

Talmudic Foundations of Mathematics

Andrew Schumann
University of Information Technology and
Management in Rzeszow
35-225 Rzeszow
Poland
andrew.schumann@gmail.com

Alexander V. Kuznetsov
Voronezh State University
Universitetskaya pl. 1
394018 Voronezh
Russia
avkuz@bk.ru

ABSTRACT

In this paper, we assume that the mathematicians in proving new significant theorem, such as Fermat's Last Theorem, deal with combining proof trees on tree forests by using the analogy as an inference metarule. In other words, the real mathematical proofs cannot be formalized as discrete sequences, but they are concurrent and can be formalized as analog processes within a space with some topological properties. For the first time, inference metarules in a topological space were proposed in the Talmud within a general Judaic approach to concurrent or even massive-parallel conclusions. The mathematician does not think sequentially like a logical automaton, but concurrently, also. Hence, we suppose that the proof technique of real mathematics cannot be formalized by discrete methods. It is just a hypothesis of the foundations of mathematics that we can use discrete tools so that mathematics can be reduced to logic. We show in the paper how the mathematical proof can be formalized just by analog computations, not discrete ones.

CCS Concepts

•Computing methodologies → Parallel computing methodologies; Artificial intelligence; •Parallel algorithms → *Massively parallel algorithms*; •Artificial intelligence → Philosophical/theoretical foundations of artificial intelligence;

Keywords

Hilbert's program; foundations of mathematics; Fermat's last theorem; Cauchy criterion; topological Cauchy-Cantor intersection theorem; proof trees

1. INTRODUCTION

The *Principia Mathematica*, a three-volume work written jointly by Alfred North Whitehead and Bertrand Russell and published in 1910, 1912, and 1913, was the first book

i.e. in fact it was the first attempt to make explicitly mathematics from the point of view of symbolic logic, that is an attempt to consider mathematical theorems as logical statements which are automatically inferred from axioms by logical inference rules. To continue and enhance this approach, David Hilbert, the German mathematician (1862–1943), put forward a new proposal for the foundation of mathematics called the *Finite Program* (or Hilbert's Program). In this proposal all of mathematics should have been formalized in axiomatic form, together with a proof by 'finitary' methods proposed by Hilbert that this axiomatization is consistent.

Now there are some basic formal theories which are regarded as start points in the foundations of mathematics. This means that these theories, in the way how it seems to mathematicians, can cover big fragments of mathematics by their extensions. For instance, it is assumed that in the foundations for number theory we should start from the five Peano's axioms, introduced by Giuseppe Peano in 1889 and now called the *Peano arithmetic* **PA**. Also, it is supposed that any set-theoretic reasoning in mathematics (like reasoning in topology) can be reduced to statements formalized in the *Zermelo-Fraenkel set theory*, constructed by mathematicians Ernst Zermelo and Abraham Fraenkel and denoted by the abbreviation **ZFC**, where **C** means axiom of choice.

To sum up, mathematicians believe still that the foundations of mathematics in the meaning of *Principia Mathematica* are possible and any correct well-done mathematical reasoning can be rewritten in a logical theory such as **ZFC**. So, they believe that all the mathematics can be reduced to a logic. Is it true indeed? Is it possible?

In mathematics there are really non-trivial theorems which are so deep that they cannot be inferred without introducing absolutely new mathematical constructions. For example, *Fermat's Last Theorem* (**FLT**) is well formulated in **PA**. Therefore this statement seems to be so simple. For the first time, **FLT** was put forward by Pierre de Fermat in 1637 in the margin of a copy of *Arithmetica*. However, this statement was proven formally only after 358 years of effort by mathematicians, namely by Andrew Wiles in 1995 [6]. The most dramatic problem of **FLT** recently is that this theorem is proved mathematically and this proof was accepted by mathematical communities, but this statement was not checked by logicians at all. It is unknown still whether there is a logical proof of **FLT**. In other words, **FLT** is not covered by any foundations of mathematics still.

As we said, **FLT** is well written in a first-order sentence of **PA**. However, it does not mean that it can be proved

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. To copy otherwise, to republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee.
BICT 2017, March 15-16, Hoboken, United States
ISBN 978-1-63190-148-5
DOI: 10.4108/eai.22-3-2017.152404
Copyright © 2017 EAI

in **PA**. After the Paris-Harrington theorem [3], it is well known that there are ever first-order arithmetic statements written in **PA** which cannot be proved in **PA**, such as the *Strong Ramsey Theorem* that can easily be proved in the second-order arithmetic from the infinite version of the standard theorem. Also, it is known that there are many other combinatorial problems that are beyond **PA**. In [2], Colin McLarty supposed that **FLT** can be proved in some higher-order extensions of **PA**, but nobody has checked it still.

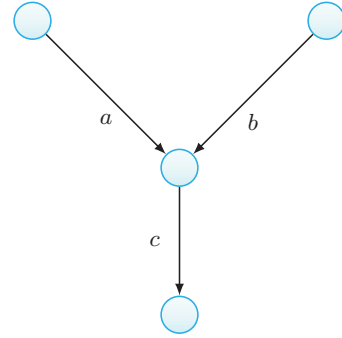
Another hypothesis of McLarty [2] is that **FLT** is beyond **ZFC**. It is quite evident taking into account the fact that cohomological number theory used by Wiles [6] is based on *Grothendieck's universes* which model **ZFC**, but the existence of a universe is not provable in **ZFC**. Grothendieck's own axiom of universes, which was added to **ZFC**, affirms that every set is contained in some universe (there is an uncountable strongly inaccessible cardinal for sets) [1]. Hence, in cohomology we deal with **ZFC+U** consisting of **ZFC** with the assumption **U** of a universe. So, **FLT** can be proved at least in **ZFC+U** or even in higher extensions, and evidently not in **ZFC**. Nevertheless, there is no formal proof still what is set theory looks like for **FLT**.

Thus, there are ever serious mathematical theorems such as **FLT** which are beyond the recent foundations of mathematics (for instance, beyond **PA** or **ZFC**). However, mathematicians and logicians unconsciously obey the quite religious faith and follow the deep-inner intuition that any mathematical theorem can be reduced to a theorem of existed symbolic logic. In symbolic logic we appeal to a *formal theory* T_i that possesses logical axioms/theorems $(a_{i1}^L, \dots, a_{in}^L)$ within a logical system L like the classical propositional logic and non-logical axioms / theorems $(a_{i1}^T, \dots, a_{im}^T)$ for defining properties of predicates and functions introduced in T_i . Then by using inference rules of L we can infer in T_i all possible provable sentences from $(a_{i1}^L, \dots, a_{in}^L) + (a_{i1}^T, \dots, a_{im}^T)$. Surely, it does work in symbolic logic indeed, but it is unknown still whether it gives something to real mathematics. In real mathematics, i.e. in a proof of deep sentences such as **FLT**, we use some axioms / presuppositions / sentences $(a_{i1}^M, \dots, a_{ik}^M)$ and the main task of the foundations of mathematics is to set up a formal theory T_i to find out an injective mapping from $(a_{i1}^M, \dots, a_{ik}^M)$ into $(a_{i1}^T, \dots, a_{im}^T)$. The problem is that, on the one hand, some sentences from $(a_{i1}^M, \dots, a_{ik}^M)$ in **FLT** can be just intuitive and not well-formulated (or even not conscious). On the other hand, their true symbolic-logical analogues in T_i can be absent from the list $(a_{i1}^T, \dots, a_{im}^T)$ and not sketched still.

