

# A MATEMATIKA FILOZÓFIÁJA

## fizikalista megközelítés

E. Szabó László

*Logika Tanszék, ELTE BTK Filozófia Intézet*

<http://phil.elte.hu/leszabo>

2009. október 5.

## Tartalomjegyzék

Mi teszi a logika következtetési szabályait „helyessé”?	4
Mi tesz egy matematikai állítást igazzá?	5
Realizmus, platonizmus, intuicionizmus . . . . .	5
A MATEMATIKA FORMALISTA FELFOGÁSA . . . . .	6
Metodológiai megjegyzések . . . . .	10
The essential difference between mathematical truth and a semantical truth in a scientific theory describing something in the world	12
Mathematical objects have no meanings	15
Thesis . . . . .	15
Arguments . . . . .	15
How is it then possible that mathematical structures prove themselves to be so expressive in the physical	

applications?	22
<b>A matematikai elméletek mint formális rendszerek</b>	<b>25</b>
Példa (Paul Lorenzen) . . . . .	25
The sentence calculus . . . . .	26
<b>A predikátum kalkulus (PC)</b>	<b>27</b>
A PC axiómái és a következtetési szabályok . . . . .	27
Elemi tételek . . . . .	32
<b>Interpretáció</b>	<b>39</b>
Egy nem teljesen helyénvaló előzetes példa . . . . .	39
Interpretáció és modell . . . . .	40
Teljességi tétel . . . . .	45
PC(=) (predikátum kalkulus identitással) . . . . .	46
Az egyenlőség axiómái . . . . .	46
Aritmetika . . . . .	47
<b>Igaz-e, hogy <math>2 + 2 = 4</math>?</b>	<b>49</b>
<b>Halmazelmélet</b>	<b>55</b>
„Naiv” halmazelmélet — formális (axiomatikus) halmazel- mélet . . . . .	55
A halmazelmélet (ZF) axiómái . . . . .	55
<b>The physicalist ontology of formal systems</b>	<b>60</b>
<b>Abstraction is a move from the concrete to the concrete</b>	<b>69</b>
<b>Induction versus deduction</b>	<b>78</b>
<b>Meta-mathematical theory</b>	<b>83</b>
Ember tervez, Isten végez! . . . . .	89
<b>Turing-gépek</b>	<b>91</b>

A Turing-gép leírása . . . . .	91
Példák elemi műveleteket végrehajtó Turing-gépekre . . . . .	93
A Turing-gépek standard leírása . . . . .	95
Egy eldönthetetlen problémaosztály („Halting problem”) . . . . .	95
Univerzális Turing-gép . . . . .	97
Halting-probléma, episztemológia, endofizika . . . . .	98

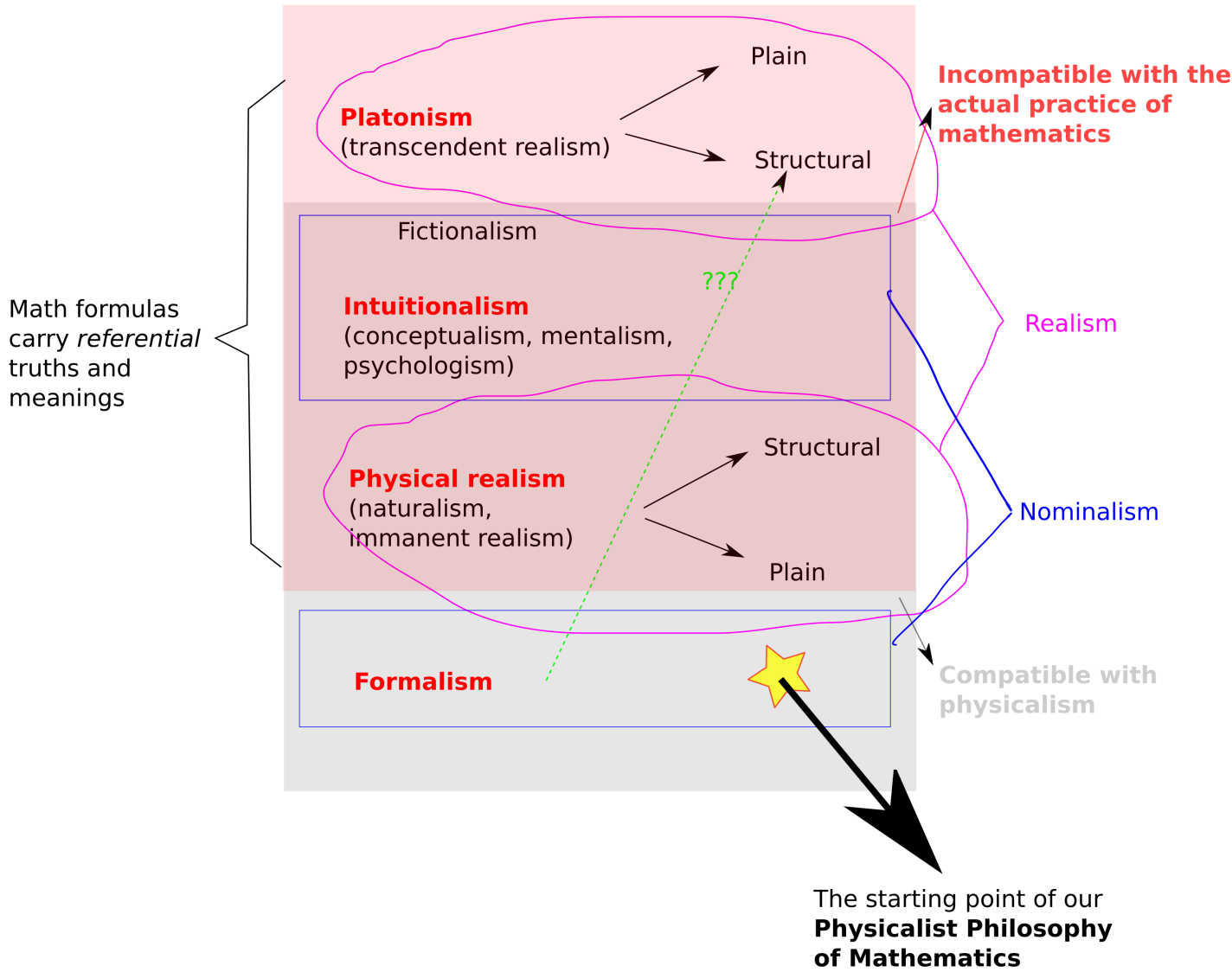
<b>Gödel inkomplettségi tétel</b>	<b>100</b>
-----------------------------------	------------

Gödel-számozás . . . . .	100
--------------------------	-----

Gödel-mondat . . . . .	101
------------------------	-----

<b>Bizonyítás és Igazság</b>	<b>104</b>
------------------------------	------------

<b>Gödel második inkomplettségi tétele</b>	<b>106</b>
--	------------



# Mi teszi a logika következtetési szabályait „helyessé”?

Elterjedt felfogás szerint az IGAZSÁG-MEGŐRZŐ TULAJDONSÁGA, vagyis, hogy igaz premisszákból igaz konklúziókra vezetnek.

Bár áttételesen beépül a racionális gondolkodás és érvelés társadalmilag/történetileg kialakult normáiba, mindenekelőtt a nyelv használatával összefüggő társadalmi normákba, s ezért úgy tűnhet, hogy semmiféle tapasztalásra nincs szükség egy következtetés helyességének megítéléséhez, ez a tulajdonság alapvetően EMPIRIKUSAN tesz-

telhető.

ha a premisszák igazak  $\Rightarrow$  a következtetések igazak  
 $\updownarrow$   $\updownarrow$   
világ tényei világ tényei

A logikai következtetés helyességének kérdése ott tűnik problematikusnak, ahol ezt a legkevésbé várnánk: A MATEMATIKÁBAN!  
Mi teszi helyessé azt a következtetést, hogy

ha az Euklideszi axiómák igazak  $\Rightarrow$  igaz, hogy  $a^2 + b^2 = c^2$

Honnan tudjuk ugyanis, hogy  $a^2 + b^2 = c^2$  igaz?!

## Mi tesz egy matematikai állítást igazzá?

### Realizmus, platonizmus, intuicionizmus

A REALIZMUS szerint (pl. J. S. Mill) *a matematikai állítások akkor igazak, ha megfelelnek a minket körülvevő fizikai valóságnak.* Más szóval, a matematika empirikus tudomány: a matematikai állítások a fizikai világ legáltalánosabb tulajdonságait fejezik ki. E felfogás fontos szerepet töltött be a matematika történetében, manapság azonban senki sem gondolja komolyan, hiszen a matematika fogalmai nincsenek közvetlen megfelelésben a valóság elemeivel, például a végtelen fogalmának semmi sem felel meg a külső (a matematikán kívüli) világban.

A MATEMATIKAI PLATONIZMUS a matematika klasszikus fogalmainak *önálló létezését tulajdonít*, függetlenül attól, gondoljuk-e azokat vagy nem, s úgy véli, a matematikai állítások igazságát pusztán e fogalmak analízisével, logikai úton fedezhetjük fel.

AZ INTUICIONISTÁK tagadják a matematikai objektumoknak – az értelemszerűen véges – *konstrukciójuktól független* létezését, ám

helyette „saját istenük” (Curry kifejezése<sup>1</sup>), az Intuíció létezésében hisznek, vagyis valami olyasmiben, ami az egyetemes emberi értelem számára *a priori adott*, garantálva ezzel a matematika *objektivitását és használhatóságát*.

REALISTÁK, PLATONISTÁK ÉS INTUICIONISTÁK mind hisznek azonban abban, hogy a matematikai állításoknak *jelentésük* van, s ha – a Hilbert-programot követve – formalizáljuk is a matematika nyelvezetét, azt azért tesszük, hogy e jelentést precízebben és tömörebben adhassuk vissza.

## A MATEMATIKA FORMALISTA FELFOGÁSA

szerint az igazság ezzel szemben az, hogy *a matematikai objektumoknak nincs jelentése*. A matematika a formális rendszerek tudománya: *Jeletet* definiálunk és *szabályokat*, melyek alapján e jeletet kombinálhatjuk. Ahogy Hilbert mondta „A matematika egy játék, melyet a papírlapra írt, jelentés nélküli szimbólumokkal játszunk, egyszerű szabályok szerint.” „Pont, egyenes és sík helyett folyamatosan mondhatnánk, asztalt, széket és söröskorsót” – mondta egy másik alkalommal az euklideszi geometriára utalva.

A matematikának semmi köze nincs a végtelen metafizikai fogalmához, és közömbös a térre, időre, valószínűségre vagy a folytonosságra vonatkozó intuíciónkkal szemben. A matematika nem produkál, és nem old meg Zénón-paradoxonokat! „Leírhatok egy jelet, mondjuk  $\alpha$ -t, és elnevezhetem az egész számok kardinalitásának. Aztán rögzíthetem a rá vonatkozó manipulációs szabályokat”, mondja Dieudonné.<sup>2</sup> Az egész finitista próbálkozás felesleges. Ha a papírra azt írom  $10^{10^{10}}$ , ez éppúgy csak egy jel, amellyel manipulálhatok, mint bármelyik más. A matematika jelenlegi gyakorlata azt mutatja, hogy minél precízebben látjuk be valamely matematikai állítás igazságát,

---

<sup>1</sup>Haskell B. Curry: *Outlines of a Formalist Philosophy of Mathematics*, North-Holland, Amsterdam 1951.

<sup>2</sup>Lásd Arend Heyting: *Intuitionism: an Introduction*, North-Holland, Amsterdam 1956.

annál nyilvánvalóbb, hogy őt kizárólag az teszi igazzá, hogy levezethető az rendszer axiómáiból a rendszerben érvényes következtetési szabályok segítségével – függetlenül attól, hogy egyébként milyen filozófiai nézeteket vall egy matematikus. Jól jellemzi a helyzetet *Jean Dieudonné*-nek, a francia Bourbaki csoport egyik vezéralakjának sokat idézett mondása :

In everyday life, we speak as Platonists, treating the objects of our study as real things that exist independently of human thought. If challenged on this, however, we retreat to some sort of formalism, arguing that in fact we are just pushing symbols around without making any metaphysical claims. Most of all, however, we want to do mathematics rather than argue about what it actually is. We're content to leave that to the philosophers.

Tehát,

1. A formalizmus lényege, hogy egy állítás bizonyításának/levezetésének létezése nem más, mint a szóban forgó állítás *igazságfeltétele*.
2. Az axiómák sem azért „igazak”, mert valamiféle referenciájuk van a valóságos (vagy valamiféle platóni) világra, hanem mert (triviálisan) levezethetők (tudniillik az axiómákból), más szóval definíció szerint igazak.
3. A matematikában az igazság fogalma általában értelmetlen, csak egy adott axiómarendszerre nézve értelmes (ahol az axiómarendszerbe a következtetési szabályokat is beleértjük). Annak a kijelentésnek, hogy „a háromszög szögeinek összege 180 fok” az igazságáról nincs értelme anélkül beszélnünk, hogy ne specifikálnánk, hogy melyik axiómarendszerben (tehát melyik geometriában) van értve.

4. A matematika története ebben a vonatkozásban nem egységes. A matematika reális interpretációja például szinte kihalt a nem-euklideszi geometriák megszületése után. Korábbi korokban elfogadottnak tekintett bizonyításokat ma nem tekintünk elfogadható, precíz formális bizonyításnak. Mint – kissé sarkítva – Russell írja Boole *Laws of Thought*-ja (1854) volt „az első könyv, amelyet matematikáról írtak”.



## 2.

Célunk tehát az, hogy megvizsgáljuk, mi tesz egy matematikai állítást igazzá. Filozófiai szempontból az alapvető probléma a következő:

If you are an incorrigible empiricist, you might encounter the following difficulties in connection with the truths of formal logic and mathematics. In Ayer's words:

For whereas a scientific generalization is readily admitted to be fallible, the truths of mathematics and logic appear to everyone to be necessary and certain. But if empiricism is correct no proposition which has a factual content can be necessary or certain. Accordingly the empiricist must deal with the truths of logic and mathematics in one of the following ways: he must say either that they are not necessary truths, in which case he must account for the universal conviction that they are; or he must say that they have no factual content, and then he must explain how a proposition which is empty of all factual content can be true and useful and surprising. ...

If neither of these courses proves satisfactory, we shall be obliged to give way to rationalism. We shall be obliged to admit that there are some truths about the world which we can know independently of experience; ...

# Metodológiai megjegyzések

- Legyen mindvégig világos: a matematikai igazság problémáját egy adott metafizikai pozíció (nyilván a saját metafizikai felfogásom) nézőpontjából tekintem végig. Ezt a pozíciót úgy fogom nevezni, hogy „fizikalizmus”.

**Physicalist =**

**Empiricism:** Genuine information about the world must be acquired by *a posteriori* means.

+

**Physicalist account of the mental:** Experiencing itself, as any other mental phenomena, including the mental processing the experiences, can be wholly explained in terms of physical properties, states, and events in the physical world.

- Mindez persze nem zárja majd ki más metafizikai álláspontok ismertetését és az azok mellett szóló érvek áttekintését
- Az előadás célja, hogy megoldjuk az Ayer által megfogalmazott problémát, anélkül azonban, hogy a legcsekélyebb mértékben fel kellene adnunk az empirista/fizikalista álláspontot.
- *My account of mathematics is philosophically (scientifically) motivated rather than “mathematically”.*  
The reason is very simple: I will claim that mathematics is a system of meaningless signs, consequently any claim *about* mathematics must be *external* to mathematics.
- *I cannot—as many philosophers of mathematics do—“take it as ‘data’ that most contemporary mathematics is correct” before first determining what this “correctness” actually means. Thus, with respect to the “philosophy first” and “philosophy last*

*if at all” dichotomy, I support the philosophy(-and-science)-first principle.*

The reason is that a big part of the corpus of contemporary mathematics is *not invariant* over the possible philosophical positions. *If Gödel is right* by saying that

after sufficient clarification of the concepts in question it will be possible to conduct these discussions with mathematical rigor and that the result then will be that...the platonistic view is the only one tenable

then it follows that *one cannot avoid questioning certain parts of mathematics from an empiricist/physicalist standpoint.*

- The central question is ‘What is mathematical truth, that is, what makes a mathematical proposition true?’.

I shall investigate the epistemological and ontological aspects of the problem. According to these two aspects I shall consistently ask the following two test questions:

Q1:           What leads us to the knowledge of the truth of a mathematical proposition?

Q2:           In what respects is the world different if a given mathematical proposition is true or false?

Investigation of the first question will lead us to a radical *formalist position*. Answering the second question, we will arrive at the *physicalist account of formal systems*.

# The essential difference between mathematical truth and a semantical truth in a scientific theory describing something in the world

A physical theory  $P$  is a formal system  $L$  + a semantics  $S$  pointing to the empirical world. In the construction of the formal system  $L$  one can *employ* previously prepared formal systems which come from mathematics and/or logic. That is, in general,  $L$  is a (first-order) system with

- some logical axioms and the derivation rules (usually the first-order predicate calculus with identity)
- the axioms of certain mathematical theories
- some physical axioms.

A sentence  $A$  in physical theory  $P$  can be true in *two different* senses:

Truth<sub>1</sub>:  $A$  is a theorem of  $L$ , that is,  $L \vdash A$  (which is a mathematical truth within the formal system  $L$ , a fact of the formal system  $L$ ).

Truth<sub>2</sub>: According to the semantics  $S$ ,  $A$  refers to an empirical fact (about the physical system described by  $P$ ).

For example, ‘The electric field strength of a point charge is  $\frac{kQ}{r^2}$ ’ is a theorem of Maxwell’s electrodynamics—one can derive the corresponding formal expression from the Maxwell equations. (This is a fact of the formal system  $L$ .) On the other hand, according to the semantics relating the symbols of the Maxwell theory to the empirical terms, this sentence corresponds to an empirical fact (about the point charges).

Truth<sub>1</sub> and Truth<sub>2</sub> are *independent concepts*, in the sense that one does not automatically imply the other. Moreover, assume that  $\Gamma$  is a set of true<sub>2</sub> sentences in  $L$ , i.e., each sentence in  $\Gamma$  refers to an empirical fact, and also assume that  $\Gamma \vdash A$  in  $L$ . It does not automatically follow that  $A$  is true<sub>2</sub>. Whether  $A$  is true<sub>2</sub> is again an empirical question. If so, then it is new empirically obtained information about the world, confirming the validity of the whole physical theory  $P = L + S$ . In **Kevin J. Davey** words:

The world is built in such a way that from certain bodies of knowledge, certain types of valid mathematical deductions take us from true claims about reality to other true claims about reality. But not all such otherwise valid mathematical deductions do this. This is a fact about the world that physicists must struggle to get their hands around — to learn which deductions are good in which contexts, and which are not.

If it turns out that  $A$  is not true<sub>2</sub>, then this information disconfirms the physical theory, *as a whole*. That is to say, one has to think about *revising one of the constituents* of  $P$ , the physical axioms, the semantics  $S$ , the mathematical axioms, or the axioms of logic or the derivation rules we applied in the derivation of  $A$ —probably in this order. (Here we can see how Quine’s *semantic holism* works. The unit of meaning is not the single sentence, but systems of sentences or even the whole of language).

We usually change one entire mathematical/logical theory for some other. For example, when we learned new empirical facts about physical space(-time), we replaced the whole Euclidean geometry with another one. (Unimportant sociological fact.)

The empirical disconfirmation of a physical theory, in which the Euclidean geometry is applied, can disconfirm the *applicability* of Euclidean geometry in the physical theory in question, *but it leaves Euclidean geometry itself intact*. In this way, one cannot disconfirm

mathematical truths like ‘formula  $a^2 + b^2 = c^2$  derives from the formulas called Euclidean axioms’. (Here we can observe how Quine’s *confirmational* holism fails. It is not the case that if an empirical finding disconfirms a physical theory which employs a given mathematical theory then it also disconfirms the mathematical theory.)

# Mathematical objects have no meanings

## Thesis

- *Mathematical objects and mathematical propositions carry no meanings.* The formulas are not about anything; they are just strings of symbols. Mathematics is a game played according to certain simple rules with *meaningless* marks on paper.
- In pure mathematics, the formulas of a formal system do not carry Tarskian truths. They are true only in the sense that they have proof in the formal system in question.

## Arguments

I apply the *verifiability theory of meaning*, in the following very weak sense: in order to determine the meaning of a scientific (mathematics included) statement we follow up how the statement in question is confirmed or disconfirmed. So the question is, how do we, finally, know that a mathematical proposition is true?

Consider the example of **the physical realist view** on mathematics.

1. If a mathematical proposition is an assertion about the physical world, then its truth-condition would be the correspondence with the physical facts. That is to say, *to decide whether a mathematical proposition is true or not, we would have to investigate the state of affairs in the physical world.* In this case, at a certain point, just like the physicist, the mathematician would have to throw down his/her pen and go to the laboratory. In **Gödel's** words:

If mathematics describes an objective world just like physics, there is no reason why inductive methods should not be applied in mathematics just the same as in physics.

But, have you ever seen a mathematician in the laboratory? Isn't it true that Gauss' proposal to *measure* (by optical instruments) the sum of the angles in a triangle is regarded as illegal from the point of view of contemporary mathematics?

— The truth of the statement  $a^2 + b^2 = c^2$  means only that  $a^2 + b^2 = c^2$  follows from the axioms, according to the derivation rules of that very formal system called Euclidean geometry.

What kind of experiment should be performed in the laboratory in order to decide whether the group-theoretical statement  $e(ee) = e$  ( $e$  is the identity element) is true or not?

— We do not refer to any experiment to decide this question. It is a true mathematical proposition in the sense that it is derived from the axioms of group theory.

*Thus, even if someone associates “meaning” to a mathematical proposition, this meaning is outside of the scope of the decisive mathematical considerations and is irrelevant from the point of view of the truth of the mathematical proposition in question.*

(This follows from the factual practice of the mathematician.)

**2.** According to the **weaker version of physical realism**, not all mathematical propositions have (empirical) meanings, only the most elementary ones (which usually constitute the axioms) expressing certain elementary facts about the physical world that are evident to everyone without laboratory experiments. The *axioms' reference to the real world is transmitted* to the more complex mathematical propositions.

