



196

Annales Universitatis Paedagogicae Cracoviensis

ISSN 2080-9751

**Studia ad Didacticam
Mathematicae Pertinentia VII
(2015)**

196

**Annales
Universitatis
Paedagogicae
Cracoviensis**

ISSN 2080-9751

**Studia ad Didacticam
Mathematicae Pertinentia VII
(2015)**

Editor-In-Chief

Piotr Błaszczak, Pedagogical University of Cracow, Poland

Vice Editor-In-Chief

Maciej Klakla, Pawel Wlodkowic University College in Płock, Poland

Assistant Editors

Ireneusz Krech, Pedagogical University of Cracow, Poland

Mirosława Sajka, Pedagogical University of Cracow, Poland

Technical Editor

Marek Janasz

Editorial Board

Ján Gunčaga, Catholic University in Ruzomberok, Slovakia

Valery A. Gusev, Moscow State Pedagogical University, Russia

Eszter Herendiné-Kónya, University of Debrecen, Hungary

Henryk Kałol, Teacher Training College in Bielsko-Biała, Poland

Leonid M. Kotlyar, Kama State Engineering and Economics Academy (INEKA), Russia

Jan Melichar, Jan Evangelista Purkyně University in Ústí nad Labem, Czech Republic

Josef Molnár, Palacký University in Olomouc, Czech Republic

Bogdan J. Nowecki, Pedagogical University of Cracow, Poland

Antoni Pardała, University of Rzeszów, Poland

Svetlana A. Rozanova, Moscow State Institute of Radio-engineering Electronics and Automation (MIREA), Russia

Ingrid Semanišínvá, Pavol Jozef Safarik University in Košice, Slovakia

Vladimir G. Shibakov, Kama State Engineering and Economics Academy (INEKA), Russia

Galina Sinkevich, Saint Petersburg State University of Architecture and Civil Engineering, Russia

Helena Siwek, Pedagogical University of Cracow, Poland

Shlomo Vinner, Hebrew University of Jerusalem, Israel

Eda Vula, University of Prishtina, Republic of Kosova

Erich Ch. Wittmann, Technical University of Dortmund, Germany

Reviewers

Tiziana Bascelli, Liceo Statale „F. Corradini”, Thiene (VI), Italy

Antoni Chronowski, Pedagogical University of Cracow, Poland

Vladimir Kanovei, Institute of Information Transmission Problems, Moscow, Russia

Semen Samsonovich Kutateladze, Novosibirsk State University, Russia

Maria Legutko, Pedagogical University of Cracow, Poland

Paolo Bussotti, University of Udine, Italy

Zdzisław Pogoda, Jagiellonian University, Cracow, Poland

Zbigniew Leśniak, Pedagogical University of Cracow, Poland

Antoni Pardała, University of Rzeszów, Poland

Janusz Czelakowski, University of Opole, Poland

Robert Sochacki, University of Opole, Poland

Jerzy Pogonowski, Adam Mickiewicz University in Poznań, Poland

Marek Ptak, University of Agriculture in Cracow, Poland

Ewa Swoboda, University of Rzeszów, Poland

Jakub Jernajczyk, Academy of Fine Arts in Wrocław, Poland

Grażyna Rygał, Jan Długosz University in Częstochowa, Poland

www.up.krakow.pl/mat/annal-dyd

e-mail: annalesd@up.krakow.pl

© Copyright by Wydawnictwo Naukowe UP, Kraków 2015

ISSN 2080-9751

Wydawca

Wydawnictwo Naukowe UP

30-084 Kraków, ul. Podchorążych 2

tel./faks: 12 662-63-83, tel. 12 662-67-56

e-mail: wydawnictwo@up.krakow.pl

<http://www.wydawnictwyoup.pl>

druk i oprawa Zespół Poligraficzny UP

Spis treści

Paolo Bussotti, <i>Leonardo Pisano called Fibonacci, between advanced mathematics, history of mathematics and mathematics education: three examples drawn from Liber Quadratorum</i>	5
Jan Górowski, Adam Łomnicki, <i>Iteracje funkcji homograficznej i równanie rekurencyjne zadane funkcją homograficzną</i>	27
Jan Górowski, Adam Łomnicki, <i>Quasi-arithmetic means</i>	35
Agata Hoffmann, <i>Wykorzystanie prac Stanisława Dróżdża jako środków dydaktycznych w nauczaniu</i>	45
Taras Kudryk, <i>Introduction to E. Nelson's internal set theory</i>	59
Bożena Olik-Pawlik, <i>O zjawisku zniekształcania obrazu matematyki</i>	79
Katarzyna Słomczyńska, <i>High School Identities</i>	91
Damian Wiśniewski, <i>Wokół liczb i szeregów harmoniczych</i>	99
Lidia Zaręba, <i>The role of visualisation in the process of generalisation</i>	111

Historia matematyki, Konferencje

Piotr Błaszczyk, <i>Nota o Lehrbuch der Algebra. Einleitung Heinricha Webers</i>	129
Heinrich Weber, <i>Podręcznik algebry. Wprowadzenie</i>	141
Piotr Błaszczyk, <i>XXX Konferencja z Historii Matematyki, Będlewo, 5–8 V 2016</i>	155

Contents

Paolo Bussotti, <i>Leonardo Pisano called Fibonacci, between advanced mathematics, history of mathematics and mathematics education: three examples drawn from Liber Quadratorum</i>	5
Jan Górowski, Adam Łomnicki, <i>Iterations of homographic functions and recurrence equations involving a homographic function</i>	27
Jan Górowski, Adam Łomnicki, <i>Quasi-arithmetic means</i>	35
Agata Hoffmann, <i>Stanisław Dróżdź's artworks as teaching aid in mathematics education</i>	45
Taras Kudryk, <i>Introduction to E. Nelson's internal set theory</i>	59
Bożena Olik-Pawlik, <i>On distorting the image of mathematics</i>	79
Katarzyna Słomczyńska, <i>High School Identities</i>	91
Damian Wiśniewski, <i>Around harmonic numbers and harmonic series</i>	99
Lidia Zaręba, <i>The role of visualisation in the process of generalisation</i>	111

History of Mathematics, Conferences

Piotr Błaszczyk, <i>Note on Heinrich Weber's Lehrbuch der Algebra. Einleitung</i>	129
Heinrich Weber, <i>Textbook of Algebra. Introduction</i>	141
Piotr Błaszczyk, <i>XXX Conference on History of Mathematics, Będlewo, 5–8 V 2016</i>	155

Paolo Bussotti

Leonardo Pisano called Fibonacci, between advanced mathematics, history of mathematics and mathematics education: three examples drawn from *Liber Quadratorum**

Abstract. In this research, three problems faced by Leonardo Pisano, called Fibonacci, are dealt with. The main purpose is to show that history of mathematics can offer interesting material for mathematics education. The approach to the use of history of mathematics in mathematics education cannot be merely historical, but adapted to the needs of the explanations in a classroom. In the course of this paper, the meaning of such assertion will be clarified. A further purpose is a trying to explore the relations history of mathematics-mathematics education-advanced mathematics. Last, but not the least, a further aim is to offer a specific series of interesting material to the teacher in order to develop stimulating lessons. The material here expounded is conceived for pupils frequenting the third and the fourth years of the high school, but, with minor modifications, it can be adapted to the fifth year of the high school and to the initial years of the scientific faculties at the university.

1. Introduction

Federigo Enriques (1871-1946) claimed on several occasions there is not a clear distinction between advanced mathematical research, history of mathematics and mathematics education.¹ Given the extreme specialization of advanced mathematics (which existed already in Enriques' epoch) and since history of mathematics and mathematics education have also become specialized fields of research, Enriques' assertion ought seem paradoxical or wrong, at all. In contrast to this, the

*Od matematyki wyższej, przez historię do dydaktyki. Trzy przykłady zaczerpnięte z *Liber Quadratorum* Leonarda z Pizy zwanego Fibonaccim

2010 Mathematics Subject Classification: Primary: 97 Mathematics Education; Secondary: 01 History and Biography

Key words and phrases: Fibonacci, Enriques, Liber Quadratorum, Relation history of mathematics-mathematics education, Paradigmatic examples, Medieval mathematics, Number theory and mathematics education, Diophantine analysis, Lagrange, Genocchi

¹ Enriques, 1921, 2003, p.12-14.

ideas expounded by Enriques hold a profound meaning and express a part of the truth, though, probably, not the whole truth.

I will refer to a specific example, drawn from Fibonacci's (1170-1245?) *Liber Quadratorum* (1225), in which a conspicuous series of ideas to catch the relations between advanced mathematics, history of mathematics and mathematics education exists. Since the main approach of this paper is mathematics-education centred and, after all, a possible concrete use of Fibonacci's technique in an educational context will be shown, the references to history of mathematics will be limited to what is strictly necessary to prove the plausibility of Enriques' thesis. Thence, this research offers a double reading-key:

- 1) A support to Enriques' conception, according to which – at least in some circumstances – mathematical research, mathematics education and history of mathematics can be tied by a thorough link;
- 2) The possible use of some techniques and methods used along history of mathematics in an education context. In particular the content of this paper can be presented to the pupils when they deal with the principle of mathematical induction, this means, in general, when they frequent the third-fourth year at the high school.

I have no claim to provide a general theory of the relations history of mathematics-mathematics education since, as I have explained in other works², I am convinced that such a general theory does not exist, but profound connections do.

2. References to the figure of Fibonacci and to his works

The figure of Leonardo Pisano, called Fibonacci, is so well known that here I will only refer to some of his main features, perhaps useful for the reader not involved in history of Medieval mathematics³: Leonardo was the son of Guglielmo Bonacci, who was *publicus scriba* (namely a delegate) of Pisa maritime republic abroad. In 1185 Guglielmo was transferred to Bugia (nowadays in Algeria) as a responsible of Pisa's commercial business in that land. The young Leonardo followed his father in Bugia. In the *Incipit* of his famous *Liber abaci* (1202, 1228)⁴ Fibonacci himself writes to have lived in Bugia and to have travelled to Egypt, Syria,

² Bussotti, 2014; Bussotti, 2015.

³ I mention here some significant works on Fibonacci without any claim to be exhaustive (the literature on Fibonacci is huge): Boncompagni, 1852; Boncompagni, 1854; Bussotti, 2008a, p. 234-249; Bussotti, 2008b, p. 43-61; Favaro, 1874; Franci, 2002; Genocchi, 1855a; Genocchi, 1855b, Genocchi, 1855c, Genocchi, 1857; Grimm, 1973; Horadam, 1991; Lüneburg, 1991; Picutti, 1979; Picutti, 1983; Pisano, Bussotti, 2015; Rashed, 1994; Rashed, 2003; Woepcke, 1854-55; Woepcke, 1860-61.

⁴ Fibonacci, 1228, 1857, p. 1. The works by Fibonacci have been published by Baldassare Boncompagni. See Fibonacci, 1228, 1857 (*Liber Abaci*); Fibonacci, 1862a, p. 227-252 (*Flos*); Fibonacci, 1862b, p. 253-283 (*Liber Quadratorum*); Fibonacci, 1862c, p. 1-224 (*Practica Geometriae*). The *Liber Abaci* has been translated into English by L.E. Sigler (Sigler, 2002); The *Liber Quadratorum* has been translated into French by P. Ver Eecke (Fibonacci, 1952) and into English by L. E. Sigler (Fibonacci, 1988). The *Practica Geometriae* has been edited and translated into English with great care by B. Hughes (Hughes, 2008).

Sicily and Byzantium to improve his knowledge of mathematics. Thence Leonardo reached a – presumably – good knowledge of Arabic mathematics. Fibonacci came in contact with the Hindu-Arabic digits and decimal positional system, imported into the Arabic countries from India, as well as with Arabic algebra and Diofantine tradition, which was alive and flourishing in the Arabic world. Fibonacci wrote a series of important works, in which his ingenious originality, the knowledge of Euclid's *Elements* and the acquisitions deriving from the Arabic mathematics converge to create a quite stimulating picture:

- A) The *Liber abaci* is surely the most famous of Fibonacci's works. It is a huge book divided into fifteen chapters, where Fibonacci introduces the calculations with the Hindu-Arabic digits, namely the four arithmetical operations, the calculations with fractions, the way to extract square and cubic roots and a numerous series of problems – in general concerning – transactions, exchange of moneys, business societies among partners and so on. Usually, but not always, these problems can be solved by system of equations of first degree. Other problems concern numbers. In the 12th chapter of the *Liber abaci*, the well known rabbit-problem is dealt with, where the Fibonacci-series is introduced. The *Liber* was a monumental work and impressed Fibonacci's contemporaries⁵. Although the introduction of the Hindu Arabic digits is documented in the Western world before Fibonacci⁶, there is no doubt that the previous works on this subject were not comparable with Fibonacci's as to completeness, competency and clearness. The *Liber* can be considered an advanced handbook, with several ideas and problems also belonging to the then high mathematics.
- B) Another significant work is the *Flos*, written around 1224-1225, in which Leonardo, answering to a problem proposed to him by the court mathematician of the Emperor Friedrich the second of Swabia, Giovanni da Palermo, showed that the equation

$$x^3 + 2x^2 + 10x = 20$$

does not have any Euclidean irrational as a solution. Here Fibonacci develops a series of arguments, which, nowadays, could be transformed into interesting didactical consideration on the limits of the use of some instruments to solve certain mathematical problems. In the specific case, the use of rule and compass, which could be connected to the three classical Greek geometrical problems: duplication of the cube, squaring of the circle, trisection of the angle and to modern algebraic approach to these questions.

Around 1220-1221 Leonardo also wrote a *Pratica Geometriae*. Though this book has some interesting aspects, it is, in my perspective, less interesting than the other works. Two further contributions by Leonardo: *the libro di*

⁵ For the references to this fascinating and difficult historiographical problem, I refer to Pisano, Bussotti, 2015. In that paper we also offer precise indications on the state of art regarding this subject.

⁶ See previous note.

merchaanti o detto di minor guisa and the *Libro sopra il 10° di Euclide* are lost.

- C) However, the most surprising and profound work by Fibonacci is the *Liber Quadratorum*. The origin of this contribution is due to a further problem posed by Giovanni da Palermo to Leonardo: to solve in whole or rational numbers the system of two equations:

$$\begin{cases} x^2 + 5 = y^2 \\ x^2 - 5 = z^2 \end{cases}$$

As a matter of fact, Leonardo proposes an entire and coherent theory - developed in the twenty propositions of the *Liber* - concerning a sequence of interrelated problems with the squares. The most interesting aspect of the *Liber* is probably that methodological: Leonardo used an approach, which is in part influenced by Euclid's geometrical tradition, in part by Diophantus' and Arabic tradition, but which is, in great part, original. As we will see, it is based on the idea to use the sums of consecutive integer numbers to solve the problems connected to the squares. Since my aim is to exploit Leonardo's method in an education perspective I refer to the literature for both the mathematical and genetic-historical problems connected to the *Liber Quadratorum*⁷. I will focus on three propositions. Each section will include the explanation of Leonardo's approach and - in a separate box - the didactical considerations. The section 5 will also present a box regarding the relations history of mathematics - mathematics education - advanced mathematics.

It is necessary to point out that, as to the figure of Leonardo, there are several questions: a) what exactly were his sources of inspirations; b) why did he succeed in introducing the Hindu-Arabic numerals in Western Europe; c) what is his heritage. In particular: how much did he directly influence the Western mathematics and mathematics education - specifically the abacus schools and the abacus treatises - in the late Middle Age. On these fascinating historical problems, I have to refer to the literature⁸, because, as already pointed out, my aim in this paper is not purely historical.

3. Squares obtained as sum of the successive odd numbers

The first proposition of the *Liber Quadratum* is a rather dense introduction to the topics dealt with in the course of the book. In the second proposition Leonardo writes:

I want to prove why the ordered series of the squares arises from the ordered collection of the odd numbers starting from 1 to the infinity.⁹

⁷ As to the *Liber Quadratorum* see specifically: Bussotti, 2008a, p. 234-249; Bussotti, 2008b, p. 43-61; Genocchi, 1855b; Genocchi, 1855c; Picutti, 1979; Rashed, 1994; Rashed, 2003; Woepcke, 1854-55. Useful references in Fibonacci, 1952 and Fibonacci, 1988.

⁸ The literature on this subject is huge. Detailed references exist in Pisano, Bussotti, 2013 and Pisano, Bussotti, 2015.

⁹ Fibonacci, 1225, 1854, p. 255. Original Latin text: "Volo demonstrare quare ex ordinata imparium collectione, ad uno incipiendo ad infinitum egrediatur ordinata series quadratorum".

Fibonacci does not claim the priority of this discovery. But – as far as we know – his way of reasoning is original. He argues like this

$$\begin{array}{ccccccccc} A & B & G^T & D^K & E^L & Z^M & I^N \\ | & | & | & | & | & | & | \\ \hline \end{array}$$

Ryc. 1.

The segments AB , BG , GD , ... represent the successive numbers, so that $AB = 1$; $BG = 2$; $GD = 3$, and so on. The exponents T , K , L , ... have a different meaning as they indicate sums: each one indicates the sum of the two previous numbers. So that T is the sum of AB and BG , thence $T = 3$; K is the sum of BG and GD , so that $K = 5$, and so on. Therefore, the exponents indicate the series of the odd numbers. Each odd number has a double representation in this figure by Fibonacci. For example, 3 is represented by GD and by T ; 5 is represented by EZ and by K , and so on. Fibonacci reasons as follows:

1. Given any number ZI , the relation

$$ZI^2 = EZ^2 + EZ + ZI \quad (1)$$

holds, being $EZ + ZI = N$. In modern terms this simply means that $(n + 1)^2 = n^2 + n + (n + 1)$. Relying upon the formula for the square of a binomial proved by Euclid in Elements, II, 4, Fibonacci had proved identity (1) in the first proposition of the *Liber Quadratorum*.

2. Applying the same formula as (1), one obtains $EZ^2 = DE^2 + DE + EZ$, where $DE + EZ = M$. This formula can be applied till reaching $BG^2 = AB^2 + AB + BG = 1 + T = 1 + 3$.
3. Therefore $GD^2 = BG^2 + BG + GD = 1 + T + K = 1 + 3 + 5$, because $BG^2 = 1 + T$ and $BG + GD = K$. It is clear that this procedure can be iterated obtaining each square as the sum of the numbers $1 + T + K + L + M + N + \dots$, namely the successive odd numbers. This proves the proposition.

Didactical considerations.

A) *The methodology.* From an educational point of view the logical and methodological considerations are important for the pupils to catch that mathematics is not only a series of results obtained in a mechanical way once and forever. It is also a set of different methods and way of reasoning, by means of which the men reached the results, often with a conspicuous work and labour. The proof of this proposition by Fibonacci is easy, but it is paradigmatic of a procedure he also applied to more complicated problems. The logic of this method has to be clearly explained to the pupils. For the explanation, let us symbolize like this:

Proposition q="every square is the sum of an initial segment of odd numbers" (this is the proposition to prove).

Proposition p=" $(n + 1)^2 = n^2 + n + (n + 1)$ ".

By the application of proposition p to any number (we indicate it by $n + 1$, Leonardo by ZI), it is possible to *descend* to 2^2 . But for 2^2 , proposition q holds, too, because $2^2 = 1 + 3$. Then the *ascending* phase begins and, by the replacements showed in the item 3., the proposition q is proved. Thence, Leonardo's methodology

has a *descending* and an *ascending* phase.

Let us now propose to the students a modern demonstration by mathematical induction:

Basis of the induction: $2^2 = 1 + 3$

Inductive step: let us suppose that $n^2 = 1 + 3 + \dots + (2n - 1)$, then

$$(n + 1)^2 = 1 + 3 + \dots + (2n - 1) + (2n + 1)$$

which proves the proposition.

The similarities between the two proofs are evident and the teacher has to point out: i) in the modern inductive proof, there is the basis $2^2 = 1 + 3$, identity, which, as seen, is fundamental in Fibonacci's argument, too; ii) the inductive step in Fibonacci's demonstration can be directly obtained by applying the proposition p and supposing that proposition q is valid until a number n . Therefore the two procedures are perfectly equivalent from a logical point of view. Then the question: why did Fibonacci add the *descending* phase, which is useless in a merely mathematical perspective? The answer to this question is very instructive because it points out the human aspect of mathematics. The main consideration concerns exactly the inductive step: in Leonardo's reasoning the function of our inductive step is replaced, in fact, by the series of descending identities

$$\left\{ \begin{array}{l} ZI^2 = EZ^2 + EZ + ZI \\ EZ^2 = DE^2 + DE + EZ \\ \dots \\ BG^2 = AB^2 + AB + BG \end{array} \right.$$

The ascending phase consists in a mere replacement of values, according to what already seen, starting from the identity $BG^2 = AB^2 + AB + BG = 1 + T$, which is the basis of the induction.