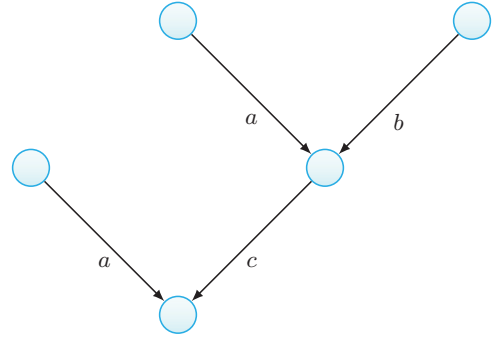
In this paper we assume that $(a_{i1}^M, \dots, a_{ik}^M)$ for serious theorems like **FLT** are not non-logical axioms / theorems which can be represented as $(a_{i1}^T, \dots, a_{im}^T)$. So, we assume that higher mathematics cannot be reduced to symbolic logic.

Let us suppose now that $(a_{i1}^M, \dots, a_{ik}^M) = (a_{i1}^M, \dots, a_{ik}^M) + (a_{i1}^L, \dots, a_{in}^L)$ and $(a_{i1}^T, \dots, a_{im}^T) = (a_{i1}^T, \dots, a_{im}^T) + (a_{i1}^L, \dots, a_{in}^L)$, i.e. they are closed under the inference rules of L . Each proof in $T_i = (a_{i1}^T, \dots, a_{im}^T)$ is a *tree* t . Let t_{e_j} be an inference rule of T_i , i.e. it is an elementary tree – it has one parent (the inferred statement) and several its direct children (axioms as premisses). Then t consists of t_{e_1}, \dots, t_{e_j} as its subtrees. Let $\{\}$ be an empty tree. Then each *edge/branch* of t can be labelled by a sentence $\alpha \in (a_{i1}^T, \dots, a_{im}^T)$ as follows: $\alpha[t']$, where t' is an elementary subtree (inference rule) of t that was used to obtain α . Then a tree can be denoted by all its branches. For in-

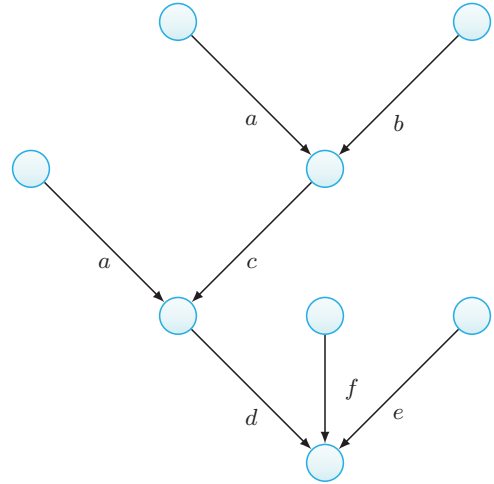
stance, the notation $t = \{c[\{a[\{\}], b[\{\}]\}]\}$ means that we have a tree consisting only of two edges/premisses $a[\{\}]$ and $b[\{\}]$ and one conclusion from them $c[\{a[\{\}], b[\{\}]\}]:$



The notation $t' = \{c[\{a[\{\}], b[\{\}]\}], a[\{\}]\}$ designates the following tree:



The formula $t'' = \{d[\{c[\{a[\{\}], b[\{\}]\}], a[\{\}]\}], f[\{\}], e[\{\}]\}$ satisfies the following tree:



Hence, in T_i we deal just with finite trees labelled in $(a_{i1}^T, \dots, a_{im}^T)$ in the way defined above.

Let an ordered sequence of trees be called a *forest*. We say a forest $s_i = \langle t_i, t'_i, t''_i, \dots \rangle$ is a *piece* of a forest $s_j = \langle t_j, t'_j, t''_j, \dots \rangle$, written $s_i \leq s_j$ if s_i can be obtained from s_j by removing nodes. In other words, there is an injective mapping from nodes of s_i to nodes of s_j that preserves the lexicographic and descendant ordering. Then a

reduction of $(a_{i1}^M, \dots, a_{im}^M)$ to $(a_{i1}^T, \dots, a_{in}^T)$ would mean that $(a_{i1}^M, \dots, a_{im}^M)$ is a piece of $(a_{i1}^T, \dots, a_{in}^T)$, i.e. $(a_{i1}^M, \dots, a_{im}^M) \preceq (a_{i1}^T, \dots, a_{in}^T)$.

In deep theorems like **FLT** we have ever a forest s_j , which is a substantial part of $(a_{i1}^M, \dots, a_{im}^M)$, such that s_j is not a piece for any forest of $(a_{i1}^T, \dots, a_{in}^T)$. So, we can distinguish two kinds of forests from $(a_{i1}^M, \dots, a_{im}^M)$:

1. a forest s_j which is a piece of some forest from $(a_{i1}^T, \dots, a_{in}^T)$, these pieces are called *logical fragments*;
2. a forest s_j which is not a piece of any forest from $(a_{i1}^T, \dots, a_{in}^T)$, this s_j is called a *new construction*.

We assume that logical fragments are well studied in the foundations of mathematics, but new constructions are not yet. The true mathematical task is to give non-trivial essential theorems containing new constructions. The mathematician starts his work with logical fragments to obtain new constructions later.

Thus, the higher mathematics is eternally to extend forests of $(a_{i1}^M, \dots, a_{im}^M)$ by proposing new constructions. Therefore, the mathematician deals not only with proofs/trees (what can be reduced completely to pieces from $(a_{i1}^T, \dots, a_{in}^T)$), but more often with forests which are irreducible to pieces of $(a_{i1}^T, \dots, a_{in}^T)$. Mathematicians know that sometimes reasoning by analogy allows them to propose really something new. They take a piece from one theory $\mathcal{T}_i = (a_{i1}^M, \dots, a_{im}^M)$ to combine it with forests of another theory $\mathcal{T}_j = (a_{j1}^M, \dots, a_{jk}^M)$ and then to obtain the next theory $\mathcal{T}_f = (a_{f1}^M, \dots, a_{fl}^M)$.

In the Talmud there was proposed a metareasoning for extending the dataset (forests) of the Torah, $(a_{i1}^M, \dots, a_{ik}^M)$, by analogies for the first time (see Section 2). It is a way how we can add new axioms to $(a_{i1}^M, \dots, a_{ik}^M)$ just looking at existing forests. In Section 3 we will show how we can use these methods in the recent foundations of mathematics. Hence, the Talmudic metareasoning in the foundations of mathematics are called by us ‘*Talmudic foundations of mathematics*’, see [4], [5]. We assume that in a real mathematical practice mathematicians use a similar metareasoning to build up new forests by looking at existing forests (let us notice that this metareasoning is considered often as a mathematical intuition).