However, as the following reflections show, this weaker understanding of physical realism suffers from the same difficulties as the stronger version.

- The axioms of set theory are often claimed to express evident truths about the “real sets” consisting of real objects of the world.



- But, what about the *axiom of choice*? It seems to express an elementary feature of the “real sets”, we are yet baffled about whether it should be added to the list of axioms or not, depending on some delicate mathematical considerations.
- It would be difficult to tell what feature of “real sets” is reflected in the *continuum hypothesis*.
- Such an unquestioned axiom as the *axiom of infinity* does not reflect any feature of “real sets” in the physical world. It rather reflects the wish of the set-theorist to have infinite sets.

So, it seems that the axioms of the most fundamental mathematical structures are chosen on the basis of *inherent mathematical reasonings*, rather than on the basis of physical facts.

- Assume, for the sake of the argument, that a kind of Truth<sub>2</sub> can be assigned to the axioms. *It would not follow that this Truth<sub>2</sub> can automatically be transmitted* to the more complex mathematical propositions derived from the axioms. It is, finally, an empirical question whether the *logical axioms and the derivation rules* we applied in the derivation preserve Truth<sub>2</sub> or not. (Reichenbach’s and Putnam’s explanation of the double-slit experiment via the violation of classical logic.)
- On the other hand, *the claim that sets and some elementary statements of set theory reflect physical facts is a description of the physical world*, that is to say, it is a physical theory. For example, whether the properties of physical objects can be modeled by sets, or not, is an empirical question. This is possible in classical physics, but it is far from obvious whether the same holds for quantum physics. If **Jauch and Piron’s description of the properties of physical objects** is correct,

then the property lattice of a quantum system does not satisfy the distributive law

$$\forall A \forall B \forall C [A \cap (B \cup C) = (A \cap B) \cup (A \cap C)] \quad (1)$$

consequently, *the properties of a quantum mechanical object cannot be identified with sets*. Thus, contrary to our intuitions about “properties”, “classes”, “having a property”, “belonging to a class”, etc., such a simple true<sub>1</sub> set-theoretic formula as (1) fails to be satisfied by the properties of a quantum mechanical object, that is to say, in this case it is not true<sub>2</sub>.

- Finally, the nonchalant allusion to the “evident” truth is problematic. As we can see in the history of modern sciences, sometimes the most “evident” concepts and laws of nature have to be revised under the pressure of new empirical findings. So, the mathematician should, at least, start by mentioning these “evident” empirical facts confirming the axioms. How are they precisely formulated? What kind of *objective* observations, experiments, or measurements lead us to these initial propositions about the physical world, etc.? However, no book, for example, on set theory ever starts with such an experimental introduction.

The reason is very simple: because such empirical/experimental introductions are not necessary. This is because *any claim about the “meaning” of a mathematical proposition, any reference to the physical world is irrelevant*. This is merely “verbal decoration” in mathematics, which can be completely ignored.

Everything I told about physical realism, *mutatis mutandis*, can be applied to **intuitionism and platonic realism**.

physical world  
⇒  
world of mental/psychological phenomena or  
platononic world/(fiction/natural langugae)

laboratory experiment  
⇒  
other means by which knowledge can be  
obtained about these other realms

But, *we never see any reference to these epistemic faculties* in mathematics! In **Dummett's** words:

[T]here is nothing in mathematics that could be described as inference to the best explanation. Above all, we do not seek, in order to refute or confirm a hypothesis, a means of refining our intuitive faculties, as astronomers seek to improve their instruments. Rather, if we suppose the hypothesis true, we seek for a *proof* of it, and it remains a mere hypothesis, whose assertion would therefore be unwarranted, until we find one.

No doubt, **Gödel** is right in arguing that

“we do have something like a perception also of the objects of set theory”

and that there is no reason

“why we should have less confidence in this kind of perception, i. e., in mathematical intuition, than in sense perception”.

But, he seems to overlook the fact that *both the sense perceptions of the external world and the internal perceptions of our intuitive ideas are irrelevant* from the point of view of the truths which mathematics

is concerned with, because they both are perceptions of that part of reality which is external to mathematics.

The only “truth” mathematics is concerned with is  $\text{Truth}_1$ , the concept of that something is being proved.

If we feel obliged to express the statements of mathematics in our everyday, no doubt referential, language, we should say that mathematics is a discipline which does not have such assertions as ‘ $a^2 + b^2 = c^2$ ’, or ‘ $e(ee) = e$ ’, or ‘ $3 + 2 = 5$ ’, but it has assertions like ‘formula  $a^2 + b^2 = c^2$  derives from the formulas called Euclidean axioms’, ‘formula  $e(ee) = e$  derives from the formulas called the axioms of group’, and ‘formula  $3 + 2 = 5$  derives from the formulas called the axioms of arithmetic’, etc.— according to a given set of logical axioms and rule(s) of inference. That is what mathematics tells us about the world. (We will see in what sense this statement is *about the world*.)

Many are unsatisfied with this simple “if-thenism”. **David Papineau** writes:

If if-thenism were true, then of course there would not be any gap between mathematical practice and mathematical truth, for mathematical truth would not answer to anything more than logical facts about which theorems followed from which axioms. And so, if if-thenism were true, there would not be any difficulties about mathematical knowledge (or at least none which did not reduce to the more general topic of logical knowledge). However, despite these attractions, it is generally agreed that if-thenism is false. The reason is that, although there are some branches of mathematics in which mathematicians are not committed to anything except exploring the consequences of postulates, there are other branches of mathematics where they are unquestionably committed to something more: for example, ... the number theorist

who says there are infinitely many prime numbers is not just saying that this follows from Peano's postulates, but that there *are* all those numbers.<sup>3</sup>

Of course, very much depends on how we understand the practice of the mathematician. I think, that in the same way like ordinary people who dream about movie stars but live with their partner, mathematicians rave about various platonic objects, but if they are seriously asked what they are confident about, they reduce their claims to mere if-thenisms. All the rest is just "folklore". And this holds not only for the more complex branches of mathematics but also for arithmetic and set theory. When the number theorist says "there are infinitely many prime numbers", or simply '7 is a prime number larger than 5' then (s)he means that all these concepts as 'larger' and 'prime', etc., are defined with formal rigor, and that the statement in question is a *theorem* within the corresponding formal framework. This is true even if (s)he proves it by simple calculations, since it was previously *proved* that the employed algorithm is correct. (I will come back to this problem after some ontological reflections.) Thus, concerning the rigorous, scientifically justified, non-folkloristic part of the claims of mathematicians, it is far from "unquestionable" that they are committed to something more than if-thenism.

---

<sup>3</sup>Papineau 1990, p. 167.

# How is it then possible that mathematical structures prove themselves to be so expressive in the physical applications?

Feynman:

I find it quite amazing that it is possible to predict what will happen by mathematics, which is simply following rules which really have nothing to do with the original thing.

There is nothing miraculous here:

- First of all, **we give too much importance** to the correspondence between mathematics and structures observed in the empirical world. **The “storehouse” of mathematics has a much larger stock of mathematical structures** than we have ever applied in describing the real world.
- It is **not mathematics alone** by which the physicist can predict what will happen, but the whole physical theory  $(L, S)$ , including physical axioms and the semantics. The physicist **tunes both the formal system  $L$  (including the syntax, the derivation rules, the logical/mathematical/physical axioms) and the semantics  $S$**  such that the theorems *derivable* from the unified logical, mathematical, and physical axioms be compatible with the empirical facts. That is to say, the **employed logical and mathematical structures in themselves need not reflect any Truth<sub>2</sub> about the real world in order to be useful.**
- Due to the empirical underdetermination of scientific theories, it can be the case that **more than one possible mathematical structure is applicable** to the same empirical reality.

- Finally, there is **no miraculous “preadaption”** involved just because certain aspects of empirical reality “fit themselves into the forms provided by mathematics”. This is simply a result of **selections** made by the physicist.

Just as there is no preadaption at work when you successfully can install kitchen units obtained from a department store in your kitchen. The rules according to which the shelves, cupboards and doors can be combined show nothing about the actual geometry of your kitchen. But the final result is that the kitchen “fits itself” to the form of the whole set, as if through a kind of preadaption. Similarly, as **G. Y. Nieuwland** points out:

From time immemorial, mankind has learnt to deploy mathematics in order to cope with the empirical world, finding helpful notions such as quantity, measure, pattern and functional dependence. For many and obvious reasons it proves expedient to order this experience into a system of knowledge, invoking notions of coherence, analogy, completion and logic. Doing so, it appears that the margins of interpretation, of whatever it is of our experience that can be organized into mathematical structure, are flexible. Such is the underdetermination of theory by the data.

Of course, I do not want to understate the **real epistemological problem** here. Namely, to what extent a statement of a physical theory applying certain mathematical structures expresses an objective feature of the real object. This long-standing epistemological problem is, however, **related to the Truth<sub>2</sub> of the statements of physical theories, which has nothing to do with what makes a mathematical proposition true, that is, with Truth<sub>1</sub>.**

Az a tézis, mely szerint abból a tényből, hogy a matematikai struktúrák nélkülözhetetlenek a fizikai elméletekben, következtethetünk a matematikai struktúrák létezésére, vagyis hogy a matematikai objektumok ontológiai státusza nem különbözik a fizikai objektumok ontológiai státuszától, gyakran a **Quine-Putnam-féle „indispensability argument”**-re épül:

- (1) We ought to have ontological commitment to all and only the entities that are indispensable to our best scientific theories.
- (2) Mathematical entities are indispensable to our best scientific theories.

- 
- (3) We ought to have ontological commitment to mathematical entities.

This means, we ought to have ontological commitment to the entities that mathematical *statements* are talking about.

Vissza fogunk térni erre az argumentumra a későbbiekben.



# A matematikai elméletek mint formális rendszerek

Általában tehát egy matematikai elmélet egy formális nyelv, amely szimbólumokat tartalmaz, szintaktikai szabályokat arra nézve, hogy ezekből a szimbólumokból hogyan lehet összetettebb un. formulákat és formula-sorozatokat előállítani, és logikai szabályokat, amelyek következtetési szabályokat mondanak ki bizonyos formulák „átalakítására”, egyikről a másikra való „áttérésre”.

## Példa (Paul Lorenzen)

**Jelek:** . Olyan stringek, amelyek két betűből állnak,  $a$  és  $b$ .

**Axiómák:**

$$L = \begin{cases} a \\ X \vdash Xb & \text{(Rule 1)} \\ X \vdash aXa & \text{(Rule 2)} \end{cases}$$

Például,

**Tétel:**  $aababb$

**Bizonyítás:**

$$\begin{array}{ccccccc} a & \vdash & ab & \vdash & aaba & \vdash & aabab & \vdash & aababb \\ (1) & & (2) & & (1) & & (1) & & \end{array}$$

# The sentence calculus

## Alphabet of symbols:

$\sim, \supset, (, ), p, q, r$ , etc.

## Well-formed formulas:

1.  $p, q, r$ , etc. are wfs.
2. If  $A, B$  are wfs. then  $(\sim A), (A \supset B)$ , are wfs.
3. All wfs. are generated by 1. and 2.

## Axiom schemes:

(L1)  $A \supset (B \supset A)$

(L2)  $(A \supset (B \supset C) \supset ((A \supset B) \supset (A \supset C)))$

(L3)  $((\sim A) \supset (\sim B) \supset (B \supset A))$

## Modus Ponens:

(MP)  $A$  and  $(A \supset B)$  implies  $B$

# A predikátum kalkulus (PC)

## A PC axiómái és a következtetési szabályok

A PC egy, a fenti értelemben vett formális nyelv +

### Következtetési szabályok

(MP)  $\phi$ -ből és  $(\phi \rightarrow \psi)$ -ből következik  $\psi$  (Modus Ponens)

(G)  $\phi$ -ből következik  $\forall x\phi$  (Generalizáció)

### Axiómák (Axióma sémák)

A következőkben,  $\phi, \psi, \chi$  formulák,  $x, y, y_1, y_2, \dots, y_n, \dots$  változók, és jelölje  $\phi(y)$  az a formulát, melyet úgy kapunk, hogy a  $\phi(x)$  formulában az  $x$  változót, annak minden szabad előfordulása esetében  $y$ -nal helyettesítjük.

(PC1)  $(\phi \rightarrow (\psi \rightarrow \phi))$

(PC2)  $((\phi \rightarrow (\psi \rightarrow \chi)) \rightarrow (\phi \rightarrow \psi) \rightarrow (\phi \rightarrow \chi))$

(PC3)  $((\neg\phi \rightarrow \neg\psi) \rightarrow (\psi \rightarrow \phi))$

(PC4)  $(\forall x(\phi \rightarrow \psi) \rightarrow (\phi \rightarrow \forall x\psi))$  ha  $x$  nem fordul elő szabadon  $\phi$ -ben.

(PC5)  $(\forall x\phi \rightarrow \phi)$  ha  $x$  nem fordul elő szabadon  $\phi$ -ben.

(PC6)  $(\forall x\phi(x) \rightarrow \phi(t))$  feltéve, hogy a  $t$  terminus szabad  $x$ -re nézve  $\phi(x)$ -ben.

PC axiómáinak halmazát úgy fogjuk jelölni, hogy  $\{PC\}$

## Megjegyzés

- Az axiómák tehát egyszerűen a nyelv kiválasztott formulái. („Alapigazságok”, stb. csak verbális dekoráció).
- Egy formális nyelv + a következtetési szabályok + néhány axióma együttesét általában *formális rendszernek* hívjuk.

**Bizonyítás, tétel.** Legyen  $\Sigma$  egy formulahalmaz (további axiómák). *Bizonyítás* formuláknak egy (véges) sorozata, úgy, hogy mind-egyik formula vagy eleme  $\{PC\} \cup \Sigma$  axiómahalmaznak, vagy a sorozatban szereplő korábbi formulából van levezetve a következtetési szabályok valamelyikének alkalmazásával. Egy bizonyítást alkotó formulasorozat elemeit *tételnek* nevezzük. Ha  $\phi$  tétel, azt így szoktuk jelölni:  $\{PC\} \cup \Sigma \vdash \phi$ .

$\vdash \phi$ . Gyakran  $\{PC\} \vdash \phi$  gyakran úgy jelölik, hogy  $\vdash \phi$ . (Mintha „semmiből” következne, de ez ne tévesszen meg minket: a PC axiómából következik!) Ennek megfelelően, ha  $\{PC\} \cup \Sigma \vdash \phi$ , azt úgy is szokás jelölni, hogy  $\Sigma \vdash \phi$  (vagyis a PC axiómái hallgatólagosan mindig oda vannak értve).

**Konzisztencia.** Formulák egy  $\Sigma$  halmazáról azt mondjuk, hogy *konzisztens*, ha nem létezik olyan  $\phi$  formula, melyre egyszerre fennállna, hogy  $\Sigma \vdash \phi$  és  $\Sigma \vdash \neg\phi$ .

**Bizonyítottan ekvivalens formulák.** Két  $\phi$  és  $\psi$  formula *bizonyítottan ekvivalens*, ha  $\vdash (\phi \leftrightarrow \psi)$ .

---

## Kis kitérő: *Kijelentéskalkulus*

Alphabet of symbols:

$\sim, \supset, (, ), p, q, r$ , etc.

Well-formed formulas:

1.  $p, q, r$ , etc. are wfs.
2. If  $A, B$  are wfs. then  $(\sim A), (A \supset B)$ , are wfs.
3. All wfs. are generated by 1. and 2.

Axiom schemes:

(SC1)  $A \supset (B \supset A)$

(SC2)  $(A \supset (B \supset C)) \supset ((A \supset B) \supset (A \supset C))$

(SC3)  $((\sim A) \supset (\sim B) \supset (B \supset A))$

Modus Ponens:

(MP)  $A$  and  $(A \supset B)$  implies  $B$

## A kijelentéskalkulus konzisztenciájának „bizonyítása”

**Definition:** A *coloring* of SC is a function  $v$  whose domain is the set of wfs. of SC and whose range is the set  $\{red, blue\}$  such that, for any wfs.  $A, B$  of SC,

(i)  $v(A) \neq v(\sim A)$

(ii)  $v(A \supset B) = \textit{blue}$  if and only if  $v(A) = \textit{red}$  and  $v(B) = \textit{blue}$

**Definition:** A wfs.  $A$  is *stably red* if for every coloring  $v$ ,  $v(A) = \textit{red}$ .

**Proposition 1:** For every formula  $A$ , if  $A$  is a theorem of SC then  $A$  is stably red.

**Proof:** Let  $A$  be a theorem. The proof is by induction on the number  $n$  of wfs. of SC in a sequence of wfs. which constitutes a proof of  $A$  in SC.

$n = 1$   $A$  is an axiom. One can easily verify that every axiom of SC is stably red.

$n > 1$  Induction hypothesis: all theorems of SC which have proofs in fewer than  $n$  steps are stably red. Assume that the proof of  $A$  contains  $n$  wfs. Now, either  $A$  is an axiom, in which case it is stably red, or  $A$  follows by (MP) from previous wfs. in the proof. These two wfs. must have the form  $B$  and  $(B \supset A)$ . But, since  $B$  and  $(B \supset A)$  are stably red, it follows from (ii) that  $A$  is stably red.

**Proposition 2:** SC is consistent.

**Proof:** As is known (nemsokára be fogjuk bizonyítani!), one can easily proof that if both  $X$  and  $\sim X$  are theorems in SC then arbitrary formula is a theorem. Consequently, if there exists at least one formula in SC which is not a theorem, then SC is consistent. By virtue of Proposition 1 one has to show that there is a formula  $Y$  in SC which is not stably red, that is, there is a coloring  $v$  such that

$v(Y) = \textit{blue}$ . Let  $Y$  be  $\sim p \supset q$ . Taking into account (i) and (ii),  $v(Y)$  is determined by  $v(p)$  and  $v(q)$ . Since  $v(Y) = \textit{blue}$  whenever  $v(p) = \textit{blue}$  and  $v(q) = \textit{blue}$ ,  $Y$  cannot be a theorem of SC.

## Formális (kétértékű) értékelés (szemantika)

### Igazságérték

Igazságérték egy olyan függvény, amelynek értelmezési tartománya a SC formálinak halmaza, értékkészlete pedig az  $\{\textit{Igaz}, \textit{Hamis}\}$  halmaz, és az alábbi tulajdonságokat elégíti ki: A PC tetszőleges két  $A, B$  formulájára

$$(i) \quad v(A) \neq v(\sim A)$$

$$(ii) \quad v(A \supset B) = \textit{Hamis} \text{ akkor és csak akkor ha } v(A) = \textit{Igaz} \text{ és } v(B) = \textit{Hamis}$$

### Tautológia

Az  $A$  formulát tautológiának nevezzük, ha tetszőleges  $v$  igazságértékfüggvényre teljesül, hogy  $v(A) = \textit{Igaz}$ .

A fenti tételekből következik, hogy az SC minden axiómája tautológia, és SC minden tétele tautológia.

---

# Elemi tételek

1. **Tétel.** *Tetszőleges  $\phi$  formulára  $\phi \rightarrow \phi$ .*

**Bizonyítás.**

1.  $\phi \rightarrow ((\phi \rightarrow \phi) \rightarrow \phi)$  [(PC1)-ből]
2.  $(\phi \rightarrow ((\phi \rightarrow \phi) \rightarrow \phi)) \rightarrow (\phi \rightarrow (\phi \rightarrow \phi)) \rightarrow (\phi \rightarrow \phi)$  [(PC2)-ből]
3.  $(\phi \rightarrow (\phi \rightarrow \phi)) \rightarrow (\phi \rightarrow \phi)$  [1. és 2. alapján (MP)-ből]
4.  $\phi \rightarrow (\phi \rightarrow \phi)$  [(PC1)-ből]
5.  $\phi \rightarrow \phi$  [4. és 3. alapján (MP)-vel]

2. **Tétel (Szintaktikai kompaktság).** *Legyen  $\Sigma$  formulák egy halmaza.  $\Sigma \vdash \phi$ , akkor és csak akkor, ha  $\Sigma$  valamely véges  $\Sigma'$  részére  $\Sigma' \vdash \phi$ .*

**Bizonyítás.** A tétel triviális következménye annak a ténynek, hogy minden bizonyítás formulák egy *véges* sorozata.

3. **Tétel.** *Ha a  $\Sigma$  formulahalmaz inkonzisztens (nem konzisztens), akkor tetszőleges formula levezethető belőle, tehát  $\Sigma \vdash \phi$  minden  $\phi$ -re.*

**Bizonyítás.** Feltevésünk szerint tehát van olyan  $\psi$  formula, hogy  $\Sigma \vdash \psi$  és ezzel együtt  $\Sigma \vdash \neg\psi$ . Legyen  $\phi$  tetszőleges. Most megadjuk  $\phi$  egy levezetését  $\Sigma$ -ból:

- (1)  $\neg\psi$  [feltétel]
- (2)  $\neg\psi \rightarrow (\neg\phi \rightarrow \neg\psi)$  [(PC1)]



$$(3) \quad \neg\phi \rightarrow \neg\psi \quad [(1), (2), (\text{MP})]$$

$$(4) \quad (\neg\phi \rightarrow \neg\psi) \rightarrow (\psi \rightarrow \phi) \quad [(\text{PC3})]$$

$$(5) \quad \psi \rightarrow \phi \quad [(3), (4), (\text{MP})]$$

$$(6) \quad \psi \quad [\text{feltétel}]$$

$$(7) \quad \phi \quad [(5), (6), (\text{MP})]$$

**4. Tétel (Dedukciótétel).**  $\Sigma \cup \{\phi\} \vdash \psi$ , és  $\psi$  levezetése nem tartalmazza  $(G)$  alkalmazását olyan  $x$  változóra nézve, amelyik szabadon fordul elő  $\phi$ -ben, akkor  $\Sigma \vdash \phi \rightarrow \psi$ .

**Bizonyítás.** Ha  $\Sigma \cup \{\phi\} \vdash \psi$ , akkor létezik olyan

$$\chi_1, \chi_2, \dots, \chi_k, \dots, \chi_n$$

formulasorozat, amelyik  $\psi$  bizonyítása. Teljes indukcióval megmutatjuk, hogy a tétel a bizonyításban szereplő minden  $\chi_k$  formulára igaz, tehát igaz  $\chi_n$ -re (azaz  $\psi$ -re) is.