Once explained the logic of Fibonacci's method, let us consider the reason why he presented such a demonstration: in the epoch in which Leonardo wrote, a set of standard methods – as mathematical induction – applicable to number theory had not yet been created. It should be clarified to the pupils that the formalization of the inductive basis and inductive step needs a certain level of symbolization and formalization:

Symbolization, because – in the case of our proof by mathematical induction, we have written the last addend of the sum, which constitutes the square of n , as $2n + 1$. This implies the capability to use a letter as n to indicate a number *in general* and the idea to write a number in function of a modulus, which is 2 in the case of $2n + 1$. These skills were acquired only in the 17th century with mathematicians as Pierre de Fermat (1601-1665) and Blaise Pascal (1623-1662), although – as we will see in section 5.1.- Fibonacci seems, in some circumstances, to have arrived close to the concept of modulus. It is not a case that mathematical induction in modern terms is present in the *Triangle Arithmetique* by Pascal. For, even if before Pascal some methods similar to mathematical induction were conceived and some scholars date back this method to a period preceding Pascal, I am convinced that the founder of modern mathematical induction was Pascal¹⁰. Fermat developed the method of infinite descent, which has similarities – but also profound differences – with mathematical induction¹¹.

Formalization, which consists in operating on the numerical forms according to the rules and formulas of our algebra. But, in Fibonacci's period, despite the conspicuous progresses made by the Arabic mathematicians, the formal calculus – we could say the literal calculus – almost did not exist. The numbers were, in general, associated to segments of straight lines or anyway, at least in the Western world, the geometrical representation played still a pivotal role. Fibonacci made considerable progresses in this sense because, for example, the idea to represent the number in two different ways within a sole proof is rather audacious: in the analysed case, i.e., 5 is represented by K , but also by $BG+GD$, according to the necessity. Furthermore, the geometrical representation is a sort of psychological-didactical support rather than a necessity intrinsic to the proof. This means that Fibonacci probably had something similar to the concept of *any number*, but his expressive means allowed him to obtain a certain level of formalization, by which it was difficult to express and conceive the necessary formal steps typical of mathematical induction. Notwithstanding, he conceived a method, which resembles mathematical induction and for which, his level of symbolization and formalization was enough. This is a very interesting lesson for the pupils because they can understand, by means of an easy example, the profound sense in which mathematics is a historical and human activity tied to the scientific context of the period: Fibonacci's method is redundant only in our eyes. In fact, it is extremely innovative.

B) *Relations with infinity*. The following considerations could be developed jointly with the teacher of philosophical and historical disciplines: the descending phase of Leonardo's proof could also be interpreted as a reluctance to resort to infinity, also in its potential form, or at least as the intention to limit as much as possible the resort to infinity. Indeed, this phase – which, in Leonardo's way of reasoning, not in ours, is quite important – operates between two finite limits: the initial number ZI , which is arbitrary, but finite and given in the specific context, and the number $BG = 2$. This scheme is completely finitary in the sense that it involves a finite segment of integer numbers in descent, from ZI to 2. One of the reasons why Fibonacci tried to tie the ascending phase, which involves the whole potentially infinite series of the integers, to the descending one, is also probably the attempt to create a logical dependence of a potentially infinite ascent from a finite descent. For, it is clear that Fibonacci had an almost dynamical vision of his procedure; a sort of movement, which, starting from an initial point (ZI) reached a minimum (2) after which the ascending phase began. Although the modern mathematical induction is not seen by us as a dynamical procedure, it is a scheme in ascent, in which the potential infinity is directly involved and there is no attempt to create a dependence of its on a finitary scheme as Fibonacci's descent. According to my interpretation, this is a proof of how the mental and psychological presuppositions can also play a role in the way in which mathematics is created and developed. The discussion on the concept of infinity in the middle ages as well as on the relations between science-mathematics and mentality of a certain epoch, are typically interdisciplinary fields, which – at least on some occasions – could be shared by the teachers of mathematics and of history or philosophy.

Now I will propose an example of a more complex proof carried out by Leonardo's method to show that its range of application is wider one could think of.

¹⁰ As to Pascal and mathematical induction see Palladino, Bussotti, 2002; Bussotti, 2004.

¹¹ A complete work on the infinite descent, where Fermat plays a fundamental role, is Bussotti, 2006.

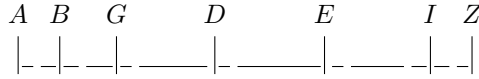
4. Leonardo's proof of the identity $(2n - 1)(2n + 1)[(2n - 1) + (2n + 1)] = 12[1^2 + 3^2 + 5^2 + \dots + (2n - 3)^2 + (2n - 1)^2]$

The 7th and 8th proposition of the *Liber Quadratorum* proves two interesting propositions by means of a technique similar to the one expounded in the previous section. In the 7th proposition Leonardo proves that

$$n(n + 1)[n + (n + 1)] = 6(1^2 + 3^2 + 5^2 + \dots + n^2),$$

while in the 8th one, he demonstrates the identity referred to in the title of this section. I will focus on proposition 8.

Fibonacci considers the series of the odd numbers. He represents them in a line so that $AB = 1$; $BG = 3$; $GD = 5$; $DE = 7$; $EZ = 9, \dots$



Ryc. 2.

Since, $DZ = DE + EZ$, Leonardo claims he has to prove that

$$DE \cdot EZ \cdot DZ = 12(AB^2 + BG^2 + GD^2 + DE^2), \quad (2)$$

which, posing $DE = 2n - 1$ and $EZ = 2n + 1$, is exactly our proposition.

Leonardo considers $DE = EI$, so that $IZ = 2$. Since the successive odd numbers are indicated in the figure, it is $GD = DE - 2$ and $GE = 2DE - 2$. Hence $GD \cdot DE = DE^2 - 2DE$. Therefore

$$CD \cdot DE \cdot GE = (DE^2 - 2DE)(2DE - 2) = 2DE^3 - 6DE^2 + 4DE \quad (3)$$

Now Fibonacci adds $12DE^2$ to the value in (3), thus obtaining

$$GD \cdot DE \cdot GE + 12DE^2 = 2DE^3 + 6DE^2 + 4DE \quad (4)$$

But $DE = EI$, thence $EZ = DE + IZ = DE + 2$ and $DZ = 2DE + 2$. This means that $DE \cdot EZ = DE^2 + 2DE$. Therefore,

$$DE \cdot EZ \cdot DZ = 2DE^3 + 6DE^2 + 4DE \quad (5)$$

That is, from (4) and (5)

$$GD \cdot DE \cdot GE + 12DE^2 = DE \cdot EZ \cdot DZ \quad (6)$$

By the same reasoning, Fibonacci constructs a descent. For, by analogous steps as those used to prove identity (6), it is possible to show that

$$BG \cdot GD \cdot BD + 12GD^2 = GD \cdot DE \cdot GE \quad (7)$$

Thus, from (6) and (7)

$$BG \cdot GD \cdot BD + 12(GD^2 + DE^2) = DE \cdot EZ \cdot DZ \quad (8)$$

By the same reasoning, it will be shown that

$$BG \cdot GD \cdot BD = AB \cdot BG \cdot AG + 12BG^2 \quad (9)$$

So that, from (8) and (9), one gets

$$AB \cdot BG \cdot AG + 12(BG^2 + GD^2 + DE^2) = DE \cdot EZ \cdot DZ \quad (10)$$

However, Fibonacci continues, since $AB = 1$, $BG = 3$ and $AG = 4$, the product $AB \cdot BG \cdot AG = 12$, or $12AB$, if one wishes to indicate the unit, as Fibonacci, by AB . Thence

$$12(AB^2 + BG^2 + GD^2 + DE^2) = DE \cdot EZ \cdot DZ \quad (11)$$

Being $AB, BG, GD, \dots$ the successive odd numbers, this proves the theorem. Obviously the reasoning works whatever the numbers DE, EZ and DZ are. Thence it is true in general.

Fibonacci adds some interesting corollaries on which I do not focus. I only add that by propositions 7 and 8, it is possible to deduce the famous formula by Archimedes on the sum of squares¹². For, from proposition 6, it follows immediately

$$1^2 + 3^2 + 5^2 + \dots + n^2 = \frac{n(n+1)(2n+1)}{6}. \quad (12)$$

Didactical considerations. The same considerations as those presented in the previous section can be developed. It is important to point out that, in this case, the particular method of Leonardo, consisting of a descent followed by an ascent, is applied to prove a far more difficult theorem, this means it is not only of historical interest, but it is a performative method, too. The teacher should point out that, in this case, even more than in the theorem presented in the previous section, the decisive inductive step is represented by the formula, which allows Fibonacci to reach the identity (10). His inductive step is in descent, whereas for us, it is in ascent. The basis of the induction, which permits the prove of the theorem - namely the identity $AB \cdot BG \cdot AG = 12$, represents, hence, the last step in Fibonacci's methods, while, in the modern mathematical induction, it is the first one.

5. The equation $\frac{x^2-y^2}{y^2-z^2} = \frac{a}{b}$ ($a > b$)

In this equation - which constitutes the 18th proposition of the *Liber Quadratorum* - every letter indicates an integer. I have translated the language of Fibonacci into the modern symbolism because, in what follows, I will not focus on the problems of symbolism and, hence, the transcription in modern symbols makes the explanation easier.

From a mathematical standpoint, this equation represents a far more difficult problem than the identities analysed in the sections 3. and 4. A complete theory,

¹² This result is obtained by Archimedes in the 10th proposition of the *Spirals*.

inside which it was solved, was discovered by Lagrange (1737-1813)¹³. Nonetheless, the considerations developed by Fibonacci are elementary – though rather refined – and they are particularly suitable to catch: a) his technique of the ”successive addends”; b) the relations history of mathematics-mathematics education-advanced mathematics.

Before entering into Fibonacci’s reasoning, it is necessary to remind the reader two easy statements he proved in the 17th proposition of the *Liber*:

- 1) Given two consecutive odd numbers $(2n + 1)$ and $(2n + 3)$, it holds

$$(2n + 3)^2 = (2n + 1)^2 + 8(n + 1) \quad (13)$$

- 2) Given two consecutive even numbers $2n$ and $(2n + 2)$, it holds

$$(2n + 2)^2 = (2n)^2 + 4(2n + 1) \quad (14)$$

5.1. The equation $\frac{x^2 - y^2}{y^2 - z^2} = \frac{a}{b}$ the three elementary cases

Fibonacci distinguishes his solution in cases.

The *first* analysed case takes places when $a = b + 1$ (in what follows it is supposed $a > b$, but the reasoning is the same in the opposite hypothesis). The solution is easy and can be obtained for every a . Fibonacci offers an example, which, however, is *paradigmatic* of a general demonstration. He considers $a = 11$; $b = 10$. Referring to proposition (13), it is

$$21^2 = 19^2 + 10 \cdot 8; \quad 23^2 = 21^2 + 11 \cdot 8$$

Thence

$$\frac{23^2 - 21^2}{21^2 - 19^2} = \frac{21^2 + 11 \cdot 8 - 21^2}{19^2 + 10 \cdot 8 - 19^2} = \frac{11}{10}.$$

Therefore $x = 23$, $y = 21$, $z = 19$ are the solutions.

To get a general proof, it is enough to pose $a = n + 1$; $b = n$. The solutions are: $x = 2n + 3$; $y = 2n + 1$; $z = 2n - 1$, because, according to proposition (13), it is

$$\frac{(2n + 3)^2 - (2n + 1)^2}{(2n + 1)^2 - (2n - 1)^2} = \frac{(2n + 1)^2 + 8(n + 1) - (2n + 1)^2}{(2n - 1)^2 + 8n - (2n - 1)^2} = \frac{n + 1}{n} = \frac{a}{b}$$

The *second* analysed case concerns the condition $a = b + 2$.

Once again, Fibonacci provides a paradigmatic example. He considers $a = 11$; $b = 13$. He adds the two numbers, obtaining 24, divides 24 by 4 and multiplies the result by 2, thus obtaining 12 (in practice he divides the sum by 2). The number 12 is the ”central” square, that is y . The other two numbers are $x = 14$, $z = 10$. For, it is

¹³ Lagrange’s fundamental work on this subject is Lagrange, 1769. For detailed explanations of Lagrange’s method, see Bussotti, 2006, p. 299-341.

$$\frac{14^2 - 12^2}{12^2 - 10^2} = \frac{12^2 + 4 \cdot 13 - 12^2}{10^2 + 4 \cdot 11 - 10^2} = \frac{13}{11}.$$

The general demonstration of Fibonacci's method relies upon the identity (14), because, posed $a = 2n + 1$; $b = 2n - 1$, obviously their sum divided by 2 is $2n$. One has

$$\frac{(2n+2)^2 - (2n)^2}{(2n)^2 - (2n-2)^2} = \frac{(2n)^2 + 4(2n+1) - (2n)^2}{(2n)^2 + 4(2n-1) - (2n-2)^2} = \frac{2n+1}{2n-1} = \frac{a}{b}.$$

What follows is very interest and shows the profoundness of Fibonacci's thought: Leonardo poses this problem, let us suppose that $a = 21$, $b = 19$. Now he requires three specific squares who solve our equation, with the further condition their bases are numbers of the form $3n$. He offers the following canon: sum the two numbers, $19 + 21 = 40$, divide (as in the previous case) by 4, obtaining 10, multiply 10 by 3, obtaining 30. This is the central square, the other two being 33 and 27. Indeed, it is

$$\frac{33^2 - 30^2}{30^2 - 27^2} = \frac{21}{19}$$

How did he reach this amazing result? Probably Fibonacci had studied, for every modulus, the differences of two consecutive squares (this is the circumstance in which Leonardo seems to have caught important aspects of the concept of modulus). Thence mod. $3n$, we have

$$(3n+3)^2 - (3n)^2 = 9(2n+1) \quad (15)$$

Therefore, let us pose $a = 2n + 1$; $b = 2n - 1$, so that $a + b = 4n$, number which, divided by 4, produces n . Let us now considers $x = 3n + 3$; $y = 3n$; $z = 3n - 3$. It will be:

$$\frac{(3n+3)^2 - (3n)^2}{(3n)^2 - (3n-3)^2} = \frac{(3n)^2 + 9(2n+1) - (3n)^2}{(3n)^2 + 9(2n-1) - (3n-3)^2} = \frac{2n+1}{2n-1} = \frac{a}{b}.$$

As Fibonacci points out, this procedure can be applied to every modulus, $4n$, $5n$, $6n$, and so on. For example mod. $5n$, the difference between two consecutive numbers $5n + 5$ and $5n$ is, obviously, $25(2n+1)^{14}$. It will be, hence, necessary to divide $a + b$ by 4 and multiply the result by 5, thus obtaining the central square, whose root is $5n$, the other two roots will be $5n + 5$ and $5n - 5$. For example, if $a = 21$, $b = 19$, the three roots are $x = 55$, $y = 50$, $z = 45$, thus obtaining

$$\frac{55^2 - 50^2}{50^2 - 45^2} = \frac{21}{19}.$$

The *third* and last elementary case takes place when both a and b are squares. Let us pose $a = x_1^2$; $b = y_1^2$. The three solutions are then given by $x = x_1^2$; $y = x_1 y_1$;

¹⁴ In general, such a difference, for any modulus kn is $k^2(2n+1)$, which is an extension of the identities (13) and (14).

$z = y_1^2$. This is the least interesting case, whose proof derives immediately by applying the properties of the proportions:

$$\begin{aligned} x_1^4 : x_1^2 y_1^2 &= x_1^2 y_1^2 : y_1^4 \\ (x_1^4 - x_1^2 y_1^2) : x_1^4 &= (x_1^2 y_1^2 - y_1^4) : x_1^2 y_1^2 \\ (x_1^4 - x_1^2 y_1^2) : (x_1^2 y_1^2 - y_1^4) &= x_1^4 : x_1^2 y_1^2 = x_1^2 y_1^2 : y_1^4 = a : b. \end{aligned}$$

Fibonacci offers the example of $a = 25$, $b = 16$. In this case, it will be $x = 25$, $y = 20$, $z = 16$.

Didactical considerations. What dealt with in this section induces several observations.

A) *Relation between proofs and numerical examples.* In order to prove his statements Fibonacci uses specific numerical values and not general symbols, hence one could thought that his proofs are not valid, but are limited to those specific cases analysed by him. In fact, this is not true because his examples are *paradigmatic*. Through this expression, I indicate an example in which the general numerical forms can be replaced to the specific numbers because all the numbers have a precise form, which can be identified (being this form $2n + 1$ or $3n$, or $5n$, or whatever it is), thus obtaining a rigorous proof. All the demonstrations examined in the present section have this feature. Therefore, Leonardo's arguments are valid in general, though apparently based on examples. I think, it is a common experience, shared by the teachers of mathematics, to observe that – especially while dealing with theorems concerning integers – that several – not all – pupils do not exactly catch the fundamental difference between a demonstration and a proposition verified for all the examined examples, but not proved. After having explained Fibonacci's argumentations, in order to clarify this difference, a teacher could consider the famous theorem by Fermat that every prime of the form $4n + 1$ is the sum of two squares. He/she could begin to decompose these primes beginning with $5 = 1^2 + 2^2$; $13 = 2^2 + 3^2$; $17 = 1^2 + 4^2$; $29 = 2^2 + 5^2$ It could be comprehensible that the pupils asked if, in this case, a replacement of the single numbers by a general form is possible, so to get a demonstration of the statement. The teacher should then explain this is not possible: the examples of Fermat's theorem are not *paradigmatic*: it is not possible – at least as far as we know until now – to deduce from them a general proof of this theorem, which, in fact, is independent from the way in which such primes are decomposed into two squares¹⁵. Therefore, a clear explanation of Fibonacci's technique jointly with the replacement of the general forms to the specific numbers used by Leonardo: i) will make it clear his reasoning works in general; ii) will be a support to clarify the difference between a paradigmatic and a not-paradigmatic example; iii) will be a support to distinguish between a non-paradigmatic example and a rigorous proof. All these aspects are fundamental for a pupil to catch the true logical foundations of mathematics as well as the nature of this discipline. History of mathematics is, in this case, a precious help for mathematics education.

B) *Solutions of undetermined equations.* Fibonacci provides one or more solutions to his equations. From a didactical point of view an interesting question arises: did he supply all the possible solutions? The answer is negative: let us consider once again the case $a = 13$; $b = 11$. Let us divide their sum by 4 and do not multiply

by 2. We get the number 6, which offers the solution $x = 7, y = 6, z = 5$, which does not exist in the *Liber Quadratorum*. Given $\frac{a}{b} = \frac{2n+1}{2n-1}$, the least y who solves the equation is $y = n$, with $x = n + 1; z = n - 1$ because $\frac{(n+1)^2 - n^2}{n^2 - (n-1)^2} = \frac{2n+1}{2n-1} = \frac{a}{b}$. This number is simply obtained dividing $a + b$ by 4, thence completely remaining within Fibonacci's methodology. These considerations have to induce the teacher to clearly specify that, given an undetermined equation and, in general a mathematical problem, there are four different questions: i) to proof if the equation has a solution; ii) to show a solution of the equation; iii) to show possible other solutions of the equation by an algorithm; iiiii) to show that the found solutions are all the solutions. These problems *ought* be connected: for example, it is clear that if one finds a solution to an equation, the question i) is automatically solved, but they are not *necessarily* connected. In the case of our equation, i. e., with $a = 2n + 1; b = 2n - 1$, Fibonacci's technique offers an infinite number of solutions, but we cannot be sure they are all the solutions. Others could exist, which are not included into Fibonacci's algorithm. The difficulties involved in the answers to the questions i)-iii) are, in general, quite different because by means of ingenious devices, which, in some circumstances, arrive at resembling a theory, it is possible to offer a solution or also an infinity of solutions. However, as I will explain more clearly in the didactical considerations added to the next section, the answer to the problem iii) for this equation is connected to the general solution of all the undetermined equations of second degree with two unknowns, problems which goes far beyond the possibilities of a mathematician living in the 12th – 13th century and far beyond the programme of mathematics in a high school. This is the connection of which I was speaking with regard to history of mathematics-mathematics education-advanced mathematics: a problem, as that of Fibonacci, starts from history, allows us a series of didactical considerations useful at high school level and is connected with an aspect of advanced mathematics, whose complete solution was offered only in the late 18th century. To follow this itinerary – though it can comprehend only a part of the mathematical details – is quite instructive for the high schools' pupils.

C) *To play with numbers*. Apart from more general considerations, the algorithm by Fibonacci is useful for the pupils to learn to deal with numbers and numerical forms. For example: given, in our equation $a = 2n + 1, b = 2n - 1$, the teacher, after having explained Fibonacci's reasoning, could ask the pupils to find the numbers of various forms: $6n, 7n, 8n, \dots$ which solve the equation. This is a useful exercise because the ability to manipulate the numbers – which is so important in mathematics – is progressively disappearing. Furthermore, to play with numbers inside an algorithm might also be enjoyable for the pupils.

