as the question of making definite statements, see Basic Definitions, 4. paragraph.
Example hypothesis: A battery will make a bulb shine if connected as follows ... . So electricity is proofed if the bulb shines.
Counterexample hypothesis: The talisman helps to stay safe. As it is unclear in which cases it applies and how the effect can be measured it is not definite.

Example for means: The wires can have any length between 5 and 100 cm. Assume the description of the experiment says, that less than 5 cm is difficult to handle and more than 100 cm may impact the performance. The researcher chooses, when executing the experiment, one possibility of this range to create an event in time and space equal in meaning to the description and different in time and space to experiments before (equal, not identical).
Counterexample for means: For a reason r1, the circuit example does only work with the exact length of 25 cm without any tolerance. This condition can never be met, because the exact value is never 25 cm, but perhabs 25.000231 cm (+/- measuring tolerance).
If the researcher can not choose within a range, he is bound to certain, infinitely defined means which never change. This would lead to magic means: Only by this certain battery or this certain wire with a certain length, not recreatable, electricity should be proven. But the researcher must be able to manipulate the battery and the wires in infinite ways within a more or less well defined range. Otherwise the experiment is not recreatable. It is not a proof of electricity, if only one single battery in the world would show a then only alleged phenomenon like electricity. It must be recreatable to be a scientific mean or result.

Axiom 7: Infinity by Induction There is a set which holds the zero set or any other element a and contains all other sets which comprise the sets containing the zero set or a. ({a},{{a}},{{{a}}},...)
An experiment must be repeatable. So it is a series of equivalent experiments, distinguished by time and space. By this axiom any correlation beside zero can turn a valid

theory in a reliable[8] one, depending on the criteria of the specific science.
Example: Batteries, wires and bulbs are available repeatably. So yet another experiment is possible.
Counterexample: If an experiment can only be done until a certain year, for example 2025, this leads to the question if there are influences which are not documented. The possibility of doing the experiment again must not face a magic constraint.
Example: Gravity can be demonstrated with any mass, especially with any handy object. So it can be demonstrated repeatedly.

Axiom of Foundation A set contains no infinitely descending (membership) sequence, so there is an element which does not have an element.
An experiment happens in a paradigma which serves as a frame. Within that frame, there is no infinite backward, descending sequence. Such a sequence is the infinite recurring of the means (things) and the methods of the experiment.
This addresses the following criticism: 1.: Not all the means are perfectly described by the experiment's description. The properties are not perfectly defined, there is no hint how to manufacture the means from the ground up. So much is left to the executor what to use, so in fact the executor has to find out how the means must be produced to proof the hypothesis. 2.: If it fails, the hypothesis may be still be hold up with regard to self-referential descriptions of the experiment. (Example: 'Set up a working circuit ...' The term 'working' already assumes and defines the successful outcome.)
Ad 1.: To any description, there is always a cultural context. In this context the description becomes clear. A caveman would not understand the description of the simple circuit. In science, the cultural context is the context of the scientific history. It's the recent history, the situation the scientist works in. So the experimenter does not have to build up all notions and methods form any absolute basic point, as there is no such point. Rather there are evident means he can use. So he does not have to know how ore is treated to get copper. Nor does he have to find all the necessary conditions along the way to build the means, like human interactions, impressions, needs or the convictions people had while creating one of the means. The axiom of foundation states that there is a reasonable size of necessary description.
Ad 2.: As there is a foundation in the chosen paradigma, the advocate of a hypothesis must give the foundation by providing available means of confirmation. The description can not be justified by claiming it was described 'as working', so it is valid, despite it doesn't work in real life.
Example: Given the minimalistic experiment with a battery, wires and a bulb. The criticism may be, that wires etc. already contain the ability to proof electricity by their definition and so the experiment is a tautology. For example, a wire is an element in this description. It is defined by the property to be a conductor. So it is already implying electricity. Any failing of the experiment would result in the ad-hoc hypothesis, that the

[8] „Reliability is the degree to which an assessment tool produces stable and consistent results.“in EXPLORING RELIABILITY IN ACADEMIC ASSESSMENT, Colin Phelan & Julie Wren,`https://www.uni.edu/chfasoa/reliabilityandvalidity.htm`, July 2005, 4th January 2015, 14:07. This is also important for any experiment, not only for testing personal metrics.

mean was not the right one.
Counterexample: Given is an aerodynamic experiment which examines the properties of the wing profile p1. The resulting forces are measured at different wind speeds and attack angles. A possible ciriticism is that the experiment is invalid because not every molecule of the air is examined. This criticism is wrong, the foundation lies in the measurable properties and controllable conditions of speeds and angles.