Tekintsük  $\chi_1$ -et.  $\chi_1$  vagy logikai axióma, vagy eleme  $\Sigma$ -nak, vagy azonos  $\phi$ -vel. Az első két esetben (PC1)-ből és (MP)-ből kapjuk, hogy  $\Sigma \vdash \phi \rightarrow \chi_1$ . Ha  $\chi_1$  azonos  $\phi$ -vel, akkor az 1. tételből triviálisan következik. Ezzel beláttuk, hogy a tétel igaz  $\chi_1$ -re.

(Indukciós feltevés) Állításunk igaz minden  $\chi_i$ -re, ha  $i < k$ .

Ennek alapján megmutatjuk, hogy igaz  $\chi_k$ -ra. Három lehetőség van:

1.  $\chi_k$  logikai axióma, vagy eleme  $\Sigma \cup \{\phi\}$ -nek. Ekkor ugyanúgy bizonyítunk, mint a  $\chi_1$  esetében.
2.  $\chi_k$ -t az (MP) alkalmazásával kaptuk valamely korábbi  $\chi_i$  és  $\chi_i \rightarrow \chi_k$  formulák alapján. Ekkor a következőképpen bizonyítunk:

$\phi \rightarrow \chi_i$  [(Indukciós feltevés)]  
 $\phi \rightarrow (\chi_i \rightarrow \chi_k)$  [(Indukciós feltevés)]  
 $(\phi \rightarrow (\chi_i \rightarrow \chi_k)) \rightarrow ((\phi \rightarrow \chi_i) \rightarrow (\phi \rightarrow \chi_k))$  [(PC2)-ből]  
 $(\phi \rightarrow \chi_i) \rightarrow (\phi \rightarrow \chi_k)$  [(MP)-ből]  
 $\phi \rightarrow \chi_k$  [(MP)-ből]

3.  $\chi_k$ -t az (G) alkalmazásával kaptuk valamely korábbi  $\chi_i$ -ből valamely  $y$  változóra vett generalizációval. Mivel a levezetés nem tartalmazza (G) alkalmazását olyan  $x$  változóra nézve, amelyik szabadon fordul elő  $\phi$ -ben,  $y$  nem jelenthet meg  $\phi$ -ben szabad változóként, hiszen a generalizációban alkalmaztuk. Tehát

$\phi \rightarrow \chi_i$  [(Indukciós feltevés)]  
 $\forall y (\phi \rightarrow \chi_i)$  [(G)-ből]  
 $\forall y (\phi \rightarrow \chi_i) \rightarrow (\phi \rightarrow \forall y \chi_i)$  [(PC4)-ből]  
 $\phi \rightarrow \forall y \chi_i$  [(MP)-ből]  
 $\phi \rightarrow \chi_k$

Ezzel a tételt bebizonyítottuk.

**5. Tétel.** *Ha  $\Sigma \cup \{\phi\} \vdash \psi$  és  $\phi$  zárt, akkor  $\Sigma \vdash \phi \rightarrow \psi$ .*

A dedukciótétel alkalmazásával további fontos és gyakran használható tételeket bizonyítunk.

**6. Tétel (Hipotetikus Szillogizmus (HS)).** *Tetszőleges  $\phi, \psi$  és  $\chi$  esetén:  $\{\phi \rightarrow \psi, \psi \rightarrow \chi\} \vdash \phi \rightarrow \chi$*

***Bizonyítás.***

(1)  $\phi \rightarrow \psi$  [feltevés]

(2)  $\psi \rightarrow \chi$  [feltevés]

(3)  $\phi$  [feltevés]

(4)  $\psi$  [(1), (3), MP]

$$(5) \quad \chi \quad [(2), (4), \text{MP}]$$

Bebizonyítottuk tehát, hogy  $\{\phi \rightarrow \psi, \psi \rightarrow \chi, \phi\} \vdash \chi$ . Végül, a dedukciótétel alkalmazásával kapjuk, hogy  $\{\phi \rightarrow \psi, \psi \rightarrow \chi\} \vdash \phi \rightarrow \chi$ .

**7. Tétel.** *Tetszőleges  $\phi$ -re és  $\psi$ -re:*  $\neg\psi \rightarrow (\psi \rightarrow \phi)$

***Bizonyítás.***

$$(1) \quad \neg\psi \rightarrow (\neg\phi \rightarrow \neg\psi) \quad [(\text{PC1})]$$

$$(2) \quad (\neg\phi \rightarrow \neg\psi) \rightarrow (\psi \rightarrow \phi) \quad [(\text{PC3})]$$

$$(3) \quad \neg\psi \rightarrow (\psi \rightarrow \phi) \quad [(1), (2), (\text{HS})\text{-tétel}]$$

**8. Tétel.** *Tetszőleges  $\phi$ -re:*  $(\neg\phi \rightarrow \phi) \rightarrow \phi$

***Bizonyítás.*** Először azt fogjuk megmutatni, hogy  $\{\neg\phi \rightarrow \phi\} \vdash \phi$ :

$$(1) \quad \neg\phi \rightarrow \phi \quad [\text{feltevés}]$$

$$(2) \quad \neg\phi \rightarrow (\neg\neg(\neg\phi \rightarrow \phi) \rightarrow \neg\phi) \quad [(\text{PC1})]$$

$$(3) \quad (\neg\neg(\neg\phi \rightarrow \phi) \rightarrow \neg\phi) \rightarrow (\phi \rightarrow \neg(\neg\phi \rightarrow \phi)) \quad [(\text{PC3})]$$

$$(4) \quad \neg\phi \rightarrow (\phi \rightarrow \neg(\neg\phi \rightarrow \phi)) \quad [(2), (3), (\text{HS})]$$

$$(5) \quad (\neg\phi \rightarrow (\phi \rightarrow \neg(\neg\phi \rightarrow \phi))) \\ \rightarrow ((\neg\phi \rightarrow \phi) \rightarrow (\neg\phi \rightarrow \neg(\neg\phi \rightarrow \phi))) \quad [(\text{PC2})]$$

$$(6) \quad (\neg\phi \rightarrow \phi) \rightarrow (\neg\phi \rightarrow \neg(\neg\phi \rightarrow \phi)) \quad [(4), (5), (\text{MP})]$$

$$(7) \quad \neg\phi \rightarrow \neg(\neg\phi \rightarrow \phi) \quad [(1), (6), (\text{MP})]$$

$$(8) \quad (\neg\phi \rightarrow \neg(\neg\phi \rightarrow \phi)) \rightarrow ((\neg\phi \rightarrow \phi) \rightarrow \phi) \quad [(\text{PC3})]$$

$$(9) \quad (\neg\phi \rightarrow \phi) \rightarrow \phi \quad [(7), (8), (\text{MP})]$$

$$(10) \quad \phi \quad [(1), (9), (\text{MP})]$$

Innen a tétel állítása a dedukciótétellel azonnal adódik.

**9. Tétel.** *Tetszőleges  $\phi$ -re:*  $\neg\neg\phi \rightarrow \phi$

**Bizonyítás.** Először azt fogjuk megmutatni, hogy  $\{\neg\neg\phi\} \vdash \phi$ :

$$(1) \quad \neg\neg\phi \quad [\text{feltevés}]$$

$$(2) \quad \neg\neg\phi \rightarrow (\neg\phi \rightarrow \neg\neg\phi) \quad [(\text{PC1})]$$

$$(3) \quad \neg\phi \rightarrow \neg\neg\phi \quad [(1), (2), (\text{MP})]$$

$$(4) \quad (\neg\phi \rightarrow \neg\neg\phi) \rightarrow (\neg\phi \rightarrow \phi) \quad [(\text{PC3})]$$

$$(5) \quad \neg\phi \rightarrow \phi \quad [(3), (4), (\text{MP})]$$

$$(6) \quad \phi \quad [(5), 8. \text{ Tétel}, (\text{MP})]$$

Innen a tétel állítása a dedukciótétellel azonnal következik.

Ezt felhasználva, adódik a fordított irányú tétel:

**10. Tétel.** *Tetszőleges  $\phi$ -re:*  $\phi \rightarrow \neg\neg\phi$

**Bizonyítás.**

$$(1) \quad \neg\neg\neg\phi \rightarrow \neg\phi \quad [9. \text{ Tétel}]$$

$$(2) \quad (\neg\neg\neg\phi \rightarrow \neg\phi) \rightarrow \phi \rightarrow \neg\neg\phi \quad [(\text{PC3})]$$

$$(3) \quad \phi \rightarrow \neg\neg\phi \quad [(1), (2), (\text{MP})]$$

A 9. és 10. Tételeket számos további tétel levezetésénél használhatjuk.

**11. Tétel.** *Tetszőleges  $\phi$ -re és  $\psi$ -re:*  $(\phi \rightarrow \psi) \rightarrow (\neg\psi \rightarrow \neg\phi)$ .

### ***Bizonyítás.***

- (1)  $\phi \rightarrow \psi$  [feltétel]
- (2)  $\neg\neg\phi \rightarrow \phi$  [9. Tétel]
- (3)  $\neg\neg\phi \rightarrow \psi$  [(1), (2), (HS)]
- (4)  $\psi \rightarrow \neg\neg\psi$  [10. Tétel]
- (5)  $\neg\neg\phi \rightarrow \neg\neg\psi$  [(3), (4), (HS)]
- (6)  $(\neg\neg\phi \rightarrow \neg\neg\psi) \rightarrow (\neg\psi \rightarrow \neg\phi)$  [(PC3)]
- (7)  $\neg\psi \rightarrow \neg\phi$  [(5), (6), (MP)]

Végül a dedukciótételt alkalmazzuk.

**12. Tétel.** *Tetszőleges  $\phi$ -re és  $\psi$ -re:  $\{\phi \rightarrow \psi, \phi \rightarrow \neg\psi\} \vdash \neg\phi$ .*

### ***Bizonyítás.***

- (1)  $\phi \rightarrow \psi$  [feltétel]
- (2)  $\phi \rightarrow \neg\psi$  [feltétel]
- (3)  $(\phi \rightarrow \psi) \rightarrow (\neg\psi \rightarrow \neg\phi)$  [(PC3)]
- (4)  $\neg\psi \rightarrow \neg\phi$  [(1), (3), (MP)]
- (5)  $\phi \rightarrow \neg\phi$  [(2), (4), (HS)]
- (6)  $(\phi \rightarrow \neg\phi) \rightarrow (\neg\neg\phi \rightarrow \neg\phi)$  [11. Tétel]
- (7)  $\neg\neg\phi \rightarrow \neg\phi$  [(5), (6), (MP)]
- (8)  $(\neg\neg\phi \rightarrow \neg\phi) \rightarrow \neg\phi$  [8. Tétel]
- (9)  $\neg\phi$  [(7), (8), (MP)]

**13. Tétel (Indirekt bizonyítás).** *Legyen  $\Sigma$  formulák egy halmaza és legyen  $\phi$  tetszőleges formula.  $\Sigma \vdash \phi$  akkor és csak akkor, ha a  $\Sigma \cup \{\neg\phi\}$  inkonzisztens.*

**Bizonyítás.** Ha  $\Sigma \vdash \phi$ , akkor  $\Sigma \cup \{\neg\phi\} \vdash \phi$ . Másrészt  $\Sigma \cup \{\neg\phi\} \vdash \neg\phi$ , tehát  $\Sigma \cup \{\neg\phi\}$  valóban inkonzisztens.

Fordítva, ha  $\Sigma \cup \{\neg\phi\}$  inkonzisztens, akkor van olyan  $\psi$ , hogy  $\Sigma \cup \{\neg\phi\} \vdash \psi$  és  $\Sigma \cup \{\neg\phi\} \vdash \neg\psi$ . Tehát, a 4. tétel miatt  $\Sigma \vdash \neg\phi \rightarrow \psi$ . (Mivel ha  $\Sigma \cup \{\neg\phi\}$  inkonzisztens,  $\psi$  mindig választható olyannak, hogy a dedukció-tétel feltételei teljesüljenek.) Hasonlóan kapjuk, hogy  $\Sigma \vdash \neg\phi \rightarrow \neg\psi$ . Alkalmazva a 12. Tételt,  $\Sigma \vdash \neg\neg\phi$ , majd a 9. Tétel felhasználásával  $\Sigma \vdash \phi$ .

**14. Tétel.** *Tegyük fel, hogy  $\Sigma \vdash \phi$  és  $\Sigma \vdash \psi$ . Ekkor  $\Sigma \vdash \phi \wedge \psi$ .*

**Bizonyítás.** A 13. tételt fogjuk alkalmazni, vagyis belátjuk, hogy a  $\Sigma \cup \{\neg(\phi \wedge \psi)\}$  inkonzisztens. Emlékezzünk,  $\phi \wedge \psi$  annak a rövidítése, hogy  $\neg(\phi \rightarrow \neg\psi)$ . Tehát azt kell belátnunk, hogy  $\Sigma \cup \{\phi \rightarrow \neg\psi\}$  inkonzisztens, ami triviálisan igaz, hiszen  $\phi \rightarrow \neg\psi$  MP-vel azonnal maga után vonja, hogy  $\Sigma \cup \{\neg(\phi \wedge \psi)\} \vdash \neg\psi$ , ugyanakkor a feltevésünkből következően  $\Sigma \cup \{\neg(\phi \wedge \psi)\} \vdash \psi$ .

Hasonlóan triviális, hogy

**15. Tétel.** *Ha  $\Sigma \vdash \phi$  vagy  $\Sigma \vdash \psi$ , akkor  $\Sigma \vdash \phi \vee \psi$ .*

**16. Tétel.** *Legyen  $x$  szabad változó a  $\phi(x)$  formulában. Legyen továbbá  $y$  egy olyan változó, amelyik nem fordul elő  $\phi(x)$ -ben, sem kötött, sem szabad formában. Ekkor*

$$\vdash \forall x\phi(x) \leftrightarrow \forall y\phi(y)$$

**Bizonyítás.**

1.  $\forall x\phi(x)$
2.  $\forall x\phi(x) \rightarrow \phi(y)$  [(PC6)]
3.  $\phi(y)$  [(MP)]

#### 4. $\forall y\phi(y)$ [(G)]

Vagyis beláttuk, hogy  $\forall x\phi(x) \vdash \forall y\phi(y)$ . A dedukció-tétel alkalmazásával tehát

$$\vdash \forall x\phi(x) \rightarrow \forall y\phi(y)$$

Teljesen hasonló módon bizonyítjuk a fordított irányt is.

**17. Tétel.** *Tetszőleges formulához létezik vele bizonyíthatóan ekvivalens prenex alakú formula.*

## Interpretáció

### Egy nem teljesen helyénvaló előzetes példa

Tekintsük a következő mondatokat a PC-ben:

$$(\sigma_1) \quad \forall x\forall y (P(x, y) \rightarrow P(x, y))$$

$$(\sigma_2) \quad (P(x, y) \wedge P(y, z)) \rightarrow P(x, z)$$

$$(\sigma_3) \quad \forall y\exists x P(x, y)$$

- Ha úgy interpretáljuk a  $P(x, y)$  két változós predikátumot, mint a valaha élt emberek halmazában (*Sic!*) értelmezett „ $x$  őse  $y$ -nak” relációt, akkor nyilvánvalóan mindhárom mondat igaz.
- Ha úgy interpretáljuk  $P(x, y)$ -et, hogy az a  $>$  reláció a természetes számok  $\mathbb{N}$  halmazán, akkor ezek a mondatok mind igazak.
- Ha úgy interpretáljuk  $P(x, y)$ -et, hogy az a  $<$  reláció az egész számok  $\mathbb{Z}$  halmazán, akkor ezek a mondatok mind igazak.

- Ha úgy interpretáljuk  $P(x, y)$ -et, hogy az a  $<$  reláció a természetes számok  $\mathbb{N}$  halmazán, akkor a  $(\sigma_1)$  és  $(\sigma_2)$  a mondatok igazak, de a  $(\sigma_3)$  hamis.

Sokan „interpretáció” alatt a fenti példához hasonlóan azt értik, hogy a formális rendszer elemeinek a fizikai világ (a platonizmus és intuicionizmus szerint a platoni illetve mentális világ) olyan elemeit feleltetjük meg, melyek valamilyen intuitív értelemben kielégítik a szóban forgó formális rendszer axiómáit. A matematikai logikában interpretáció alatt mást értünk.

## Interpretáció és modell

**Interpretáció.** Egy PC-ben értelmezett formális rendszer interpretációja egy

$$\mathcal{A} = \langle U, R_1^{n_1}, R_2^{n_2}, \dots, f_1^{m_1}, f_2^{m_2}, \dots \rangle$$

struktúra, ahol

- $U$  egy nem üres halmaz, melyet az interpretáció univerzumának fogunk nevezni.
- $R_1^{n_1}, R_2^{n_2}, \dots$  az  $U$ -n értelmezett  $n_1, n_2, \dots$  argumentumos relációk, melyeket a formális rendszer  $n_1, n_2, \dots$  argumentumos  $P_1^{n_1}, P_2^{n_2}, \dots$  predikátumainak feleltetünk meg.
- $f_1^{m_1}, f_2^{m_2}, \dots$  olyan  $\underbrace{U \times U \times \dots \times U}_{m_1} \rightarrow U$ ,  
 $\underbrace{U \times U \times \dots \times U}_{m_2} \rightarrow U$ , stb. típusú függvények, melyek a formális rendszerben előforduló  $m_1, m_2, \dots$  argumentumos függvényjeleket reprezentálják.



**Szereposztás (értékelés).** A formális rendszerben előforduló  $t_1, t_2, \dots$  individum változókhoz és individum konstansokhoz rendre hozzárendeljük  $U$ -nak valamelyik elemét. (Több változóhoz is rendelhetjük ugyanazt az elemét  $U$ -nak.) Egy ilyen szereposztást röviden a következőképpen fogunk jelölni:  $[u_1, u_2, \dots]$

**Teljesítés.** Most definiáljuk egy  $\phi$  formula *teljesülését* az  $\mathcal{A}$  interpretációban egy adott  $[u_1, u_2, u_3, \dots]$  szereposztás mellett. Ezt úgy fogjuk jelölni, hogy

$$\mathcal{A} \models \phi [u_1, u_2, u_3, \dots]$$

Felhasználva, hogy a nyelv helyesen képzett formuláit hogyan építjük fel (lásd a ?? bekezdést), a definíciót a következő módon adjuk meg:

1.  $\mathcal{A} \models P_i^n (t_1, t_2, \dots, t_n) [u_1, u_2, u_3, \dots]$  akkor és csak akkor, ha az  $[u_1, u_2, u_3, \dots]$  szereposztásnak megfelelően a  $t_1, t_2, \dots, t_n$  terminusoknak megfelelően  $u_{t_1}, u_{t_2}, \dots, u_{t_n}$  elemekre fennáll a  $P_i^n$  predikátumnak megfelelő  $R_i^n$  reláció, tehát

$$R_i^n (u_{t_1}, u_{t_2}, \dots, u_{t_n}) \tag{2}$$

Értelemszerűen azt is megengedjük (összhangban a terminus definíciójával), hogy egy  $t_k$  terminus függvénykifejezés legyen, tehát pl. legyen  $t_k$  a  $f^2(x_1, x_2)$  kifejezés. Ekkor az adott szereposztásban az  $x_1$  és  $x_2$  változókat az univerzum valamely  $u_{x_1}$  és  $u_{x_2}$  eleme reprezentálja. Az  $f^2$  2-argumentumos függvényjelet pedig valamilyen  $\tilde{f} : U \times U \rightarrow U$  függvény. Ekkor a (2) relációban az  $u_{t_k}$  helyére az  $\tilde{f}(u_{x_1}, u_{x_2})$  kifejezést, azaz az  $\tilde{f}$  függvénynek az  $u_{x_1}, u_{x_2}$  helyen felvett értékét írjuk.

2.  $\mathcal{A} \models \neg\phi [u_1, u_2, u_3, \dots]$  akkor és csak akkor, ha nem igaz, hogy  $\mathcal{A} \models \phi [u_1, u_2, u_3, \dots]$ .

3.  $\mathcal{A} \models \phi \rightarrow \psi [u_1, u_2, u_3, \dots]$  akkor és csak akkor, ha vagy  $\mathcal{A} \models \neg\phi [u_1, u_2, u_3, \dots]$  vagy  $\mathcal{A} \models \psi [u_1, u_2, u_3, \dots]$ .
4.  $\mathcal{A} \models \forall y\phi(x_1, x_2, \dots, x_n, y) [u_1, u_2, \dots, u_n]$  akkor és csak akkor, ha minden  $[u_1, u_2, \dots, u_n, w]$  értékelésre (ahol  $u_1, u_2, \dots, u_n$  fix)  $\mathcal{A} \models \phi [u_1, u_2, \dots, u_n, w]$ .

Ezzel egy formula teljesülésének fogalmát konstruktíve megadtuk.

**Igaz  $\mathcal{A}$ -ban.** Ha egy  $\mathcal{A} \models \phi [u_1, u_2, u_3, \dots]$  minden  $[u_1, u_2, u_3, \dots]$  értékelés (szereposztás) esetén, akkor azt mondjuk, hogy  $\phi$  formula *igaz  $\mathcal{A}$ -ban*, és azt írjuk, hogy  $\mathcal{A} \models \phi$ . Ha  $\phi$  mondat, azaz nem tartalmaz szabad változót, akkor  $\mathcal{A} \models \phi$  minden olyan esetben ha  $\mathcal{A} \models \phi [u_1, u_2, u_3, \dots]$  tetszőleges  $[u_1, u_2, u_3, \dots]$  értékelés esetén ( $[u_1, u_2, u_3, \dots]$ -nek nincs jelentősége).

**Univerzálisan igaz.** Ha tetszőleges  $\mathcal{A}$  interpretációra  $\mathcal{A} \models \phi$ , akkor azt mondjuk, hogy  $\phi$  *univerzálisan igaz*, és ezt úgy jelöljük, hogy  $\models \phi$ .