5.2. The equation $\frac{x^2 - y^2}{y^2 - z^2} = \frac{a}{b}$ the general solution

To deal with the general solutions, Fibonacci offers a series of arguments, which can be summarized in the following reasoning: if it is possible to decompose $a + b$ or a multiple of this sum $k(a + b)$ into a series of $m + n$ consecutive addends, such that the first m represent kb and the other n the same multiple ka of a , then the problem is solved. For, let us consider the series of successive addends

$$h, h + 1, h + 2, \dots, h + m, h + m + 1, h + m + 2, \dots, h + m + n \quad (16)$$

¹⁵ For detailed explanations on the $4n + 1$ primes-theorem, see Bussotti, 2006.

with

$$h + (h + 1) + (h + 2) + \dots + (h + m) = kb$$

$$h + m + 1) + (h + m + 2) + \dots + (h + m + n) = ka$$

Then the three roots are $z = 2h - 1$, $y = 2(h + m) + 1$; $x = 2(h + m + n) + 1$ because:

$$\begin{aligned} (2h + 1)^2 &= (2h - 1)^2 + 8h \\ [2(h + m + 1) + 1]^2 &= [2(h + m) + 1]^2 + 8(h + m + 1) \\ (2h + 3)^2 &= (2h + 1)^2 + 8(h + 1) \\ [2(h + m + 2) + 1]^2 &= [2(h + m + 1) + 1]^2 + 8(h + m + 2) \\ &\dots\dots\dots \\ [2(h + m) + 1]^2 &= [2(h + m) - 1]^2 + 8(h + m) \\ [2(h + m + n) + 1]^2 &= [2(h + m + n - 1) + 1]^2 + 8(h + m + n) \\ [2(h + m) + 1]^2 &= (2h - 1)^2 + 8(h + h + 1 + \dots + h + m) \\ [2(h + m + n) + 1]^2 &= [2(h + m) + 1]^2 + 8(h + m + 1 + \dots + h + m + n) \end{aligned}$$

Thus obtaining

$$\begin{aligned} \frac{x^2 - y^2}{y^2 - z^2} &= \frac{[2(h + m + n) + 1]^2 - [2(h + m) + 1]^2}{[2(h + m) + 1]^2 - (2h - 1)^2} = \\ \frac{[2(h + m) + 1]^2 + 8(h + m + 1 + \dots + h + m + n) - [2(h + m) + 1]^2}{(2h - 1)^2 + 8(h + h + 1 + \dots + h + m) - (2h - 1)^2} &= \frac{a}{b} \end{aligned}$$

The question is: is it always possible to find a decomposition as (16)? Fibonacci does not offer a general answer, but a method by trials and errors, which let us induce to think he was aware that such decomposition is always possible. However, this is a purely historical problem¹⁶. What is very interesting from a didactical standpoint is that the solutions to our equation are easy if the difference $a - b$ is an odd number or a number of the form 2^n or $2^n(2m + 1)$, $n > 1$. In these cases, a direct numerical application of the technique of the *consecutive addends* offers a solution. By the way, the numerical examples of Fibonacci belong to these two cases. While, if $a - b = 2(2m + 1)$, a different demonstration is necessary. Thence, I will divide the argument into two cases; 1) $a - b = (2n + 1)$ or $a - b = 2^n$ or $a - b = 2^n(2m + 1)$, $n > 1$; 2) $a - b = 2(2m + 1)$.

First case: $a - b = (2n + 1)$ or $a - b = 2^n(2m + 1)$. Let first consider $a - b = (2n + 1)$. I offer the following canon, which is not exactly present – in this form – in Fibonacci's work, but strictly tied to his general approach:

- a) Let us consider the difference $a - b = d$;
- b) Multiply a by d and b by d ;
- c) Let us consider the sum $a + b = s$ and the number $ds = d(a + b) = da + db$;

¹⁶ See Genocchi, 1855b, p. 348-353.

- d) Let us now decompose da and db in the sum of d consecutive addends, of which a and b are the central addends. This is always possible because d is an odd number. This construction ensures that in the decomposition of da and db in consecutive addends, the places between b and a are exactly $d = a - b$, hence we have a decomposition of $d(a + b)$ in a series of consecutive addends of which the first d represent da and the last d represent db . Therefore the equation is solved.

An example will completely clarify the situation: let the equation $\frac{x^2 - y^2}{y^2 - z^2} = \frac{14}{5}$ be given. In this case $a - b = 9$, Thence we have to decompose $9 \cdot 5$ and $9 \cdot 14$ into nine consecutive addends, of which 5 and 14 are the central addends, thus obtaining the required decomposition of $ds = 9(5 + 14)$, which is

$$1 + 2 + 3 + 4 + 5 + 6 + 7 + 8 + 9 \mid + 10 + 11 + 12 + 13 + 14 + 15 + 16 + 17 + 18$$

The first nine addends are the decomposition of $9 \cdot 5$, the last nine of $9 \cdot 14$. Therefore, using the same symbols as those adopted for the general method, it is $h = 1$; $m = 8$, $n = 9$. The three roots are hence $z = 2h - 1 = 1$; $y = 2(h + m) + 1 = 19$; $x = 2(h + m + n) + 1 = 37$. Indeed, it is

$$\frac{37^2 - 19^2}{19^2 - 1^2} = \frac{14}{5}$$

Second case: $d = a - b = 2^l$ ($l > 1$, because the cases $n = 0, 1$ have already been dealt with). If $d = 4$, the problem is quite easy because $a + b$ can, then, be decomposed into four consecutive addends, of which the first two are equal to b and the second two are equal to a . If $n > 2$, it is always possible – as it is easy to prove – to decompose $2^{l-2}a$ and $2^{l-2}b$ in a sum of 2^{l-1} consecutive addends, in which the numbers $\frac{a-1}{2}$, $\frac{a+1}{2}$; $\frac{b-1}{2}$, $\frac{b+1}{2}$ are the central addends of the respective sums.

Example: $\frac{x^2 - y^2}{y^2 - z^2} = \frac{27}{11}$. It is $27 - 11 = 16 = 2^4$, thence $l = 4$. It is hence necessary to decompose $4 \cdot 11$ and $4 \cdot 27$ into 8 consecutive addends, in which the central pairs are respectively

$$\left(\frac{b-1}{2}, \frac{b+1}{2} \right) = (5, 6)$$

and

$$\left(\frac{a-1}{2}, \frac{a+1}{2} \right) = (13, 14).$$

The decomposition is thence:

$$2 + 3 + 4 + \mid 5 + 6 \mid + 7 + 8 + 9 \mid \mid + 10 + 11 + 12 + \mid 13 + 14 \mid + 15 + 16 + 17 + 18$$

Where one vertical bar indicates the central pairs of every decomposition and two vertical bars the step from the decomposition of $4 \cdot 11$ to that of $4 \cdot 27$. Thus: $h = 2$; $m = 7$, $n = 8$. The three roots are hence $z = 2h - 1 = 3$; $y = 2(h + m) + 1 = 19$; $x = 2(h + m + n) + 1 = 35$. We have:

$$\frac{35^2 - 19^2}{19^2 - 3^2} = \frac{27}{11}$$

If $d = a - b = 2^l(2q + 1)$ ($l > 1$), the reasoning is analogous. The difference is that $2^{l-1}(2q + 1)a$ and $2^{l-2}(2q + 1)b$ will be decomposed in the sum of $2^{l-1}(2q + 1)$ consecutive addends. I provide an example: let us suppose that $\frac{a}{b} = \frac{17}{5}$. It is $a - b = 12 = 3 \cdot 2^2$, $2q + 1 = 3$, $2^{l-2} = 1$. This means that $2^{l-2}(2q + 1)a = 3 \cdot 17 = 51$ and $2^{l-2}(2q + 1)b = 3 \cdot 5 = 15$ will be decomposed into 6 consecutive addends, whose central pairs are 8, 9 and 2, 3. The decomposition is hence

$$0 + 1 + |2 + 3| + 4 + 5| + 6 + 7 + |8 + 9| + 10 + 11$$

So that it is $z = 1$, $y = 11$, $x = 23$.

The particularity of this method is that we have always $n = m + 1$ because the equimultiples of a and b , ka and kb , are decomposed into the same number of addends, which is not necessarily true in the general conditions of solution.

This method cannot be applied if the difference $a - b$ is of the form $2(2n + 1)$. In this case, there is no particular device connected to a numerical canon: it is necessary to offer a general solution. Fibonacci – as told above – offers a method by trials and errors, which works because of the following considerations due to Angelo Genocchi¹⁷: Let us suppose (see (16)) that

$$h, h + 1, h + 2, \dots, h + m, h + m + 1, h + m + 2, \dots, h + m + n$$

can be constructed so that kb is the sum of the first m terms and ka of the last n . It is hence

$$\begin{cases} hm + \frac{m(m+1)}{2} = kb \\ (h + m)n + \frac{n(n+1)}{2} = ka \end{cases}$$

Therefore

$$\begin{cases} 2k = \frac{mn(m+n)}{am-bn} \\ 2h + 1 = \frac{2bmn + bn^2 - am^2}{am-bn} \end{cases}$$

Since a and b are coprime, the equation $am - bn = 1$ has solutions, for example $(m, n) = (b+1, a)$; $(m, n) = (b, a-1)$ and so on. This means that the decomposition (16) is possible and that hence our equation has always solutions. In particular

$$\begin{cases} z = 2h + 1 = 2bmn + bn^2 - am^2 \\ y = 2(h + m) + 1 = bn^2 + am^2 \\ x = 2(h + m + n) + 1 = 2amn - bn^2 + am^2 \end{cases}$$

Didactical considerations.

Relations history of mathematics-mathematics education-advanced mathematics.

1. The general solution of Fibonacci's equation is an ideal example to detect the relations between history of mathematics, mathematics education and advanced

¹⁷ Genocchi, 1855a, p. 186-189; Genocchi, 1855b, p. 348-353.

mathematics: let us start from the case in which the difference $a-b$ is an odd number or a number, which can be divided by 4 or a bigger power of 2. I have offered a general explanation, but, in this case, the method of the *paradigmatic examples* still works. The explanation why the examples are paradigmatic is a little more complicated than in the cases in which $a-b = 1$ or 2, but the teacher can offer it without resorting to any general formula. It is enough to show the canon and to give an explanation in words, which is quite intuitive. If the difference $a-b$ is of the form $2(2n+1)$ – and only in this case – my canon cannot be applied. It is hence necessary to show, as Genocchi did, that the series (16) can be anyway obtained by a general and formal reasoning. This is instructive from an educative standpoint because the pupils have a tangible example how much different can be the difficulties behind an apparently similar problems. It is significant the teacher explains why my canon works when the difference $a-b$ is, for example, of the form $2^2(2n+1)$, but not when it is $2(2n+1)$. Nuances, but nuances are the essence of mathematics. And history plays an important role in this context, because there is a historiographical counterpart to the previous didactical question: Fibonacci, probably, knew the canon, but is it possible to interpret, as Genocchi seems to do, his method by trials and errors as an evidence of the fact the he knew the form of the general solution? This is a historiographical question, to which it is not possible to answer in this paper. Notwithstanding, I cared about pointing out the parallelism between some historiographical and educative questions. The system by which Genocchi proves that a general solution exists is tied to an appropriate symbolism, but the reasoning is, in itself, rather easy. The same consideration holds if we look for other solutions to Fibonacci's equations (as Fibonacci himself highlights), however the request to find rigorously all the solutions drives to a field, which is completely different, because no reasoning structured as Genocchi's is of help. It is possible to show that this problem can be reduced to the integral solutions of the equation $At^2 + B = u^2$ (each letter indicates an integer), whose general theory is due to Lagrange (see note 13) and drives to a sector of the theory of numbers which, for a long period, belonged to advanced mathematics. Obviously, it is not possible to deal with Lagrange's solution in a high school, but it is possible the students catch how a single problem can open the doors on interesting educative perspectives, profound historical questions and subjects concerning high mathematics.

2. To complete the canon when the difference $a-b$ is an odd number or a number of the form $2^n(2m+1)$, $n > 1$, it is necessary to add a consideration, which can be interestingly exploit in an educative perspective: let us consider the equation $\frac{x^2-y^2}{y^2-z^2} = \frac{22}{5}$. Since $22-5 = 17$, which is an odd number, it is possible to decompose $17 \cdot 5$ and $17 \cdot 22$ into the sum of seventeen consecutive addends, of which 5 and 17 are the central addends. Nevertheless, here the canon works perfectly only if we admit negative numbers. The decomposition is this:

$$-3 - 2 - 1 - 0 + 1 + 2 + 3 + 4 + 5 + 6 + 7 + 8 + 9 + 10 + 11 + 12 + 13 + |14 + 15 + 16 + 17 + 18 + 19 + \dots + 30$$

Namely:

$$4 + 5 + 6 + 7 + 8 + 9 + 10 + 11 + 12 + 13 + |14 + 15 + 16 + 17 + 18 + 19 + \dots + 30$$

So that the roots are $z = 7$; $y = 27$, $x = 61$.

The question is from a historical point of view: could Fibonacci resort to a canon, which, though elementary, works if one admits negative numbers? The answer

should be negative because it is problematic to admit Fibonacci was able to operate with negative numbers, albeit Genocchi hypothesizes, in reference to another problem dealt with by Fibonacci¹⁸, he was, or, at least, he had the idea of negative numbers. Whatever the answer is, it is interesting to point out a particularity of the relation history of mathematics-mathematics education: history of mathematics is a fundamental support for mathematics education, even though the solutions to a certain problem, presented by the teacher in the classroom, might not exactly coincide with those likely offered in the historical period in which the problem was posed. It is important that: 1) history provides stimulating material for mathematics education; 2) the reference to history induces to address the subjects by means of a more critical and, at the same time, intuitive approach, which is different from those connoting most of handbooks; 3) the creativity of the pupils can be stimulated, so that they learn to play with numbers (and this paper deals with such a problem) and with the geometrical figures, the two ingredients, which, for centuries and centuries, were at the basis of both advanced mathematics and mathematical education and which nowadays risk to go lost for a more abstract approach based on the structures and their laws.

6. Conclusion

I hope to have clarified – although by means of a sole example – why Enriques' opinion on the lack of a very board-line between history of mathematics, mathematics education and advanced mathematics is plausible, even though not necessarily true in any aspect. To support this thesis I have considered an example which could be used at the third-fourth year in the high schools. Others could be conceived for the university level.

Rather than to summarize the content of this paper, I prefer to suggest seven lessons to develop in a classroom in order to present the material here proposed.

- Lesson 1 (after the explanation of the principle of mathematical induction): Fibonacci's method to prove that each square is the sum of a series of consecutive addends. Comparison between this method and the mathematical induction. The importance of a good mathematical symbolism. The concept of generality (a letter which indicate any number).
- Lesson 2: further examples of Fibonacci's method and comparison with modern demonstrations by mathematical induction.
- Lesson 3 and 4: interdisciplinary lessons with the teachers of history and philosophy (in the nations and schools where this discipline is taught). The problem of infinity. How the relation of the man with mathematical infinity has changed from Fibonacci's time to nowadays. Relation with the infinity in the history of philosophy. These subjects will not be – this is obvious – completely analysed in two lessons. But two lessons are sufficient to offer stimulating and interesting ideas to the pupils.

¹⁸ Genocchi, 1855b, p. 172-175.

- Lesson 5: the equation $\frac{x^2-y^2}{y^2-z^2} = \frac{a}{b}$, when $a - b = 1$ or 2. Relation between *paradigmatic examples* and rigorous proofs. Relation between paradigmatic examples and empirical examples.
- Lesson 6 and 7: $\frac{x^2-y^2}{y^2-z^2} = \frac{a}{b}$ in the other cases. The technique of the successive addends. Its range of applicability. Extension of this technique when it is necessary to introduce negative numbers. Cases in which this technique cannot be applied. The general algorithm. *Very important*: the teacher should point out the limits inside which a method can be used. Each method has its ray of applicability. Point out that apparently similar problems need different procedures of solution.

It is easy to think of exercises connected to the methods dealt with here.
For example: to solve the equations

$$\frac{x^2 - y^2}{y^2 - z^2} = \frac{43}{11}; \quad \frac{x^2 - y^2}{y^2 - z^2} = \frac{41}{11}; \quad \frac{x^2 - y^2}{y^2 - z^2} = \frac{17}{5}.$$

References

- Boncompagni, B.: 1852, *Della vita e delle opere di Leonardo Pisano, matematico del secolo decimoterzo*, Tipografia delle Belle Arti, Roma.
- Boncompagni, B.: 1854, *Intorno ad alcune opere di Leonardo Pisano matematico del secolo decimoterzo*, Tipografia delle Belle Arti, Roma.
- Bussotti, P.: 2004, Il Triangulus Arithmeticus di Pascal, *Nuova Secondaria* **2**, 75-81.
- Bussotti, P.: 2006, *From Fermat to Gauss: indefinite descent and methods of reduction in number theory*, Rauner Verlag, Augsburg.
- Bussotti, P.: 2008a, *Fibonacci und sein Liber Quadratorum. In Kaiser Friedrich II. (1194-1250). Welt und Kultur des Mittelmeerraums*, Philip von Zabern, Mainz am Rhein.
- Bussotti, P.: 2008b, *Problems and methods at the origin of the theory of numbers*, Accademia Pontaniana, Napoli.
- Bussotti, P.: 2014, The possible relations between history of mathematics and mathematics education, in Science and Technology Education for the 21st Century. Research and Research Oriented Studies. Proceedings of the 9th IOSTE Symposium for Central and Eastern Europe, Hradek Králové, *Gaudeamus* 29-41.
- Bussotti, P.: 2015, Differential calculus: the use of Newton's Methodus Fluxionum et Serierum infinitarum in an education context, *Problems of Education in the 21th Century* **69**, 39-65.
- Enriques, F.: 1921, 2003, *Insegnamento dinamico, con scritti di Franco Ghione e Mauro Moretti*, La Spezia, Agorà.
- Favaro, A.: 1874, Notizie storiche sulle frazioni continue dal secolo decimoterzo al decimosettimo, *Bullettino di bibliografia e di storia delle scienze matematiche e fisiche* **VII**, 451-589.
- Fibonacci, L.: 1228, 1857, Liber Abbaci, *Scritti di Leonardo Pisano matematico del secolo decimoterzo pubblicato da Baldassarre Boncompagni. Volume I*, Tipografia delle Scienze matematiche e fisiche, Roma.

- Fibonacci, L.: 1862a, Flos, *Scritti di Leonardo Pisano matematico del secolo decimoterzo pubblicato da Baldassarre Boncompagni. Volume II (Practica Geometriae ed opuscoli)*, Tipografia delle Scienze matematiche e fisiche, Roma, 227-252.
- Fibonacci, L.: 1862b, Liber Quadratorum, *Scritti di Leonardo Pisano matematico del secolo decimoterzo pubblicato da Baldassarre Boncompagni. Volume II (Practica Geometriae ed opuscoli)*, Tipografia delle Scienze matematiche e fisiche, Roma, 253-283.
- Fibonacci, L.: 1862c, Practica Geometriae, *Scritti di Leonardo Pisano matematico del secolo decimoterzo pubblicato da Baldassarre Boncompagni. Volume II (Practica Geometriae ed opuscoli)*, Tipografia delle Scienze matematiche e fisiche, Roma, 1-224.
- Fibonacci, L.: 1952, *Le Livre de nombres carres*, edited by Paul ver Eecke, Blanchard, Paris. (as Leonarde de Pise).
- Fibonacci, L.: 1988, The Book of Squares, translation into English of Fibonacci's Liber Quadratorum by L.E. Sigler, *The Fibonacci Quarterly* **26**(4).
- Franci, R.: 2002, Il Liber abaci di Leonardo Fibonacci. La matematica nella società e nella cultura, *Bollettino dell'Unione Matematica Italiana* **8**, 293-328.
- Genocchi, A.: 1855a, Intorno a tre scritti inediti di Leonardo Pisano pubblicati da Baldassare Boncompagni. Nota, *Annali di scienze matematiche e fisiche* **6**, 115-120.
- Genocchi, A.: 1855b, Sopra tre scritti inediti di Leonardo Pisano pubblicati da B. Boncompagni. Note analitiche di Angelo Genocchi, *Annali di scienze matematiche e fisiche* **6**, 161-185, 219-251, 273-320, 345-362.
- Genocchi, A.: 1855c, Passages of letters by genocchi to boncompagni, *Annali di scienze matematiche e fisiche* **6**, 129-134, 186-194, 195-205, 206-209, 251-253, 254-257, 257-259.
- Grimm, R. E.: 1973, The autobiography of leonardo pisano, *The Fibonacci Quarterly* **11**(1), 99-104.
- Horadam, A. F.: 1991, Fibonacci's Mathematical Letter to Master Theodorus, *The Fibonacci Quarterly* **29**(2), 103-107.
- Hughes, B.: 2008, *Fibonacci's De Practica Geometrie*, Springer, New York.
- Lagrange, J. L.: 1769, Sur la solution de problème indéterminés du second degré, in Mémoires de l'Académie royale des Sciences et Belles-Lettres de Berlin, t. XXIII, *Œuvres de Lagrange*, Vol. 2, Gauthier-Villars, Paris, 375-535.
- Lüneburg, H.: 1991, Fibonacci's aufsteigende Kettenbrüche, ein elegantes Werkzeug mittelalterlicher Rechenkunst, *Sudhoffs Archiv* **75/76**, 129-139.
- Palladino, D., Bussotti, P.: 2002, Il principio di induzione 1. Storia e ruolo, *Nuova Secondaria* **2**, 41-51.
- Picutti, E.: 1979, Il libro dei quadrati di Leonardo Pisano e i problemi di analisi indeterminata nel Codice Palatino 557 della Biblioteca Nazionale di Firenze, *Physis* **21**, 195-339.
- Picutti, E.: 1983, Il "Flos" di Leonardo Pisano. Traduzione e commento, *Physis* **25**, 293-387.
- Pisano, R., Bussotti, P.: 2013, On popularization of Scientific Education in Italy between 12th and 16th Century, *Problems of Education in the 21th Century* **57**, 90-101.
- Pisano, R., Bussotti, P.: 2015, Fibonacci and the Reception of the Abacus Schools in Italy. Mathematical Conceptual Streams and their Changing Relationship with Society. *Almagest* **6**, 2, 127-165.