Counterexample: The physicist p gives an example: 'Imagine a space in which ...' A critic says that an imagined space may have different properties than a real space. But as the properties of the imagined space are defined by the physic, the physicist can reject this critic by the axiom of foundation.
However, the axiom does not guarantee that an imagined space has the same properties as a real one. It only guarantees that thought experiments can not be rejected, just because they are virtual. 'Virtual' just means that there is a sufficiently small size of the description to have the human intellect understanding the mechanism. Of course the validity and reliability is not as good as a real world experience, because in empirical science, empirical knowledge is valued higher than knowledge gained by conclusion.

Axiom of Replacement If the elements x of a set m are replaced by x', then there is a set m' containing x'.
The means that make up the experiment (any set) can be changed, so the mechanism can be interrupted at any technical possible way. An experiment can become another experiment, typically a variant or an uninteresting or an invalid set. The possibility of the experiment to become invalid (having another value as the former set) is necessary part of any experiment. The replacement also assures that an experiment can be variated. The replacement also ensures that paribus ceteris rule to assure the possibility to vary the experiment to confirm the hypothesis: An element can be replaced by another (variant of execution of an experiment), but this execution is still the same experiment. The axiom of replacement also guarantees that a contrast sample can be part of an experiment.

Example: The wire in a circuit can be replaced by other wires.
Example: The wires used in lab 1 can be used to help out in lab 2 if they have same specifications.
Counterexample: A simulation which can be only done by 1 computer in the world, due to the specific needs of the simulation does not fulfill the axiom of replacement. So to derive valid knowledge from this simulation, the computer must be replaceable. Until then the outcome is just more plausible, not safe empirical knowledge.
Counterexample: An experimenter claims to have found a a certain proof. But he says only he himself is able to do the experiment.

Axiom of Pairing If there is an element a and an element b, there exists a set which comprises a and b, but nothing else.

The existence of an elementary set of 2 elements means, that any 2 elements have a property which they have in common. $\mathfrak{E}(element)$ of each of them is true.
The identfiying of 2 elements is always possible, if any 2 elements exist in any set. Nevertheless, the particular pairing is not necessarily known. This axiom is a basic demand for a non-magic approach, any 1 or 2 elements can be made to an object of research.
As a result of this, the things must be 'comparable' in a wider sense. Any thing can be compared to another on any property. The property $\mathfrak{E}(element)$ can be as general as 'being observed at a timespan' or 'looking alike'. They both just have to satisfy the $\mathfrak{E}(element) = true$ criteria. So any 2 things have something in common and are therefore comparable.
Example: The researcher R1 knows that researcher R2 also uses the same type of measurement device. These devices form a set.
Counterexample: The researcher R1 tells researcher R2 that his measurement device is not comparable to the other researcher ones, because it is of a different make. This is a non-scientific statement.
Example: An holothurian is comparable with ice dancing: It both is an expression of life on earth. Being an expression of life is valid property, so it forms a set. This axiom holds despite the cases in which two things have 'nothing in common' in the sense of everyday's speech.

In experiments, the means used must be independent of their environments. Validity of an experiment must be the same in the following process: If 2 labs (A, B) do the same experiment, they can send their equipment to a 3rd lab C and (after recalibrating) the 3rd lab must be able to confirm the results. The equipment can be randomly chosen from lab A or B. The means of the 3rd lab are an arbitrary subset of the powerset of $\{A, B\}$. If C is already equipped, the means can be replaced.
Of course the arbitrary pairing is also possible with means which lead to the failing of an experiment. Together with axiom of replacement, this is a necessary option.
Example: The wires used in lab 1 and the battery used in lab 2 can be used in lab 3 for the same experiment. The lab 3 makes up a set which comprises the wires and the battery. If lab 3 does not comfirm the results despite following the receipt (hypothesis), it is not confirmed and is rejected or needs refinement.

Citation suggestion: Andreas Reif, ZFC-Empirical, 2015, [version or full date], url:[see your browser].
Document history: V.1.0 at 15th August 2015.