**Példa.** Legyen  $\mathcal{A} = \langle W, A \rangle$ , ahol  $W$  a valaha élt emberek halmaza, és  $A$  az „őse” reláció. Vegyük pl. a  $\exists xP(x, y)$  formulát.  $\mathcal{A} \models \exists xP(x, y) [v]$  akkor és csak akkor, ha létezik olyan  $w$  ember, hogy  $\mathcal{A} \models P(x, y) [w, v]$ . Ez akkor és csak akkor áll fenn, ha  $A(w, v)$ . De ez tetszőleges  $v$  esetén igaz, hogy tudniillik van olyan  $w$ , akire  $A(w, v)$ . Tehát  $\mathcal{A} \models \exists xP(x, y) [v]$  minden lehetséges  $v$ -re, ezért  $\mathcal{A} \models \exists xP(x, y)$ , azaz  $\exists xP(x, y)$  igaz  $\mathcal{A}$ -ban.

Ezzel szemben, nyilván  $\mathcal{N} \not\models \exists xP(x, y)$ , ahol  $\mathcal{N} = \langle \mathbb{N}, < \rangle$ .

**18. Tétel.** *A PC axiómái univerzálisan igazak.*

**Bizonyítás.** pl. (PC6)-ra: Tegyük fel hogy hogy valamilyen  $\mathcal{A}$  interpretációban a változók valamilyen  $[u_1, u_2, u_3, \dots]$  értékelése esetén, (PC6) nem igaz. Ez akkor és csak akkor lehetséges, ha  $\mathcal{A} \models \forall x\phi(x) [u_1, u_2, u_3, \dots]$  ugyanakkor  $\mathcal{A} \not\models \phi(y) [u_1, u_2, u_3, \dots]$ . De ez ellentmondás, hiszen ha az  $y$  változó az értékelésben valamely  $u_i$ -nek felel meg, az előző formula éppen azt állítja, hogy a  $\phi$  reláció fennáll minden lehetséges  $u_i$  mellett.

**HF.** Bizonyítsuk be a tételt a többi axiómára is.

**Egy formulahalmaz modellje.** Legyen  $\Sigma$  formulák egy halmaza PC-ben, és legyen az  $\mathcal{A}$  interpretáció olyan, hogy  $\mathcal{A} \models \phi$  minden  $\phi \in \Sigma$  esetén. Ekkor azt mondjuk, hogy  $\mathcal{A}$  a  $\Sigma$  egy *modellje*.

**19. Tétel.** *Legyen  $\mathcal{A}$  egy tetszőleges interpretáció. Ha  $\mathcal{A} \models \phi$  és  $\mathcal{A} \models \phi \rightarrow \psi$ , akkor  $\mathcal{A} \models \psi$*

**Bizonyítás.** Legyen  $[u_1, u_2, u_3, \dots]$  tetszőleges értékelés.  $\mathcal{A} \models \phi [u_1, u_2, u_3, \dots]$  és  $\mathcal{A} \models (\phi \rightarrow \psi) [u_1, u_2, u_3, \dots]$ . A teljesülés (implikációra vonatkozó) definíciójánál fogva: vagy  $\mathcal{A} \models \neg\phi [u_1, u_2, u_3, \dots]$ , ami feltevésünk szerint lehetetlen, vagy  $\mathcal{A} \models \psi [u_1, u_2, u_3, \dots]$ . Mivel ez tetszőleges értékelésre igaz, a tételt bebizonyítottuk.

**20. Tétel.** *Legyen  $\mathcal{A}$  egy tetszőleges interpretáció.  $\mathcal{A} \models \phi$  akkor és csak akkor, ha  $\mathcal{A} \models \forall x\phi$ .*

**Bizonyítás.** Tegyük fel, hogy  $\mathcal{A} \models \phi$ . Ekkor  $\mathcal{A} \models \phi [u_1, u_2, u_3, \dots]$  tetszőleges  $[u_1, u_2, u_3, \dots]$  értékelésre, tehát  $\mathcal{A} \models \phi [u_1, \dots, u_i, \dots]$  minden olyan értékelésre is, ahol az  $x$  változónak megfelelő  $u_i$  elemet változtatjuk csak, a többit fixen tartjuk. Tehát  $\mathcal{A} \models \forall x\phi [u_1, u_2, u_3, \dots]$  minden értékelésre, azaz  $\mathcal{A} \models \forall x\phi$ .

Fordítva, ha  $\mathcal{A} \models \forall x\phi$ , akkor  $\mathcal{A} \models \forall x\phi [u_1, u_2, u_3, \dots]$  tetszőleges  $[u_1, u_2, u_3, \dots]$  értékelésre. Mivel az összes értékelést úgy is megkapjuk, ha előbb vesszünk egy értékelést és az  $x$ -nek megfelelő  $u_i$  elemet variáljuk, majd vesszük az összes ilyen,  $\mathcal{A} \models \phi [u_1, \dots, u_i, \dots]$  minden lehetséges  $[u_1, u_2, u_3, \dots]$  esetén, tehát  $\mathcal{A} \models \phi$ .

**21. Tétel.** *Legyen  $PC(\Sigma)$  a  $PC$  egy tetszőleges  $\Sigma$ -kiterjesztése, és legyen  $\mathcal{A}$  egy tetszőleges interpretáció. Ha a  $\Sigma$  axiómalista minden formulája igaz  $\mathcal{A}$ -ban, akkor  $\mathcal{A}$  egy modellje  $PC(\Sigma)$ -nak, abban az értelemben, hogy minden olyan  $\phi$  formulára, melyre  $\Sigma \vdash \phi$ , fennáll, hogy  $\mathcal{A} \models \phi$ .*

**Bizonyítás.** Tekintsünk egy tetszőleges  $\phi$  formulát, melyre  $\Sigma \vdash \phi$ . Ez azt jelenti, hogy létezik  $\phi$ -nek bizonyítása. Legyen a bizonyítás egy  $n$  formulából álló formulasorozat. Most teljes indukcióval megmutatjuk, hogy  $\phi$  igaz  $\mathcal{A}$ -ban.

1.  $n = 1$ .  $\phi$  axióma, tehát igaz  $\mathcal{A}$ -ban.
2.  $n > 1$ . Indukciós hipotézis: A bizonyítandó állítás igaz minden olyan  $\phi$  tételre (azaz  $\Sigma \vdash \phi$  formula esetében), amelynek bizonyítása maximum  $n - 1$  lépésből áll.
3. Ekkor igaz az  $n$  lépésből álló bizonyítással rendelkező  $\phi$ -re is. Ugyanis a következő esetek lehetségesek:
  - (a)  $\phi$  maga is axióma, tehát  $\mathcal{A} \models \phi$ .
  - (b)  $\phi$  a (MP)-ből (modus ponens) következik, mondjuk valamilyen korábbi  $\chi_i$  és  $\chi_i \rightarrow \phi$  felhasználásával. Mármint  $\chi_i$  és  $\chi_i \rightarrow \phi$  mindketten olyan  $\Sigma$ -ból levezethető tételek, amelyek bizonyítása maximum  $n - 1$  lépésből áll, tehát a 19. tétel következtében  $\mathcal{A} \models \phi$ .

- (c) Hasonlóan, ha  $\phi$  a (G) (generalizáció) alkalmazásával következik valamely korábbi  $\chi_i$  formulából, akkor a 20. tétel következtében  $\mathcal{A} \models \phi$ .

## Teljességi tétel

**22. Tétel (Teljességi tétel).** *Egy  $\phi$  formula akkor és csak akkor bizonyítható PC-ben (vagyis csak a PC axiómáiból), ha univerzálisan igaz. Szokásos jelöléseinket használva,  $\vdash \phi$  akkor és csak akkor, ha  $\models \phi$ .*

### *Bizonyítás.*

**1.**  $\vdash \phi \Rightarrow \models \phi$ . Mint már bebizonyítottuk, a PC axiómái univerzálisan igazak. A 21. tételből következően tehát PC minden tétele univerzálisan igaz.

**Fontos következmény.** A predikátum kalkulus konzisztens. Ugyanis ha nem volna az, tehát  $\vdash \phi$  és  $\vdash \neg\phi$  egyszerre állna fenn, akkor ebből következne, hogy  $\models \phi$  és  $\models \neg\phi$ , azaz lenne olyan  $\mathcal{A}$  interpretáció és olyan értékelés, hogy egyszerre  $\mathcal{A} \models \phi [u_1, \dots, u_i, \dots]$  és nem  $\mathcal{A} \models \phi [u_1, \dots, u_i, \dots]$ .

**2.**  $\models \phi \Rightarrow \vdash \phi$ . Ez akkor teljesül, ha abból, hogy  $\phi$  nem tétel, következik, hogy nem univerzálisan igaz. Vagyis azt kell megmutatnunk, hogy ha  $\not\models \phi$ , akkor  $\neg\phi$ -nek létezik modellje.  $\neg\phi$ -nek ugyanis csak akkor létezik modellje, ha  $\phi$  nem univerzálisan igaz. Az 13. tétel következtében, ha  $\not\models \phi$ , akkor a  $\{\neg\phi\}$  egy elemű formulahalmaz konzisztens. Ezért, a Gödel–Henkin-tétel következtében – melyet az alábbiakban fogunk bizonyítani – létezik modellje. Márpedig ha ez igaz, akkor ebben a modellben  $\phi$  hamis, tehát  $\phi$  nem univerzálisan igaz. Tehát  $\models \phi$ -ből következik  $\vdash \phi$ , és ezzel a tételt bebizonyítottuk.

Természetesen, most következnek a Gödel–Henkin-tétel.

**23. Tétel (Gödel–Henkin teljességi tétel).** *Ha egy  $\Sigma$  formula-halmaz konzisztens, akkor létezik modellje, azaz létezik olyan  $\mathcal{A}$  interpretáció, hogy  $\mathcal{A} \models \phi$  minden  $\phi \in \Sigma$  formulára.*

## PC(=) (predikátum kalkulus identitással)

Az előzőekben megismert predikátum kalkulust egy további predikátum-jellel egészítjük ki. Legyen  $E$  („ugyanaz mint”, „egyenlő”) egy kétváltozós predikátum.  $E$  tulajdonságait a következő axiómák hozzáadásával rögzítjük:

### Az egyenlőség axiómái

$$(E1) \quad E(x, x)$$

$$(E2) \quad E(t, s) \rightarrow E(f^n(u_1, u_2, \dots, t, \dots, u_n), f^n(u_1, u_2, \dots, s, \dots, u_n))$$

$$(E3) \quad E(t, s) \rightarrow (\phi(u_1, u_2, \dots, t, \dots, u_n) \rightarrow \phi(u_1, u_2, \dots, s, \dots, u_n))$$

Kényelmesebb jelölés:  $x = y \equiv E(x, y)$

**HF..** Mutassuk meg, hogy  $E$  tranzitív és szimmetrikus.

# Aritmetika

$$(A1) \quad \neg(0 = sx)$$

$$(A2) \quad (sx = sy) \rightarrow (x = y)$$

$$(A3) \quad x + 0 = x$$

$$(A4) \quad x + sy = s(x + y)$$

$$(A5) \quad x \cdot 0 = 0$$

$$(A6) \quad x \cdot sy = (x \cdot y) + x$$

$$(A7) \quad (P(0) \wedge \forall x (P(x) \rightarrow P(sx))) \rightarrow \forall x P(x)$$

ahol természetesen  $x = y$ ,  $x + y$  illetve  $x \cdot y$  az  $E(x, y)$ ,  $+(x, y)$  illetve  $\cdot(x, y)$  helyett áll, ahol  $E$  az egyenlőség predikátum,  $+$  és  $\cdot$  pedig függvények.  $s$  szemléletes jelentése az „hozzáadunk egyet” művelet,  $P$  pedig egy tetszőleges egyváltozós predikátum. (A7)-et az *indukció axiómasémájának* is szokás nevezni.

Vezessük be a  $1, 2, 3, \dots$  jeleket a következő individuális konstansok jelölésére:

$$1: \quad s0$$

$$2: \quad ss0$$

$\vdots$

$$k: \quad \underbrace{s \dots s}_{k \text{ darab}} 0$$

$\vdots$

**NB.**

A jelölésekben használjuk a számokat és írunk olyat, hogy „ $k$ -darab”, stb. Vegyük észre, hogy ezek csak kényelmi, tipográfiai eszközök, és nem történik lényegi hivatkozás valamilyen „előzetesen ismert aritmetikára”.



# Igaz-e, hogy $2 + 2 = 4$ ?

**24. Tétel.**  $\{aritmetika\} \vdash 2 + 2 = 4$  a  $PC(=)$ -ben.

**Bizonyítás.** A jelölések definícióját alapul véve tehát azt kell bizonyítanunk, hogy  $ss0 + ss0 = ssss0$ :

1.  $ss0 + ss0 = s(ss0 + s0)$  [(A4)-ből]
2.  $s(ss0 + s0) = ss(ss0 + 0)$  [(A4)-ből]
3.  $ss0 + ss0 = ss(ss0 + 0)$  [1. és 2. alapján (E3) és (MP) felhasználásával]
4.  $ss0 + 0 = ss0$  [(A3)-ből]
5.  $ss0 + ss0 = ssss0$  [3. és 4. alapján (E3) és (MP) felhasználásával]

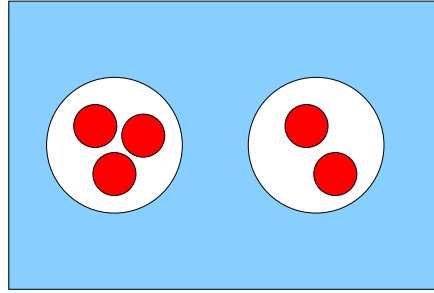
**NB..**

- Gyakran olvashatunk az irodalomban olyan gondolatmeneteket, amelyek a „szándékolt interpretációról” szólnak. Természetesen, lehet valamilyen intuíciónk előzetesen arról, hogy az axiomatikusan felépítendő matematikai struktúrától mit várunk. De ennek szigorú, elméleti, matematikai értelemben nyilván nem lehet semmiféle jelentősége. (A matematikában egyébként is számtalanszor elfogadjuk formális elméleti gondolatmenetek útján nyert konklúziókat, melyek esetleg ellentmondanak a „józan észnek”, vagy az előzetes intuitív várakozásainknak. Gondoljunk például arra, hogy intuitíve több racionális számnak kellene lennie, mint egész számnak, mégis elfogadjuk a formális bizonyítást, hogy a két halmaz számossága azonos.) Összegezve tehát, az aritmetika az, amit most axiomatikusan felépítünk!

- Természetesen lehet arról beszélni, hogy egy axiomatikusan felépített aritmetika hasznos matematikai struktúra-e számunkra, abban az értelemben, hogy használható-e a világ leírásában, vagyis a fizikai elméletekben. Tehát az aritmetika axiomatikus felépítése során lehet az a szándékunk, hogy egy olyan struktúrát hozzunk létre, amely majd alkalmas lesz — egy megfelelő fizikai elmélet részeként — annak leírására, hogy hogyan működik a pénztárgép, vagy alkalmazható lesz abban a fizikai elméletben, amelyet egy juhász használ a nyájba tartozó juhok nyilvántartására, stb.
- Félreértések elkerülése érdekében felhívjuk a figyelmet arra, hogy az itt alkalmazott jelölések eltérnek a tankönyvekben szokásos jelölésektől. Az itt számokkal jelölt  $1, 2, \dots$ , és számoknak, tehát „egynek”, „kettőnek”, stb. nevezett individuum konstansok rendszerint valamilyen megkülönböztető jelölést kapnak,  $\bar{1}, \bar{2}, \bar{3}, \dots$  (lásd Crossley), vagy  $0^{(1)}, 0^{(2)}, 0^{(3)}$ , (lásd Hamilton), stb. És rendszerint nem is nevezik őket számoknak, hanem „számjegyeknek”, „számneveknek” (numerals, numeral terms), megkülönböztetésül az „igazi” számoktól, azaz valamilyen értelemben már előzetesen létező számelmélet számfogalmától, melyeknek a fenti értelemben vett axiomatikus aritmetika valamiféle „axiomatizált” elmélete. Az itt szorgalmazott felfogás szerint azonban az aritmetika az, amit itt axiomatikusan megadunk. Nincsenek „aritmetikai igazságok” mások, mint amiket az axiomatikus aritmetikában az axiómákból levezethetünk. Semmi okunk tehát arra, hogy éppen azt jelöljük valami mással, ami van, és azt jelöljük  $1, 2, 3, \dots$ -mal, ami nincs!

It is completely meaningless to talk about “intuitive arithmetic”, “naive set theory”, “intended interpretation”, and the like, or to differentiate “numbers” from “numerals”, or use the phrase “axiomatization of ...”, etc. For instance, from the point of view of physicalism, there can be no arithmetic in mathematics before axiomatic arithmetic has

been established. We do not “know” in advance that “ $3 + 2 = 5$ ” until the corresponding axiomatic theory is constructed in which this formula exists, until this formula is derived from the arithmetical axioms.



Counting using your fingers, or observing that

- (A) When the apple-counter equipment interacting with the apples on plate I shows figure  $\boxed{3}$  and the apple-counter interacting with the content of plate II shows figure  $\boxed{2}$  then the apple-counter interacting with apples on the table shows figure  $\boxed{5}$ .

are observations about the physical world outside of the realm of mathematics. These physical observations have nothing to do with the mathematical truth of formula  $3 + 2 = 5$ . It does not follow from these observations that  $3 + 2 = 5$  because this formula is a formula of arithmetic. And the only truth mathematics (arithmetic) is concerned with is  $\text{Truth}_1$ . We cannot infer the truth of formula  $3 + 2 = 5$  from the physical observations of apples, plates and tables. This is not because “the general arithmetical truth  $3 + 2 = 5$  is obtained by abstraction from many other similar observations”. It is because we can infer its truth from a *completely different physical* observation: the observation of a sequence of formulas constituting a proof of formula  $3 + 2 = 5$  within arithmetic.

- (PC1)  $(\phi \rightarrow (\psi \rightarrow \phi))$
- (PC2)  $((\phi \rightarrow (\psi \rightarrow \chi)) \rightarrow (\phi \rightarrow \psi) \rightarrow (\phi \rightarrow \chi))$
- (PC3)  $((\neg\phi \rightarrow \neg\psi) \rightarrow (\psi \rightarrow \phi))$
- (PC4)  $(\forall x (\phi \rightarrow \psi) \rightarrow (\phi \rightarrow \forall x\psi))$  feltéve, hogy  $\phi$  nem tartalmazza  $x$  szabad előfordulását.
- (PC5)  $(\forall x\phi \rightarrow \phi)$  ha  $x$  nem fordul elő szabadon  $\phi$ -ben
- (PC6)  $(\forall x\phi(x) \rightarrow \phi(y))$  feltéve, hogy amikor az  $x$  szabad előfordulásait  $y$ -nal helyettesítjük, akkor  $y$  szabad változó  $\phi(y)$ -ban.
- (MP)  $\phi, (\phi \rightarrow \psi)$ -ből következik  $\psi$  (modus ponens)
- (G)  $\phi$ -ből következik  $\forall x\phi$  (generalizáció)
- (E1)  $E(x, x)$
- (E2)  $E(t, s) \rightarrow E(f^n(u_1, u_2, \dots, t, \dots, u_n), f^n(u_1, u_2, \dots, s, \dots, u_n))$
- (E3)  $E(t, s) \rightarrow (\phi(u_1, u_2, \dots, t, \dots, u_n) \rightarrow \phi(u_1, u_2, \dots, s, \dots, u_n))$
- (A1)  $\neg(0 = sx)$
- (A2)  $(sx = sy) \rightarrow (x = y)$
- (A3)  $x + 0 = x$
- (A4)  $x + sy = s(x + y)$
- (A5)  $x \cdot 0 = 0$
- (A6)  $x \cdot sy = (x \cdot y) + x$
- (A7)  $(P(0) \wedge \forall x (P(x) \rightarrow P(sx))) \rightarrow \forall x P(x)$
- 1  $sss0 + ss0 = s(sss0 + s0)$
- 2  $s(sss0 + s0) = ss(sss0 + 0)$
- 3  $sss0 + ss0 = ss(sss0 + 0)$
- 4  $sss0 + 0 = sss0$
- 5  $sss0 + ss0 = sssss0$

Of course, one can construct a *physical theory* describing a physical system consisting of apples, plates and a table. It would consist of a formal system  $L$  and semantics  $S$  which relates some elements of

the formal system  $L$  to observed phenomena. Of course, we observe the situation described in (A) many times and regard it as a law-like regularity:

- (B) Whenever the apple-counter interacting with the apples on plate I shows figure  $\boxed{3}$  and the apple-counter interacting with the apples on plate II shows figure  $\boxed{2}$  then the apple-counter interacting with the apples on the table shows figure  $\boxed{5}$ .

This regularity can be expressed in the physical theory  $(L, S)$  in the following way: The formal system  $L$  contains the axioms of the first-order predicate calculus with identity,  $PC(=)$ , and the axioms of arithmetic. Semantics  $S$  relates the figures  $\boxed{3}$ ,  $\boxed{2}$  and  $\boxed{5}$  shown by the counters to the symbols 3, 2 and 5 (where 3 is an abbreviation for  $sss0$ , etc.) and the regularity (B) will be related with formula  $3 + 2 = 5$ .

**NB.** When I say that the semantics “relates” some elements of the observed phenomena with some symbols in  $L$ , I do not mean a map, or function or relation of mathematical character. **That would be a categorical mistake!** Maps, functions and relations are mathematical objects, existing only in a formal system (incorporating set theory). Such a semantical relationship is always embodied in causal concatenations, relating the formal system with other physical systems. You see the figure  $\boxed{3}$  on the instrument and put your finger on the symbol 3 on the paper. Simultaneously the neural configuration of your brain or—let me put it in a more Searlean way—the whole physical state of your brain and body is such that you have the impression in your mind that the two things are germane to each other.

From physicalist point of view, the involvement of human activity does not make an essential difference. In order to avoid any

confusion about the role of human mind in this procedure, one can always imagine a robot doing the job. Thus, imagine a robot (a computer with all the required peripherals) designed (programmed) in the following way: he observes the figure shown by the apple-counter interacting with the apples on plate I. He identifies the observed figure with a definite symbol in  $L$ , namely, with one of the numbers (“numerals”, if you like, I do not make a distinction), say 3. He observes the figure on the other counter and puts his finger on number 2. He also observes the figure shown by the apple-counter interacting with the apples on the whole table, and identifies this figure with, say, 5. Then the robot checks on whether  $3 + 2 = 5$  is a theorem in  $L$ , that is, in arithmetic, or not. Since  $\vdash_L 3 + 2 = 5$ , he reports that the phenomenon he observed is compatible with the physical theory  $(L, S)$ . And so on.