- Rashed, R.: 1994, Fibonacci et les Mathématiques arabes, *Micrologus* **2**, 145-160.
- Rashed, R.: 2003, Fibonacci et le Prolongement Latin des Mathématiques Arabes, *Bollettino di storia delle scienze* **XXIII**, 55-73.
- Sigler, L.: 2002, *Fibonacci's Liber Abaci*, Springer, New York.
- Woepcke, F.: 1854-55, Sur le traité des nombres carrés de Léonard de Pise, retrouve et publié par M. le prince Balthasar Boncompagni, *Journal de mathématiques pures et appliquées*, 1re série, **20**, 54-62.
- Woepcke, F.: 1860-61, *Recherches sur plusieurs ouvrages de Léonard de Pise. III- Traduction d'un fragment anonyme sur la formation des triangles rectangles en nombres entiers et d'un traité sur le meme sujet de Abou Dia'far Mohammed ben Alogain*, Imprimerie des Sciences Mathématiques et Physiques, Rome.

University of Udine
Via T. Petrarco
e-mail palobussotti66@gmail.com

Jan Górowski, Adam Łomnicki

Iterations of homographic functions and recurrence equations involving a homographic function*

Abstract. The formulas for the m -th iterate ($m \in \mathbb{N}$) of an arbitrary homographic function H are determined and the necessary and sufficient conditions for a solution of the equation $y_{m+1} = H(y_m)$, $m \in \mathbb{N}$ to be an infinite n -periodic sequence are given. Based on the results from this paper one can easily determine some particular solutions of the Babbage functional equation.

1. Preliminaries

The recurrence equations involving a homographic function were studied in (Graham, Knuth, Patashnik, 2002). The authors stated that the only known examples of such equations possessing periodic solutions are

$$y_{m+1} = 2i \sin \pi r + \frac{1}{y_m}, \quad m \in \mathbb{N},$$

where r is a rational number such that $0 \leq r < \frac{1}{2}$.

Various approaches to the sequences given by the recurrence equation

$$y_{m+1} = H(y_m), \quad m \in \mathbb{N}, \tag{1}$$

where H is a homographic function may be found in (Koźniewska, 1972; Levy, Lessman, 1966; Uss, 1966; Wachniczy, 1966).

In this paper we prove formulas determining all solutions of (1). We also give the necessary and sufficient conditions for a solution of (1) to be periodic.

We also determine some particular solutions of the Babbage functional equation

*Iteracje funkcji homograficznej i równanie rekurencyjne zadane funkcją homograficzną

2010 Mathematical Subject Classification: Primary: 26A18; 37B20; 39B22

Key words and phrases: iterations of homographic functions, recurrence equation, periodic sequences

$$\varphi^m(x) = x, \quad x \in X, \quad (2)$$

where m is an arbitrary fixed integer. Recall that ψ^n for $n \in \mathbb{N}$ denotes the n -th iterate of a function $\psi: X \rightarrow X$, i.e. $\psi^0 = \text{Id}_X$ and $\psi^n = \psi \circ \psi^{n-1}$ for integer $n \geq 1$.

Some results concerning (2) may be found in (Kuczma, 1968). In particular the following

THEOREM 1 (KUCZMA, 1968, P. 291)

If φ is a meromorphic solution of equation (2), then

$$\varphi(x) = \frac{a'x + b'}{c'x + d'}$$

for some $a', b', c', d' \in \mathbb{C}$.

THEOREM 2 (KUCZMA, 1968, P. 291)

If $L(x) = \alpha x + \beta$, where $\alpha \neq 0$, then φ satisfies (2) if and only if $L^{-1} \circ \varphi \circ L$ does so.

THEOREM 3 (KUCZMA, 1968, P. 291)

Let $K_{-1} = 0$, $K_0 = 1$, $K_m = \gamma K_{m-1} + \delta K_{m-2}$ for $m \in \mathbb{N}_+$ and let $S(x) = \gamma + \frac{\delta}{x}$, where $\gamma, \delta \in \mathbb{C}$ and $\delta \neq 0$, then

$$S^m(x) = \frac{K_mx + \delta K_{m-1}}{K_{m-1}x + \delta K_{m-2}} \quad \text{for } m \in \mathbb{N}_+. \quad (3)$$

Let

$$H(x) := \frac{ax + b}{cx + d}, \quad \text{where } a, b, c, d \in \mathbb{C}, \quad c \neq 0, \quad ad - bc \neq 0 \quad (4)$$

In the sequel we assume that the domain of H is the set D defined as follows

$$D := \bigcap_{m \in \mathbb{N}_+} D_{H^m},$$

where D_{H^m} denotes the domain of H^m and $\mathbb{N}_+ := \mathbb{N} \setminus \{0\}$. We also set $H^0 := \text{Id}_D$.

Let $H: D \rightarrow \mathbb{C}$ be a function given by (4) and let

$$y_0 = x_0 \quad \text{and} \quad y_{m+1} = H(y_m) \quad \text{for an } x_0 \in D \text{ and } m \in \mathbb{N}.$$

Notice that

$$y_m = H^m(x_0) \quad m \in \mathbb{N} \quad (5)$$

and $x_0, H(x_0), H^2(x_0), \dots, H^{m-1}(x_0) \in D$.

Based on the theory of recurrence linear equations of order 2 with constant coefficients (Koźniewska, 1972, p. 59) we get

LEMMA 1

Let $K_{-1} = 0$, $K_0 = 1$, $K_m = \gamma K_{m-1} + \delta K_{m-2}$ for $m \in \mathbb{N}_+$, $\delta \neq 0$ and let $\Delta = \gamma^2 + 4\delta$. Then for $m \in \mathbb{N} \cup \{-1\}$,

$$1^o \ K_m = (m+1)\left(\frac{\gamma}{2}\right)^m, \text{ if } \Delta = 0,$$

$$2^o \ K_m = \frac{1}{\sqrt{\Delta}} \left(\left(\frac{\gamma+\sqrt{\Delta}}{2} \right)^{m+1} - \left(\frac{\gamma-\sqrt{\Delta}}{2} \right)^{m+1} \right), \text{ if } \Delta \neq 0,$$

where $\sqrt{\Delta}$ denotes one of the complex square roots of Δ .

A consequence of Lemma 1 is

LEMMA 2

If $K_{-1} = 0$, $K_0 = 1$, $K_m = \gamma K_{m-1} + \delta K_{m-2}$ for $m \in \mathbb{N}_+$, $\delta \neq 0$, $\gamma, \delta \in \mathbb{R}$ and $\Delta = \gamma^2 + 4\delta$, then for $m \in \mathbb{N} \cup \{-1\}$,

$$1^o \ K_m = (m+1)\left(\frac{\gamma}{2}\right)^m, \text{ if } \Delta = 0,$$

$$2^o \ K_m = \frac{1}{\sqrt{\Delta}} \left(\left(\frac{\gamma+\sqrt{\Delta}}{2} \right)^{m+1} - \left(\frac{\gamma-\sqrt{\Delta}}{2} \right)^{m+1} \right), \text{ if } \Delta > 0,$$

$$3^o \ K_m = (\sqrt{-\delta})^m \cos \frac{m\pi}{2}, \text{ if } \gamma = 0 \text{ and } \Delta < 0,$$

$$4^o \ K_m = (\sqrt{-\Delta})^m (\cos m\psi + \cot \psi \sin m\psi), \text{ if } \gamma \neq 0 \text{ and } \Delta < 0,$$

where ψ is the principal value of an argument of the complex number $\frac{\gamma}{2} + i\frac{\sqrt{-\Delta}}{2}$.

2. Periodic solutions of the recurrence equation

DEFINITION 1

An infinite sequence $(y_m)_{m \in \mathbb{N}}$ is called periodic with period n (or n -periodic), where $n \in \mathbb{N}$, $n \geq 1$, if $y_{m+n} = y_m$ for every $m \in \mathbb{N}$.

Consider equation (1) with the initial condition $y_0 = x_0$, where $H: D \rightarrow \mathbb{C}$ is a function defined by (4) and $x_0 \in D$. By (5) we get

LEMMA 3

Let $H: D \rightarrow \mathbb{C}$ be a function defined by (4) and let $n \geq 2$ be a fixed integer. Every solution of (1) is periodic with period n if and only if

$$H^n = \text{Id}_D. \quad (6)$$

Proof. Assume that for some integer $n \geq 2$ equation (6) holds, then by (5) for every $m \in \mathbb{N}$ we have

$$y_{m+n} = H^{m+n}(x_0) = H^m(H^n(x_0)) = H^m(x_0) = y_m,$$

where $y_0 = x_0 \in D$. For the converse suppose that every solution of (1) is n -periodic. Let $x_0 \in D$, so $H^m(x_0) \in D$ for every $m \in \mathbb{N}$. Put $y_m := H^m(x_0)$, $m \in \mathbb{N}$. The sequence $(y_m)_{m \in \mathbb{N}}$ satisfies (1), so it is n -periodic. Thus

$$H^n(x_0) = y_n = y_0 = H^0(x_0) = x_0.$$

Hence (6) holds.

Observe that Lemma 3 holds true if H is an arbitrary function with a proper domain satisfying (2).

THEOREM 4

Let $S: D' \rightarrow \mathbb{C}$ be a function defined as $S(x) = \gamma - \frac{\delta}{x}$, where $\gamma, \delta \in \mathbb{C}$, $\delta \neq 0$ and $D' := \bigcap_{m \in \mathbb{N}_+} D_{S^m}$, where D_{S^m} denotes the domain of S^m . Every sequence $(y_m)_{m \in \mathbb{N}}$ satisfying the following recurrence relation

$$y_{m+1} = S(y_m), \quad m \in \mathbb{N} \quad (7)$$

is 2-periodic if and only if $\gamma = 0$.

Proof. In view of Lemma 3 it follows that every sequence $(y_m)_{m \in \mathbb{N}}$ satisfying (7) is 2-periodic if and only if

$$S^2(x) = \gamma + \frac{\delta x}{\gamma x + \delta} = \text{Id}_{D'}(x), \quad x \in D'.$$

Which is equivalent to the fact that $\gamma = 0$.

Now we prove the following results.

THEOREM 5

Let S be as in Theorem 4, $\Delta = \gamma^2 + 4\delta$ and let $n \in \mathbb{N}$ be such that $n \geq 3$.

- (i) If every sequence $(y_m)_{m \in \mathbb{N}}$ satisfying (7) is n -periodic, then $\Delta \neq 0$ and $\delta = \frac{-\gamma^2}{4 \cos^2 \frac{k\pi}{n}}$ for some $k \in \{1, 2, 3, \dots, n-1\}$.
- (ii) If $k \in \{1, 2, 3, \dots, n-1\}$ and $\gamma^2 + 4\delta \neq 0$ and $4\delta \cos^2 \frac{k\pi}{n} = -\gamma^2$, then every sequence $(y_m)_{m \in \mathbb{N}}$ satisfying (7) is n -periodic.

Proof. To show (i) observe that by Lemma 3 we get

$$S^n(x) = x, \quad x \in D'. \quad (8)$$

By Theorem 3 and Lemma 1, (8) is equivalent to the following conditions

$$\begin{aligned} \frac{K_n x + \delta K_{n-1}}{K_{n-1} x + \delta K_{n-2}} - x \frac{K_{n-1} x + \delta K_{n-2}}{K_{n-1} x + \delta K_{n-2}} &= 0, \quad x \in D', \\ \frac{\gamma K_{n-1} x + \delta K_{n-2} x + \delta K_{n-1} - K_{n-1} x^2 - \delta K_{n-2} x}{K_{n-1} x + \delta K_{n-2}} &= 0, \quad x \in D', \\ \frac{K_{n-1}(-x^2 + \gamma x + \delta)}{K_{n-1} x + \delta K_{n-2}} &= 0, \quad x \in D', \\ K_{n-1} &= 0, \\ \Delta \neq 0 \quad \text{and} \quad \left(\frac{\gamma + \sqrt{\Delta}}{2} \right)^n &= \left(\frac{\gamma - \sqrt{\Delta}}{2} \right)^n, \\ \Delta \neq 0 \quad \text{and} \quad (\gamma + \sqrt{\Delta})^n &= (\gamma - \sqrt{\Delta})^n, \end{aligned}$$

$$\Delta \neq 0 \text{ and } \exists k \in \{1, \dots, n-1\} : \gamma + \sqrt{\Delta} = (\gamma - \sqrt{\Delta}) \left(\cos \frac{2k\pi}{n} + i \sin \frac{2k\pi}{n} \right). \quad (9)$$

Now notice that

$$\gamma + \sqrt{\Delta} = (\gamma - \sqrt{\Delta}) \left(\cos \frac{2k\pi}{n} + i \sin \frac{2k\pi}{n} \right)$$

is equivalent to the following conditions

$$\begin{aligned} \sqrt{\Delta} \left(1 + \cos \frac{2k\pi}{n} + i \sin \frac{2k\pi}{n} \right) &= \gamma \left(\cos \frac{2k\pi}{n} + i \sin \frac{2k\pi}{n} - 1 \right), \\ 2\sqrt{\Delta} \cos \frac{k\pi}{n} \left(\cos \frac{k\pi}{n} + i \sin \frac{k\pi}{n} \right) &= 2\gamma \sin \frac{k\pi}{n} \left(i \cos \frac{k\pi}{n} - \sin \frac{k\pi}{n} \right), \\ 2\sqrt{\Delta} \cos \frac{k\pi}{n} \left(\cos \frac{k\pi}{n} + i \sin \frac{k\pi}{n} \right) &= 2\gamma i \sin \frac{k\pi}{n} \left(\cos \frac{k\pi}{n} + i \sin \frac{k\pi}{n} \right). \end{aligned} \quad (10)$$

Thus condition (9) is equivalent to

$$\begin{aligned} \Delta \neq 0 \text{ and } \exists k \in \{1, \dots, n-1\} \quad \sqrt{\Delta} \cos \frac{k\pi}{n} &= \gamma i \sin \frac{k\pi}{n}, \\ \Delta \neq 0 \text{ and } \exists k \in \{1, \dots, n-1\} \quad \sqrt{\Delta} &= \gamma i \tan \frac{k\pi}{n}, \\ \Delta \neq 0 \text{ and } \exists k \in \{1, \dots, n-1\} \quad \Delta &= -\gamma^2 \tan^2 \frac{k\pi}{n}, \\ \Delta \neq 0 \text{ and } \exists k \in \{1, \dots, n-1\} \quad \delta &= \frac{-\gamma^2}{4 \cos^2 \frac{k\pi}{n}}, \end{aligned}$$

which completes the proof of (i).

For the implication (ii) consider two cases:

- a. n is an even number and $k = \frac{n}{2}$,
- b. n is an even number and $k \neq \frac{n}{2}$ or n is odd and $k \in \{1, 2, 3, \dots, n-1\}$.

In the case a, we get $\gamma = 0$ and according to Theorem 4 every sequence satisfying (7) is 2-periodic and hence n -periodic.

For the case b, notice that for every $k \in \{1, 2, 3, \dots, n-1\}$, $\cos \frac{k\pi}{n} \neq 0$ we have $\Delta \neq 0$ and $\delta = \frac{-\gamma^2}{4 \cos^2 \frac{k\pi}{n}}$ which yields $\Delta = -\gamma^2 \tan^2 \frac{k\pi}{n}$. Denote by $\sqrt{\Delta}$ the number $\gamma i \tan \frac{k\pi}{n}$, thus $\sqrt{\Delta} \cos \frac{k\pi}{n} = \gamma i \sin \frac{k\pi}{n}$ which is equivalent to (10). Now reversing the reasoning from the case (i) – from condition (10) to (8) (without condition (9)) – finishes the proof.

The results obtained above will be now applied to examine the sequences defined by (1).

THEOREM 6

If $H: D \rightarrow \mathbb{C}$ is a function given by (4) and $(K_{-1}, K_0, K_1, \dots)$ is a sequence defined in Theorem 3 for which $\gamma = a + d$ and $\delta = bc - ad$, then

$$H^m(x) = \frac{1}{c} \frac{cK_mx + dK_m + \delta K_{m-1}}{cK_{m-1}x + dK_{m-1} + \delta K_{m-2}} - \frac{d}{c} \quad \text{for } m \in \mathbb{N}_+, L(x) \in D. \quad (11)$$

Proof. Let $\gamma = a + d$, $\delta = bc - ad$, $L(x) = \frac{x}{c} - \frac{d}{c}$ and $S(x) = (L^{-1} \circ H \circ L)(x)$, $L(x) \in D$. We have

$$\begin{aligned} L^{-1}(x) &= cx + d, \\ (H \circ L)(x) &= \frac{ax + bc - ad}{cx}, \\ S(x) &= (L^{-1} \circ H \circ L)(x) = a + d + \frac{bc - ad}{x} = \gamma + \frac{\delta}{x}, \end{aligned}$$

for $L(x) \in D$, where $\delta \neq 0$. By Theorem 3 we obtain

$$S^m(x) = \frac{K_mx + \delta K_{m-1}}{K_{m-1}x + \delta K_{m-2}} \quad \text{for } m \in \mathbb{N}_+, L(x) \in D.$$

Now observe that

$$S^m = L^{-1} \circ H^m \circ L \quad \text{for } m \in \mathbb{N}_+,$$

thus

$$H^m = L \circ S^m \circ L^{-1} \quad \text{for } m \in \mathbb{N}_+,$$

which gives (11).

Lemma 3 and Theorem 2 yield

THEOREM 7

Let $H: D \rightarrow \mathbb{C}$ be a function defined by (4), $L(x) = \frac{x}{c} - \frac{d}{c}$, $S(x) = (L^{-1} \circ H \circ L)(x)$, $L(x) \in D$ and let $n \geq 2$ be a fixed integer. Then every solution of (1) is n -periodic if and only if

$$S^n = \text{Id}_D.$$

3. Examples

From the proof of Theorem 5 it follows that condition

$$\gamma^2 + 4\delta \neq 0 \quad \text{and} \quad \exists k \in \{1, \dots, n-1\} \quad \sqrt{\gamma^2 + 4\delta} = -\gamma i \tan \frac{k\pi}{n} \quad (12)$$

is equivalent to the fact that every sequence $(y_m)_{m \in \mathbb{N}}$ satisfying equation

$$y_{m+1} = \gamma + \frac{\delta}{y_m}$$

or equation

$$y_{m+1} = \frac{ay_m + b}{cy_m + d},$$

where $a + d = \gamma$ and $bc - ad \neq 0$, is n -periodic with $n \geq 3$.

Moreover, it is easy to find numbers γ, δ satisfying (12) and a, b, c, d - solutions of the system $a + d = \gamma, bc - ad \neq 0$. Namely, for $n = 3, \gamma = 1, \delta = -1$ (12) is fulfilled and numbers $a = 2, b = -3, c = 1, d = -1$ satisfy the system $a + d = 1, bc - ad = -1$, thus in view of Theorem 5 and Lemma 3 the following functions

$$S(x) = 1 + \frac{-1}{x}, \quad H(x) = \frac{2x - 3}{x - 1}$$

fulfil the Babbage equation $\varphi^3(x) = x$ (which can be directly checked).

Now let $n = 4$, for $\gamma = 2, \delta = -2$ (12) holds true. Let $a = 3, b = -5, c = 1$ and $d = -1$, then $a + d = 2, bc - ad = -2$. Similarly as above we get that the mappings

$$S(x) = 2 + \frac{-2}{x}, \quad H(x) = \frac{3x - 5}{x - 1}$$

satisfy equation $\varphi^4(x) = x$.

References

- Graham, R. L., Knuth, D. E., Patashnik, O.: 2002, *Matematyka konkretna*, PWN, Warszawa.
- Koźniewska, J.: 1972, *Równania rekurencyjne*, PWN, Warszawa.
- Kuczma, M.: 1968, *Functional Equations in a Single Variable*, Monogr. Math. 46, PWN-Polish Scientific Publishers, Warszawa.
- Levy, H., Lessman, F.: 1966, *Równania różnicowe skończone*, PWN, Warszawa.
- Uss, P.: 1966, Rekurencyjność inaczej, *Gradient* **2**, 102-106.
- Wachniccy, K. E.: 1966, O ciągach rekurencyjnych określonych funkcją homograficzną, *Gradient* **5**, 275-288.

*Instytut Matematyki
Uniwersytet Pedagogiczny
ul. Podchorążych 2
PL-30-084 Kraków
e-mail jangorowski@interia.pl
e-mail alomnicki@poczta.fm*

*Jan Górowski, Adam Łomnicki***Quasi-arithmetic means***

Abstract. We present a list of geometric problems with solutions that lead to known or less known means. We also prove, by elementary means, some property for so-called quasi-arithmetic means. We use the proved result to justify some inequalities between the means.

1. Introduction

Let $J \subset \mathbb{R}$ denote the open interval or respectively closed or half-closed. The sets $\mathbb{R}$, $\mathbb{R}_+ := (0, +\infty)$ and $\mathbb{R}_+ \cup \{0\}$ will be also considered as intervals.

One of the most general definition of a mean is the following

DEFINITION 1

Every function $d: J \times J \rightarrow J$ satisfying

- (i) $\forall a, b \in J \quad \min\{a, b\} \leq d(a, b) \leq \max\{a, b\}$,
- (ii) d is a increasing function with respect to each variable

is called a mean.