2. WHAT IS TALMUDIC LOGIC?

Usually, the term ‘Talmudic logic’ means the Judaic hermeneutic rules (Hebrew: *middot*, מידות) [4], first formulated as a special hermeneutics by Hillel in the 1st century B.C. (he proposed the 7 rules), by Rabbi Ishmael in the 2nd century A.D. (he established the 13 rules for inferring the law, *halakhah*), and by Rabbi Eliezer ben Jose HaGelili in the same 2nd century A.D. (he examined the 32 rules for inferring the holy stories, *haggadah*). As a result, in the Talmud we face a logical point of view in respect to the Torah: all the Biblical statements are considered ‘particular’ (פרט) or ‘general’ (כלל) implying some conclusions by using hermeneutic rules. A Biblical statement, φ_{part}^i , is regarded *particular*, if its verification $|\varphi_{part}^i|$ (i.e. all denotates of φ_{part}^i) is a subset of a verification $|\varphi_{gen}^j|$ for another statement, φ_{gen}^j , namely: $|\varphi_{part}^i| \subseteq |\varphi_{gen}^j|$. In this case the implication $\varphi_{part}^i \rightarrow \varphi_{gen}^j$ is true. Hence, we can draw appropriate proof trees by using *modus ponens* and other classical inference rules in the way

of Section 1, if we define all the last particulars of the Torah as axioms of the Talmud: $(a_{i1}^{part}, \dots, a_{in}^{part})$.

Now, let us consider an example from *Exodus* 22 and let us try to build up a proof tree for the notion ‘responsibility for the property of his neighbour’, just basing on the text of the Torah. We have the following statements mentioned in this chapter of the Holy Book:

G “He is responsible for the property of his neighbour.”

G1 “The neighbour’s property is given for safekeeping for free of charge.”

G2 “The neighbour’s property is given for safekeeping for money.”

G3 “The neighbour’s property is borrowed.”

G1P1 “The safekeeping for free of charge is stolen. He should swear that he did not lay his hand upon his neighbor’s property.”

If a man shall deliver unto his neighbour money or stuff to keep, and it be stolen out of the man’s house; if the thief be found, let him pay double. If the thief be not found, then the master of the house shall be brought unto the judges, to see whether he have put his hand unto his neighbour’s goods (KJV, *Exodus*, 22:7-8).

G2P1 “The safekeeping for money is stolen. He should pay the loss.”

G2P2 “The safekeeping for money is destroyed for natural reasons. He should swear that he did not lay his hand upon his neighbor’s property.”

If a man deliver unto his neighbour an ass, or an ox, or a sheep, or any beast, to keep; and it die, or be hurt, or driven away, no man seeing it: Then shall an oath of the Lord be between them both, that he hath not put his hand unto his neighbour’s goods; and the owner of it shall accept thereof, and he shall not make it good. And if it be stolen from him, he shall make restitution unto the owner thereof (KJV, *Exodus*, 22:10-12).

G3P2 “The borrowed is destroyed for natural reasons. He should pay the loss.”

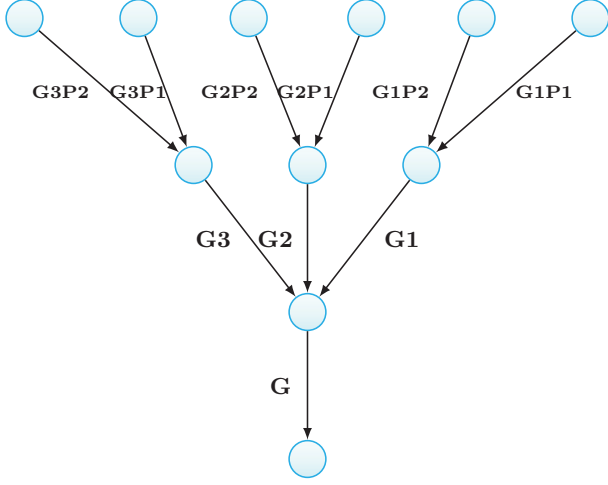
And if a man borrow ought of his neighbour, and it be hurt, or die, the owner thereof being not with it, he shall surely make it good. But if the owner thereof be with it, he shall not make it good: if it be an hired thing, it came for his hire (KJV, *Exodus*, 22:14-15).

As we see, we deal here with the four last particulars (i.e. Talmudic axioms): (**G1P1**, **G2P1**, **G2P2**, **G3P2**). However, the data set for ‘responsibility for the property of his neighbour’ is not complete – we know nothing about the following two particulars (axioms) which are supposed also:

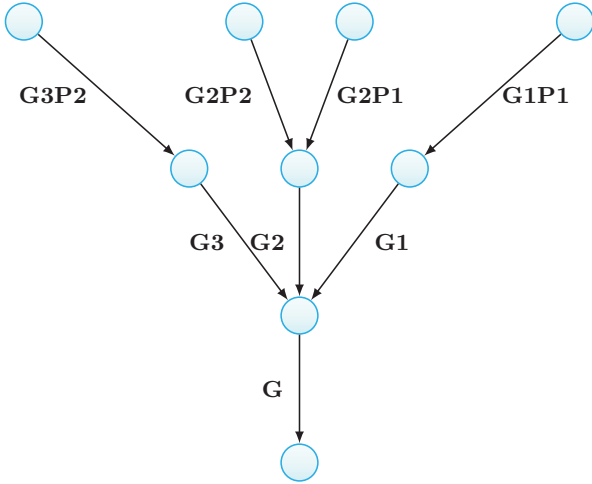
G1P2 “The safekeeping for free of charge is destroyed for natural reasons. What should he do?”

G3P1 “The borrowed is stolen. What should he do?”

In the picture form, the complete data set must be seen as follows:



Nevertheless, in the Torah an appropriate data set is sketched in the following manner, i.e. it is absolutely incomplete for inferring ‘responsibility for the property of his neighbour’:



Thus, we need to find out possible ways for defining the two new axioms, **G1P2** and **G3P1**. But how? Do not worry. In the *Bava Metzia*, the second of the first three tractates of the Babylonian Talmud in the order of *Nezikin* (“damages”), chapter 8, there is a metareasoning for this purpose – the inference metarule for defining new axioms, called *qal wa-homer* (קל וחמר). This rule occurs among the 7 rules of Hillel, as well as among the 13 rules of Rabbi Ishmael and the 32 rules of Rabbi Eliezer ben Jose HaGelili. Let us take all the particulars for **G2**, because they are complete. Between **G2P1** and **G2P2** there is a strong ordering relation: $\mathbf{G2P2} \lesssim \mathbf{G2P1}$. Indeed, **G2P2** means that he is free of paying and **G2P1** means that he should pay the loss. Hence, we assume that all last particulars (**G1P1**, **G1P2**, **G2P1**, **G2P2**, **G3P1**, **G3P2**) are partially ordered by a relation ‘ \lesssim or =’ denoted by \lesssim , where = means the same payment. Then **G1P2** is defined by *qal wa-homer* as follows:

$$\mathbf{G1P2} = \min(\mathbf{G2P2}, \mathbf{G1P1}).$$

From this it follows that

G1P2 “The safekeeping for free of charge is destroyed for natural reasons. He should swear that he did not lay his hand upon his neighbor’s property.”