This example, I believe, gives an intuition about how arithmetic works within a physical theory that describes real phenomena. *But, the success of such a physical theory—within its range of applicability, of course—does not entitle us to claim that “there are numbers” as entities—different from “numerals”. Nor can we say that “ $3 + 2 = 5$ ” is a truth about them, that we can know before proving the corresponding formula in (“axiomatic”) arithmetic.  $3 + 2 = 5$  is only a formula in arithmetic, and the only truth it holds is its  $\text{Truth}_1$ : that is it is a theorem of arithmetic.*

Finally, it is worth while emphasizing that not only does the empirical fact (B) not imply that  $3 + 2 = 5$  is a theorem of arithmetic, but, vice versa, the arithmetical theorem  $3 + 2 = 5$  in itself does not imply the physical fact (B). Moreover, the arithmetical theorem  $3 + 2 = 5$  in itself does not even imply the physical hypothesis, that (B) is true. Such a physical prediction only follows from the corresponding physical theory  $(L, S)$  incorporating arithmetic.

# Halmazelmélet

## „Naiv” halmazelmélet — formális (axiomatikus) halmazelmélet

Szokás azt mondani, hogy azért kell a halmazelméletet „axiomatizálni”, mert a „naiv” halmazelméletben bizonyos paradoxonok fogalmazhatók meg, és az axiomatikus módszerrel ezek kiküszöbölhetők. Természetesen nem ezért kell megadnunk a halmazelmélet axiomatikus elméletét, hanem azért, hogy egyáltalán legyen halmazelmélet. Más szóval, nincs „naiv” halmazelmélet! (Legfeljebb abban a didaktikai értelemben, ha egy tankönyvben bevezetünk néhány halmazelméleti fogalmat és kimondunk néhány halmazelméleti tételt, anélkül, hogy megadnánk ezek bizonyítását.)

Ernst Zermelo (1905) és Abraham Fraenkel (1920) után, a halmazelmélet itt tárgyalt axiomatikus elméletét ZF-nek szokás nevezni.

A halmazelméletet a PC(=)-ben adjuk meg. Az egyenlőségén kívül a nyelv tartalmazni fog egy kétváltozós predikátumot,  $\in$  („eleme”).

## A halmazelmélet (ZF) axiómái

(ZF1)  $\exists x \forall u \neg (u \in x)$   
(*üres halmaz axióma*) Mivel ez az axióma garantálja az üres halmaz létezését, bevezetjük az üres halmaz jelölésére a  $\emptyset$  jelet.

(ZF2)  $\forall x \forall y (\forall u (u \in x \leftrightarrow u \in y) \leftrightarrow x = y)$   
(*meghatározottsági axióma*) Az axióma azt fejezi ki, hogy egy halmazt egyértelműen meghatározza, hogy mely halmazok az elemei.

**Jelölés.**  $u \subseteq v$  a következő formula rövidítése:

$$\forall x (x \in u \rightarrow x \in v)$$

(ZF3)  $\forall x \forall y \exists z \forall u (u \in z \leftrightarrow u = x \vee u = y)$   
(*páraxióma*) Vagyis két halmazból lehet képezni egy olyan halmazt, amelynek ők az elemei.

**Jelölés.** Az axióma által garantált  $z$  halmazt szokás a következőképpen jelölni:  $\{x, y\}$

(ZF4)  $\forall x \exists y \forall z (z \in y \leftrightarrow \exists u (u \in x \wedge z \in u))$   
(*az unió axiómája*)

**Jelölés.** Azt az objektumot, amelynek létezését (ZF4) garantálja  $\cup x$ -el fogjuk jelölni. Pl. két halmaz uniójára, vagyis az  $\cup \{x, y\}$  halmazra bevezetjük az  $x \cup y$  jelölést.

(ZF5)  $\forall x \exists y \forall z (z \in y \leftrightarrow z \subseteq x)$   
(*a hatványhalmaz axiómája*)

**Jelölés.** Az axióma által garantált  $y$  halmazt szokás  $2^x$ -el jelölni.

(ZF6)  $\forall x \exists! y \phi(x, y) \rightarrow \forall z \exists! u \forall v (v \in u \leftrightarrow \exists o (o \in z \wedge \phi(o, v)))$ ,  
ahol  $\phi(x, y)$  tetszőleges két szabad változót tartalmazó formula, melyben feltesszük, hogy a  $\forall v$  és  $\forall o$  kvantifikációk nem fordulnak elő.  
(*helyettesítési axiómaséma*)

(ZF7)  $\exists x (\emptyset \in x \wedge \forall y (y \in x \rightarrow y \cup \{y\} \in x))$  ( $\{y\}$  a rövidítése az  $\{y, y\}$ -nak)  
(*a végtelen halmaz axiómája*)

(ZF8)  $\forall x (\neg x = \emptyset \rightarrow \exists y (y \in x \wedge \neg \exists z (z \in y \wedge z \in x)))$   
(*regularitási axióma*) Vagyis, hogy minden nem üres  $x$



halmaznak van olyan eleme, amely diszjunkt  $x$ -től. Ezzel elérjük azt, hogy egyetlen halmaz sem lehet eleme önmagának.

(ZF1)–(ZF8) elégséges ahhoz, hogy a matematika egy jelentős részét felépítsük. Pl. a természetes számok egy modelljét a következő halmazokból álló univerzumon adhatjuk meg:

0	$\emptyset$
1	$\{\emptyset\}$
2	$\{\emptyset, \{\emptyset\}\}$
3	$\{\emptyset, \{\emptyset, \{\emptyset\}\}\}$
:	

## ZFC

(AC) Tetszőleges nem üres  $x$  halmazhoz létezik olyan  $y$  halmaz, amelyre igaz, hogy  $x$  minden elemével pontosan egy közös eleme van.

## Kontinuum Hipotézis

(CH) Valós számokból álló tetszőleges végtelen halmaz vagy megszámlálható számosságú, vagy kontinuum számosságú.

E két utolsó axiómát illetően kérdések merültek fel. Le lehet-e vezetni ezeket a (ZF1)–(ZF8)-ból? És ha nem, konzisztens módon hozzávehető-e az alap ZF rendszerhez, külön-külön, és együtt? Ezekre a kérdésekre részben 1938-ban Gödel egyik munkájában, majd később (1963) Cohen munkáiban kaptunk választ. Gödel megmutatta, hogy ha a ZF konzisztens, akkor (AC) és (CH) konzisztens

módon hozzávehető az axiómarendszerhez. Cohen azt mutatta meg, hogy sem (AC) illetve (CH), sem a negáltjaik nem vezethetők le a ZF-ből, tehát független axiómákról van szó. (Egymástól is függetlenek.)

## NB

1. Felmerül a kérdés, hogy a matematikai logika kifejtésekor honnan vannak halmazok és azokon értelmezett relációk. Pontosabban, hogy honnan vesszük, hogy azok az állítások, amelyeket az interpretációt jelentő halmazelméleti struktúrákra vonatkozóan teszünk, igazak. Ezek a fejezetek most váltak teljessé, azzal, hogy megadtuk a halmazelmélet axiómáit. Például, a teljesítés fogalmának definíciójában,  $\mathcal{A} \models P(x_1, x_2)[u_1, u_2]$  akkor és csak akkor, ha  $\{ZF\} \vdash \{u_1, u_2\} \in R$ . Természetesen a modell-elméleti szemantika még ezzel sem teljesen problémamentes. Hiszen az interpretálandó elsőrendű nyelv elemei és a halmazelméleten belül, mint másik elsőrendű nyelven belül definiált, az interpretációt nyújtó struktúra elemei közötti megfeleltetés nincs valamely formális, elsőrendű nyelv keretei között megadva, hanem a metanyelv, ha tetszik a köznapi magyar nyelv segítségével van elmesélve. Valamint a teljesülés definíciójában valójában nem a ZF-ben levezethető állításokra hagyatkozunk, hanem a ZF-ről szóló metamatematikai állításokra. Világosan látszik ez a „ $\mathcal{A} \models \neg P(x_1, x_2)[u_1, u_2]$  akkor és csak akkor, ha  $\{ZF\} \not\vdash \{u_1, u_2\} \in R$ ” definícióban.

2. Első pillantásra bizarrnak tűnhet, a halmazelméletnek a modelljeiről beszélni, hiszen ez azt jelenti, hogy a halmazelméletnek a halmazelméletben adjuk meg az interpretációját. Valójában itt nincs semmi probléma, és formálisan ugyanúgy járunk el, mint más axiómarendszer esetében.

**3.** Az axiómarendszerrel kapcsolatban gyakran teszik fel a kérdést: „Van-e valami, ami a szóban forgó axiómákat kielégíti?” Sőt, azt is meg szokás kérdezni, hogy „Azok a dolgok, amelyeknek az axiómáiról van szó, kielégítik-e ezeket az axiómákat? És azt is, hogy „Vajon csak azok a dolgok tesznek-e eleget a szóban forgó axiómáknak, amelyeknek szándékunk szerinti axiómáiról van szó?” Ezek értelmetlen kérdések. Említettük már, hogy értelmetlen „szándékolt interpretációról” és valaminek az „axiomatizálásáról” beszélni, továbbá nincs „standard aritmetika”, amelyet „axiomatizálunk” és nincs „naiv halmazelmélet”, amelyet „axiomatizálunk”. A szigorú értelemben vett matematika számára ezek az elméletek akkor léteznek, ha megadjuk a megfelelő axiomatikus felépítését, méghozzá az itt megismert PC(=)-ben. És ezek az elméletek semmi egyéb, mint amelyeket ilyen módon axiomatikusan megadunk. Ontológiai értelemben értelmetlen olyan „dolgokról” beszélni, amelyek „eleget tesznek” ezeknek az axiómáknak, vagy amelyek „tulajdonságait” ezek az axiómák „tükrözik”. Amik léteznek, azok egyszerűen azok a jelek és azok a szintaktikai szabályok (mechanizmusok), amelyek az adott deduktív rendszerben használatosak.

# The physicalist ontology of formal systems

Thesis:

The objective fact expressed by a mathematical proposition is a fact of the formal system itself, that is, a fact about the physical signs and the mechanical rules according to which the signs can be combined.

It is a common belief that philosophy of mathematics must take account of our impression that mathematical truth is a reflection of fact. As Hardy expresses this constraint,

[N]o philosophy can possibly be sympathetic to a mathematician which does not admit, in one manner or the other, the immutable and unconditional validity of mathematical truth. Mathematical theorems are true or false; their truth or falsity is absolute and independent of our knowledge of them. In *some* sense, mathematical truth is a part of objective reality.<sup>4</sup>

In this section, my aim is to determine what this objective fact actually is.

As we have seen in the previous sections, mathematical propositions, contrary to the propositions of physical theories, are not about anything outside of mathematics, neither in the physical world nor a platonic world—even if the latter were not disqualified for various philosophical reasons. Therefore, this fact can only be a fact of inside the realm of mathematics. More exactly, taking into account that the only means of obtaining reliable knowledge about this fact is mathematical proof, it must be a fact of the realm inside of the scope of formal derivations. I shall argue that this fact is a fact of

---

<sup>4</sup>Hardy 1929.

the formal system itself, that is, a fact about the physical signs and the mechanical rules according to which the signs can be combined.

In my—perhaps prejudiced<sup>5</sup>—reading, Hilbert’s position was very close to this view:

Kant already taught—and indeed it is part and parcel of his doctrine—that mathematics has at its disposal a content secured independently of logic and hence can never be provided with a foundation by means of logic alone; that is why the efforts of Frege and Dedekind were bound to fail. Rather, as a condition for the use of logical inferences and performance of logical operations, something must already be given to our faculty of representation [*in der Vorstellung*], certain extralogical concrete objects that are intuitively [*anschaulich*] present as immediate experience prior to all thought. If logical inference is to be reliable, it must be possible to survey these objects completely in all their parts, and the fact that they occur, that they differ from one another, and that they follow each other, or are concatenated, is immediately given intuitively, together with the objects, as something that neither can be reduced to anything else nor requires reduction. This is the basic philosophical position that I consider requisite for mathematics and, in general, for all scientific thinking, understanding, and communication. And in mathematics, in particular, what we consider is the concrete signs themselves, whose shape, according to the conception we have adopted, is immediately clear and

---

<sup>5</sup>It is an arguable historical question how Hilbert actually considered the “concrete signs themselves” as intuitively present as immediate experience prior to all “logical inferences”. In some readings, Hilbert’s views are compatible with a kind of structural/conceptual realism/platonism. (Cf. Isaacson 1994.) Anyhow, I shall formulate an argument against the view that mathematics has anything to do with some “abstract structure” over and above the physical signs and physical mechanisms constituting the formal system in question.

recognizable.<sup>6</sup>

Sometimes it is argued that symbolism is merely a “convenient shorthand writing” to register the results obtained by *thinking*. To be sure, according to the physicalist account of the mental, it is not important how much “thinking” is involved into the formal manipulations. It is worth while to mentioning, however, that thinking actually plays a marginal role in formal derivations—from the point of view of the truth conditions of a mathematical proposition. I am not concerned about the context of discovery but about the context of justification of a mathematical truth. The discovering of a new conjecture and the discovering of a proof of a conjecture or the definition of a new concept no doubt require the faculty of creative thinking. But, the justification of a mathematical truth, that is, to test that a given sequence of formulas constitutes a proof of a proposition, does not. The more strictly formalized the proof is, the less creative thinking involved. Consequently, the test of the truth conditions of a mathematical proposition is indifferent to complex creative thinking. For instance, when we perform a formal derivation on paper, since each step of manipulation is governed by strict rules, human beings could be replaced by trained animals, robots, etc. Actually, the test whether a given sequence of formulas constitutes a proof of a proposition can be performed by a Turing machine.

Hilbert still emphasizes that **the complete process of symbolic operations must be surveyable to us**. This was a very common idea at his time. However, in the contemporary mathematics there are derivations which are not surveyable by the human mind—we cannot observe the whole derivation process, only the outcome of the process, the proved theorem. This is the case, for example, in the proof of the four-color theorem,<sup>7</sup> where certain steps of the proof are performed through very complex computer manipula-

---

<sup>6</sup>Hilbert 1967, p. 376.

<sup>7</sup>See Tymoczko 1979.

tions. Sometimes even the theorem obtained through the derivation process is not surveyable. It often happens, for example, that the result of a symbolic computer language manipulation is a formula printed on dozens of pages which is completely incomprehensible to the human mind.

So far we have focused on **the epistemological aspects** of the problem of mathematical truth. The results of our investigation **can be summarized in the following observations.**

- Regarding the truth conditions of a mathematical proposition, contrary to the propositions of physical theories, we do not need to refer to the world outside of the formal system in which they are formulated.
- Testing whether a given string of signs is proof of a mathematical statement is completely determined by the formal system itself, no matter whether and to what extent human mental faculties are involved in the mechanical procedure of derivation.
- The process of derivation **that leads us to the *knowledge of the truth of a mathematical proposition is nothing but the *truth-condition**** of the mathematical proposition in question.

**The *ontology of formal systems*** is crystal-clear: marks, say ink molecules diffused among paper molecules, more exactly, their interaction with the electromagnetic field illuminating the paper, or something like that. The rules according by which the marks are written on the paper form a strict mechanism which is, or easily can be, encoded in the regularities of real physical processes. From the point of view of the truth conditions of a mathematical proposition, human activity in this process plays a marginal role. Moreover, the marks and rules can be of an entirely different nature, like, for

instance, the cybernetic states of a computer, supervening on the underlying physical processes.

Sometimes one executes simple formal derivations also **in the head**.<sup>8</sup> However, *from the point of view of the physicalist interpretation of mind* this case of formal manipulation does not principally differ from any other cases of derivation processes. If the signs and the rules of a formal system can be embodied in various physical states/processes, why not let them be embodied in the neuro-physiological, biochemical, biophysical brain configurations—more exactly, in the physical processes of the human brain? If this is the case, that one of the paths—as some rationalists believe, the only path—to trustworthy knowledge, the deductive/logical thinking, can be construed as a mere complex of physical (brain) phenomena, without any reference to the notions of “meaning” and “intentionality”, or the vague and untenable concept of the acausal “global” supervenience on the physical,<sup>9</sup> then this is, actually, a very strong argument for physicalism.

That is to say, **mathematical truths are revealed to us only through real physical processes**. From this point of view we must agree with the quantum computer theorists David Deutsch, Artur Ekert, and Rossella Lupacchini:

Numbers, sets, groups and algebras have an autonomous reality quite independent of what the laws of physics decree, and the properties of these mathematical structures can be just as objective as Plato believed they were (and as Roger Penrose now advocates). But they are revealed to us only through the physical world. It is only physical objects, such as computers or human brains, that ever give us glimpses of the abstract world of mathematics.<sup>10</sup>

---

<sup>8</sup>Much more rarely than one would think. Even in the simplest cases, a proper formal derivation is much too complex to be executable in one’s head.

<sup>9</sup>Cf. Chalmers 1996, pp. 33–34

<sup>10</sup>Deutsch *et al.* 2000, p. 265.



:

It seems that we have no choice but to recognize the dependence of our mathematical *knowledge* (though not, we stress, of mathematical truth itself) on physics, and that being so, it is time to abandon the classical view of computation as a purely logical notion independent of that of computation as a physical process.<sup>11</sup>

(Note that they still maintain a platonic concept of truth in logic and pure mathematics as independent of any contingent facts. The reason is the distinction they draw between knowledge and truth. They do not recognize what I emphasized above that the existence of a physical process of derivation that leads us to the *knowledge* of the truth of a mathematical proposition is nothing but the *truth-condition* of the mathematical proposition in question.)

In order to determine **what kind of objective fact is reflected in a mathematical truth**, we apply our ontological test question Q2: *In what respects is the world different if a given mathematical proposition is true or false?* In other words, what kind of state of affairs in the world determine whether a mathematical proposition is true or false? Now we can answer this question. As the above epistemological analysis shows, the truth of a mathematical proposition is determined by a fact of the formal system in which the proposition is formulated. That is to say, it reflects a fact of the physical system consisting of the signs and the mechanism of derivation in question; *a mathematical proposition is nothing but an ordinary scientific statement of the existence of a proof, that is, the existence of a particular physical process in the formal system in question.* The mathematical proposition ‘ $3 + 2 = 5$ ’, which actually means formula ‘ $3 + 2 = 5$  derives from the formulas called the axioms of arithmetic’, is nothing but—this is the physicalist step—the scientific assertion

---

<sup>11</sup>Deutsch *et al.* 2000, p. 268.

that there exists a proof-process in the formal system called arithmetic, the result of which is the formula  $3+2 = 5$ . This is an ordinary scientific assertion, just as the assertion of the chemist about the existence of the process  $2H_2 + O_2 \rightarrow 2H_2O$ . In this way, mathematical truths express objective facts (of the physical world inside of the formal system as physical system). They are synthetic, *a posteriori*, not necessary, and not certain, they are fallible, but have contingent factual content, as any similar scientific assertion.

‘ $3 + 2 = 5$ ’

actually means that

‘formula  $3 + 2 = 5$  derives from the formulas called the axioms of arithmetics’

which is nothing but

the assertion that there exists a proof-process in the formal system called arithmetics, the result of which is the formula  
 $3 + 2 = 5$ .

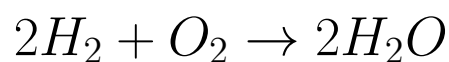
This *is not* a linguistic object!

the usual formalist step

This *is* a linguistic object!

[the physicalist step](#)

This is a usual scientific assertion, just as the assertion of the chemist about the existence of the process



According to this view, a mathematical proposition can be true before anybody can prove it. This simply refers to the normal situation in sciences, that things exist in the world that have not been discovered yet. It is true that process  $2H_2 + O_2 \rightarrow 2H_2O$  exists in the world even if the chemist has not discovered the existence of this process yet. Actually, it would be better to give an example of a reaction in the chemistry of man-made materials such as plastic. The laws of nature predetermine whether a certain chemical process is possible or not, even if nobody has initiated such a process yet. But this kind of independence of the concept of objective truth of mathematical statement from the concept of ‘having been proved’

does not entitle us to claim—as the platonists do—that there is a “truth” in mathematics which is different from the concept of ‘being a theorem in a formal system’. In this sense, the statement of Goldbach’s conjecture is objectively true or false; it is an objective fact of the formal system, as a physical system, that there exists proof for such a statement or not. But this objective truth or falsity has nothing to do with such a platonic concept of truth and falsity as “either it is the case for all even numbers that it is the sum of two primes, or there exists an even number which is not”. The reason is that the phrase that something “is the case for all natural numbers” is meaningless. Not because of the infinity involved in this phrase, but because there is no such a realm of “natural numbers” where the state of affairs may or may not correspond with such a statement.

Not only is the ontological picture so obtained uniform but we have a uniform semantics for all discourse: mathematical and non-mathematical—just as the different sorts of mathematical realism aim for. In order to achieve this semantical uniformity under the umbrella of the Tarskian theory, we only had to take a small step. We had to recognize that mathematical statements, if expressed in everyday language, have the form of ‘ $\Sigma$  implies  $X$ ’ instead of ‘ $X$ ’. Our epistemological/methodological investigations concluded that only the ‘ $\Sigma$  implies  $X$ ’-sentences are scientifically justified *statements* in mathematics.  $X$  is not a statement. It is merely a formula of a formal system and it should *not* be regarded as a linguistic object at all—in the ordinary sense of language. Consequently it has no meaning, it does not refer to anything, it is not a carrier of Tarskian truth. Just like we do not assign meaning and truth or falsity to a dishwasher or a brick, since they are not linguistic objects. One has to recognize that ‘ $X$ ’-sentences merely belong to the sloppy jargon of the mathematician and they are actually used as abbreviations for the corresponding ‘ $\Sigma$  implies  $X$ ’-sentences. If not, then they are negligible verbal decorations. ‘ $\Sigma$  implies  $X$ ’-sentences do have mean-

ings and carry Tarskian truths about real physical entities of clear ontology.