In (Aczél, 1948) and (Kitagawa, 1934) it was proved that under some additional conditions on d there exists a strictly monotone function g defined on J such that

$$d(a, b) = g^{-1}(pg(a) + qg(b)), \quad a, b \in J,$$

for some $p, q \in (0, 1)$ such that $p+q = 1$. Such means will be called quasi-arithmetic means.

For the purposes of this paper we modify Definition 1.

*Średnie quasi-arytmetyczne

2010 Mathematical Subject Classification: Primary: 26E60, 26D07

Key words and phrases: quasi-arithmetic means, inequalities involving means, extended mean values, means in geometry

DEFINITION 2

Let $T = \{(x, y) \in J \times J : x \geq y\}$. Every function $d: T \rightarrow J$ satisfying conditions (i) and (ii) of Definition 1 is said to be a mean.

In the sequel by a mean we understand a function in a sense of Definition 2.

There is a wide literature on means, some information may be found in (Aczél, 1948; Aczel, Dhombres, 1989; Galwani, 1927; Głazowska, Jarczyk, Matkowski, 2002; Górowski, Łomnicki, 2010; Kitagawa, 1934; Kołgomorov, 1930; Leach, Sholander, 1978; Leach, Sholander, 1983; Witkowski, 2009).

2. Geometric problems leading to means

Let us assume that a quadrangle $ABCD$ (see Fig. 1.) is a trapezium such that $AB \parallel CD$, $|AB| = a$, $|DC| = b$ and according to Fig. 1. $EF \parallel AB$, $|EF| = d$ and $DA' \parallel CB$. Denoting $\lambda = \frac{|AE|}{|ED|}$ we express d as a function of λ . By The Intercept

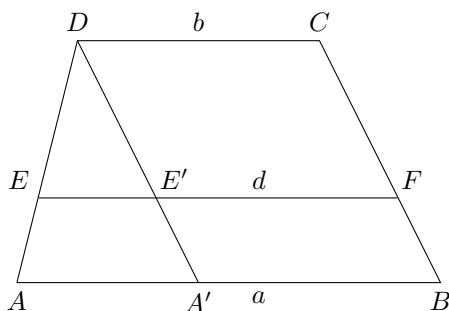


Fig. 1.

Theorem we get

$$\frac{|AA'|}{|EE'|} = \frac{|AD|}{|ED|} = \frac{|AE| + |ED|}{|ED|} = \lambda + 1.$$

Hence

$$|EE'| = \frac{|AA'|}{\lambda + 1} = \frac{a - b}{\lambda + 1} \quad \text{and} \quad d = \frac{a - b}{\lambda + 1} + b = \frac{a + \lambda b}{\lambda + 1}$$

and thus

$$d = \frac{a + \lambda b}{\lambda + 1}. \quad (1)$$

Now we formulate some geometric problems leading to means. Notice the well known problems P1-P4.

PROBLEM 1

Find the length d of the segment EF in the $ABCD$ (see Fig 1.) if

P1. E, F are the midpoints of the segments AD and BC , respectively;

P2. the diagonals AC and BD and the segment EF intersect at a point;

- P3.** the trapezes $ABFE$ and $EFCD$ are similar;
- P4.** the areas of the trapezes $ABFE$ and $EFCD$ are equal;
- P5.** the volumes of the solids of revolution obtained by rotating $ABFE$ and $EFCD$ around the line EF are equal;
- P6.** the volumes of the solids of revolution obtained by rotating $ABFE$ around the line AB and $EFCD$ around the line DC are equal;

It is easy to see that the solution of P1 is $d = \frac{a+b}{2}$.

Denote by S the intersection point of the diagonals AC and BD and the segment EF (problem P2). Then by Intersection Theorem we get

$$\lambda = \frac{|AE|}{|ED|} = \frac{|AS|}{|SC|} = \frac{|AB|}{|DC|} = \frac{a}{b}.$$

thus

$$d = \frac{2ab}{a+b}.$$

For the problem P3 notice that since the trapezes $ABFE$ and $EFCD$ are similar we obtain

$$\frac{|AE|}{|ED|} = \frac{|AB|}{|EF|} = \frac{|EF|}{|DC|},$$

hence $d^2 = ab$, and $d = \sqrt{ab}$.

To solve P4 denote by h_1, h_2 the altitudes of the trapezes $ABFE$ and $EFCD$, respectively. Let P denotes the area of the trapezium $ABFE$ (also trapezium $EFCD$). Then

$$\lambda = \frac{|AE|}{|ED|} = \frac{h_1}{h_2} = \frac{P}{a+d} \cdot \frac{d+b}{P} = \frac{d+b}{a+d}.$$

This and (1) give

$$d = \sqrt{\frac{a^2 + b^2}{2}}.$$

Now (problem P5) let h_1, h_2 be defined as above, then

$$\lambda^2 = \frac{h_1^2}{h_2^2} = \frac{\pi h_1^2}{\pi h_2^2}. \quad (2)$$

On the other hand, the volumes of the solids of revolution obtained by rotating $ABFE$ and $EFCD$ around the line EF are equal

$$\pi h_1^2 d + \frac{2}{3} \pi h_1^2 (a-d), \quad \pi h_2^2 b + \frac{1}{3} \pi h_2^2 (d-b), \quad (3)$$

respectively. From (2) and (3) we get

$$\lambda^2 = \frac{\pi h_1^2 (d + \frac{2}{3}(a-d))}{\pi h_2^2 (b + \frac{1}{3}(d-b))} \cdot \frac{b + \frac{1}{3}(d-b)}{d + \frac{2}{3}(a-d)} = \frac{b + \frac{1}{3}(d-b)}{d + \frac{2}{3}(a-d)} = \frac{d+2b}{d+2a},$$

which by (1) yields

$$d = \frac{2(a^2 + ab + b^2)}{3(a + b)} = \frac{\frac{a^3 - b^3}{3}}{\frac{a^2 - b^2}{2}}. \quad (4)$$

Finally, for the solution of P6 observe that

$$V_1 = \pi h_1^2 d + \frac{1}{3} \pi h_1^2 (a - d) \quad \text{and} \quad V_2 = \pi h_2^2 b + \frac{2}{3} \pi h_2^2 (d - b),$$

where h_1, h_2 denote the altitudes of the trapezes $ABFE$ and $EFCD$, resp., and V_1, V_2 are the volumes of the solids of revolution obtained by rotating $ABFE$ around the line AB and $EFCD$ around the line DC , resp. Similarly as above we get

$$\lambda^2 = \frac{b + 2d}{a + 2d},$$

thus

$$d = \sqrt{\frac{a^2 + ab + b^2}{3}} = \sqrt{\frac{a^3 - b^3}{3(a - b)}}. \quad (5)$$

Observe that (4) is one of the means introduced by Leach and Sholander in (Leach, Sholander, 1978), and (5) is a Stolarsky's mean from (Kolgomorov, 1930).

PROBLEM 2

Consider three pairwise homothetic squares with side length $a > d > b$, see Fig 2. Find d in terms of a and b .

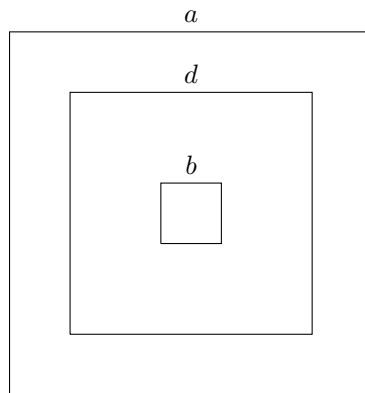


Fig. 2.

Let $\lambda = \frac{a^2 - d^2}{d^2 - b^2}$. Then

$$d = \sqrt{\frac{a^2 + \lambda b^2}{1 + \lambda}}.$$

This yields the following relationships:

$$\begin{aligned}
 d &= \sqrt{\frac{a^2 + b^2}{2}} && \text{for } \lambda = 1, \\
 d &= \sqrt{ab} && \text{for } \lambda = \frac{a}{b}, \\
 d &= \frac{a + b}{2} && \text{for } \lambda = \frac{3a + b}{a + 3b}, \\
 d &= \frac{2ab}{a + b} && \text{for } \lambda = \frac{a^2(a + 3b)}{b^2(3a + b)}, \\
 d &= \sqrt{\frac{a^3 + b^3}{a + b}} && \text{for } \lambda = \frac{b}{a}.
 \end{aligned}$$

3. Quasi-arithmetic means

Considerations from the previous section imply that, under some assumptions on a function $g: J \rightarrow \mathbb{R}$ and λ , it is worth to consider the function d_g^λ defined on $T = \{(x, y) \in J \times J : x \geq y\}$ and given by

$$d_g^\lambda(a, b) = g^{-1} \left(\frac{g(a) + \lambda g(b)}{1 + \lambda} \right), \quad (a, b) \in T. \quad (6)$$

we prove now the following result.

THEOREM 1

If $g: J \rightarrow \mathbb{R}$ is a strictly monotonic function, continuous on J and λ is a non-negative real number, then d_g^λ given by (6) is a mean (in a sense of Definition 2).

Proof. It is easy to see that d_g^λ is well defined. Indeed, if for some $(a, b) \in T$ and some $\lambda \in [0, +\infty)$ we had

$$\frac{g(a) + \lambda g(b)}{1 + \lambda} - g(x) \neq 0 \quad \text{for every } x \in J,$$

then since g is continuous,

$$\frac{g(a) + \lambda g(b)}{1 + \lambda} - g(x) > 0 \quad \text{for every } x \in J \quad (7)$$

or

$$\frac{g(a) + \lambda g(b)}{1 + \lambda} - g(x) < 0 \quad \text{for every } x \in J. \quad (8)$$

From (7) we obtain the following system of inequalities

$$\frac{g(a) + \lambda g(b)}{1 + \lambda} - \frac{(1 + \lambda)g(a)}{1 + \lambda} > 0 \quad \text{and} \quad \frac{g(a) + \lambda g(b)}{1 + \lambda} - \frac{(1 + \lambda)g(b)}{1 + \lambda} > 0,$$

which leads to a contradiction. The similar argument can be applied to (8).

The task is now to show that

$$\min\{a, b\} \leq d_g^\lambda(a, b) \leq \max\{a, b\}, \quad (a, b) \in T. \quad (9)$$

Observe that if g is a strictly increasing function, then so is g^{-1} and (9) is equivalent to

$$g(b) \leq \frac{g(a) + \lambda g(b)}{1 + \lambda} \leq g(a),$$

$$(1 + \lambda)g(b) \leq g(a) + \lambda g(b) \leq (1 + \lambda)g(a),$$

where the last inequality holds true. Similar argument applies to the case when g is strictly decreasing.

Finally, we prove that d_g^λ is an increasing function with respect to each variable. Fix $a \in J$ and suppose that g is strictly increasing. Let $b_1, b_2 \in J$ be such that $b_1 > b_2$ and $a \geq b_1$, then $g(b_1) > g(b_2)$, $\lambda g(b_1) \geq \lambda g(b_2)$, $g(a) + \lambda g(b_1) \geq g(a) + \lambda g(b_2)$ and in a consequence $d_g^\lambda(a, b_1) \geq d_g^\lambda(a, b_2)$. Now fix $b \in J$ and assume that $a_1 > a_2 \geq b$ for arbitrary $a_1, a_2 \in J$. We have $g(a_1) > g(a_2)$, $g(a_1) + \lambda g(b) > g(a_2) + \lambda g(b)$ and $d_g^\lambda(a_1, b) \geq d_g^\lambda(a_2, b)$.

For a strictly decreasing g the proof runs similarly.

DEFINITION 3

Let g satisfies the assumptions of Theorem 1. Every function defined by (6) will be called a mean generated by pair (g, λ) .

THEOREM 2

Let d_g^λ be a mean on a set T generated by pair (g, λ) . A function $\psi: [0, +\infty) \rightarrow \mathbb{R}^{T \setminus \{(a, a): a \in J\}}$ defined by

$$\psi(\lambda) = \bar{d}_g^\lambda,$$

where $\bar{d}_g^\lambda$ is a restriction of d_g^λ to the set $T \setminus \{(a, a) : a \in J\}$ is strictly decreasing.

Proof. Fix $a, b \in J$ such that $a > b$ and put

$$\phi(\lambda) := \frac{g(a) + \lambda g(b)}{1 + \lambda}, \quad \lambda \in [0, +\infty).$$

It follows that

$$\phi'(\lambda) := \frac{g(b) - g(a)}{(1 + \lambda)^2},$$

thus ϕ is strictly increasing (resp. strictly decreasing) if g is strictly decreasing (resp. strictly increasing). Hence for $\lambda_1 < \lambda_2$ we have $d_g^{\lambda_1}(a, b) > d_g^{\lambda_2}(a, b)$ and $\bar{d}_g^{\lambda_1} > \bar{d}_g^{\lambda_2}$, which completes the proof.

4. Means generated by the identity function

Suppose that $g = \text{Id}_{\mathbb{R}_+}$, where $\text{Id}_{\mathbb{R}_+}(x) = x$ for $x \in \mathbb{R}_+$, then (6) becomes

$$d_{\text{Id}_{\mathbb{R}_+}}^\lambda(a, b) = \frac{a + \lambda b}{1 + \lambda}, \quad (a, b) \in \mathbb{R}_+ \times \mathbb{R}_+.$$

Some of the means of this kind appeared in Problem 1. (problems P1-P6).

Now using Theorem 2 we establish some inequalities between means generated by pair $(\text{Id}_{\mathbb{R}_+}, \lambda)$. Fix $a, b \in \mathbb{R}_+$ such that $a > b$, then

$$\frac{a}{b} > \sqrt{\frac{a}{b}} > 1 > \frac{\sqrt{\frac{a^2+b^2}{2}} + b}{\sqrt{\frac{a^2+b^2}{2}} + a},$$

which yields the following relation between the harmonic, geometric, arithmetic and quadratic mean of a and b ,

$$\frac{2ab}{a+b} < \sqrt{ab} < \frac{a+b}{2} < \sqrt{\frac{a^2+b^2}{2}}.$$

Moreover, the means from problems P5 and P6 are greater than the arithmetic mean. Indeed, solving problem P5 we proved that for $a > b$,

$$\lambda^2 = \frac{2b + d_{\text{Id}_{\mathbb{R}_+}}^\lambda(a, b)}{2a + d_{\text{Id}_{\mathbb{R}_+}}^\lambda(a, b)}$$

which means that

$$\lambda < 1 \quad \text{and} \quad \frac{2a^3 - b^3}{3a^2 - b^2} > \frac{a+b}{2}.$$

By a similar argument, from equality obtained in the solution of problem P6,

$$\lambda^2 = \frac{b + 2d_{\text{Id}_{\mathbb{R}_+}}^\lambda(a, b)}{a + 2d_{\text{Id}_{\mathbb{R}_+}}^\lambda(a, b)}$$

it follows that for $a > b$,

$$\sqrt{\frac{a^3 - b^3}{3(a-b)}} > \frac{a+b}{2}.$$

To end this section let us remark that for arbitrary fixed $a, b \in \mathbb{R}_+$ such that $a > b$ we have

$$\left(\frac{a}{b}\right)^\mu > 1 \quad \text{for } \mu > 0 \quad \text{and} \quad \left(\frac{a}{b}\right)^\mu < 1 \quad \text{for } \mu < 0,$$

thus for $\lambda = \left(\frac{a}{b}\right)^\mu$,

$$d_{\text{Id}_{\mathbb{R}_+}}^\lambda(a, b) = \frac{ab^\mu + ba^\mu}{a^\mu + b^\mu} < \frac{a+b}{2} \quad \text{for } \mu > 0$$

and

$$d_{\text{Id}_{\mathbb{R}_+}}^\lambda(a, b) = \frac{ab^\mu + ba^\mu}{a^\mu + b^\mu} > \frac{a+b}{2} \quad \text{for } \mu < 0.$$

5. Some other generated means

In this section we consider means generated by pair (g, λ) , where $g: \mathbb{R}_+ \rightarrow \mathbb{R}$ is a power or a logarithmic function.

Let $g(x) = x^\nu$, $x \in \mathbb{R}_+$, $\nu \in \mathbb{R} \setminus \{0\}$. Then

$$d_g^\lambda(a, b) = \left(\frac{a^\nu + \lambda b^\nu}{1 + \lambda} \right)^{\frac{1}{\nu}}, \quad (a, b) \in \{(x, y) : x \in \mathbb{R}_+, x \geq y\}.$$

By Theorem 2 it follows that for $a > b$ and $\mu > 0$ we have $\left(\frac{a}{b}\right)^\mu > 1$ and

$$d_g^\lambda(a, b) < d_g^1(a, b),$$

hence for $\lambda = \left(\frac{a}{b}\right)^\mu$,

$$\left(\frac{a^\nu b^\mu + a^\mu b^\nu}{a^\mu + b^\mu} \right)^{\frac{1}{\nu}} < \left(\frac{a^\nu + b^\nu}{2} \right)^{\frac{1}{\nu}}.$$

Similarly, for $\mu < 0$ we get

$$\left(\frac{a^\nu b^\mu + a^\mu b^\nu}{a^\mu + b^\mu} \right)^{\frac{1}{\nu}} > \left(\frac{a^\nu + b^\nu}{2} \right)^{\frac{1}{\nu}}.$$

Now suppose that $g(x) = \ln x$, $x \in \mathbb{R}_+$. We have

$$d_g^\lambda(a, b) = \exp \frac{\ln a + \lambda \ln b}{1 + \lambda} = (ab^\lambda)^{\frac{1}{1+\lambda}}, \quad (a, b) \in \{(x, y) : x \in \mathbb{R}_+, x \geq y\}.$$

Setting again $\lambda = \left(\frac{a}{b}\right)^\mu$, $\mu \in \mathbb{R} \setminus \{0\}$ we get

$$d_g^\lambda(a, b) = a^{\frac{b^\mu}{a^\mu + b^\mu}} b^{\frac{a^\mu}{a^\mu + b^\mu}}$$

and from Theorem 2 the following inequalities

$$a^{\frac{b^\mu}{a^\mu + b^\mu}} b^{\frac{a^\mu}{a^\mu + b^\mu}} < \sqrt{ab} \quad \text{for } \mu > 0,$$

$$a^{\frac{b^\mu}{a^\mu + b^\mu}} b^{\frac{a^\mu}{a^\mu + b^\mu}} > \sqrt{ab} \quad \text{for } \mu < 0.$$

Notice that every strict inequality obtained by Theorem 2 for $(a, b) \in \{(x, y) \in \mathbb{R}_+ \times \mathbb{R}_+ : x \geq y\}$ if replaced by a its corresponding non-strict inequality holds true for $(a, b) \in \mathbb{R}_+ \times \mathbb{R}_+$.

References

- Aczél, J.: 1948, On mean values, *Bull. Amer. Math. Soc.* **54**(4), 392-400.
- Aczel, J., Dhombres, J.: 1989, *Functional equation in several variables*, Vol. 31, Cambridge Univ. Press, Cambridge-New York-Rochelle-Melbourne-Sydney.
- Galwani, L.: 1927, Dei limiti a cui tendono alcune media, *Boll. Un. Math. Ital.* **6**, 173-179.

- Głazowska, D., Jarczyk, W., Matkowski, J.: 2002, Arithmetic mean as a linear combination of two quasi-arithmetic means, *Publ. Math. Debrecen* **61**, 455-467.
- Górowski, J., Łomnicki, A.: 2010, O średnich, *Ann. Univ. Paed. Cracov. Studia ad Didacticam Math. Pertinentia* **3**, 55-66.
- Kitagawa, T.: 1934, On some class of weighted means, *Proc. Phys.-Math. Soc. Japan* **16**(3rd series), 117-126.
- Kołgomorov, A.: 1930, Sur la notion de la moyenne, *Atti Accad. Naz. Lincei* **12**(6), 388-391.
- Leach, E., Sholander, M.: 1978, Extended mean values, *Amer. Math. Monthly* **85**, 84-90.
- Leach, E., Sholander, M.: 1983, Extended mean values ii, *J. Math. Appl.* **92**, 207-223.
- Witkowski, A.: 2009, Comparison theorem for two-parameter means, *Math. Inequal. Appl.* **12**, 11-20.

*Instytut Matematyki
Uniwersytet Pedagogiczny
ul. Podchorążych 2
PL-30-084 Kraków
e-mail jangorowski@interia.pl
e-mail alomnicki@poczta.fm*

Agata Hoffmann

Wykorzystanie prac Stanisława Dróždźa jako środków dydaktycznych w nauczaniu*

Abstract. As far as teaching practice in mathematics is concerned, we use different methods, forms and tools, which enable pupils to acquire both theoretical knowledge and practical skills more efficiently. Stanisław Dróždź (1939-2007) was a concrete poet, who, in his work, used not only words, but also visual art. Poetry, art and mathematics are apparently distant domains, yet when I saw Dróždź's works, I noticed the opportunity to use them as original tools in teaching mathematics. I will show his chosen works and different ways of using them in mathematical education. Some of them could be used as illustrations of various concepts (for example at early stages of maths education). Others could be starting points in discovering and exploring some regularities. There are also some which could inspire pupils to look into certain mathematical problems. Whilst creating his concept-shapes, Stanisław Dróždź used similarities and differences of various situations.