Analogically, **G3P1** is defined by *qal wa-homer* thus:

$$\mathbf{G3P1} = \min(\mathbf{G2P1}, \mathbf{G3P2}).$$

Then it is inferred that

G3P1 “The borrowed is stolen. He should pay the loss.”

So, the main goal of *qal wa-homer* is to add new axioms for the Torah data sets to make the proof trees more symmetrical: in the same tree t with only one root all subtrees must bear the same number of edges. For instance, in the tree

$$t = \{\mathbf{G}[\{\mathbf{G1}[\{\mathbf{G1P1}\{\}\}], \mathbf{G2}[\{\mathbf{G2P1}\{\}, \mathbf{G2P2}\{\}\}], \mathbf{G3}[\{\mathbf{G3P2}\{\}\}]\}],$$

we had the subtree $t' = \{\mathbf{G2}[\{\mathbf{G2P1}\{\}, \mathbf{G2P2}\{\}\}]\}$ with the number of edges 2 that exceeds the numbers of edges of all other subtrees. Then we must add new branches to each subtree where the number of edges is less than 2.

DEFINITION 1 (INFERENCE METARULE *qal wa-homer*). Let a tree t have i subtrees: t_1, \dots, t_i and let N_j be a number of axioms ($a_{j1}^{part}, \dots, a_{jn}^{part}$) for each $j = \overline{1, i}$ and let all axioms for each subtree be lexicographically ordered in the same way. Assume that $\max_{j=1}^i N_j = n$. This means that there exists l such that $1 \leq l \leq i$ and $N_l = n$. Let us suppose now that there exists $k \in \{1, \dots, i\}$ such that $k \neq l$ and $N_k = m < n$. Then we can add new axioms by *qal wa-homer* to complete the axioms of t_k up to the set ($a_{k1}^{part}, \dots, a_{kn}^{part}$):

1. the set ($a_{l1}^{part}, \dots, a_{ln}^{part}$) is partially ordered by \lesssim and the order \lesssim of ($a_{o1}^{part}, \dots, a_{on}^{part}$) is the same for any subtree t_o of t such that $N_o = n$;
2. the set ($a_{k1}^{part}, \dots, a_{km}^{part}$) is partially ordered by the same \lesssim , but the order can be different;
3. the item $a_{k(m+1)}^{part} = \min(a_{km}^{part}, a_{l(m+1)}^{part})$;
4. the item $a_{k(m+p)}^{part} = \min(a_{k(m+p-1)}^{part}, a_{l(m+p)}^{part})$ for each $p = \overline{1, n - m}$.

According to this metarule, we can have the same number of axioms for each subtree of the same tree t . In other words, for all subtrees t_1, \dots, t_i of t we have the axioms ($a_{11}^{part}, \dots, a_{1n}^{part}, a_{21}^{part}, \dots, a_{2n}^{part}, \dots, a_{i1}^{part}, \dots, a_{in}^{part}$) due to *qal wa-homer*.

In Section 1 we have said that the task of every true mathematician is to extend a set of mathematical axioms. In the statements like **FLT** we exceed the set of existing axioms (i.e. we put proofs outside of the foundation of mathematics). And we assume that the true mathematicians appeal to some inference metarule to obtain new axioms for

proving their non-trivial sentences. In other words, they deal not with mechanical proofs from existing axioms within the foundations of mathematics, but they combine different trees to expand the set of possible axioms beyond any foundations of mathematics. This way of proving is called by us *Talmudic* because of the priority of the Talmudic logic in proposing some inference metarules for defining axioms. So, let us generalize definition 1 as follows.

DEFINITION 2 (INFERENCE METARULE \mathcal{I}). *Let a tree t has i subtrees: t_1, \dots, t_i and let N_j be a number of axioms $(a_{j1}^M, \dots, a_{jn}^M)$ for each $j = \overline{1, i}$. Assume that $\max_{j=1}^i N_j = n$. This means that there exists l such that $1 \leq l \leq i$ and $N_l = n$. Let us suppose now that there exists $k \in \{1, \dots, i\}$ such that $k \neq l$ and $N_k = m < n$. Then we can add new axioms by the inference metarule \mathcal{I} to complete the axioms of t_k up to the set $(a_{k1}^M, \dots, a_{kn}^M)$:*

1. *the sets $(a_{l1}^M, \dots, a_{ln}^M), \dots, (a_{o1}^M, \dots, a_{on}^M)$ are partially ordered by \approx for all subtrees t_1, \dots, t_o with the number of axioms $N_l = \dots = N_o = n$;*
2. *the set $(a_{k1}^M, \dots, a_{km}^M)$ is partially ordered by the same \approx , but the order can be different;*
3. *the item $a_{k(m+p)}^M = \square_{\mathcal{I}}(a_{k1}^M, \dots, a_{k(m+p-1)}^M, a_{l1}^M, \dots, a_{l(m+p)}^M, \dots, a_{o1}^M, \dots, a_{o(m+p)}^M)$ for each $p = \overline{1, n-m}$, where $\square_{\mathcal{I}}$ is a Boolean function.*

Thus, in the *Talmudic foundations of mathematics* we transform one space of mathematical proof trees into another space by some inference metarules. By analogy how in the Talmud we follow a transformation from one space of Biblical particular-and-general to another one. The proof trees supposed in the Torah are not complete and by the Talmudic inference metarules such as *gal wa-homer* we can make trees more symmetric so that their subtrees must have the same number of branches at the end.

3. METAREASONING IN MATHEMATICAL ANALYSIS

Let us consider an example from mathematical analysis to show that a mathematician, extending a mathematical horizon, operates with forests and appeals to metarules for combining different trees, indeed. Great mathematicians do not prove theorems in the way like a logical automaton does it. For example, Augustin-Louis Cauchy, the French mathematician of the 19th century who became a founder of modern mathematical analysis, put forward an axiom now called Dedekind Completeness for the expansion of mathematics. Due to this axiom, some properties of real numbers (non-logical axioms describing real numbers) can have transmitted to complex numbers, vectors, and even infinite sequences by a mathematical analogy of the Talmudic *gal wa-homer* rule.

3.1 Cauchy Criterion

Let us denote an infinite sequence (i.e. a countable set) of objects $x_1, x_2, \dots, x_k, \dots$ by $\{x_n\}$. Then some basic definitions of the Cauchy approach are as follows:

DEFINITION 3. $x_n \xrightarrow{\mathbb{R}} x \stackrel{\text{def}}{=} \{x_n\} \subset \mathbb{R}$ *converges to $x \in \mathbb{R}$ what is defined thus: $\{x_n\} \subset \mathbb{R} \wedge \exists(x \in \mathbb{R}) \forall(\varepsilon > 0) \exists(N_\varepsilon) \forall(n \geq N_\varepsilon)[|x_n - x| < \varepsilon]$.*

Let us denote definition 3 by **Conv** $(\{x_n\}; \mathbb{R}, |\cdot|)$.