Let me make this still clearer. Assume that  $\Sigma$  is a collection of formulas  $Y_1, Y_2, \dots, Y_m$ . Not only is  $X$  merely a formula of the formal system without meaning and Tarskian truth but the same holds for the formulas  $Y_1, Y_2, \dots, Y_m$ —even if they are axioms. Since  $Y_1, Y_2, \dots, Y_m$  are not linguistic objects either.

The ‘ $\Sigma$  implies  $X$ ’-sentences are the statements of mathematics that can be indispensable to physical theories—since they are no other statements. Let us accept, for the sake of the argument,<sup>12</sup> the Quine-Putnam indispensability argument:

- (1) We ought to have ontological commitment to all and only the entities that are indispensable to our best scientific theories.
- (2) Mathematical entities are indispensable to our best scientific theories.

- 
- (3) We ought to have ontological commitment to mathematical entities.

This means, we ought to have ontological commitment to the entities that mathematical *statements* are talking about. These statements are talking about formal systems, strings of signs, and about derivation processes relating certain strings ( $\Sigma$ ) with some other string of signs ( $X$ ), and so on. All these entities do exist, and to top it all they exist in the physical world.

This picture is entirely compatible with the other fact that (an originally non-linguistic) objects like  $X$  can also be indispensable for

---

<sup>12</sup>There have been many objections to both premises of the argument. See Field 1980, Maddy 1992; 1997, and Sober 1993. See also my objections in points ??–??.

a physical theory  $(L, S)$ . They can be indispensable as *formulas* in the formal system  $L$ , and they may carry physical Truths<sub>2</sub> according to the semantics  $S$ . For example, if  $X \equiv \forall q \exists z q = z \cdot e$  is a formula expressing the (empirically confirmed) physical statement (according to  $S$ ) that all electric charge is a multiple of the elementary charge, then—by virtue of the indispensability argument—we ought to have ontological commitment to the strings like  $X$  and the derivation processes in  $L$ , the electric charge, the elementary charge, and we may also have commitment to the reality of the property of electric charges expressed by the formula in question. Again, all these entities are accommodated in the physical world.

Since  $X$  is not a linguistic object in the ordinary sense of language, the quantifiers eventually used in  $X$  should not be “interpreted” in the usual way. For example, the arithmetical formula  $\exists n n > 17$  should not be interpreted as an ontological statement of the existence of an entity which is a natural number and larger than 17. Actually it shouldn't be interpreted at all, because it has no meaning, since it is just a formula, a string of physical signs in a formal system.

## Abstraction is a move from the concrete to the concrete

Many philosophers of mathematics, while admitting that formal systems are “represented” in the form of physical signs and mechanical rules, still assume that there is something behind this physical representation, an “abstract structure” that is “represented”. Sometimes we find the same ambivalent views in the formalist school. Curry writes:

... although a formal system may be represented in various ways, yet the theorems derived according to the spec-

ifications of the primitive frame remain true without regard to changes in representation. There is, therefore, a sense in which the primitive frame defines a formal system as a unique object of thought. This does not mean that there is a hypostatized entity called a formal system which exists independently of any representation. On the contrary, in order to think of a formal system at all we must think of it as represented somehow. But when we think of it *as* formal system we abstract from all properties peculiar to the representation.<sup>13</sup>

What does such an “abstraction” actually mean? Let us first consider what we obtain if we abstract from some unimportant, peculiar properties of a physical system  $Z$ . In accordance with what we said about the physical theories, we obtain a theory  $P = L + S$  about  $Z$ , that is, a formal system  $L$  with a semantics  $S$  relating the marks of the formal system to the (important) empirical facts of the physical system  $Z$ —where  $L$  is a formal system in the mind, or in a computer, etc. Now, the same holds if the physical system is a formal system (a “representation of a formal system”, in Curry’s terminology)  $Z = L_1$ . Through the abstraction we obtain a theory  $L_2 + S$  describing some important properties of the system  $L_1$ . That is, instead of an “abstract structure” we obtain another formal system  $L_2$  “represented somehow”—in Curry’s expression.

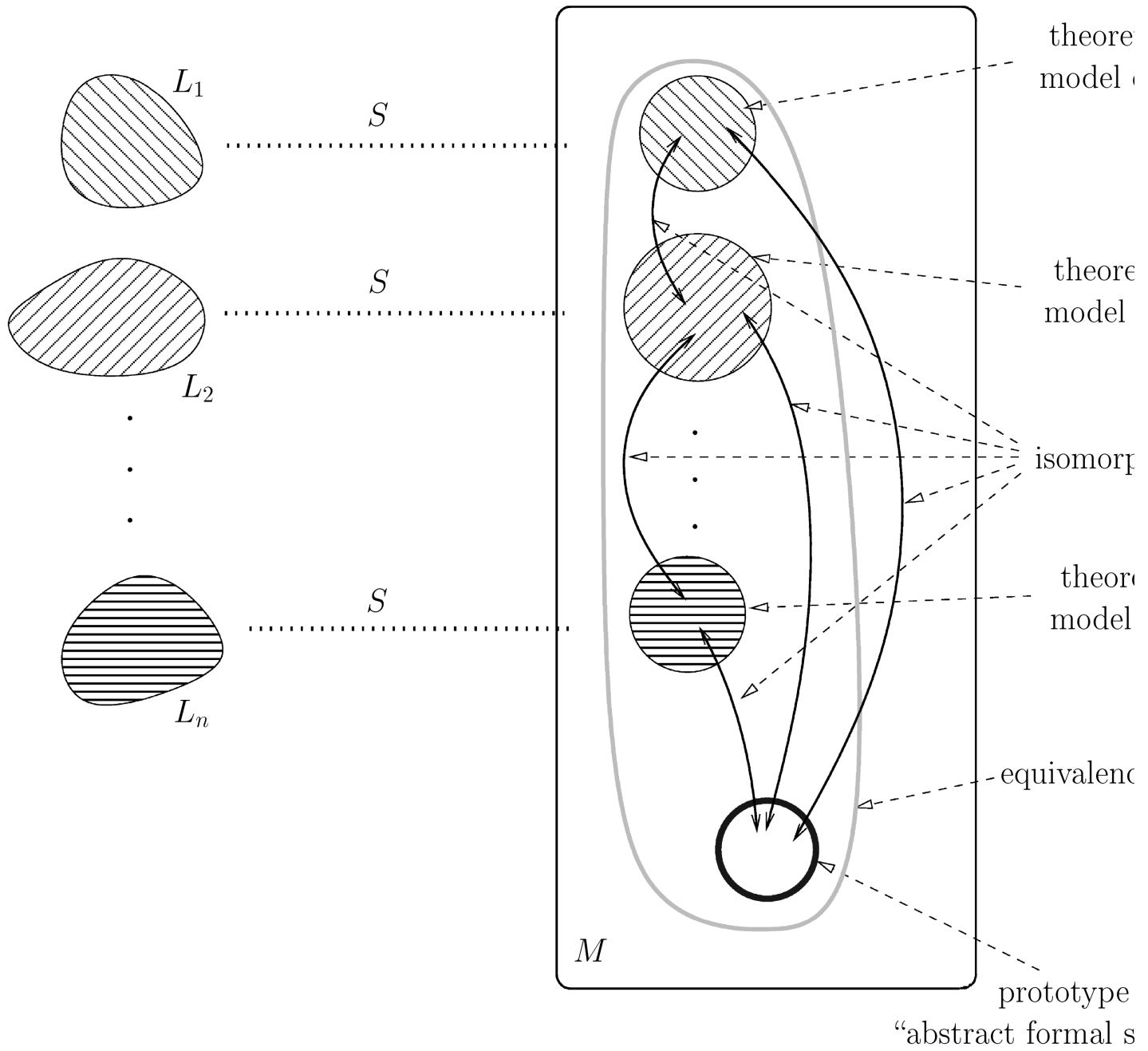
Thus, formal systems are always “flesh and blood” physical systems. These concrete physical systems should not be regarded as physical representations of some abstract formal systems. There are no such abstract things over and above the physically existing formal systems.

By the same token, one cannot obtain an “abstract structure” as an “equivalence class of isomorphic formal systems” or something like that, since in order to think of such things as “isomorphism”, “equiv-

---

<sup>13</sup>Curry 1951, p. 30.

alence”, “equivalence class” at all we must think of them as living in a formal system “represented somehow”. **For it is a categorical mistake to talk about “isomorphism” between two physical objects.** To compare two formal systems  $L_1$  and  $L_2$  we have to use a meta-mathematical theory capable of describing both  $L_1$  and  $L_2$ . That is to say we have to have a physical theory  $(M, S)$  where  $M$  is a third formal system and the semantics  $S$  points partly to  $L_1$ , which is a part of the physical world, and partly to  $L_2$ , which is another part of the physical world.



Since "isomorphism", "equivalence class", etc. are set-theoretic concepts,  $M$  must be a formal system containing set theory. Formal



systems  $L_1, L_2, \dots, L_n$  are simultaneously represented in the physical theory  $(M, S)$ . Only in  $M$  it is meaningful to say that the theoretical models of  $L_1, L_2, \dots, L_n$  are isomorphic and constitute an equivalence class. Only in  $M$  we can define the prototype of these structures, which can be regarded as an “abstract mathematical structure”. And, more importantly, all these mathematical objects live in the formal system  $M$ , in a “flesh and blood” formal system existing in the physical world.

This is neither a nominalistic view nor an attack on *scientific* realism. When a satisfactorily confirmed physical theory claims that a physical object has a certain property adequately described by means of a formal system, then this reflects—with or without the “foot-stamp” of the true realist—a real feature of physical reality. When many different physical objects display a similar property that is describable by means of the same formal system, then we may generalize and claim that these physical objects all possess the feature in question. This will be a true general feature of the group of objects in question, described by means of a formal system as a real physical system.

But, this realist commitment does not entitle us to claim that “abstract structures” exist over and above the real formal systems of physical existence. Again, the reason is that if we tried to consider such an “abstract structure” as a feature of the formal system itself, or as a general feature of many similar formal systems, then we would obtain another, “flesh and blood”, formal system of physical existence.

This observation, in conjunction with our previous observation that the role of human faculties in the formal machinery establishing the truth of a mathematical sentence is limited and unessential, is not a surprise from a physicalist point of view. For even if mathematical objects are “all in someone’s mind”, they would be nothing but physical processes going on in our brain and body. However, **independen-**

dently of the general physicalist account of the mental, **the simple fact that abstraction is unable to go beyond the “physically represented” formal systems** is a very strong argument against structuralism and concept Platonism.

In brief: if, according to Frege’s account,<sup>14</sup> the abstract things are items in the “third realm”, which are non-mental and non-sensible, then mathematics has nothing to do with “abstract” things. This conclusion is in contradiction with the generally accepted view according to which mathematical objects constitute paradigm cases of abstract entities. The confusion is caused by the misunderstanding of the following facts:

- (a) Mathematical truths are independent from the state of affairs in that part of the mathematician’s external world, which is also external to the formal system in question. (Let us call this part of the world realm *A*.) That is, mathematical truths are independent of the realm traditionally described by physical theories. Therefore mathematical truths seem to be spaceless and timeless.
- (b) Mathematical truths are independent of that part of the mathematician’s internal world, which is external to the formal system in question. (Let us call this part of the world realm *B*.) Mathematical truths are intersubjective.
- (c) Within the framework of a physical theory, a mathematical truth may correspond to a fact of realm *A*. This correspondence is depending on the concrete physical theory and its faithfulness is an empirical question. Similarly, a mathematical truth can correspond to an idea in realm *B*. This idea, however, is subjective and may vary from person to person. As it follows from (a) and (b), mathematical truths are over and above the realms *A* and *B*.

---

<sup>14</sup>Frege 1968.

Now, observation (c) is widely misinterpreted as saying that mathematical truths are truths about “abstract entities” existing over and above the “concrete representations” in both the internal and the external worlds. However, this claim obviously overlooks the fact that realm  $A$  is not identical with the external world and realm  $B$  is not identical with the internal world. What is missing from  $A$  and  $B$  is just what mathematical truths refer to, the formal systems themselves constituting a particular part of the real world, either in its external or in its internal parts—it does not make essential difference in our physicalist framework. And, as we have seen, abstraction does not lead outside this realm of concrete physical entities.

To sum up, a formal system is a physical system, the marks of the formal system are embodied in different phenomena related to the system and the derivation rules are embodied in those regularities that govern the system’s behavior. A mathematical derivation, making a mathematical proposition true, is nothing but a physical process going on in the formal system, and a theorem is the output of the process. To prove a theorem is nothing but to *observe* a derivation process in a formal system—that is, to observe a physical process in a physical system. That is all! In this physicalist ontological picture there are no “mathematical structures”, as abstract thoughts, which are “represented” in the various formal systems.

Thus, physicalism—including the physicalist account of the mental—completes the formalist foundation of mathematics and removes the last residues of Platonism. The physicalist ontology of mathematical truth makes it completely pointless in mathematics to introduce a concept of truth different from that of being proved. Mathematical proposition, as a formula in a formal system, does not carry meaning and semantic truth. At the same time, however, it corresponds to a physical fact. By this correspondence, a  $\text{true}_1$  mathematical proposition reflects a truth in the usual sense of Tarski’s semantical theory of truth, just like a  $\text{true}_2$  sentence in a physical

theory. Namely, it reflects a fact, a physical fact of the formal system itself. In this way, indeed, “mathematical truth is a part of objective reality”.

This is the way I propose to “naturalize mathematics”. In this way, **mathematical knowledge is *not conventional***—except the choice of the topic itself, there is nothing conventional in the statement ‘ $\Sigma$  implies  $X$ ’. **It is *not trivial***—sometimes it is highly non-trivial whether  $\Sigma$  implies  $X$ . **It is *not perfect, not a priori, and not certain***. Just like non-mathematical sciences, **mathematics delivers to us *knowledge of contingent facts about a particular part of the physical world***. Formal systems constitute this particular part of the physical world. This is what we can call “mathematical reality”, and mathematicians rightly think themselves as scientists, exploring the intricacies of mathematical reality

Since there are no “abstract formal systems” over and above the physically existing systems of signs and derivation processes, in order to simplify our further considerations, we make the following stipulation without the loss of generality:

**Stipulation 1** *A formal system is a machine (like a computer) which has the following behavior: when it is started it prints out the list of axioms and derivation rules of the system in question and then it prints out a sequence of formulas constituting a proof of a theorem and stops.*

In order to remind the reader of this stipulation I shall sometimes call a formal system a **formal machine**. I do not want to specify what will happen if we start the machine again. In general it produces another sequence of theorems and stops.

An important consequence of this stipulation is that a statement is a “mathematical statement” only if it is a string printed out by the corresponding formal machine. In other words, it is not a math-

ematically true statement if its proof involves faculties beyond the formal machine itself.

# Induction versus deduction

The physicalist-formalist approach has interesting and important consequences in the philosophical analysis of Gödel's theorems and other foundational questions of mathematics, as I shall discuss in the next chapters. In the last part of this philosophical introduction we confine our attention to an important epistemological consequence of the physicalist ontology of formal systems.

It is a long tradition in the history of philosophy that—in **Leibniz's** words:

There are ... two kinds of truths: those of reasoning and those of fact. The truths of reasoning are necessary and their opposite is impossible; the truths of fact are contingent and their opposites are possible.<sup>15</sup>

According to this tradition, one cannot justify a general statement about the world by *induction*. *Deduction*, contrary to induction, provides secure confidence because it is based on pure reasoning, without referring to empirical facts. According to the key idea of rationalism, cognition is an independent source of trustworthy knowledge. Moreover, it is the only secure source of knowledge, the rationalists say, because cognition is the only source of *necessary* truth, while experience cannot deliver to us necessary truths, i. e., truths completely demonstrated by reason.

Let us leave aside the epistemological valuation of knowledge we obtain through inductive inference and consider in more detail the problem of deduction. As Ayer pointed out, the empiricist encounters the following difficulties: it must be either assumed that the truths of formal logic and mathematics are not necessary truths, in which case one must account for the universal conviction that they are; or one must say that they have no factual content, and then it must

---

<sup>15</sup>Rescher 1991, p. 21.



be explained how a proposition which is empty of all factual content can be true and useful and surprising.

According to the *physical realist*, mathematical and logical truths are not certain and not necessary, since they are nothing but generalizations of our fundamental experiences about the physical world, and, as such, they are, admittedly, fallible.

*Logical empiricists*, on the contrary, did not reject the necessity and certainty of mathematical and logical truths. According to their solution, analytical truths do not refer to the facts of reality. For we cannot obtain more information through deductive inference than that already contained in the premises. In other words, according to the logical empiricism, there are no synthetic *a priori* statements.

*Popper's falsification principle* also accepts the necessity and certainty of mathematical and logical truths. This is the basis of the principal distinction between induction and deduction. Similarly, this principal distinction between the “trustworthy deductive inference” and the “always uncertain inductive generalization” is the fundamental tenet upon which the widely accepted *hypothetico-deductive* and *Bayesian* theories of science are built up, seemingly eliminating the problem of induction.

Now, from the standpoint of our *physicalist ontology of formal systems*, we have arrived at the conclusion that **mathematical and logical truths are *not necessary and not certain, but they do have factual content referring to the real world.***

For “deduction” is a concept which is meaningful only in a given formal system. On the other hand, as we have seen, a formal system is nothing but a physical system, and derivation is a physical process. The knowledge of a mathematical truth is the knowledge of a property of the formal system in question—the knowledge of a fact about the physical world. *The formal system is that part of physical reality to which mathematical and logical truths refer.*

It must be emphasized that this reference to the physical world

is of a nature completely different from that assumed, for example, by Mill in his physical realist philosophy of mathematics. In the terminology we introduced with respect to physical theories, the formal statements still do not have any reference to the real world in the sense of the truth-conditions of  $\text{Truth}_2$ , since mathematics does not provide us with a semantics directed from the formal system to the outside world. When we are talking about the empirical character of mathematical truths, we are still talking about  $\text{Truth}_1$ , namely we assert that even  $\text{Truth}_1$  is of empirical nature, the factual content of which is rooted in our experiences with respect to the formal system itself. Mathematics is, in this sense, an empirical science.

The knowledge we obtain through a deductive inference is nothing but an empirical knowledge we obtain through the observation of the derivation process within the formal system in question. In other words, ***deduction is a particular case of induction***. Consequently, *the certainty of mathematics, that is the degree of certainty with which one can know the result of a deductive inference, is the same as the degree of certainty of our knowledge about the outcomes of any other physical processes.*

For example, the reason why the truth of the height theorem is uncertain is not that our knowledge about the properties of “real triangles” is uncertain, as Mill takes it,<sup>16</sup> but rather that our knowledge about the deductive (physical) process, the outcome of which is the height theorem, is uncertain, no matter how many times we repeat the observation of this process.

In order to explain the universal conviction that mathematical truths are necessary and certain, notice that there are many elements of our knowledge about the world which *seem* to be necessary and certain, albeit they have been obtained from inductive generalization. If we need a shorter stick, we break a long one. We are “sure” about

---

<sup>16</sup>The same kind of fallibility appears in Maddy’s theory of naturalised mathematical intuition (Maddy 1990, p. 268).



the outcome of such an operation: the result is a shorter stick. This regularity of the physical world is known to us from experiences.

The certainty of this knowledge is, however, no less than the certainty of the inference from the Euclidean axioms to the height theorem. Mathematical and logical truths are considered necessary and certain for the following two reasons:

1. Usually formal systems are simple and stable physical systems.
2. The knowledge of mathematical truths does not require observations of the world external to mathematics.

To sum up, the physicalist approach resolves Ayer's problem in the following way:

- Mathematical and logical truths **express objective facts** of a particular part of the physical world, namely, the facts of the formal systems themselves. They are **synthetic, a posteriori, not necessary, and not certain, they are fallible**, but have contingent factual content, as any similar scientific assertion.
- The fact that
  - a) the formal systems usually are simple physical systems of **stable** behaviour and
  - b) that the knowledge of mathematical and logical truths **does not require observations** of the world **external** to the formal systems

explains why mathematical and logical truths **appear to everyone to be necessary, certain and a priori**.

Empiricism is not challenged by the alleged necessary truths delivered by mathematical and logical reasoning. On the contrary, consequent physicalism can resolve the long-standing debate surrounding

the truth-of-reasoning versus truth-of-facts dichotomy. **Mathematical and logical truths are nothing but knowledge obtained through inductive generalization from experiences with respect to a particular physical system, the formal system itself.** Since mathematical and logical derivations are reasonings *par excellence*, one must conclude that reasoning does not deliver to us necessary truths. **Reasoning is, if you like, a physical experiment. Therefore, *the certainty available in inductive generalization is the best of all possible certainties!***

# Meta-mathematical theory

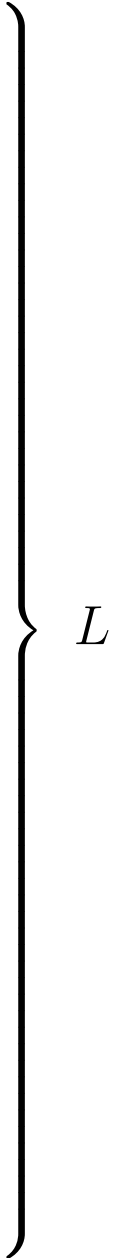
A meta-mathematical theory is a theory *describing* a formal system. This fact in itself provides compelling reason to follow Hilbert's careful distinction between mathematics (i.e., a system of meaningless signs) and meta-mathematics (meaningful statements *about* mathematics).<sup>17</sup> That is to say, a meta-mathematical theory is a theory describing a physical system. Consequently, **meta-mathematical theories are physical theories**. All the truths that a meta-mathematical theory can tell us about its object are of the type Truth<sub>2</sub>. And, since there is no *a priori* physical truth, there is no *a priori* meta-mathematical truth. This means that **no feature of a formal system can be "proved" mathematically**. (Not to mention that, according to the physicalist account of mathematical truth, there is no *a priori* mathematical truth either.) Genuine information about a formal system must be acquired by *a posteriori* means, that is, by observation of the formal system and, as in physics in general, by inductive generalization.