Searching for similarities and differences using analogies and contrast also makes it possible to build mathematical concepts in pupils' minds. By putting together concept and shape, Dróždź worked in a way similar to practice in mathematical education - while defining concepts, we give them a name and (very often) a symbol, and we try to visualize them by giving them a "shape". Using Stanisław Dróždź's works, it is possible to build yet another bridge between theory and practice, particularly for pupils to whom maths is not a life-long passion.

1. Wstęp

Środkiem dydaktycznym nazywamy: „przedmioty materialne umożliwiające usprawnienie procesu nauczania-uczenia się i uzyskanie optymalnych osiągnięć szkolnych” (Okoń, 2007) lub „przedmioty, które dostarczają uczniom określonych

*Stanisław Dróždź's artworks as teaching aid in mathematics education

2010 Mathematics Subject Classification: Primary: 97A80; Secondary: 97A20, 00A66

Key words and phrases: theory, practice, teaching aids, working methods

bodźców sensorycznych oddziałujących na ich wzrok, słuch, dotyk itd., ułatwiając im bezpośrednie i pośrednie poznanie rzeczywistości” (Kupisiewicz, 2009). Oba sformułowania odwołują się do obiektów, które użyte są tak, że usprawniają nabywanie nowej wiedzy czy umiejętności lub pogłębiają zrozumienie już posiadanych. W praktyce nauczania nie sposób odłączyć przedmiotu od funkcji, którą ma pełnić. Wymieniane są następujące funkcje środków dydaktycznych występujące w procesie nauczania-uczenia się: „1) *upogłdowanie procesu kształcenia*, 2) *ułatwienie procesów myślowych*, 3) *pomoc w wykonywaniu przez uczniów ćwiczeń w zdobywaniu sprawności praktycznego działania*, 4) *eksponowanie materiałów wywołujących przeżycia uczniów*” (Okoń, 2007). Obecnie dysponujemy szerokim wachlarzem środków dydaktycznych dostępnych w specjalistycznych ofertach i to z każdej z wymienianych ich kategorii (słownych, wzrokowych, technicznych-wzrokowych, słuchowych i wzrokowo-słuchowych, a także automatyzujących proces dydaktyczny) (Okoń, 2007). Warto jednak nie zapominać o własnej kreatywności i w procesie nauczania wykorzystywać obiekty, z którymi się zetknęliśmy, jeśli dostrzeżemy w nich możliwość sensownego użycia jako środków dydaktycznych. Tak było w przypadku prac Stanisława Dróżdża, które miałam możliwość poznać. Zarysuję niektóre, dostrzeżone przeze mnie, możliwości zastosowania wybranych jego prac w nauczaniu matematyki.

2. Upogłdowanie procesu kształcenia

Pierwszą z wymienianych funkcji środków dydaktycznych jest *upogłdowanie procesu kształcenia*.



Ilustracja 1. *między* (za szybami) / *between* (behind the glazings), 1994, Muzeum Współczesne / Contemporary Museum, Wrocław © Anna Dróżdż

Uczniowie z poziomu edukacji wczesnoszkolnej mają zazwyczaj problemy z opisywaniem wzajemnych relacji pomiędzy obiektami. Zademonstrowanie i omówienie z uczniami niektórych pojęciokształtów stworzonych przez Stanisława Dróżdża, np. *między* (za szybami) / *between* (behind the glazings), 1994 (ilustracja 1) może być punktem wyjścia do samodzielnych, kreatywnych działań uczniów, by idąc za ideą

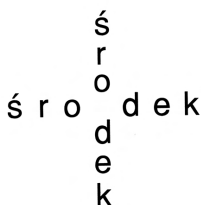
poety spróbować stworzyć „samorealizujące się” (Łubowicz, 2014) obiekty (np. ilustracja 2). Na wystawie stałej pokazującej prace Stanisława Dróždza w Muzeum Współczesnym Wrocław można zobaczyć więcej tego typu prac.

ZANIKANIE

Ilustracja 2. Pomysł i wykonanie Agata Hoffmann

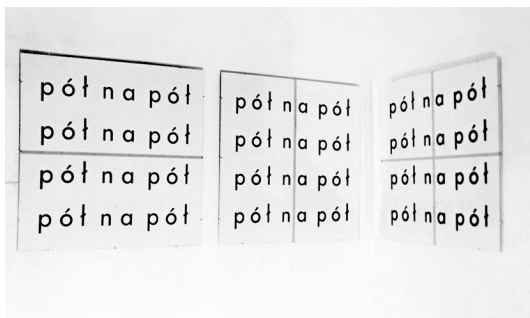


Ilustracja 3. Bez tytułu (trójkąt) / untitled (triangle), 2006 © Anna Dróždź
(<http://www.drozd.art.pl/> / <http://www.drozd.art.pl/>)

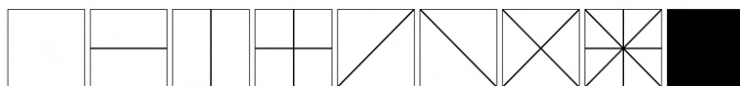


Ilustracja 4. Bez tytułu (środek) / untitled (center), 1998 © Anna Dróždź
(<http://www.drozd.art.pl/>)

Uczniowie będący na innych poziomach edukacyjnych mogą spróbować wykorzystać ideę opisu kształtu obiektu, używając jego nazwy i to zarówno w sposób „dosłowny” - jak to zostało przedstawione w pracy bez tytułu (trójkąt) / untitled (triangle), 2006 (ilustracja 3), jak i abstrakcyjny - jak w pracy bez tytułu (środek) / untitled (center), 1998 (ilustracja 4), gdzie choć „środek” nie został „narysowany”, nie mamy wątpliwości co do jego istnienia.



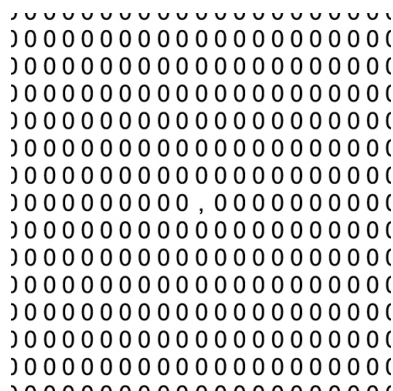
Ilustracja 5. Bez tytułu (pół na pół) / untitled (half by half), 1998,
Galeria Potocka / Potocka Gallery © Anna Dróždź (<http://www.drozd.art.pl/>)



Ilustracja 6. *Sytuacja semiotyczna / Semiotic situation*, 2006

© Anna Drózdź (<http://www.drozd.art.pl/>)

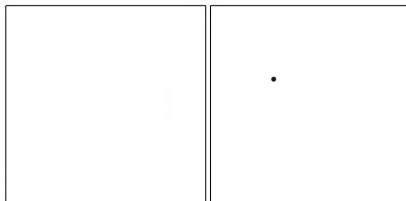
Poeta w pracy bez tytułu (*pół na pół*) / *untitled (half by half)*, 1998 (ilustracja 5) pokazał, jak całość (cztery razy napisane „pół na pół”) została najpierw na dwa różne sposoby podzielona na pół i jak z tego wyszło „pół na pół”, czyli ćwierć. Również tę ideę możemy dostrzec w części pracy *Sytuacja semiotyczna / Semiotic situation*, 2006 (ilustracja 6). Drugi i trzeci kwadrat (licząc od lewej strony) przedstawione w tym dziele są podzielone na pół odcinkami zawartymi w osiach symetrii kwadratu przechodzącymi przez środki jego odpowiednich boków, a czwarty pokazuje nałożenie tych dwóch sytuacji. Analogicznie, piąty i szósty kwadrat są podzielone na pół jego przekątnymi, a siódmy pokazuje nałożenie tych dwóch sytuacji. Wykorzystując tę pracę możemy pójść dalej, bo ósmy kwadrat możemy potraktować jako nałożenie sytuacji z czwartego i siódmego kwadratu (i możemy poprosić uczniów o interpretację tego, co w ten sposób otrzymamy). Zastanawiając się nad rolą pierwszego i ostatniego kwadratu w tej sekwencji możemy spróbować znaleźć interpretację całości bez żadnego podziału (kwadrat biały) i całości z np. nieskończenie wieloma podziałami (kwadrat czarny).



Ilustracja 7. *Continuum*, 1973 © Anna Drózdź (<http://www.drozd.art.pl/>)

W procesie kształcenia często spotykamy się z sytuacjami, w których uczniowie napotykają trudności i to nie tylko dotyczące rozumienia samych pojęć, ale i sensownego ich używania. Gdy poznają ułamki zapisane w postaci dziesiętnej, bardzo często nie wiedzą, które to zera mogą „bezkarnie” opuścić, a które muszą zostawić, by nie zmienić wartości liczby. Praca *Continuum*, 1973 (ilustracja 7) może służyć jako punkt wyjścia do rozmowy na ten temat. Tytuł sugeruje, że mamy do czynienia z czymś, co trwa. I rzeczywiście, przedstawiony, „źle wykadrowany” (obcięty „przez” obiekty) obraz sugeruje, że mamy do czynienia tylko z fragmentem pewnej większej całości, która... rozciąga się w nieskończoność. Bazując na doświadczeniu, że gdy zapisujemy dużą liczbę i jej zapis nie mieści się nam w jednej linijce, to kontynuujemy jej zapis w następnej linijce, możemy przyjąć, że przedstawiony

zapis przedstawia liczbę. Przyjmując, że w każdą stronę, w nieskończoność „ciągną się” te same znaki (czyli 0), to choć zapisana jest cała płaszczyzna, to wartość przedstawionej tu liczby to... 0. Rozmowa dotycząca interpretacji tego dzieła realizuje też czwartą funkcję, którą mają spełniać środki dydaktyczne - wywoływanie przeżyć uczniów. Rozmowa na temat tego, co by było, gdyby np. *nie* wszystkie znaki przed przecinkiem były zerami, czy też nie wszystkie znaki po przecinku *byłyby* zerami, dopełniają treści merytorycznych dotykanego zagadnienia.



Ilustracja 8. Bez tytułu (punkt) / untitled (point), 2006 © Anna Dróždź
(<http://www.drozd.art.pl/>)

Kontrastowe zestawienia bardzo często wykorzystujemy w nauczaniu. W pracy bez tytułu (punkt) / untitled (point), 2006 (ilustracja 8) możemy mieć do czynienia z takim kontrastowym zestawieniem, ale też możemy przeanalizować je głębiej. Jeśli zadamy pytania, np. *Czy rzeczywiście po lewej stronie nie ma punktów? Czy po prawej jest tylko jeden punkt? Czy jest to tylko wyróżnienie? Dlaczego jest ono takie duże?*, to praca ta może nam posłużyć jako punkt wyjścia do głębszego zrozumienia pojęcia punktu.



Ilustracja 9. Bez tytułu (koło) / untitled (circle), 1971, Galeria Foksal /
Gallery Foksal, Warszawa (Stanisław Dróždź *Pomysły*) © Anna Dróždź

Innego rodzaju kontrast został wykorzystany w pracy bez tytułu (koło) / untitled (circle), 1971 (ilustracja 9). Zestawienie zaprezentowanego kształtu różnego od tego, które opisuje użyte słowo, przy jednoczesnym takim zapisie tego słowa, by czytający wykonał ruch kreślący zarys opisywanego obiektu może być początkiem rozmowy na temat istoty pojęcia koła.

3. Ułatwienie procesów myślowych

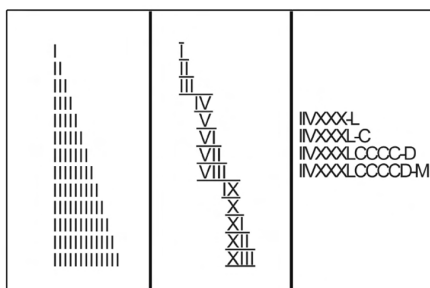
Drugą z wymienianych funkcji środków dydaktycznych jest *ułatwienie procesów myślowych*.

Przedstawię tu trzy prace Stanisława Drózdza, które można wykorzystać do przeprowadzenia logicznych rozumowań.

	białe
białe	czarne
czarne	

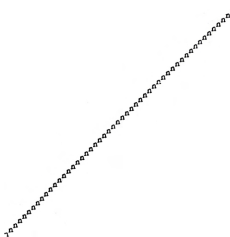
Ilustracja 10. Bez tytułu (białe - czarne) / untitled (white - black), 1970
© Anna Drózdź (<http://www.drozdz.art.pl/>)

Pierwsza z nich to praca bez tytułu (białe - czarne) / untitled (white - black), 1970 (ilustracja 10) z postawionymi pytaniami: *Czy w prostokątach znajdujących się w lewym górnym oraz w prawym dolnym rogu jest coś napisane? Jeśli tak, to co i dlaczego? Jeśli nie, to dlaczego?*



Ilustracja 11. Bez tytułu (cyfry rzymskie) / untitled (Roman numerals), 2006
© Anna Drózdź (<http://www.drozdz.art.pl/>)

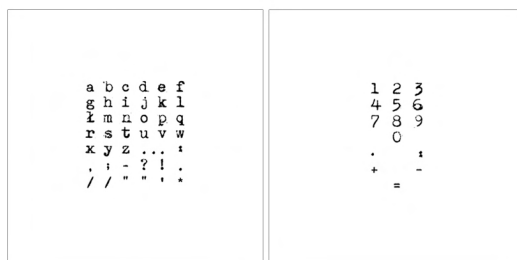
Druga to praca bez tytułu (cyfry rzymskie) / untitled (Roman numerals), 2006 (ilustracja 11) z pytaniami: *Czy trzy części tego dzieła są ze sobą jakoś powiązane? Jeśli tak, to jak? Jeśli nie, to dlaczego?*



Ilustracja 12. Bez tytułu (potęgowanie) / untitled (exponentiation), 1981
© Anna Drózdź (<http://www.drozdz.art.pl/>)

Trzecia praca to bez tytułu (potęgowanie) / untitled (exponentiation), 1981 (ilustracja 12), a pytania to: *Co może oznaczać ten zapis? Jak jest związany z potęgowaniem? Jak zmieni się znaczenie tego zapisu, gdy zmienimy zbiór, do którego należy n ? Dlaczego?*

Każda z przedstawionych tu sytuacji, choć odwołuje się do innych zagadnień (zauważania analogii w ułożeniach sekwencyjnych, zapisu liczb w systemie rzymskim czy zapisu potęgi), pozwala na przedstawienie logicznego rozumowania.



Ilustracja 13. *Język i matematyka / Language and Mathematics*, 1968

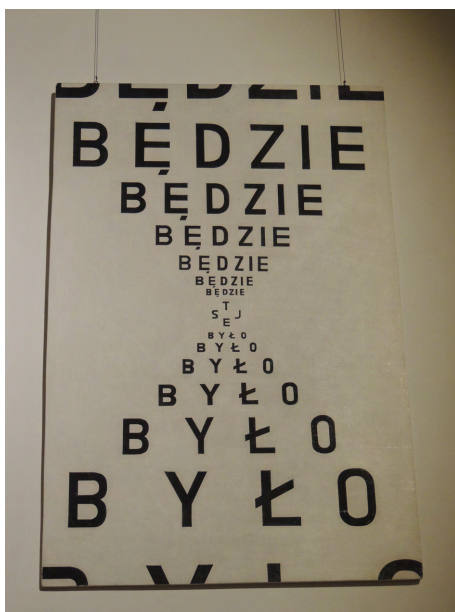
© Anna Dróždź (<http://www.drozdz.art.pl/>)

Do wyrażenia tego, co nas nurtuje, używamy symboli. Ostatnie dzieło, które chcę w tej części przedstawić, to *Język i matematyka / Language and Mathematics*, 1968 (ilustracja 13). Autor zestawiał symbole używane w języku i w matematyce. Przeanalizowanie tego zestawienia, jak i każdej z części osobno, może pozwolić uczniom na zastanowienie się np. nad tym: *Czy wszystkie wypisane tu znaki pełnią takie same funkcje? Jeśli tak, to jaka ona jest? Jeśli nie, to czym się one różnią? Czy zostały tu wymienione wszystkie znane ci, używane w danej dziedzinie symbole? Jeśli nie, to jakie zostały pominięte i dlaczego? Czy istnieją takie symbole, które występują w obu częściach? Jeśli tak, to czy oznaczają to samo?* Po przeprowadzonej analizie (która na każdym poziomie edukacyjnym będzie przebiegała inaczej), używane symbole i konteksty ich użycia staną się dla uczniów czytelniejsze. Rozmowę związaną z tym dziełem możemy też potraktować jako punkt wyjścia do zbadania historii przedstawionych symboli.

4. Ćwiczenie sprawności praktycznego działania

Trzecią z wymienianych funkcji środków dydaktycznych jest *pomoc w wykonywaniu przez uczniów ćwiczeń w zdobywaniu sprawności praktycznego działania*. *Praktyczne działanie* możemy rozumieć na wiele sposobów. Między innymi może to być wykorzystanie matematyki w praktyce życia codziennego (prywatnego czy też zawodowego), może to być dostrzeżenie matematyki w tym, co nas otacza (w dziełach natury czy wytworach pracy rąk ludzkich), a może to być też użycie fizycznych aktywności do zbadania problemu matematycznego.

W tej części przedstawię wybrane prace Stanisława Dróždza, w których można dostrzec wykorzystanie matematyki. Nauczyciel może wykorzystać je do postawienia zadania i (w zależności od sposobu użycia) do jego rozwiązania.



Ilustracja 14. *Klepsydra (było, jest, będzie) / The Hourglass (it was, it is, it will be)*, 1967, Muzeum Narodowe / National Museum, Wrocław
© Anna Dróżdż

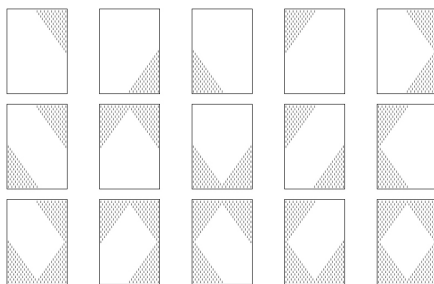
Praca *Klepsydra (było, jest, będzie) / The Hourglass (it was, it is, it will be)*, 1967 (ilustracja 14) znajduje się w Muzeum Narodowym we Wrocławiu. Rozmowa dotycząca interpretacji tego dzieła realizuje czwartą funkcję, którą mają spełniać środki dydaktyczne - wywoływanie przeżyć uczniów. Autor, bardzo prosto i trafnie, pokazuje w niej, jak znikomy jest ten fragment rzeczywistości, w którym żyjemy, a jednocześnie jak jest on wyjątkowy. Zadanie matematyczne dotyczy kolejnych podejść poety do tej pracy.



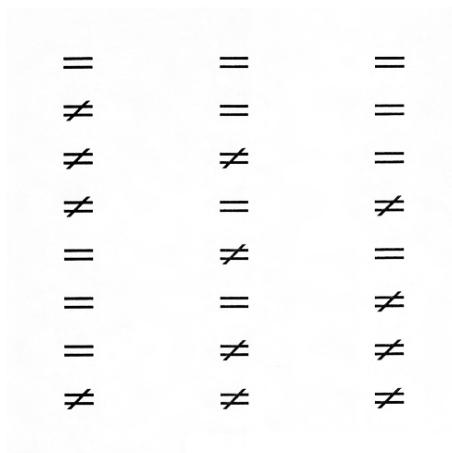
Ilustracja 15. *Klepsydra (było, jest, będzie) / The Hourglass (it was, it is, it will be)*, 1967-1990, fragment pracy złożonej z 54 plansz, Muzeum Współczesne / Contemporary Museum, Wrocław © Anna Dróżdż

W Muzeum Współczesnym Wrocław możemy zobaczyć to dzieło w postaci rozbudowanej (ilustracja 15). Dodając do tego przedstawienia pytanie: *Dlaczego*

cały przedstawiony cykl prac składa się z 54 plansz? mamy bardzo ciekawy problem kombinatoryczny. Praktyczne podejście do teoretycznej sytuacji matematycznej realizuje się tu według mnie poprzez użycie konkretnej pracy artystycznej, jako sytuacji wyjściowej, pozwalającej zastosować i uzasadnić użycie odpowiedniego modelu teoretycznego - *Który model opisuje sposób znalezienia liczby prac w całym cyklu (permutacji, kombinacji czy wariacji, z powtórzeniami czy bez)?*



Ilustracja 16. *"i" (fragments)/ "and" (fragments)*, 1970-1997
© Anna Dróždź (<http://www.drozd.art.pl/>)



Ilustracja 17. *Bez tytułu (równa się, nie równa się) / untitled (equal to, not equal to)*, 1971-1972 © Anna Dróždź (<http://www.drozd.art.pl/>)



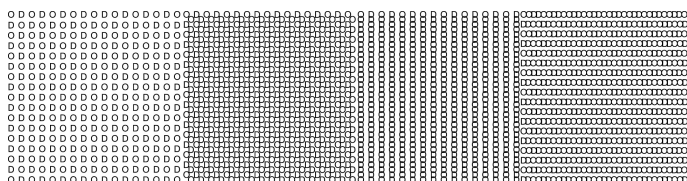
Ilustracja 18. *Alea iacta est*, 2003, 50. Biennale w Wenecji, Pawilon Polski
© Anna Dróždź (<http://www.drozd.art.pl/>)

gra w kości zasady według Stanisława Drózdza

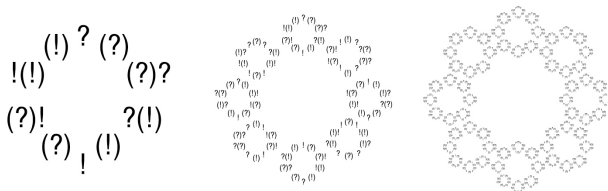
jest 46.656 możliwości układu 6 kostek
wszystkie kombinacje zostały umieszczone na ścianie
weź kości leżące na stole i rzuć
zestaw je razem w jednym rzędzie
spróbuj odnaleźć ten układ na ścianie
jeżeli Ci się to uda - wygrasz, jeżeli nie - przegrywasz,
wygrałem ja

Ilustracja 19. Instrukcja - część pracy *Alea iacta est*, 2003,
50. Biennale w Wenecji, Pawilon Polski © Anna Drózdź
(<http://www.drozd.art.pl/>)

W tym zakresie możemy wykorzystać jeszcze inne prace Stanisława Drózdza, jak np. *i (fragments)/and (fragments)*, 1970-1997 (ilustracja 16) czy bez tytułu (równa się, nie równa się) / *untitled (equal to, not equal to)*, 1971-1972 (ilustracja 17), ale kulminacją jest praca *Alea iacta est*, 2003 (ilustracja 18), z którą autor reprezentował Polskę na 50. Biennale w Wenecji. Tę ostatnią pracę możemy wykorzystać na dwa sposoby. Tak jak omawianą wcześniej *Klepsydre*, podając instrukcję (ilustracja 19), którą autor zamieścił na wystawie w wielu językach. Można też (nie podając instrukcji) zapytać: *Iloma kośćmi do gry zostały wyłożone ściany tego pawilonu, jeśli każde z możliwych wyników rzutu sześcioma klasycznymi kośćmi do gry można tam było znaleźć?* Rozważając różne zasady, których można użyć organizując „wykładanie”, odpowiedź nie musi być jedna.