DEFINITION 4. $\{x_n\} \subset \mathbb{R}$ *satisfies the Cauchy condition iff $\{x_n\} \subset \mathbb{R} \wedge \forall(\varepsilon > 0) \exists(N_\varepsilon \in \mathbb{N}) \forall(m, n \in \mathbb{N}) [n \geq N_\varepsilon \wedge m \geq N_\varepsilon] [|x_n - x_m| < \varepsilon]$.*

Definition 4 is denoted by **CC** $(\{x_n\}; \mathbb{R}, |\cdot|)$

AXIOM 1 (DEDEKIND COMPLETENESS). $\forall(A \subseteq \mathbb{R} | A \neq \emptyset) \forall(B \subseteq \mathbb{R} | B \neq \emptyset) [\forall(a \in A) \forall(b \in B) [a \leq b] \rightarrow \exists(\xi \in \mathbb{R}) \forall(a \in A) \forall(b \in B) [a \leq \xi \leq b]]$.

The completeness axiom is denoted by **Ax3** $(A, B, a, b; \mathbb{R}, |\cdot|)$.

DEFINITION 5. *A sequence $[a_n, b_n]$ of closed intervals of \mathbb{R} is called a nested closed intervals system iff $\forall(i \in \mathbb{N}) [a_i \in \mathbb{R} \wedge b_i \in \mathbb{R} \wedge [a_i, b_i] \supset [a_{i+1}, b_{i+1}]]$.*

Let us denote definition 5 by **SES** $(\{[a_n, b_n]\}; \mathbb{R}, |\cdot|)$.

THEOREM 1. *This theorem called the Cauchy-Cantor's intersection theorem consists of two parts:*

1. *Existence of the common point (this property is denoted by **ECP** $(\{[a_n, b_n]\}; \mathbb{R}, |\cdot|)$):*

$$\forall(\{[a_n, b_n]\} \subset 2^{\mathbb{R}}) [\mathbf{SES}(\{[a_n, b_n]\}; \mathbb{R}, |\cdot|) \rightarrow \exists(\xi \in \mathbb{R}) \forall(i \in \mathbb{N}) [\xi \in [a_i, b_i]]]$$

2. *Uniqueness of the common point (this property is denoted by **SCP** $(\{[a_n, b_n]\}; \mathbb{R}, |\cdot|)$):*

$$\mathbf{ECP}(\{[a_n, b_n]\}; \mathbb{R}, |\cdot|) \wedge b_n - a_n \xrightarrow{\mathbb{R}} 0 \rightarrow \exists!(\xi \in \mathbb{R}) \forall(i \in \mathbb{N}) [\xi \in [a_i, b_i]]]$$

PROOF. (1) *The existence of the common point:*

$$A_0 = \{a_n\}, B_0 = \{b_n\};$$

$$\forall(a \in A_0) \forall(b \in B_0) [a \leq b], \mathbf{Ax3}(A, B, a, b; \mathbb{R}, |\cdot|) \rightarrow$$

$$\mathbf{Ax3}(A_0, B_0, a, b; \mathbb{R}, |\cdot|);$$

$$\mathbf{Ax3}(A_0, B_0, a, b; \mathbb{R}, |\cdot|) \leftrightarrow$$

$$\exists(\xi \in \mathbb{R}) \forall(a \in A_0) \forall(b \in B_0) [a \leq \xi \leq b].$$

The last inequality means that ξ is a common point of the nested closed intervals system.

(2) *The uniqueness of the common point.*

Assume a contrary:

$$\forall(n \in \mathbb{N}) \exists(\xi, \xi' \in [a_n, b_n]) [\xi \neq \xi'] \rightarrow$$

$$\forall(n \in \mathbb{N}) [|\xi - \xi'| \leq b_n - a_n];$$

$$b_n - a_n \xrightarrow{\mathbb{R}} 0 \text{ iff } \forall(\varepsilon > 0) \forall(n > N_\varepsilon) [b_n - a_n < \varepsilon];$$

$$\varepsilon = \frac{1}{2} |\xi - \xi'| > 0 \rightarrow |\xi - \xi'| < \frac{1}{2} |\xi - \xi'|.$$

It is a contradiction. \square

The statement of theorem 1 is denoted by **CCP** $(\{\alpha_n\}; \mathbb{R}, |\cdot|)$ where $\alpha_n = [a_n, b_n]$ is a closed interval.

DEFINITION 6. A sequence $\{x_n\}$ is called bounded iff

$$\exists(C)\forall(n \in \mathbb{N})[|x_n| \leq C].$$

Definition 6 is denoted by $\mathbf{B}(\{x_n\}; \mathbb{R}, |\cdot|)$.

THEOREM 2 (BOLZANO–WEIERSTRASS). Each bounded sequence $\{x_n\} \subseteq \mathbb{R}$ has a convergent subsequence $\{x_{n_k}\} \subseteq \mathbb{R}$, $\exists(x \in \mathbb{R})[x_{n_k} \xrightarrow{\mathbb{R}} x]$.

PROOF. The fact that $\{x_n\}$ is bounded means that

$$\exists(C)\forall(n \in \mathbb{N})[|x_n| \leq C] \rightarrow$$

$$\forall(n \in \mathbb{N})[x_n \in [-C, C] \subset \mathbb{R}].$$

We apply the following algorithm notated as **BWAlgo**:

1. $a_0 = -C, b_0 = C$.
2. For all $i \in \mathbb{N} \setminus \{0\}$ let us divide $[a_i, b_i]$ into two equal segments, and choose one of them which has an infinite number of members of the sequence, denote it by $[a_{i+1}, b_{i+1}]$.
3. We obtain a sequence $\{[a_n, b_n]\}$ of closed intervals, i.e. there is $\mathbf{SES}(\{[a_n, b_n]\}; \mathbb{R}, |\cdot|)$. All closed intervals in this sequence contain an infinite number of members of the sequence $\{x_k\}$. Lengths of this closed intervals converge to zero: $|b_m - a_m| = \frac{|b_0 - a_0|}{2^m} = \frac{2C}{2^m} = \frac{C}{2^{m-1}}$, $\frac{C}{2^{m-1}} \xrightarrow{\mathbb{R}} 0$.
4. Choose a sequence $x_{k_m} \in [a_m, b_m]$, $m \in \mathbb{N}$ with the following condition: $\forall(i \in \mathbb{Z})[k_{i-1} < k_i]$.

From this it follows that $[[\mathbf{CCP}(\{[a_n, b_n]\}; \mathbb{R}, |\cdot|) \rightarrow \exists!(\xi \in \mathbb{R})\forall(m \in \mathbb{N})[\xi \in [a_m, b_m]]] \rightarrow \forall(m \in \mathbb{N})[|x_{k_m} - \xi| \leq |b_m - a_m|] \rightarrow |x_{k_m} - \xi| \xrightarrow{\mathbb{R}} 0$. \square

Theorem 2 is denoted by $\mathbf{BWT}(\{x_n\}; \mathbb{R}, |\cdot|)$.

THEOREM 3 (TRIANGLE INEQUALITY). $\forall(x, y \in \mathbb{R})[|x + y| \leq |x| + |y|]$.

The proof of theorem 3 for \mathbb{R} is obvious. Let us denote its statement by $\mathbf{Tr}(x, y; \mathbb{R}, |\cdot|)$.