So, a properly construed meta-mathematical theory has the same structure as physical theories in general. Let  $L$  denote the object formal system described by the meta-theory in question. The meta-mathematical theory consists of two components, a formal system  $M$  and a semantics  $S$ . Now, in order to make any *prediction* using the meta-mathematical theory  $(M, S)$  about the properties of the formal system  $L$ , first one has to confirm that  $(M, S)$  is a faithful theory of  $L$ . And, as in the case of any other physical theory, **there is no way to confirm this faithfulness other than to confirm it empirically**.

---

<sup>17</sup>See Nagel and Newman 1958, p. 28.

An important consequence of this fact is that one cannot “mathematically prove” a property of a formal system such as, for example, its consistency. Just as one cannot “mathematically prove” the conservation of energy in a certain physical processes. Let me illustrate this with a well known and simple example of the so called “absolute proof of the consistency of sentence calculus”.<sup>18</sup>

<p><b>Alphabet of symbols:</b></p> <p><math>\sim, \supset, (, ), p, q, r,</math> etc.</p> <p><b>Well-formed formulas:</b></p> <ol style="list-style-type: none"> <li>1. <math>p, q, r,</math> etc. are wfs.</li> <li>2. If <math>A, B</math> are wfs. then <math>(\sim A), (A \supset B),</math> are wfs.</li> <li>3. All wfs. are generated by 1. and 2.</li> </ol> <p><b>Axiom schemes:</b></p> <p>(L1)            <math>A \supset (B \supset A)</math></p> <p>(L2)            <math>(A \supset (B \supset C) \supset ((A \supset B) \supset (A \supset C)))</math></p> <p>(L3)            <math>((\sim A) \supset (\sim B) \supset (B \supset A))</math></p> <p><b>Modus Ponens:</b></p> <p>(MP)           <math>A</math> and <math>(A \supset B)</math> implies <math>B</math></p>	
--	--

The system  $L$  is called consistent if there is no formula  $X$  such  $\vdash_L X$  and  $\vdash_L \sim X$ . The standard “absolute proof” of the consistency of  $L$  goes as follows:

---

<sup>18</sup>See Nagel and Newman 1958, pp. 45–56, or any mathematical logic text-book, for example, Hamilton 1988, Chapter 2.



**Definition:** A *coloring* of  $L$  is a function  $v$  whose domain is the set of wfs. of  $L$  and whose range is the set  $\{red, blue\}$  such that, for any wfs.  $A, B$  of  $L$ ,

$$(i) \quad v(A) \neq v(\sim A)$$

$$(ii) \quad v(A \supset B) = blue \text{ if and only if } v(A) = red \text{ and } v(B) = blue$$

**Definition:** A wfs.  $A$  is *stably red* if for every coloring  $v$ ,  $v(A) = red$ .

**Proposition 1:** For every formula  $A$ , if  $A$  is a theorem of  $L$  then  $A$  is stably red.

**Proof:** Let  $A$  be a theorem. The proof is by induction on the number  $n$  of wfs. of  $L$  in a sequence of wfs. which constitutes a proof of  $A$  in  $L$ .

$n = 1$   $A$  is an axiom. One can easily verify that every axiom of  $L$  is stably red.

$n > 1$  Induction hypothesis: all theorems of  $L$  which have proofs in fewer than  $n$  steps are stably red.

Assume that the proof of  $A$  contains  $n$  wfs. Now, either  $A$  is an axiom, in which case it is stably red, or  $A$  follows by (MP) from previous wfs. in the proof. These two wfs. must have the form  $B$  and  $(B \supset A)$ . But, since  $B$  and  $(B \supset A)$  are stably red, it follows from (ii) that  $A$  is stably red.

**Proposition 2:**  $L$  is consistent.

**Proof:** As is known, one can easily proof that if both  $X$  and  $\sim X$  are theorems in  $L$  then arbitrary formula is a theorem. Consequently, if there exists at least one formula in  $L$  which is not a theorem, then  $L$  is consistent. By virtue of Proposition 1 one has to show that there is a formula  $Y$  in  $L$  which is not stably red, that is, there is a coloring  $v$  such that  $v(Y) = blue$ . Let  $Y$  be  $\sim p \supset q$ . Taking into account (i) and (ii),  $v(Y)$  is determined by  $v(p)$  and  $v(q)$ . Since  $v(Y) = blue$  whenever

Now, a properly formulated meta-mathematical theory of  $L$ , will be a formal system with a semantics,  $(M, S)$ , where semantics  $S$  points to the empirical facts of the object formal system  $L$ . Like in other physical theories, the formal system  $M$  is generated from logical, mathematical and physical axioms. Assume that the meta-theory  $(M, S)$  is strong enough to accommodate something like the “absolute proof of consistency” of  $L$ . In the proof we talk about “*functions* over the formulas of  $L$ ”. ‘Function’ is a mathematical concept which is meaningful only in set theory. Also, we “prove *by induction*”, which requires either set theory or arithmetic. Therefore, we assume that the formal system  $M$  contains set theory (say  $ZF$ ), and, consequently,  $M$  also contains the first order predicate calculus with equality.

One may think that there are some vicious circularities here, but this is not so. The object system  $L$  must be regarded as an entirely autonomous formal system, not as a particular part of the predicate calculus contained in  $M$ <sup>19</sup>. What concerns me is an entirely different problem.

The object formal system  $L$  stands as a physical system which has to be described by a physical theory  $(M, S)$ . The elements of the alphabet, the complex strings, the derivation processes, etc., must be somehow represented in  $M$ . Then we introduce theoretical concepts like “stably red”, which expresses a structural property of formulas of  $L$ . “To be a theorem of  $L$ ” is another concept we define in  $M$  expressing another possible property of a formula of  $L$ . Then, we prove a theorem in  $M$  saying that if a formula of  $L$  is a theorem of  $L$  then it is stably red. What is important is that this claim is just a theoretical description of the system  $L$ , which *may be correct or incorrect*. There is no way to decide whether it is correct or not other than testing it empirically. Whether a formula is stably red or not

---

<sup>19</sup>I postpone the question whether we need to be sure about the consistency of the formal system  $M$  in order to use it in a physical (meta-mathematical) theory.

is an empirical question. We must observe the formula in question and analyze its structure. Whether a formula is a theorem or not is another empirical question. Consequently, whether such a statement of the theory  $(M, S)$  as ‘Every theorem of  $L$  is a stably red formula.’ is correct or not is an empirical question—in spite of the fact that a corresponding statement can be derived in  $M$ . (Truth<sub>2</sub> does not follow from Truth<sub>1</sub>!) We *observe* that whenever a formula of  $L$  is a theorem of  $L$ , it is stably red. From these observations, through *inductive generalization*, we arrive at the conclusion that this law is empirically confirmed. And similar observations confirm the whole theory  $(M, S)$  describing  $L$ .

To avoid any misunderstandings, I do not question the statement that sentence calculus is consistent. I believe in it as I believe in the conservation of energy. That is to say, what I claim is that the epistemological status of this statement is the epistemological status of a physical law, which is not the same as the epistemological status of a mathematical theorem. To be sure, both a mathematical theorem and a physical law are confirmed by physical observation. But, the mathematical theorem we derive in  $M$  is confirmed by the physical observation of a sequence of formulas in the formal system  $M$ , constituting a proof of the theorem of  $M$ , while the meta-mathematical statement about  $L$  is confirmed by observations of physical facts of the object formal system  $L$ .



# Ember tervez, Isten végez!

1. *Lehet-e tudni, hogy a Merkúr hol lesz két év múlva?*

Nem. Alkalmazhatjuk a fizika törvényeit és kiszámíthatjuk, hol lesz. Az eredmény azonban csak egy hipotézis, amelyet a korábbi tapasztalat alapján, induktív általánosítással nyerünk.

2. *Tudhatom-e előre azt, hogy egy általam tervezett berendezés a jövőben hogyan viselkedik?*

Nem. *Tervezni* annyit tesz, mint a világ egy porcikáját az  $A$  helyzetbe hozni és – a korábbi tapasztalat alapján induktív általánosítással nyert hipotézisnek megfelelően – remélni, hogy  $B$  történik.

3. *Lehet-e azt előre tudni, hogy egy általunk tervezett komputer egy általunk írt programmal elindítva, mit fog kinyomtatni?*

Nem. Az általunk tervezett komputerből és az általunk írt programot tartalmazó adathordozóból álló rendszer nem más mint a világ egy porcikája valamely  $A$  állapotba hozva, s csak remélhetjük, hogy a korábbi tapasztalatra épített hipotézisnek megfelelő  $B$  dolog fog történni.

4. *Lehet-e azt előre tudni, hogy egy általunk kitalált (megalkotott) formális rendszernek milyen tulajdonságai vannak?*

Nem. A formális rendszer mindig egy fizikai rendszer formájában testesül meg. A rá vonatkozó jóslataink episztemológiai státusza semmiben nem különbözik más fizikai rendszerek viselkedésére vonatkozó jóslataink episztemológiai státuszától. Ha van egy „elméletünk”, mely szerint a szóban forgó formális rendszernek ilyen és ilyen tulajdonságot kell mutatnia, nem tudhatjuk *a priori*, hogy az elméletünk helyes. Elméletünk konfirmálásának egyetlen lehetséges módja – csakúgy, mint minden más fizikai rendszer leírására szolgáló elmélet esetében –, hogy

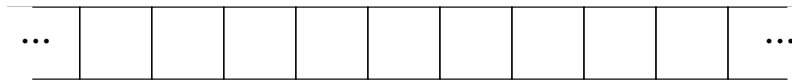
*megfigyeljük*, a rendszer viselkedése valóban olyan-e, mint ahogyan azt az elméletünk megjósolja.

# Turing-gépek

A Turing-gép fogalma és elmélete a mechanikus kiszámíthatóság koncepcióját kívánja megragadni a matematikában.

## A Turing-gép leírása

A gépnek van egy szalagja, amely kis négyzetekre van osztva.



Egy olvasó fej egyszerre egyetlen négyzet tartalmát tudja beolvasni, vagy átírni. Továbbá tud a szalagon egy kockával előre vagy hátra lépni. A gép egy meghatározott ábécét használ:  $S_0, S_1, \dots, S_n$ , ahol  $S_0$  megállapodás szerint az üres kockának felel meg. Feltételezzük, hogy a gépnek véges sok belső állapota lehetséges:  $q_0, q_1, \dots, q_m$ . Továbbá, hogy a gép egy adott pillanatban a pillanatnyi belső állapota és az éppen beolvasott négyzet tartalma által *egyértelműen determinált módon* teszi meg a következő lépést. Ez a lépés a következők egyike lehet:

- (i) megváltoztatja a beolvasott kockában beírt szimbólumot
- (ii) egy kockával jobbra lép
- (iii) egy kockával balra lép

Mint ebből kiderül, a gép működése egyértelműen megadható a következő fajta négyesekből álló véges táblázat segítségével:

	Állapot	Beolvasott	Akció	Új állapot	
(i)	$q_i$	$S_j$	$S_k$	$q_l$	átírás
(ii)	$q_i$	$S_j$	$R$	$q_l$	lépés jobbra
(iii)	$q_i$	$S_j$	$L$	$q_l$	lépés balra

A gép működésének determinisztikus jellege abban nyilvánul meg, hogy nincs két négyes, amelyik ugyanazzal a ⟨állapot, jel⟩ párral kezdődne. Ha a gép egy olyan ⟨állapot, jel⟩ párhoz érkezik, amelyhez nem tartozik négyes, akkor megáll.

Azt a szituációt, melyben a  $q_k$  állapotú gép egy adott jellel ellátott kockáját olvassa be a szalagnak

			$\downarrow q_k$				
$\dots$	$S_{i_0}$	$S_{i_1}$	$S_{i_1}$	$S_{i_3}$	$S_{i_4}$	$S_{i_5}$	$\dots$

a következőképpen fogjuk jelölni:

$$\dots S_{i_0} S_{i_1} q_k S_{i_2} S_{i_3} S_{i_4} S_{i_5} \dots$$

Nevezzük az ilyen stringet *szituáció stringnek*. Például, tegyük fel, a gép a következő instrukciókat kapja:

$$\begin{aligned} q_1 S_1 L q_2 \\ q_2 S_2 L q_2 \end{aligned}$$

A szalag nem üres része mondjuk a következő:

$$S_1 S_2 S_2 S_1 S_2 \dots S_1$$

és a  $q_1$  állapotú gép éppen a második  $S_1$ -et fogja beolvasni. Vagyis

$$S_1 S_2 S_2 q_1 S_1 S_2 \dots S_1$$

Ekkor a  $q_1 S_1 L q_2$  négyesnek megfelelő műveletet hajtja végre, és a következő szituáció fog előállni:

$$S_1 S_2 q_2 S_2 S_1 S_2 \dots S_1$$

Ekkor a  $q_2 S_2 L q_2$  instrukció szerint azt kapjuk, hogy

$$S_1 q_2 S_2 S_2 S_1 S_2 \dots S_1$$

majd

$$q_2 S_1 S_2 S_2 S_1 S_2 \dots S_1$$

és mivel egyetlen instrukció sem kezdődik  $q_2 S_1$ -gyel, a gép megáll.

# Példák elemi műveleteket végrehajtó Turing-gépekre

## Egy $S_j$ jel keresése

Állapot	Beolvasott	Akció	Új állapot
$q_0$	$S_0$	$R$	$q_0$
$q_0$	$S_1$	$R$	$q_0$
$\vdots$			
$q_0$	$S_{j-1}$	$R$	$q_0$
$q_0$	$S_j$	$S_j$	$q_1$
$q_0$	$S_{j+1}$	$R$	$q_0$
$\vdots$			
$q_0$	$S_n$	$R$	$q_0$

A gép megáll amikor  $S_j$ -t talál.

## Mozogjon jobbra és mindenre tegyen vesszőt

$$\left. \begin{array}{cccc} q_0 & S' & R & q_0 \\ q_0 & S & S' & q_0 \end{array} \right\} \text{ az abc minden } S, S' \text{ jelére}$$

Ha azt akarjuk, hogy a gép megálljon egy adott jelnél, például  $\square$ -nál, akkor a táblázatból eltávolítjuk azokat a sorokat, amelyek második eleme  $\square$ .

Nyilvánvalóan semmi akadályja annak, hogy több elemi műveletet végrehajtani képes Turing-gépet egy komplexebb Turing-géppé rakjunk össze. Hogy a gépeket egymástól megkülönböztessük, az állapotaikat kell megfelelően átnevezni. Pl. a fenti két gépből, készítsünk olyan Turing-gépet, amelyik jobbra mozogva megkeres az első  $S_j$ -t, majd onnantól fogva mindenre tesz egy vesszőt! A második gép állapotait átnevezzük, még hozzá éppen úgy, hogy legyen a második gép  $q_0$  állapota az első gép  $q_1$  állapota. Tehát az összetett gép táblázata

$q_0$	$S_0$	$R$	$q_0$
$q_0$	$S_1$	$R$	$q_0$
$\vdots$			
$q_0$	$S_{j-1}$	$R$	$q_0$
$q_0$	$S_j$	$S_j$	$q_1$
$q_0$	$S_{j+1}$	$R$	$q_0$
$\vdots$			
$q_0$	$S_n$	$R$	$q_0$
$q_1$	$S'_j$	$R$	$q_1$
$q_1$	$S_j$	$S'_j$	$q_1$

**HF.** Minden jelet tegyen egy kockával jobbra. [Trükk: a gép úgy tud emlékezni egy információra, hogy egy az információnak megfelelő állapotban van. (Vagyis a Turing-gép egy Markov-folyamat!)]

**HF.** A gép a szalagon egy egyesekből álló blokkot lemásol a szalag üres helyére.

## Eldönthető problémaosztály

Tegyük fel, hogy valamely kérdéseknek/problémáknak egy osztálya megfogalmazható egy véges  $abc$  segítségével úgy, hogy felvihető egy Turing-gép szalagjára. (A szokásos meghatározás szerint)  $Q$  típusú problémáknak egy osztálya *kiszámítható (eldönthető, megoldható)*, ha létezik olyan  $M$  Turing-gép, amely – alkalmazva az osztályba tartozó tetszőleges  $Q$  kérdésre – az 1-en áll meg, ha a  $Q$ -ra adott válasz IGEN és  $\square$ -n, ha a válasz NEM.

PL. Legyen  $Q$  az a kérdés, hogy adott három természetes szám esetén,  $(a, b, c)$ , igaz-e, hogy  $c$  az  $a$  és  $b$  legnagyobb közös osztója? Ennek eldöntésére, ismert egyszerű algoritmus alapján, könnyen konstruálható olyan Turing-gép, amely ezt a fenti értelemben eldönti.

# A Turing-gépek standard leírása

Mivel a Turing-gépek véges számú szimbólumot használnak, az általánosság csorbítása nélkül feltehetjük, hogy ezek a jelek a  $\square, 1, 1', 1'', \dots$ . Az állapotokat is jelölhetjük a  $q, q', q'', \dots$  jelekkel. A gép működését megadhatjuk tehát a  $\square, 1, ', q, R, L$  jelekből álló stringek segítségével, pl. a

$$\begin{array}{cccc} q_0 & 1 & R & q_1 \\ q_1 & 1'' & 1' & q_2 \end{array}$$

utasításokból álló táblázat megadható egyértelműen a következő stringgel:

$$q 1 R q' q' 1'' 1' q''$$

Sőt, mindent kifejezhetünk a  $\square, 1, 1', 1'', \dots$  abc-vel:

$$\begin{array}{l} \square \leftrightarrow \square \\ 1 \leftrightarrow 1 \\ ' \leftrightarrow 1' \\ q \leftrightarrow 1'' \\ R \leftrightarrow 1''' \\ L \leftrightarrow 1'''' \end{array}$$

Az  $M$  Turing-gép működését meghatározó táblázatot tehát egyetlen  $\square, 1, 1', 1'', 1''', 1''''$ -stringgel megadhatjuk. Ezt a jelsorozatot  $[M]$ -mel fogjuk jelölni, és a Turing-gép *standard leírásának* nevezük.

## Egy eldönthetetlen problémaosztály („Halting problem”)

Tekintsük a következő kérdést:

$Q_M$  : Megáll-e az  $M$  Turing-gép egy  $\square$  jelen, ha az  $[M]$  jelsorozatra alkalmazzuk?

A  $Q_M$  kérdés egyértelműen megadottnak tekinthető az  $[M]$  megadásával. Jelölje  $\{Q_M\}_M$  az ilyen kérdések osztályát, ahol  $M$  tetszőleges Turing-gépet jelöl. Arra keresünk választ, vajon létezik-e olyan algoritmikus eljárás, magyarul olyan Turing-gép, amely képes megválaszolni minden a  $\{Q_M\}_M$  osztályba tartozó kérdést.

Képzeljünk el egy  $S$  gépet, amelyik teljesíti ezt a feladatot, tehát beolvassa az  $[M]$  stringet és 1-en áll meg, ha a válasz a  $Q_M$  kérdésre IGEN, és  $\square$ -n, ha a válasz NEM. A probléma, hogy hogyan viselkedik ez a gép — melynek tetszőleges  $[M]$ -re működnie kellene —, ha az inputja éppen  $[S]$ ? Ha  $S$  megáll az 1-en, az azt jelenti, hogy a  $Q_S$  kérdésre a válasz IGEN, azaz az  $S$  a  $\square$ -n áll meg ha  $[S]$ -ra alkalmazzuk. És fordítva, ha  $S$  a  $\square$ -n áll meg, az azt jelenti, hogy a  $Q_S$  kérdésre a válasz NEM, tehát az  $S$  gép nem áll meg a  $\square$ -n. Mindkét esetet egyfajta ellentmondásnak szokás tekinteni, és a szokásos konklúzió az, hogy ilyen  $S$  gép nem létezik. Más szóval, hogy a  $Q_M$  problémaosztály *algoritmikusan nem megoldható, eldönthetetlen*.

**NB.** Könnyen gondolhatjuk, hogy a probléma abból származik, hogy a gép az IGEN és NEM válaszokat az 1 és  $\square$  jeleken való megállással közli, és hogy más lenne a helyzet, ha a gép a  $\square$ -n állna meg, ha a válasz IGEN és 1-en, ha NEM. Ez azonban nem igaz. Ha létezik ilyen  $T$  gép, akkor könnyen konstruálható egy másik gép, amely a  $T$  IGEN jelzését 1-be, a NEM jelzését  $\square$ -ba konvertálja, s a két gép kombinációja megint egy olyan Turing-gép lenne, ami eleget tesz az eredeti feltételeinknek, s a fenti argumentum alkalmazható, tehát nem létezik ilyen gép, következésképpen nem létezik  $T$  sem.

**NB.** Könnyen belátható az is, hogy a helyzeten semmit sem változtat, ha a  $Q_M$  kérdést másképpen kódoljuk, hiszen mindig találni olyan Turing-gépet, amelyik a  $M$  egy tetszőleges másik kódolását az  $[M]$  stringbe konvertálja, és viszont.



# Univerzális Turing-gép

Azt gondolhatnánk, hogy a probléma abból fakad, hogy nem lehetséges olyan gépet konstruálni, amely képes átfogni az összes lehetséges Turing-gép működését. Belátható azonban, hogy ilyen gép létezik. *Univerzális Turing-gépnek* nevezzünk egy olyan  $U$  gépet, amely képes arra, hogy beolvassa egy tetszőleges gép  $[M]$  standard leírását, és beolvasva egy tetszőleges  $[P]$  kódját a szalagnak szimulálja  $M$  működését a  $P$  szalag-tartalom mellett. (Annak leírását, hogy egy ilyen univerzális Turing-gép hogyan működik, lásd Crossley 40-41 oldal.)

## A „Halting” probléma univerzális Turing-gépre

Vizsgáljuk tehát azt a szituációt, amikor az  $U$  univerzális Turing-gép az  $M$  gépet fogja szimulálni, amikor az az  $[M]$  stringre van alkalmazva. Más szóval,  $U$  számára adott a

$$W_M = * [M] * * [[M]] *$$

string, mint input. (A  $*$  csak segéd jel, amely jelzi a gép számára, hogy mettől meddig terjed egy egybefüggő része a beolvasott stringnek.) Tekintsük a következő kérdést:

$Q_W$  : Megáll-e az  $U$  univerzális Turing-gép egy  $\square$  jelen, ha a  $W$  jelsorozatra alkalmazzuk?