Ilustracja 20. *Czasoprzestrzennie (OD-DO) / Temporally - Spatially (FROM-TO)*, 1969-1993, fragment pracy złożonej z 82 plansz
© Anna Drózdź (<http://www.drozd.art.pl/>)



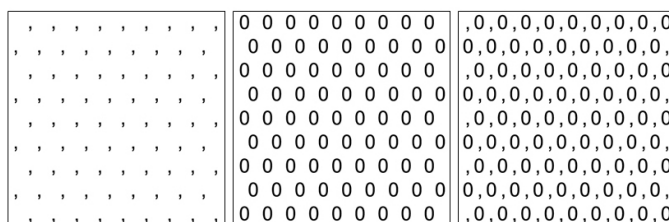
Ilustracja 21. Bez tytułu (wykrzykniki i znaki zapytania) / *untitled (exclamation, question marks)*, 2001, fragment pracy złożonej z 13 plansz
© Anna Drózdź (<http://www.drozd.art.pl/>)

Prace *Czasoprzestrzennie (OD-DO) / Temporalny - Spatially (FROM-TO)*, 1969-1993 (ilustracja 20) oraz bez tytułu (wykrzykniki i znaki zapytania) / *untitled (exclamation, question marks)*, 2001 (ilustracja 21) możemy wykorzystać do

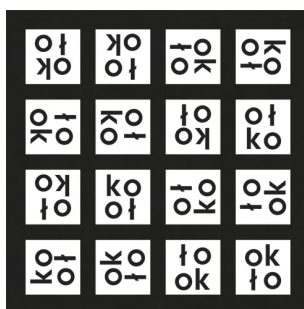
znajdowania i opisywania przekształceń geometrycznych, z którymi mielibyśmy do czynienia, gdybyśmy chcieli przekształcić jedną pracę w drugą. Tu już, w zależności od wieku i umiejętności uczniów, możemy praktyczne podejście do teoretycznych sytuacji matematycznych ograniczyć do znalezienia odpowiedniego przekształcenia i opisanie go znaną symboliką lub też, wyposażając uczniów w odpowiednie kopie np. na foliach, fizycznie pozwolić im znaleźć odpowiednie przekształcenia i spróbować je opisać.



Ilustracja 22. *Triptyk* (Niepewność - Wahanie - Pewność) / *Triptych* (Uncertainty - Hesitation - Certainty), 1967 © Anna Dróżdż (<http://www.drozdz.art.pl/>)



Ilustracja 23. Bez tytułu (przecinki, zera) / untitled (commas, zeros), 2006 © Anna Dróżdż (<http://www.drozdz.art.pl/>)



Ilustracja 24. Bez tytułu (koło) / untitled (circle), 1971-1972, © Anna Dróżdż, Stanisław Dróżdż, początekoniec. *Pojęciokształty. Poezja konkretna. Prace z lat 1967-2007*

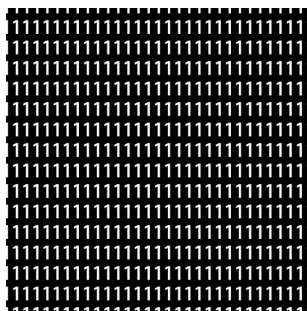
Do tego celu możemy też wykorzystać inne prace poety np. *Triptyk* (Niepewność - Wahanie - Pewność) / *Triptych* (Uncertainty - Hesitation - Certainty), 1967 (ilustracja 22), bez tytułu (przecinki, zera) / untitled (commas, zeros), 2006 (ilustracja 23) czy też bez tytułu (koło) / untitled (circle), 1971-1972 (ilustracja 24).

Jedna z wersji ostatniej z wymienionych prac znajduje się w przestrzeni miejskiej we Wrocławiu, więc rozważania można przeprowadzić w plenerze, mając pracę przed sobą.

5. Ekspozowanie materiałów wywołujących przeżycia uczniów

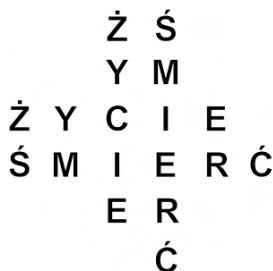
Czwartą z wymienionych funkcji środków dydaktycznych jest *ekspozowanie materiałów wywołujących przeżycia uczniów*.

O realizacji tej funkcji z wykorzystaniem prac Stanisława Dróždza już wspominałam przy okazji wcześniej omawianych dzieł *Continuum*, 1973 (ilustracja 7) i *Klepsydra* (było, jest, będzie) / *The Hourglass* (it was, it is, it will be), 1967 (ilustracja 14), odnieść ją można również do wspomnianej pracy *Tryptyk* (Niepewność - Wahanie - Pewność) / *Triptych* (Uncertainty - Hesitation - Certainty), 1967 (ilustracja 22), ale wspomnę jeszcze o dwóch utworach, w których widzę wykorzystanie istoty obiektów matematycznych do opisu sytuacji życiowych.



Ilustracja 25. *Samotność / Loneliness*, 1967 © Anna Dróždź
(<http://www.drozdz.art.pl/>)

Pierwsza praca to *Samotność / Loneliness*, 1967 (ilustracja 25). Widzimy w niej, jak liczba 1, która z jednej strony reprezentuje całość, ale z drugiej strony oznacza również osobność, poprzez powielenie jej z zachowaniem wyraźnej spacji, staje się dziełem - wspaniałą interpretacją tytułowej samotności (wyalienowania, braku komunikacji).



Ilustracja 26. *Bez tytułu (życie - śmierć) / untitled (life - death)*, 1969
© Anna Dróždź (<http://www.drozdz.art.pl/>)

Druga praca to *bez tytułu (życie - śmierć) / untitled (life - death)*, 1969 (ilustracja 26). Wyrazy opisujące oba stany zaczynają się równocześnie - wskazuje

to na jednoczesność istnienia ich obu. Ale oba wyrazy zapisane są w ten sposób, że wyznaczają nie pokrywające się, proste równoległe – co z kolei wskazuje na ich rozłączność, niezależność. Dodatkowo autor zapisał ten stan używając dwóch różnych kierunków - pionowego i poziomego. Tym samym, zostały zasugerowane kąty proste, które mogą symbolizować przecinanie się losów różnych istnień, ale w bardzo „uporządkowany” sposób (proste prostopadłe jako jedyne dzielą płaszczyznę na cztery równe części). Jednocześnie perspektywa, z której możemy tę pracę oglądać, zmienia to, co widzimy (a nie jest to bez znaczenia). Gdy przyglądamy się jej bardzo blisko, jesteśmy w stanie dostrzec tylko wybrane fragmenty. W miarę zwiększania odległości, coraz lepiej ogarniamy całość, dostrzegamy zależności. Z dostatecznie dużej odległości widzimy już tylko zarys dwóch prostych prostopadłych.

Na zaprezentowanych przykładach widać, że środkami dydaktycznymi mogą okazać się różne obiekty. Dla osób odwiedzających Wrocław może być to o tyle cenne, że dwie plenerowe realizacje prac poety możemy zobaczyć spacerując po mieście: jeden z wariantów *Klepsydry* - na gmachu Muzeum Współczesnego Wrocław oraz bez tytułu (koło) - na placu Nowy Targ przed Urzędem Miejskim.

Podziękowania

Składam serdeczne podziękowania pani Annie Dróżdż za udostępnienie możliwości skorzystania ze zdjęć prac Stanisława Dróżdża i skontaktowanie z osobami zajmującymi się jego pracami oraz pani Ewie Trojanowskiej za udzielenie cennych wskazówek związanych z opisem prac i historią ich powstawania.

Literatura

- Dawidek-Gryglicka, M.: 2012, *Odprysk poezji. Stanisław Dróżdż mówi / A Piece of Poetry. Conversations with Stanisław Dróżdż*, Korporacja Ha!art - Narodowe Centrum Kultury, Kraków.
- Kupisiewicz, C.: 2012, *Dydaktyka. Podręcznik akademicki*, IMPULS, Kraków.
- Kupisiewicz, C., Kupisiewicz, M.: 2009, *Słownik pedagogiczny*, Wydawnictwo Naukowe PWN, Warszawa.
- Okoń, W.: 1998, *Wprowadzenie do dydaktyki ogólnej*, Wydawnictwo Akademickie „Żak”, Warszawa.
- Okoń, W.: 2007, *Nowy słownik pedagogiczny*, Wydawnictwo Akademickie „Żak”, Warszawa.
- Trojanowska, E., (red.): 2014, *Stanisław Dróżdż, Pomysły / Ideas*, Muzeum Współczesne Wrocław.
- Łubowicz, E.: 2009, Rzeczywistość jest tekstem. O pojęciokształatach Stanisława Dróżdża / Reality is a Text. On 'Concept-Shapes' by Stanisław Dróżdż, w: *Stanisław Dróżdż, początekoniec. Pjęciokształty. Poezja konkretna. Prace z lat 1967-2007 / beginend. Concept-Shapes. Concrete Poetry. Works 1967-2007*, Ośrodek Kultury i Sztuki we Wrocławiu, Fine Grain, Wrocław.
- Łubowicz, E.: 2014, *Matematyka w poezji Dróżdża*, skrót wykładu *Poezja konkretna Stanisława Dróżdża* wygłoszonego w 2010 roku na posiedzeniu Studium Generale

UWr. Pełna wersja została opublikowana w t. XV materiałów Studium pt. *Człowiek, kultura, historia*. Przedruk za zgodą Komitetu Naukowego SG i autorki.
<http://www.matematyka.wroc.pl/doniesienia/matematyka-w-poezji-drozdza>.

*Instytut Matematyczny
Uniwersytet Wrocławski
pl. Grunwaldzki 2/4
PL-50-384 Wrocław
e-mail Agata.Hoffmann@math.uni.wroc.pl*

*Taras Kudryk***Introduction to E. Nelson's internal set theory***

Abstract. An axiomatic approach to Non-standard Analysis by E. Nelson is presented in a simplified form. The main aim of the article is strictly the popularization of NSA, and not its foundations. No special preparation in mathematical logic is required from the reader but it is assumed that he (she) is familiar with elementary calculus and linear algebra.

Introduction

Non-standard Analysis is an enlargement of Standard or Ordinary Mathematics. It provides a strictly logical foundation of Leibniz's infinitesimals. It enables us to better understand and simplify a lot of Standard Mathematics. It also advances new mathematical problems.

In Non-standard Analysis (NSA in short), the old problem of the substantiation of differential and integral calculus with the application of infinitesimals was solved. This problem seemed to be unsolvable in the times of G. Leibniz and L. Euler. NSA has changed the face of the whole of Mathematics: it is a new mathematical outlook. It is necessary to emphasise that NSA does not object or contradict the Ordinary Mathematics (OM in short). NSA extends and supplements OM. This means that all objects which exist in OM also exist in NSA, and all statements which are true in OM remain true in NSA. NSA often simplifies OM and makes it more transparent. NSA states new mathematical theorems and problems.

NSA was created mainly by A. Robinson (1960), who developed H. Hahn's, T. Skolem's, A. Malcev's, E. Hewitt's, and J. Łoś's ideas. Afterwards, W.A.J. Luxemburg proposed an ultrafilter approach. In this article, we present an axiomatic approach to NSA, due to E. Nelson (Nelson, 1977), which is less difficult to learn and apply. Our exposition, contrary to that of Nelson, is not always

*Wprowadzenie do teorii zbiorów wewnętrznych E. Nelsona

2010 Mathematics Subject Classification: Primary: 03H05; Secondary: 26E35

Key words and phrases: Nonstandard Models in Analysis, Internal Set Theory

strictly logical. Our aim is strictly the popularization of NSA, and not its foundations. Other axiomatic approaches were presented in the seventies by K. Hrbáček (Hrbáček, 1979, 1987) and T. Kawai (Kawai, 1981). In our opinion, Nelson's approach is the best for many applications.

Interested readers may find a continuation of this presentation in the same spirit in (Lyantse, Kudryk, 1997). There we provided a more complete picture of NSA. Good supplements to the book are (Albeverio, Fenstad, Høegh-Krohn, Lindstrøm, 1986, Cutland, 1983, Cutland, 1988, Davis, 1977, Diener, 1983, Diener, Reeb, 1989, Goldblatt, 1998, Lutz, Goze, 1981). In the references, we also listed other books which influenced our thinking. Furthermore, we would recommend looking at more recent publications such as (Hrbáček, Lessman, O'Donovan, 2015) and (Kanovei, Reeken, 2004). The latter is the most comprehensive source on various modern non-standard axiomatic theories, including IST.

Some exercises are left to the reader (through courtesy of the author).

1. Sets

Let us begin with a remark that all what is determined and investigated in Mathematics can be considered as a set. (We are not going to pursue a class path here.) For instance, the number zero may be identified with the empty set $\emptyset$, 1 can be defined as the one-element set $\{0\}$. Whenever a natural number n is defined as a set, we can consider $n + 1$ as $n \cup \{n\} = \{0, 1, \dots, n\}$. We consider each function and relation as a set by identifying them with their graphs, and so on.

2. Standard, internal, and external sets

The base of OM (= ordinary mathematics = standard mathematics) is ZFC (= Zermelo–Fraenkel Set Theory with the Axiom of Choice); see e.g. (Luxemburg, Robinson, 1972) or (Robert, 1988).

Sets which are uniquely determined in ZFC are said to be **standard**.

Thus $0, 1, 2, \dots, 10^{99}, \dots, 1/2, \sqrt{2}, \pi, \dots, +, <, \dots, \sin, \dots, \mathbb{N}, \mathbb{R}, \mathbb{C}, \dots, C[0, 1], \dots, L_2(\mathbb{R}), \dots$ are all standard.

In NSA, elements of standard sets are said to be **internal**.

Thus, in NSA, if $A \in B$ and B is standard, then A can be nonstandard, but it must be internal. This means that for such A (*for arbitrary internal A*) *all laws of OM hold*. But due to NSA's methods of the construction of subsets (parts) of a standard set, some statements of OM are not true for these subsets.

If $A \subseteq B$ and B is standard, then A is said to be **external**.

If there is a statement which is true for all standard sets but not true for A , then A is said to be **strictly external**. See below for examples.

REMARK 1

If an A is an element of a set B which is standard, internal, or external, then it must be internal. The cause is that each external set is a part of a standard one, and elements of standard sets are internal (which is sometimes repeated for

pedagogical reasons). Furthermore, in order for IST to be a closed theory, we ought to assume that the elements of internal sets are internal.

Denote by $\mathbf{S}$, $\mathbf{I}$, $\mathbf{V}$ universes (i.e. totalities) of standard, internal, and external sets respectively. We have

$$\mathbf{S} \subset \mathbf{I} \subset \mathbf{V}.$$

These universes are not sets, but proper classes. We shall see later that $\mathbf{S} \setminus \mathbf{I}$ and $\mathbf{V} \setminus \mathbf{I}$ are not empty.

REMARK 2

In this article, we accept the following agreement: "A set" always means "an internal set". If X is a set, then 2^X is the (internal) set of all internal parts of X . If X, Y are sets, then Y^X denotes the (internal) set of all functions f with $\text{dom } f = X$ and $\text{ran } f \subset Y$. Exceptional cases are possible in which it is clear from context that a set under consideration is standard or external.

3. Might of the Word; the word "standard"

The Bible asserts that the Word has the ability to create (see *John 1, 1-4*). Mathematics permanently demonstrates this capacity of words. For instance, take the preposition "between." Let a be a straight line, and A, B its different points. Due to the word "between" we can create the segment $AB := \{X \in a: X \text{ is between } A \text{ and } B\}$. Next, we can define the straight ray $r_{AB} := \{X \in a: A \text{ is not between } X \text{ and } B\}$ and convert r_{AB} into an ordered set by $(\forall X, Y \in r_{AB})(X < Y \iff X \text{ is between } A \text{ and } Y)$, and so on.

We can also notice that the same word plays a fundamental role in J. Conway's theory of surreal numbers (see Conway, 1976). Unfortunately, the absolute arithmetical continuum created by Conway is unfitting for the needs of the analysis for the time being.

EXERCISE 1

Let a be a straight line of the plane Σ . Define with the help of "between" two halfplanes Σ_a^1 and Σ_a^2 , bounded by a . Hint: Use the concept of a segment which is descended from "between".

Without exaggeration, one can say that the entire OM is created by two words: "a set" and "to belong". If we wish to extend the OM, we ought to invent a new word. To this end, Robinson has invented the adjective "standard," which came into general use. It is very important to keep in mind that *OM does not know what is "standard."* Therefore, *the laws of OM do not control the property "to be standard"* and the properties which are derived from it.

Let us consider some simple applications of the notion "standard".

4. Classification of reals

It is convenient and comfortable to agree the following. The formula " $st(x)$ " denotes " $x \in \mathbf{S}$ ", i.e. " x is standard"; " $\forall^{st} x p(x)$ " denotes " $(\forall x)(st(x) \Rightarrow p(x))$ ",

i.e. “for any standard x $p(x)$ ”; “ $\exists^{st} x p(x)$ ” denotes “ $(\exists x) (st(x) \wedge p(x))$ ”, i.e. “for a standard x $p(x)$ ”.

Let $x \in \mathbb{R}$, then “ $x \approx 0$ ” denotes “ $(\forall^{st} n \in \mathbb{N}) (|x| < 1/n)$ ”, i.e. “ x is **infinitesimal**”. For $x, y \in \mathbb{R}$ “ $x \approx y$ ” means “ $x - y \approx 0$ ”, i.e. “ x is near y ”. We write “ $x \approx \infty$ ” for “ $(\forall^{st} n \in \mathbb{N}) (|x| > n)$ ”, and “ $x \approx +\infty$ ” or “ $x \approx -\infty$ ” if, respectively, “ $x \approx \infty \wedge x > 0$,” or “ $x \approx \infty \wedge x < 0$.” If $x \approx \infty$, we say that x is **infinitely large**. At last, notation “ $x \gg 0$ ” and “ $|x| \ll \infty$ ” are equivalent to “ $x > 0 \wedge x \not\approx 0$ ” and “ $x \not\approx \infty$ ” respectively. If $|x| \ll \infty$, we say that x is a **limited** number, and if $0 \ll |x| \ll \infty$, x is an **appreciable** number.

REMARK 3

It can be shown that the continuity of a standard function f at a standard point a is strictly equivalent to the implication “ $x \approx a \implies f(x) \approx f(a)$ ”. This is an example which demonstrates how to use our new concepts.

5. Existence of infinitesimals; the special idealization principle (I_0)

To ensure the existence of infinitesimals, we accept the following axiom:

(I_0) *A standard set is infinite if (and only if) it contains a non-standard element.*

“Non-standard” means “is not standard”. We write “ $\neg st(x)$ ” for “ $x \notin \mathbf{S}$ ”. It can be shown that (I_0) is a special case of the general principle of idealization (I).

COROLLARY 1

There exist non-standard natural, rational, real numbers.

Indeed, the sets $\mathbb{N}$, $\mathbb{Q}$, $\mathbb{R}$ are standard and infinite. The reader can point out many other examples, for instance: there exist non-standard vectors, transformations, groups, spaces, and so on.