THEOREM 4 (CAUCHY CRITERION). $(\{x_n\} \subset \mathbb{R} \wedge x_n \xrightarrow{\mathbb{R}} x) \leftrightarrow \mathbf{CC}(\{x_n\}; \mathbb{R}, |\cdot|)$

PROOF. *Necessity.* If $\{x_n\}$ converges to x , then $\forall(\varepsilon > 0)\exists(N_\varepsilon \in \mathbb{N})\forall(n > N_\varepsilon)[|x_n - x| < \varepsilon]$. Hence, $\forall(\varepsilon > 0)\exists(N_\varepsilon \in \mathbb{N})\forall(m > N_\varepsilon)[|x_m - x| < \varepsilon]$. Then $\forall(\varepsilon > 0)\exists(N_\varepsilon \in \mathbb{N})\forall(n, m > N_\varepsilon)[|x_n - x| < \varepsilon \wedge |x_m - x| < \varepsilon]$.

Thus, $\forall(\varepsilon > 0)\exists(N_\varepsilon \in \mathbb{N})\forall(n, m > N_\varepsilon)[|x_n - x| < \varepsilon \wedge |x_m - x| < \varepsilon] \wedge \mathbf{Tr}(x_n - x, x_m - x; \mathbb{R}, |\cdot|)$ implies that $\forall(\varepsilon > 0)\exists(N_\varepsilon \in \mathbb{N})\forall(n, m > N_\varepsilon)[|x_n - x_m| < 2\varepsilon]$, then $\forall(\varepsilon > 0)\exists(N_\varepsilon \in \mathbb{N})\forall(n, m > N_\varepsilon)[|x_n - x_m| < \varepsilon]$.

Sufficiency. $\mathbf{CC}(\{x_n\}; \mathbb{R}, |\cdot|) \rightarrow \forall(\varepsilon > 0)\exists(N_\varepsilon)\forall(m > N_\varepsilon)[|x_m - x_{N_\varepsilon}| < \varepsilon]$. If $C = \max\{|x_1|, |x_2|, |x_3|, \dots, |x_{N_\varepsilon-1}|, |x_{N_\varepsilon} - \varepsilon|, |x_{N_\varepsilon} + \varepsilon|\}$, then $\forall(n \in \mathbb{N})[|x_n| \leq C]$. This means that $\{x_n\}$ is bounded. So, $\{x_n\}$ is bounded and $\mathbf{BWT}(\{x_n\}; \mathbb{R}, |\cdot|)$, then from $\exists(\{x_{n_k}\} \subseteq \{x_n\})\exists(x \in \mathbb{R})[x_{n_k} \xrightarrow{\mathbb{R}} x]$ it follows that $\exists(\{x_{n_k}\} \subseteq \{x_n\})\exists(x \in \mathbb{R})[|x_{n_k} - x| < \varepsilon]$.

Thus, $\mathbf{CC}(\{x_n\}; \mathbb{R}, |\cdot|)$ implies that $\forall(\varepsilon > 0)\exists(N_\varepsilon \in \mathbb{N})\forall(m, k \in \mathbb{N})[n \geq N_\varepsilon \wedge m \geq N_\varepsilon][|x_{n_k} - x_m| < \varepsilon]$. Then $\forall(\varepsilon > 0)\exists(N_\varepsilon \in \mathbb{N})\forall(m, k \in \mathbb{N})[n \geq N_\varepsilon \wedge m \geq N_\varepsilon][|x - x_m| \leq |x_{n_k} - x_m| + |x - x_{n_k}| < 2\varepsilon]$. \square

The Cauchy criterion is denoted by $\mathbf{CCrit}(\{x_n\}; \mathbb{R}, |\cdot|)$.

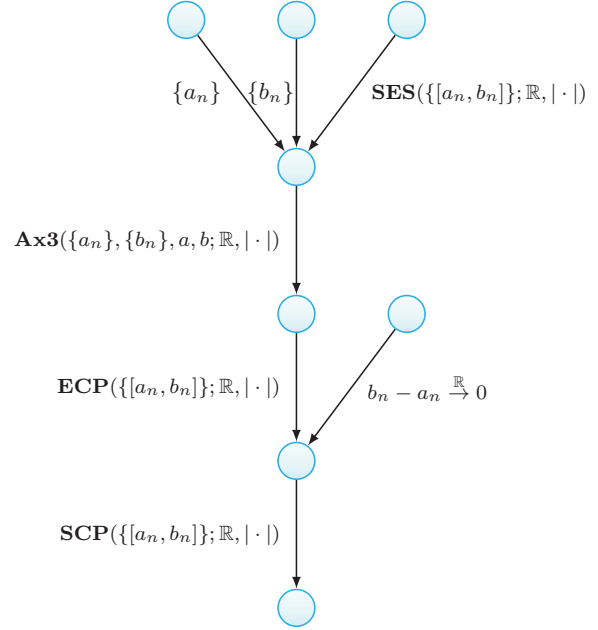
3.2 Qal Wa-Homer for the Cauchy Criterion

Let us fix the most important steps in the proofs of the previous subsection.

The proof of theorem 1 can be represented as the following tree:

$$t_1 = \{\mathbf{CSP}\{\{\mathbf{ECP}\{\{\mathbf{Ax3}\{\{a_n\}\{\}, \{b_n\}\{\}, \mathbf{SES}\{\{\}\}\}\}\}, b_n - a_n \xrightarrow{\mathbb{R}} 0\{\}\}\},$$

the same is in the graph form:

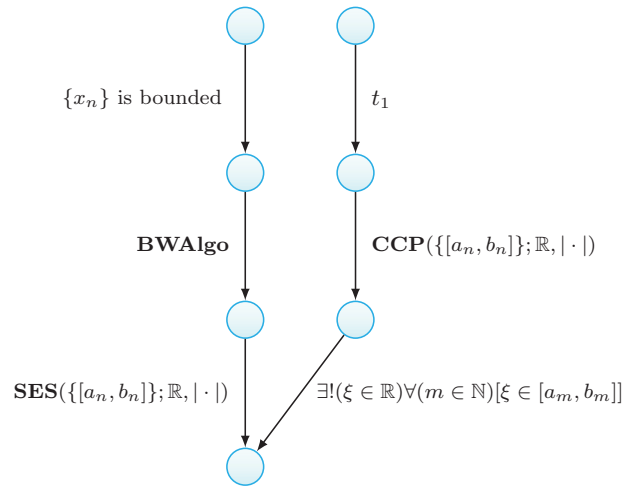


The proof tree for theorem 2 is built up as follows:

$$t_2 = \{\mathbf{SES}\{\{\mathbf{BWAlgo}\{\{x_n\} \text{ is bounded } \{\}\}\}\},$$

$$\exists!(\xi \in \mathbb{R})\forall(m \in \mathbb{N})[\xi \in [a_m, b_m]]\{\{\mathbf{CCP}[t_1]\}\},$$

which is pictured in the following manner:



Let us define now the proof tree for theorem 4:

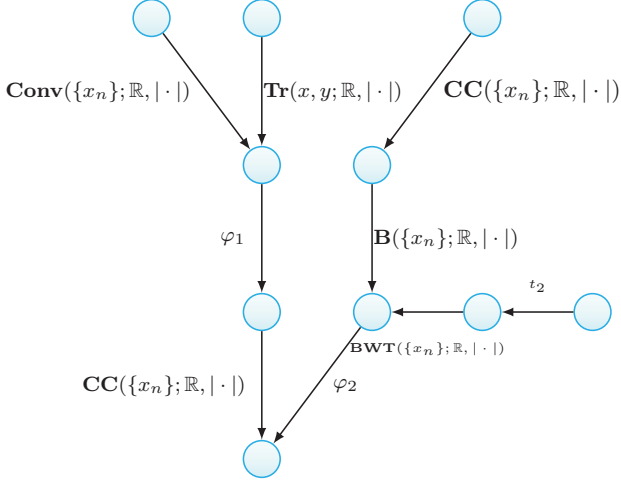
$$t_4 = \{\mathbf{CC}\{\{\forall(\varepsilon > 0)\exists(N_\varepsilon \in \mathbb{N})\forall(n, m > N_\varepsilon)[|x_n - x| < \varepsilon] \wedge$$

$$|x_m - x| < \varepsilon [\{\mathbf{Conv}\{\cdot\}\}, \mathbf{Tr}\{\cdot\}],$$

$$\forall(\varepsilon > 0) \exists(N_\varepsilon \in \mathbb{N}) \forall(m, k \in \mathbb{N} | n \geq N_\varepsilon \wedge$$

$$m \geq N_\varepsilon) [|x_{n_k} - x_m| < \varepsilon] [\{\mathbf{B}\{\mathbf{CC}\{\cdot\}\}\mathbf{BWT}[t_2]\}],$$

with the following graph:



where $\varphi_1 = \forall(\varepsilon > 0) \exists(N_\varepsilon \in \mathbb{N}) \forall(n, m > N_\varepsilon) [|x_n - x| < \varepsilon \wedge |x_m - x| < \varepsilon]$; $\varphi_2 = \forall(\varepsilon > 0) \exists(N_\varepsilon \in \mathbb{N}) \forall(m, k \in \mathbb{N} | n \geq N_\varepsilon \wedge m \geq N_\varepsilon) [|x_{n_k} - x_m| < \varepsilon]$.

Let us define an ordering relation among non-logical axioms, \prec , as follows: $A \prec B$ iff an axiom A is contained in the proof three of B . Hence, for t_4 we obtain:

$$[\mathbf{Ax3}(A, B, a, b; \mathbb{R}, |\cdot|) \prec \mathbf{CCP}(\{\alpha_n\}; \mathbb{R}, |\cdot|) \prec$$

$$\mathbf{BWT}(\{x_n\}; \mathbb{R}, |\cdot|) \prec \mathbf{CCrit}(\{x_n\}; \mathbb{R}, |\cdot|)] \wedge$$

$$[\mathbf{Tr}(x, y; \mathbb{R}, |\cdot|) \prec \mathbf{CCrit}(\{x_n\}; \mathbb{R}, |\cdot|)].$$

Theorem 4 holds for real numbers. Nevertheless, this theorem can be extended to some other numbers by a *gal wahomer*:

1. A sequence of complex numbers $\{x_n + iy_n\} \subset \mathbb{C}$ can be defined in the way of $\mathbf{CCrit}(\{x_n + iy_n\}; \mathbb{C}, |\cdot|_c)$ where $|x + iy|_c = \sqrt{x^2 + y^2}$, i.e. it is sufficient for us to replace all occurrences of \mathbb{R} to \mathbb{C} and all occurrences of the metrics $|\cdot|$ to the metrics $|\cdot|_c$ in the proof statements of the theorem.
2. Also, sequences of vectors from \mathbb{R}^s can be formulated as $\mathbf{CCrit}(\{x_n\}; \mathbb{R}^s, \|\cdot\|_{\mathbb{R}^s})$ where

$$\|(v^1, \dots, v^s)\|_{\mathbb{R}^s} = \left(\sum_{i=1}^s (v^i)^2 \right)^{1/2}.$$

So, we can generalize theorem 4 (the Cauchy criterion) due to the fact that there is an analogue for the Dedekind completeness axiom (**Ax3**) (defined in real numbers) that holds for aforesaid types of numbers which can be represented as some tuples of real numbers. In this case some new inequalities $a_i \leq \xi_i \leq b_i$, $i = 1, \dots, n$ for $a = (a_1, \dots, a_n)$, $b = (b_1, \dots, b_n)$, $\xi = (\xi_1, \dots, \xi_n)$ take place instead of the one inequality of **Ax3**.

We can formulate the same generalization by using definition 2. Assume that we have some axioms $(a_1^{\mathbb{R}}, \dots, a_N^{\mathbb{R}})$ for real numbers, some axioms $(a_1^{\mathbb{C}}, \dots, a_M^{\mathbb{C}})$ for complex numbers, and some axioms $(a_1^{\mathbb{V}}, \dots, a_L^{\mathbb{V}})$ for vectors. We know that $N > M$ (respectively, $N > L$), because in the beginning, \mathbb{R} was studied better than \mathbb{C} (respectively, better than \mathbb{V}). Let us take a subset of axioms $(a_1^{\mathbb{R}}, \dots, a_n^{\mathbb{R}})$ for real numbers, which express the real metrics. So, $a_n^{\mathbb{R}}$ is **Ax3**. Now let us define a partial ordering relation \lesssim on $(a_1^{\mathbb{R}}, \dots, a_n^{\mathbb{R}})$. Let \lesssim be a standard implication. For instance, $\mathbf{Conv}(\{x_n\}; \mathbb{R}, |\cdot|) \rightarrow \mathbf{Ax3}$ means that $\mathbf{Conv}(\{x_n\}; \mathbb{R}, |\cdot|) \lesssim \mathbf{Ax3}$. Suppose that \lesssim is extended to an ordering relation on the set of axioms $(a_1^{\mathbb{C}}, \dots, a_m^{\mathbb{C}})$, where $m < n$, from which we are going to infer the complex metrics. By definition 2 we can add new axioms for \mathbb{C} , for example: $a_n^{\mathbb{C}} = [[\mathbf{Conv}(\{x_n\}; \mathbb{R}, |\cdot|) \wedge \mathbf{CC}(\{x_n\}; \mathbb{R}, |\cdot|)] \rightarrow \mathbf{Ax3}] \rightarrow [[\mathbf{Conv}(\{x_n\}; \mathbb{C}, |\cdot|_c) \wedge \mathbf{CC}(\{x_n\}; \mathbb{C}, |\cdot|_c)] \rightarrow \mathbf{Ax3}]$. We have just added this axiom by *gal wahomer* (see definition 2).

3.3 Cauchy criterion in the functional analysis

By the end of the 19th century the mathematical objects for which applying **Ax3** seems quite natural have been exhausted. The worlds of real numbers, complex numbers, vectors, and even infinite sequences have already been studied enough and mathematicians wanted something more, especially because there were problems of variation calculus that works with maps of maps. It has been assumed for a long period that the spaces of maps usually are infinite and so different from the well-studied finite-dimensional spaces in their basic properties. Further, in the minds of founders of the functional analysis there were born metric spaces X , where you can select a distance (metric) $\rho : X^2 \rightarrow \mathbb{R}$ among elements, but it is impossible to determine a module (norm) of elements in any reasonable way.