Legyen  $\{Q_W\}_W$  az ilyen kérdések osztálya. Létezik-e Turing-gép, amelyik képes megválaszolni a  $\{Q_W\}_W$  osztályba tartozó összes kérdést? Válaszunk az, hogy nem.

**25. Tétel.** A  $\{Q_W\}_W$  problémaosztály nem eldönthető.

**Bizonyítás.**  $\{Q_W\}_W$  nyilván tartalmazza az olyan  $Q_W$  kérdéseket is, ahol  $W = W_M$ . Mivel ilyenkor  $U$  szimulálja  $M$  működését,  $U$  akkor és csak akkor áll meg  $\square$ -n ha  $M$   $\square$ -n áll meg, ha  $M$ -et  $\lceil M \rceil$ -re alkalmazzuk. Más szóval, a  $\{Q_W\}_W$  problémaosztály csak akkor lehetne eldönthető, ha a  $\{Q_M\}_M$  problémaosztály eldönthető lenne.

## Halting-probléma, episztemológia, endofizika

1. Elkerülendő a félreértéseket, amelyekkel a populáris irodalomban gyakran találkozunk, ne gondoljunk többet a fenti eldönthetlenségi tételek mögé, mint amit valójában bizonyítottunk! *A tételek nem azt állítják, hogy az adott problémaosztályba tartozó problémák nem dönthetők el algoritmikusan!* A tételek azt állítják, hogy *nem létezik egyetlen* algoritmus (Turing-gép), amely az osztályba tartozó összes kérdést meg tudja válaszolni.

2. A „Halting” probléma tárgyalásánál felbukkan az „önreferencia” motívuma. Világosan kell látni azonban, hogy ennek nincs különösebb jelentősége, és semmi köze nincs a „megismerhető-e a világ, amelynek mi is részei vagyunk” jellegű *endofizikai* problémához és más episztemológiai kérdéshez. Egyáltalán, a szóban forgó matematikai tételek mögött semmiféle metafizikai mélység nincs. Világos ugyanis, hogy csak akkor vonhatunk le a Turing-gépekről mondotakból bármiféle metafizikai/episztemológiai következtetést, ha a Turing-gépre mint valóságos fizikai objektumra gondolunk. Márpedig kérdéses, hogy mi történik egy valóságos fizikai komputerrel olyan esetben, amikor a  $* \lceil M \rceil ** \lceil \lceil M \rceil \rceil *$  stringet olvassa be, vagyis olyan utasítások összességét, melynek alapján egyszerre kellene neki igent és nemet mondania, amit nyilván nem tud. Ugyanúgy, mint ahogy kérdéses mi történik egy autóval, ha egyszerre nyomjuk a féket és a gázt.

**3.** Mit is jelent az, ha egy Turing-gép „képes” megválaszolni egy kérdést. Mikor értelmes azt mondani, hogy a Turing-gépbe bevitt input egy adott kérdésnek a „reprezentációja”? Nyilvánvaló, hogy a reprezentáció nem merülhet ki egy egyszerű megfeleltetésben. Még Újév hajnalán is képesek vagyunk kölcsönösen egyértelmű megfeleltetést definiálni a kiöntött ólom formája és a valóságos világban a dolgok ilyen vagy olyan állása között, mégsem gondolhatjuk komolyan, hogy az ólomöntés kimenetele „megválaszolja” azt a kérdést, hogy a nagylány férjhez megy-e az új évben vagy sem! Tehát ahhoz, hogy azt monthassuk, hogy a Turing-gép viselkedésének egy  $A$  típusa a valós világban a dolgok állásának valamely  $a$  típusát reprezentálja, az szükséges, hogy objektíve fennálljon, hogy a Turing-gép akkor és csak akkor mutatja az  $A$  viselkedést, ha  $a$  fennáll, de legalábbis erős korreláció álljon fenn  $A$  és  $a$  között. Márpedig nem lehetséges egy ilyen korreláció tényét *a priori* állítani.

# Gödel inkomplettségi tétel

## Gödel-számozás

**Egy formula Gödel-száma.** Az aritmetikában használt jelekhez számokat rendelünk:

0	$\Leftrightarrow$	1
s	$\Leftrightarrow$	2
+	$\Leftrightarrow$	3
·	$\Leftrightarrow$	4
=	$\Leftrightarrow$	5
(	$\Leftrightarrow$	6
)	$\Leftrightarrow$	7
,	$\Leftrightarrow$	8
x	$\Leftrightarrow$	9
	$\Leftrightarrow$	10
$\neg$	$\Leftrightarrow$	11
$\wedge$	$\Leftrightarrow$	12
$\exists$	$\Leftrightarrow$	13

A változók jelölésére használjuk a  $x_1, x_2, x_3, \dots$  jeleket. Tekintsük az aritmetika egy formuláját, például

$$+(s(s(0)), s(s(0))) = s(s(s(s(0))))$$

Ehhez a következőképpen rendelünk számot:

$$\begin{array}{cccccccc} + & ( & s & ( & s & \dots & ) & ) & ) \\ 3 & 6 & 2 & 6 & 2 & \dots & 7 & 7 & 7 \\ 2 & 3 & 5 & 7 & 11 & \dots & 113 & 127 & 131 \\ & & 2^3 & 3^6 & 5^2 & 7^6 & 11^2 & \dots & 113^7 127^7 131^7 \end{array}$$

Világos, hogy ezzel a módszerrel minden  $\phi$  formulához egyértelműen hozzárendeltünk egy számot. Ezt a számot a  $\phi$  formula *Gödel-számának* nevezzük, és  $[\phi]$ -vel fogjuk jelölni. Természetesen, nem

minden természetes szám Gödel-száma valamilyen formulának. De ha az, a prímfelbontás egyértelműsége miatt, egyértelmű, hogy milyen formula Gödel-számáról van szó.

**Egy formula sorozat Gödel-száma.** Legyen  $\phi_1, \phi_2, \phi_3, \dots$  formulák egy sorozata. A formulasorozathoz rendelt Gödel-szám:  $2^{\lceil \phi_1 \rceil} 3^{\lceil \phi_2 \rceil} 5^{\lceil \phi_3 \rceil} \dots$ . Világos, hogy a prímfelbontás egyértelműsége miatt egy formulasorozat Gödel-száma egyértelműen meghatározza, hogy milyen formulák sorozatáról van szó.

## Gödel-mondat

Tekintsük a következő meta-elméleti (tehát az aritmetikáról szóló) predikátumot:

$Pf^M(x, y)$ : az  $x$  Gödel-számú formulasorozat az  $y$  Gödel-számú formula bizonyítása.

Most kicsit bonyolítsuk meg:

$Pf^M(x, y, z)$ :  $x$  azon formula bizonyításának a Gödel-száma, melyet az  $y$  Gödel-számú, egy szabad változót tartalmazó formulából kapunk, úgy, hogy a változó helyére a  $z$  számot (individuum változót) helyettesítjük.

A  $Pf^M(x, y, z)$  állításra úgy tekinthetünk, mint számok közötti relációra, vagyis az állítás akkor és csak akkor igaz, ha a megfelelő reláció fennáll. Ha mondunk három számot,  $x, y, z$ , akkor egyszerű aritmetikai algoritmusoknak a (természetesen bonyolult) sorozatával eldönthető, hogy a  $Pf^M(x, y, z)$  mondat igaz-e vagy hamis. Hiszen számok prímfelbontását, és más hasonló aritmetikai műveleteket kell ehhez elvégezni. (A megfelelő reláció rekurzíve megadható.) Ennek alapján megmutatható, hogy létezik olyan  $Pf(x, y, z)$  formulája az

aritmetikának, amelyre

$$\begin{array}{ll} \{\text{aritmetika}\} \vdash Pf(x, y, z) & \text{ha } Pf^M(x, y, z) \text{ igaz} \\ \{\text{aritmetika}\} \vdash \neg Pf(x, y, z) & \text{ha } Pf^M(x, y, z) \text{ hamis} \end{array}$$

Tekintsük most a  $\neg\exists x Pf(x, y, y)$  formulát az aritmetikában. Legyen ennek a formulának a Gödel-száma  $g$ . A következő mondatot Gödel-mondatnak szokás nevezni:

$$\neg\exists x Pf(x, g, g)$$

és  $G$ -vel fogjuk jelölni.

**26. Tétel.** *Sem  $G$ , sem  $\neg G$  nem vezethető le az aritmetikában, tehát*

$$\begin{array}{l} \{\text{aritmetika}\} \not\vdash G \\ \{\text{aritmetika}\} \not\vdash \neg G \end{array}$$

**Bizonyítás.** Tegyük fel, hogy  $G$  bizonyítható, vagyis hogy  $\{\text{aritmetika}\} \vdash \neg\exists x Pf(x, g, g)$ . Legyen a bizonyítást alkotó formulasorozat Gödel-száma  $m$ . Tekintettel arra, hogy a  $g$  Gödel-számú formula a  $\neg\exists x Pf(x, y, y)$ , ez azt jelenti, hogy  $m$  a Gödel-száma azon formula bizonyításának, melyet úgy kapunk, hogy a  $g$  Gödel-számú formulában a változó helyére  $g$ -t helyettesítünk. Azaz, a  $Pf^M(m, g, g)$  mondat igaz, más szóval az  $m, g, g$  számokra fennáll a megfelelő reláció, vagyis  $\{\text{aritmetika}\} \vdash Pf(m, g, g)$ , ami ellentmondás.

Most tegyük fel, hogy  $\neg G$  bizonyítható, tehát  $\{\text{aritmetika}\} \vdash \neg\neg\exists x Pf(x, g, g)$ , vagyis  $\{\text{aritmetika}\} \vdash \exists x Pf(x, g, g)$ . De az az első részben éppen azt bizonyítottuk, hogy  $\{\text{aritmetika}\} \not\vdash \neg\exists x Pf(x, g, g)$ , más szóval, hogy nincs olyan formulasorozat, amely bizonyítása lenne  $\neg\exists x Pf(x, g, g)$ -nek. Ez azonban azt jelenti, hogy 1 nem Gödel-száma egy megfelelő bizonyításnak, vagyis  $Pf^M(1, g, g)$

hamis, hasonlóan,  $Pf^M(2, g, g)$  hamis, és így tovább. Következésképpen,

$$\begin{aligned}\{\text{aritmetika}\} &\vdash \neg Pf(1, g, g) \\ \{\text{aritmetika}\} &\vdash \neg Pf(2, g, g) \\ &\vdots\end{aligned}$$

ami ellentmondás, ha feltesszük, hogy az aritmetika  $\omega$ -konzisztens, ami azt jelenti, hogy nem létezik olyan egy szabad változót tartalmazó  $\phi(x)$  formula, amelyre egyszerre fennállna, hogy

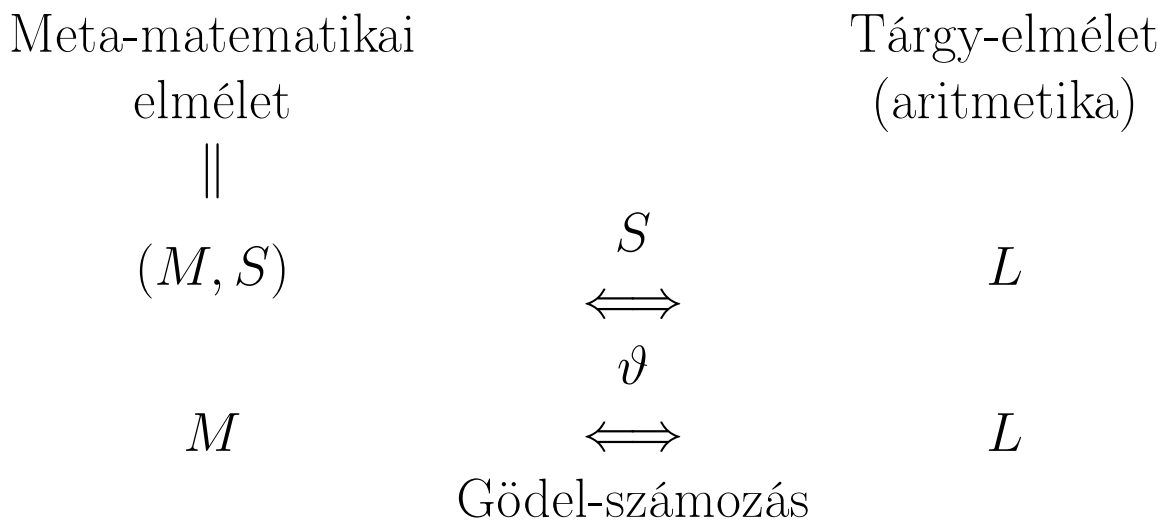
$$\begin{aligned}\{\text{aritmetika}\} &\vdash \exists x\phi(x) \\ \{\text{aritmetika}\} &\vdash \neg\phi(1) \\ \{\text{aritmetika}\} &\vdash \neg\phi(2) \\ \{\text{aritmetika}\} &\vdash \neg\phi(3) \\ &\vdots\end{aligned}$$

A tétel alapján tehát azt mondhatjuk, hogy az aritmetika vagy nem  $\omega$ -konzisztens, vagy létezik benne olyan mondat, amelyre az áll, hogy sem ő, sem a negáltja nem bizonyítható.

**NB.** Gödel-féle eredeti bizonyítás kis módosításával sikerült gyengébb feltétel mellett is bebizonyítani a tételt, nevezetesen, hogy ha az aritmetika *konzisztens*, akkor létezik benne olyan mondat, hogy sem ő sem a negáltja nem bizonyítható.

# Bizonyítás és Igazság

"Röviden, Gödel megmutatta, hogy a bizonyítás az igazságnál gyengébb fogalom, függetlenül a használt axiómarendszerétől." — írja Hofstadter a *Gödel, Escher, Bach* c. művében. Vitatkoznunk kell ezzel a széles körben elterjedt nézettel, noha a tétel jelentésének egy ilyenfajta értelmezése nem áll távol Gödel platonista nézeteivel. Tekintsük át újra a Gödel-tétel bizonyításának sémáját:



Vagyis, adott egy meta-matematikai elmélete az  $L$  formális rendszernek. Ez azt jelenti, hogy adott egy másik formális rendszer  $M$  és egy szemantika  $S$ , ami  $M$ -et és  $L$ -et összeköti. Például olyan mondatokat tudunk mondani  $M$ -ben, mint „a  $\phi$  formula  $L$ -ben nem bizonyítható”, amely az  $L$  egy tulajdonságát hivatott állítani. Jelöljük az egyszerűség kedvéért ezt a mondatot  $nb(\phi)$ -vel. Az ilyen és hasonló mondatoknak van egy Igazság<sub>2</sub> értelemben vett igazsága az  $(M, S)$ -ben. Vagyis egy  $M$ -beli formula akkor igaz<sub>2</sub> <sup>$M$</sup> , ha az  $S$  szemantika értelmében ő egy olyan állítás  $L$ -ről, amely tényszerűen fennáll  $L$ -re. Például,  $nb(\phi)$  akkor igaz<sub>2</sub> <sup>$M$</sup> , ha nem létezik  $\phi$ -nek bizonyítása  $L$ -ben, más szóval, ha nem igaz, hogy  $\phi$  igaz<sub>1</sub> <sup>$L$</sup> .

A tétel bizonyításában ezek után megjelenik egy másik leképezés is, a Gödel-számozás által generált  $\vartheta$  leképezés (Gödel-izomorfizmus). Gyakran tévesen azt állítják, hogy  $\vartheta$  „megőrzi az igazságot”, vagyis, hogy ha  $\alpha$  igaz<sub>2</sub> <sup>$M$</sup> , akkor  $\vartheta(\alpha)$  igaz  $L$ -ben. Más szóval, hogy Gödel-zseniális trükkje éppen az volt, hogy az aritme-



tikáról szóló meta-matematikai elméletet reprezentálta magában az aritmetikában.

Erről azonban nincs szó. Vegyük észre, hogy a bizonyításban nem is használtuk ki, hogy  $\vartheta$  egy igazság-megőrző izomorfizmus lenne. Csupán azt tettük fel (láttuk be), hogy speciálisan a  $Pf^M(x, y, z)$  típusú meta-elméleti mondatokon az. Vagyis, hogy ha  $Pf^M(x, y, z) \text{ Igaz}_2^M$ , akkor  $\{\text{aritmetika}\} \vdash Pf(x, y, z)$ , ahol  $Pf(x, y, z)$  a  $\vartheta(Pf^M(x, y, z))$  formulát jelöli, és ha  $Pf^M(x, y, z)$  nem  $\text{Igaz}_2^M$ , akkor  $\{\text{aritmetika}\} \vdash \neg Pf(x, y, z)$ .

Téves tehát minden olyan megfogalmazás, hogy a  $G$  Gödel-mondat, vagyis a  $\neg \exists x Pf(x, g, g)$  aritmetikai mondat azzal a meta-elméleti jelentéssel bír, hogy „ $G$  (vagyis saját maga) nem bizonyítható  $L$ -ben”. (Vagyis, hogy  $nb(G)$ .) Ezt csak akkor mondhatnánk, ha valóban megadtunk volna egy olyan igazság-megőrző leképezést  $M$ -ből  $L$ -be, amelyik kiterjed  $nb(G)$ -re is és amelyre igaz, hogy  $\vartheta(nb(G)) = G$ . Éppen a bizonyított tétel teszi ezt lehetetlenné: Ha ugyanis,  $nb(G)$  egy igaz meta-matematikai állítás, akkor a neki megfelelő  $G = \vartheta(nb(G))$  formula  $\text{Igaz}_1^L$ , azaz bizonyítható. De a tétel szerint nem bizonyítható, tehát a kondíciók valamelyike nem lehet igaz.

Természetesen, ezzel együtt az is értelmetlen, hogy „a  $G$  mondat igaz, hiszen azt állítja, hogy ő nem bizonyítható, és — minthogy bebizonyítottuk, hogy nem bizonyítható — igazat állít.”

**NB.** Gyakori felvetés, hogy a tételben levezetett állítás önmagában is paradox. Az tudniillik, hogy van olyan mondata az aritmetikának, amelyre az áll, hogy sem ő sem a negáltja nem bizonyítható. Nem kétséges, hogy a matematikai platonista számára ez az tény zavarba ejtő. Minthogy a matematika tanítása erősen platonista szemléletet alakít ki már gyermekkorban, sokan gondolják úgy, hogy mivel a Gödel-mondat az aritmetika egy mondata, egy számokról tett kijelentés, szükségképpen vagy igaz vagy hamis. Nem tehetünk mást,

mint hogy hangsúlyozzuk: aritmetika az, amit itt axiomatikusan megadtunk. És az mondható „igaznak” az aritmetikában, amit az adott rendszerben bizonyítani lehet.

## Gödel második inkomplettségi tétele

Az aritmetika akkor és csak akkor konzisztens, ha  $\{\text{aritmetika}\} \not\vdash 0 = 1$ . Ha ugyanis  $\{\text{aritmetika}\} \vdash 0 = 1$ , akkor (A1)-ből és (A2)-ből azonnal a negáltja is következik, tehát a rendszer inkonzisztens. Másfelől, ha a rendszer inkonzisztens, akkor a 3. tételből következően bármilyen mondat levezethető, így az is, hogy  $0 = 1$ . Az aritmetika konzisztenciája tehát ekvivalens azzal az állítással, hogy nem vezethető le a  $0 = 1$  mondat. Legyen ennek a mondatnak a Gödel-száma  $k$ .

Tekintsük most a következő, *Consis*-nek nevezett mondatot:  $\forall x \neg Pf(x, k)$ . Bonyolult bizonyítással levezethető az aritmetikában, hogy  $Consis \rightarrow G$  ahol  $G$  a  $\neg \exists x Pf(x, g, g)$  Gödel-mondatot jelöli. Ha tehát *Consis* levezethető lenne az aritmetikában, azaz  $\{\text{aritmetika}\} \vdash Consis$ , akkor a  $\{\text{aritmetika}\} \vdash Consis \rightarrow G$ -ből (MP)-vel azonnal következne  $G$ , melyről viszont bebizonyítottuk, hogy nem levezethető. Vagyis igaz a következő tétel:

**27. Tétel (Gödel II. inkomplettségi tétel).** *A Consis mondat nem vezethető le az aritmetikában.*

**NB.**

A tételt rendszerint úgy interpretálják, hogy az aritmetika konzisztenciáját nem lehet magában az aritmetikában bizonyítani. Ez az interpretáció azonban problematikus: Nem igaz, hogy a *Consis* mondat, vagyis a  $\forall x \neg Pf(x, k)$  a „Nem vezethető le a  $0 = 1$  formula az aritmetikában”, vagy a vele ekvivalens „Az aritmetika konzisztens”

meta-elméleti mondatot reprezentálja. Ugyanis semmilyen aritmetikai mondat nem tudja ezt a meta-matematikai mondatot reprezentálni, így tehát a *Consis* sem.

Tegyük fel ugyanis, hogy azt a meta-elméleti állítást, hogy „az  $L$  rendszer konzisztens”, valóban képesek vagyunk reprezentálni az  $L$  rendszer egy  $\psi$  mondatával. Ez azt kell hogy jelentse, hogy a szóban forgó mondat akkor és csak akkor  $\text{Igaz}_1^L$ , vagyis akkor és csak  $\vdash_L \psi$ , ha a rendszer konzisztens. De ez nem lehetséges, hiszen ha fel is tesszük, hogy a konzisztens rendszerben a  $\psi$  mondat valóban levezethető, ez semmiben nem különbözik attól a helyzettől, ha a rendszer inkonzisztens, hiszen a 3. tétel értelmében az inkonzisztens rendszerben bármi levezethető, így az állítólagos konzisztenciát reprezentáló  $\psi$  mondat is. Vagyis nem lehet igaz, hogy  $\psi$  akkor és csak akkor  $\text{Igaz}_1^L$ , ha  $L$  konzisztens, azaz  $\psi$  nem reprezentálhatja az „az  $L$  rendszer konzisztens” meta-matematikai mondatot.