Immediately, a question arose: what properties should X and ρ have in order to satisfy

$$\mathbf{CCrit}(\{x_n\}; X, \rho) = \text{true?}$$

It is clear that in this case we have to replace all expressions of the form $|x - y|$ by $\rho(x, y)$ at all occurrences in

$$\mathbf{CCrit}(\{x_n\}; \mathbb{R}, |\cdot|)$$

Obviously, the field of real numbers is a special case of metric space with the following metric: $\rho_{\mathbb{R}}(x, y) = |x - y|$. At the same time, the Cauchy criterion holds not in every metric space, for example, it does not hold in the space of continuous functions $C[0, 1]$ with the metric $\rho_{L_1}(x, y) = \int_0^1 |x(t) - y(t)| dt$. Cauchy himself, of course, knew nothing about the metric spaces and could hardly suspect that there will be the following definition:

DEFINITION 7. The metric space (X, ρ) where

$$\forall(\{x_n\} \subset X) [\mathbf{CCrit}(\{x_n\}; X, \rho) \leftrightarrow \exists(x \in X) [x_n \xrightarrow{X} x]],$$

is called complete.

The idea was to define the property $\mathbf{Tr}(x, y; X, \rho)$, natural for $(\mathbb{R}, |\cdot|)$, on the elements of (X, ρ) , by transforming this property from the derived theorem to the preexisted axiom, also by transforming $\mathbf{CCrit}(\{x_n\}; X, \rho)$ from the theorem to the axiom and by expecting that the final object will

behave like numbers with almost the same structural theorems. Surprisingly, it became true! For example, for \mathbb{R} there exists the compactness of the bounded and closed set and for the complete metric space of functions, continuous on $[0, 1]$, with the norm $\|x\|_C = \max_{t \in [0,1]} |x(t)|$, also there exists the compactness of the bounded set (with an additional condition of equicontinuity) – it is the claim of the Arzelà–Ascoli theorem, denoted by **AAT**, the analogue of **BWT**. The Cauchy–Cantor intersection theorem, i.e. **CCP**, gets its counterpart in the form of the topological Cauchy–Cantor intersection theorem, denoted by **TCCP**, for the closed non-empty nested subsets of X .

Hence, the mathematicians of the end of the 19th century proposed to replace the axiom **Ax3** by the theorem **CCrit**($\{x_n\}; X, \rho$) in the proof trees and, as a result, they changed all proof sequences invented by the mathematicians of the early 19th century. They did it to extend the mathematical limits in building new proof trees. For instance, if we define an ordering relation among non-logical axioms, \prec , as follows: $A \prec B$ iff an axiom A is contained in the proof tree of B , then the reasoning proposed by the mathematicians of the late 19th century is as follows:

$$\begin{aligned} & [\mathbf{CCrit}(\{x_n\}; X, \rho) \prec \mathbf{AAT}(\{x_n\}; X, \rho)] \wedge \\ & [\mathbf{Tr}(x, y; X, \rho) \prec \mathbf{AAT}(\{x_n\}; X, \rho)]; \\ & [\mathbf{CCrit}(\{x_n\}; X, \rho) \prec \mathbf{TCCP}(X, \rho)] \wedge \\ & [\mathbf{Tr}(x, y; X, \rho) \prec \mathbf{TCCP}(X, \rho)]. \end{aligned}$$

To sum up, it is important to point out that if the mathematicians followed the standard foundations of mathematics in fact, they would not change the proof sequence. But they did it indeed, because it is significant for drawing new trees by inference rules defined in definition 2. So, the new theorems **AAT**($\{x_n\}; X, \rho$) and **TCCP**(X, ρ) allow the mathematicians to apply definition 2 in their reasoning more often and in more cases.

4. CONCLUSIONS

We have just tried to show that the mathematicians deal not with a logical way of automatic proving from some axioms, but with combining proof trees on tree forests by using the analogy as an inference metarule. For the first time, such metarules were proposed in the Talmud within a general Judaic approach to concurrent or even massive-parallel conclusions, see [4], [5]. The mathematician does not think sequentially like a logical automaton, but concurrently, also.

In the logical foundations of mathematics there are two approaches in drawing computer-assisted proofs: (i) automated theorem proving (i.e. proving mathematical theorems by computer programs) and (ii) automated proof checking (i.e. using computer programs for checking proofs for correctness). There are many objections for these approaches. For instance, for (i) one the main objections is that these methods do not give new and useful concepts in mathematics in fact, but they present just a long gloomy calculation. For (ii) one of the main objections is that these methods can check just very simple theorems. There are no even insights how to check **FLT** by computer programs.

In our opinion, the most significant problem of existing logical foundations of mathematics is that a mathematical

proof is considered a discrete process that can be formalized by discrete methods. However, it is only a hypothesis that mathematics can be reduced to logic and the mathematical thinking is discrete. We can assume that it is not so and a mathematical proof is an operation in a space of proof trees with some topological properties. As a result, the mathematical proof can be formalized just by analog computations, not discrete ones. The meaning of mathematical proofs is to transform one space of proof trees to another space with inducing new topological properties. For example, in the Talmud this transformation means that each branch in proof trees with one root should have the same number of subbranches and inferring allows us to construct additional subbranches to make the trees more symmetric. In mathematics the goal of proofs is quite similar and it is to extend the mathematical limits to make proof trees in forests more symmetric, too.

Hence, we suppose that computer-assisted proofs can be based on some analog computations involving topological properties of proof trees.

Acknowledgment

For the first time, these ideas were expressed as a joke in the Andrew Schumann’s Russian-speaking blog:

<http://minski-gaon.livejournal.com/99986.html> (3 September 2016);

<http://minski-gaon.livejournal.com/100791.html> (21 September 2016).

5. REFERENCES

- [1] M. Artin, A. Grothendieck, and J.-L. Verdier, *Theorie des topos et cohomologie et ale des schemas, Seminaire de Geometrie Algebrique du Bois-Marie*, 4, Springer-Verlag, 1972.
- [2] Colin McLarty, What does it take to prove Fermat’s last theorem? Grothendieck and the logic of number theory, *Bulletin of Symbolic Logic* 16(3):359-377, 2010.
- [3] Paris, J. and Harrington, L. A Mathematical Incompleteness in Peano Arithmetic, in: *Handbook for Mathematical Logic* (Ed. J. Barwise). Amsterdam, Netherlands: North-Holland, 1977.
- [4] Schumann, A. Preface, *History and Philosophy of Logic*, 32(1):1-8, 2011.
- [5] Schumann, A. Qal wa-homer and Theory of Massive-Parallel Proofs, *History and Philosophy of Logic*, 32(1):71-83, 2011.
- [6] A. Wiles, Modular elliptic curves and Fermat’s Last Theorem, *Annals of Mathematics*, vol. 141, 1995, pp. 443-551.