It is useful to formulate (I_0) in a more exact way. As it is known, a set E is said to be *finite* if there exist a number $n \in \mathbb{N}$ and a bijective (that is one-to-one) transformation f which sends E onto $\{1, 2, \dots, n\}$. In this case, we write $n = \text{card } E$, and n is said to be the *cardinality* or the *quantity of elements* of E . We also write “ $\text{fin } E$ ” for “ E is finite.” The number $n = \text{card } E$ (which is independent of the choice of f) can be non-standard, but if we write $\text{card } E = n$, we assume that the function f , which is a bijection $E \rightarrow \{1, 2, \dots, n\}$, is *internal*: $f \in \mathbf{I}$ (i.e. its graph is an internal set). We introduce the unary predicate “ fin ” as follows:

$$\text{“fin}(x)\text{”} \equiv \text{“}x \text{ is finite.”}$$

The special principle of idealization is as follows:

$$(\forall x \in \mathbf{S}) (\neg \text{fin}(x) \iff (\exists y \in x) \neg st(y)), \quad (I_0)$$

where $\neg a$ means the negation of a .

Apply to (I_0) the following laws of Logic: $\neg \neg a \equiv a$, $a \Leftrightarrow b \equiv \neg a \Leftrightarrow \neg b$, $\neg(\exists x) p(x) \equiv (\forall x) \neg p(x)$. We get

$$(\forall x \in \mathbf{S}) (\text{fin}(x) \iff (\forall y \in x) st(y)). \quad (I_1)$$

This is the *second form* of the special principle of idealization: *a standard set is finite if and only if all its elements are standard.*

PROPOSITION 1

A natural number is infinite if and only if it is non-standard.

Proof. Let $n \in \mathbb{N}$ be standard. Then $\mathbb{N}_n := \{1, \dots, n\}$ is a well-defined set of OM. Therefore, it is standard. Since $\mathbb{N}_n$ is finite, namely $\text{card } \mathbb{N}_n = n \in \mathbb{N}$, each of its elements is standard (see (I_1)). But $(\forall p \in \mathbb{N}) (p \leq n \iff p \in \mathbb{N}_n)$. Since

$$(\forall n, p \in \mathbb{N}) (st(n) \wedge p < n \implies st(p)),$$

we have

$$(\forall p \in \mathbb{N}) (\neg st(p) \iff (\forall^{st} n \in \mathbb{N}) (p > n)),$$

and we see that

$$(\forall p \in \mathbb{N}) (\neg st(p) \iff p \approx \infty).$$

REMARK 4

For $x \in \mathbb{R}$, the formula $|x| \ll \infty$ does not imply that x is standard. For example, the segment $[0, 1]$ is standard and infinite, therefore (see (I_0)) it contains some non-standard numbers.

EXERCISE 2

Prove that any $n \in \mathbb{Z}$ is limited if and only if it is standard.

PROPOSITION 2

There exist infinitesimal real numbers different from 0.

Proof. Let $\omega \in \mathbb{N}$ and ω be non-standard. Then (see (I_0)) $\omega \approx \infty$, i.e. $(\forall^{st} n \in \mathbb{N}) (\omega > n)$. Therefore $(\forall^{st} n \in \mathbb{N}) (1/\omega < 1/n)$, i.e. $1/\omega \approx 0$. ►

6. The first examples of strictly external sets

In OM, we often use the following subset construction. Let A be a set, and $p(\cdot)$ a property (predicate). Then $B := \{x \in A : p(x)\}$ is the set of all x from A which have the property $p(\cdot)$. In NSA, we also apply this construction. For instance, we set

$$^{st}A := \{x \in A : st(x)\}.$$

Thus ^{st}A denotes the totality of the standard elements of A . According to section 2, if A is standard or internal, ^{st}A is external as a part of A .

Recall that an external set is said to be *strictly external* if it is non-internal. In other words, it is not subordinate to the laws of OM.

PROPOSITION 3

The set $\mathbb{N} \setminus ^{st}\mathbb{N} := \{n \in \mathbb{N} : \neg st(n)\}$ is strictly external.

Proof. In OM, the *least number principle* is known: each non-empty set of natural numbers contains the least number. We have $\mathbb{N} \setminus {}^{st}\mathbb{N} \neq \emptyset$. Set $n := \min(\mathbb{N} \setminus {}^{st}\mathbb{N})$ and suppose that $n \approx \infty$. Then $n - 1$ is also unlimited, therefore, $n - 1 \in \mathbb{N} \setminus {}^{st}\mathbb{N}$, contrary to the definition of n . But the inequality $n \ll \infty$ is also impossible because $n \ll \infty \implies st(n)$, contrary to $n \in \mathbb{N} \setminus {}^{st}\mathbb{N}$. ►

EXERCISE 3

Prove that ${}^{st}\mathbb{N}$ is strictly external. Hint: ${}^{st}\mathbb{N}$ is bounded by each $\omega \approx \infty$, but $\max({}^{st}\mathbb{N})$ does not exist.

REMARK 5

With the help of the word “standard,” we have constructed not only a new kind of numbers but also a new kind of sets, namely strictly external ones. This once more presents the power of words which are happily selected.

WARNING 1

Nelson’s IST knows only internal sets: in IST “a set” means “an internal set.” Thus, the Internal Set Theory does not know what an external set is. For instance, in IST there is no ${}^{st}\mathbb{N}$, ${}^{st}\mathbb{Q}$, ${}^{st}\mathbb{R}$, and so on. What is the ground for manipulation with such (and other) external sets? The matter is that IST knows external formulas. In IST, we don’t understand what ${}^{st}\mathbb{N}$ is, but we understand what $x \in {}^{st}\mathbb{N}$ means. $x \in {}^{st}\mathbb{N}$ is just an abbreviation for $x \in \mathbb{N} \wedge st(x)$. Another example: let A, B be standard sets, and $p(x), q(x)$ be external formulae. External sets $A_p := \{x \in A : p(x)\}$ and $B_q := \{x \in B : q(x)\}$ are outside of IST. Therefore, $A_p \cup B_q$ cannot be formed in IST. But the formula $x \in A_p \cup B_q$ makes sense. It is an abbreviation of $(x \in A \wedge p(x)) \vee (x \in B \wedge q(x))$.

EXERCISE 4

What about $A_p \cap B_q$, $A_p \setminus B_q$, $A_p \times B_q$, $B_q^{A_p}$? Let A be external. What is $\{A\}$?

EXERCISE 5

For any $n \in \mathbb{N}$, denote by $f(n)$ the first $k \in \mathbb{N}$, such that the interval $]\frac{k-1}{n}, \frac{k}{n}]$ contains a standard real number. Why is this definition of the function f not correct in the framework of IST?

7. Formulae

Mathematical formulae express mathematical statements. For instance, formulae “ $5 + 3 = 8$ ”, “ $2 < 1$ ” are sentences (of Arithmetics), the former is *true*, the latter is *false*. The formulae “ $x + 3 = 8$ ”, “ $x < 1$ ” are not sentences, for they are not false or true. But they become sentences if we replace x by a fixed real number. Formulae such as “ $x + 3 = 8$ ”, “ $x < 1$ ” are *predicates*, they express some *property* (for example “ $x < 1$ ” is the property of real numbers “to be less than 1”). The variable x , which occurs in a formula $p(x)$ in which after it is replaced by an (admissible) constant, it results a sentence (true or false), is said to be *free*. Thus, x is a free variable in “ $x + 3 = 5$ ”. If a formula is denoted by $p(x, y)$, it means that there are two free variables. To get a sentence, we ought to replace both x and y

by (admissible) constants. Evidently, the quantity of free variables can be larger. For instance, in the equation $3x - 2y + 5z = 1$ (of a plane in the space $\mathbb{R}^3$) we have three free variables.

From the formula $p(x)$, one can obtain a sentence (true or false) not only by replacing x by a constant, but also by *bounding* it by the quantifier $\forall \equiv$ "for all", or $\exists \equiv$ "for some". Thus, $(\forall x \in \mathbb{R})(x + 3 = 8)$ and $(\exists x \in \mathbb{R})(x + 3 = 8)$ contain x as a *bounded* (apparent) variable. They are sentences: the first is false, the second is true.

In NSA, we deal with *internal* and *external* formulae. The basic external formula is $st(x)$ (read "x is standard" for $st(x)$). Each formula, in which " $st(x)$ " occurs explicitly or implicitly, is also said to be *external*. *Internal* formulae are those in which " $st(x)$ " does not occur in any way (explicitly or implicitly). Obviously, we only mean formulae which are well-constructed by the laws of Mathematics and Logic.

Let us consider some examples. The formulae $x \approx 0$, $|x| \ll \infty$, $x \gg 0$ are all external. For instance, $|x| \ll \infty$ is an abbreviation for $(x \in \mathbb{R}) \wedge (\exists n \in \mathbb{N})st(n) \wedge (|x| < n)$ which contains " $st(x)$." At the same time, the formulae $x = 0$, $x > 0$, $|x| < 1$ are internal. Let a, b be straight lines. Then, $a \parallel b$, $a \perp b$ are internal sentences. If values of variables x, y are straight lines, then $x \perp a$, $x \parallel a$, $x \perp y$, $x \parallel y$ are internal formulae.

A formula can contain variables and constants. For instance, the formula $a < x < y < b$ contains two variables x, y , and two constants a, b ; traditionally, variables are denoted by $x, y, z, \dots, u, v, w, \dots$ and constants by $a, b, c, \dots, k, l, \dots$; $m, n, p, \dots$ are reserved for variables which range over the set $\mathbb{N}$ of natural numbers.

We accept the following **definition**: An internal formula containing only standard constants is said to be a *standard formula*. For instance, the formula $0 < x < \varepsilon$ is internal. It is standard if ε denotes a standard number, but if ε denotes e.g. an infinitesimal number, it is non standard.

8. The transfer principle

According to Nelson's approach to NSA (the main work of Nelson on this subject is IST = Internal Set Theory = Idealization + Standardization + Transfer; (see Lyantse, Kudryk, 1997) we need to add only three new axioms to OM (based on ZFC). One of them is *the transfer principle* (T). We express it by using the following statement:

$$(\exists x) p(x) \implies \exists^{st} x p(x); \quad (\text{T})$$

here $p(x)$ denotes an arbitrary *standard* formula. Therefore, $p(x)$ in (T) does not involve $st(x)$ as a subformula, and all constants in $p(x)$ are standard. The domain of action of the quantifier $\exists$ in (T) is the universe **I** of internal sets. Thus, the principle (T) can be rewritten in the following form:

$$(\exists x \in \mathbf{I}) p(x) \implies (\exists x \in \mathbf{S}) p(x).$$

Since $\mathbf{S} \subset \mathbf{I}$, we can replace " $\implies$ " with " $\iff$ ". Therefore, (T) is, in essence, the statement

$$(\exists x \in \mathbf{I}) p(x) \iff (\exists x \in \mathbf{S}) p(x). \quad (\text{T})$$

Let us explain the content of (T) in other words. Let $p(x)$ be a well-defined standard formula with a free variable x which ranges over $\mathbf{I}$. Let $dom_t p$ be a part of $\mathbf{I}$, such that $p(x)$ is true for $x \in dom_t p$. (T) says that *if $dom_t p$ is not empty then $dom_t p \cap \mathbf{S}$ is not empty* as well.

8.1. The characterization of standardness

Now we can repeat in a more rigorous way what we have said in Section 2. Remember that one reads “ $\exists!x$ ” as “there is a unique x .” “ $\exists!^{st}x$ ” means “there is a unique standard x .” Let us formulate the principle (Ch) (characterization of standardness):

$$(\exists!x \in \mathbf{I}) p(x) \wedge p(x_0) \implies st(x_0). \quad (\text{Ch})$$

Indeed, *if there exists a unique (internal) x for which the standard formula $p(x)$ is true, then, by (T), it must be standard.*

Let us consider some examples.

EXAMPLE 1

The empty set $\emptyset$ is standard.

The formula $p(x) \equiv (\forall y)(y \notin x)$ is standard (it contains no constant). The statement $p(\emptyset)$ is true: $(\forall y)(y \notin \emptyset)$. This defines $\emptyset$ uniquely. To prove it, let us recall the extensionality principle of OM:

$$(A = B) \iff (\forall x)(x \in A \iff x \in B).$$

Now, if both $\emptyset_1$ and $\emptyset_2$ are empty, then for any x both $x \in \emptyset_1$ and $x \in \emptyset_2$ are false. Therefore, the formula $x \in \emptyset_1 \iff x \in \emptyset_2$ is true.

EXAMPLE 2

If A is standard, then so is $B := \{A\}$.

Really, consider the formula $p(x) = (\forall y)(y \in x \iff y = A)$. This formula is standard, because A is standard, and B is a unique set, for which $p(x)$ is true.

In the same way we can show that

EXAMPLE 3

If $A, B, C, \dots$ are standard, then so are $\{A\}, \{A, B\}, \{A, B, C\}, \dots$

EXAMPLE 4

If A, B are standard, then the ordered pair (A, B) is standard.

Indeed, according to Kuratowski, $(A, B) = \{\{A\}, \{A, B\}\}$.

EXAMPLE 5

If A, B are standard, then $A \cup B, A \cap B, A \setminus B, A \times B$ are standard.

For instance, $A \times B$ is a *unique* set, for which $p(A \times B)$ is true, where $p(x)$ is the following *standard* formula, $(\forall y)(y \in x) \iff (\exists a \in A)(\exists b \in B)(y = (a, b))$.

EXAMPLE 6

The numbers $0, 1, 2, 3, \dots$ are standard.

This follows from Example 1 to Example 4, because $0 := \emptyset$, $1 := \{0\}$, $2 := \{0, 1\}$, $3 := \{0, 1, 2\} \dots$

EXAMPLE 7

The set $\mathbb{N}$ of natural numbers is standard.

Indeed, $\mathbb{N}$ is a unique x which satisfies the standard formula $p(x) \wedge q(x) \wedge r(x)$, where $p(x) \equiv (1 \in x)$, $q(x) \equiv (\forall y)(y \in x \implies y \cup \{y\} \in x)$, $r(x) \equiv (\forall z)(p(z) \wedge q(z) \implies x \subseteq z)$.

REMARK 6

Instead of $y \cup \{y\}$, one usually writes $y + 1$; $r(x)$ means that x is a minimal set which satisfies $p(x)$ and $q(x)$.

EXERCISE 6

Prove that the set $\mathbb{Z}$ of all whole numbers is standard.

EXAMPLE 8

Let A be a standard set, and $p(x)$ a standard formula. Then $B := \{x \in A : p(x)\}$ is standard.

B is uniquely determined by the standard formula $p(x) \equiv (\forall y) (y \in x \iff y \in A \wedge p(y))$.

EXAMPLE 9

The set $\mathbb{Q}$ of all rational numbers is standard.

Really, any rational number can be *uniquely* represented as y/x , where $x \in \mathbb{N}$, $y \in \mathbb{Z}$, and x, y have no common divisors. Thus, $\mathbb{Q}$ is a part of $\mathbb{N} \times \mathbb{Z}$ defined by a standard condition.

EXAMPLE 10

If A, B are standard, then the set B^A of all functions f such that $\text{dom } f = A$, $\text{im } f \subseteq B$, is standard.

Indeed, B^A is a part of the standard set $A \times B$ which consists of all f satisfying the standard formula $f \subseteq A \times B \wedge \forall x \in A \exists! y \in B (x, y) \in f$.

EXERCISE 7

1° Prove that the set $\mathbb{R}$ of all real numbers is standard. *Hint:* represent $\mathbb{R}$ as $\{r \in \mathbb{Q}^{\mathbb{N}} : p(r)\}$, where $p(r)$ is a suitable standard formula.

2° Let A be a standard bounded set of real numbers. Prove that $\inf A$ and $\sup A$ are standard.

- 3° Let $f \in \mathbb{R}^{\mathbb{R}}$ be a standard function. Suppose that the equation $f(x) = 0$ has only a finite number of roots. Prove that each such root is standard. Hint: apply (I₁) to $\{x \in \mathbb{R} : f(x) = 0\}$. Does the equation $\sin x = 0$ have non-standard roots?
- 4° If E is a standard finite set, then $\text{card } E$ is a standard natural number. Why?
- 5° Let A, B be standard sets, and $A \cap B \neq \emptyset$. Prove that A and B have a common standard element. Prove that the common point of two standard straight lines is standard.
- 6° Consider the Cauchy problem $y' = f(x, y)$, $y|_{x=x_0} = y_0$. Provide a condition for its solution to be standard.

PROPOSITION 4

- 1° A standard function takes a standard value at any standard point.
- 2° The inverse image of a standard set under a standard transformation is standard.

Proof is left as an exercise for the reader.

9. The second form of the transfer principle

Apply to the principle (T) the following law of logic $(a \Leftrightarrow b) \Leftrightarrow (\neg a \Leftrightarrow \neg b)$. Note that if the formula $p(x)$ is standard, then so is $\neg p(x)$. This way, we get

$$\forall^{st} x p(x) \Leftrightarrow \forall x p(x), \quad (T')$$

where $p(x)$ is an *arbitrary standard* formula. This (T') is the second form of the transfer principle. More strictly, (T') is as follows:

$$(\forall x \in \mathbf{S}) p(x) \Leftrightarrow (\forall x \in \mathbf{I}) p(x). \quad (T')$$

Thus, *if each standard x has a standard property p , then each internal x has this property*. In other words, as long as we consider standard properties only, we cannot differentiate standard objects from non-standard (internal) ones. Standard properties are their common properties.

Now let us consider some applications.

9.1. Non-standard extensionality principle

It states that *a standard set A is uniquely determined by the totality ${}^{st}A$ of its standard elements*. Namely,

$$(\forall A, B \in \mathbf{S}) (A = B \Leftrightarrow {}^{st}A = {}^{st}B).$$

(We recall that ${}^{st}A$ is the external set $\{x \in A : st(x)\}$).

Proof. By the extensionality principle of OM, $A = B$ if (and only if) $\forall x p(x)$, where $p(x) \equiv (x \in A) \Leftrightarrow (x \in B)$. If A and B are standard, this $p(x)$ is standard. (T') tells us that $\forall^{st} x p(x)$ is sufficient for $A = B$, i.e. ${}^{st}A = {}^{st}B$ is sufficient for $A = B$. ►

EXERCISE 8

- 1° Prove that if A, B are standard, then $A \subset B$, whenever ${}^{st}A \subset {}^{st}B$.
- 2° Give a counterexample for $({}^{st}A = {}^{st}B) \implies (A = B)$ with non-standard A, B .

9.2. The uniqueness of standard functions

Let X, Y be standard sets and $f \in Y^X$ be a standard function. Such f is uniquely determined by its values at standard points. To be more exact, if $f_1, f_2 \in {}^{st}(Y^X)$, then

$$(\forall {}^{st}x \in X) (f_1(x) = f_2(x)) \implies (\forall x \in X) (f_1(x) = f_2(x)). \quad (1)$$

Proof. If f_1, f_2 are standard, then $f_1(x) = f_2(x)$ is a standard formula, and, by (T') , we have (1). ►

COROLLARY 2

A standard sequence $u \in \mathbb{R}^{\mathbb{N}}$ can be defined in no way as follows

$$u_n = \begin{cases} f(n) & \text{for } n \ll \infty, \\ g(n) & \text{for } n \approx \infty, \end{cases}$$

where f, g are different standard functions.

REMARK 7

The transfer principles may be extended to formulae with an arbitrary quantity of free variables. To indicate explicitly the standardness of these formulae, we write

$$\begin{aligned} & \forall {}^{st}t_1 \dots \forall {}^{st}t_r \exists x_1 \dots \exists x_n p(x_1, \dots, x_n, t_1, \dots, t_r) \implies \\ & \implies \forall {}^{st}t_1 \dots \forall {}^{st}t_r \exists {}^{st}x_1 \dots \exists {}^{st}x_n p(x_1, \dots, x_n, t_1, \dots, t_r), \end{aligned} \quad (T)$$

$$\begin{aligned} & \forall {}^{st}t_1 \dots \forall {}^{st}t_r \forall {}^{st}x_1 \dots \forall {}^{st}x_n p(x_1, \dots, x_n, t_1, \dots, t_r) \implies \\ & \implies \forall {}^{st}t_1 \dots \forall {}^{st}t_r \forall x_1 \dots \forall x_n p(x_1, \dots, x_n, t_1, \dots, t_r), \end{aligned} \quad (T')$$

where $p(x_1, \dots, x_n, t_1, \dots, t_r)$ is an arbitrary *internal* formula which contains no constants, with free variables $x_1, \dots, x_n, t_1, \dots, t_r$.

9.3. The uniqueness of a standard relation

Let X, Y be standard sets, and $\mathcal{R}_1, \mathcal{R}_2$ be standard relations between elements of X and Y . (This means that $\mathcal{R}_i \subseteq X \times Y$, $i = 1, 2$; as usually we write $x\mathcal{R}_iy$ for $(x, y) \in \mathcal{R}_i$.) Then

$$\begin{aligned} & (\forall {}^{st}(x, y) \in X \times Y)(x\mathcal{R}_1y \iff x\mathcal{R}_2y) \\ & \iff (\forall (x, y) \in X \times Y)(x\mathcal{R}_1y \iff x\mathcal{R}_2y). \end{aligned}$$

Proof is left as an exercise for the reader.

EXAMPLE 11

Define $\mathcal{R} \subset \mathbb{R}^2$ by $x\mathcal{R}y \equiv x = y$ for standard x, y and $x\mathcal{R}y \equiv x \leq y$ for non-standard x, y . Then, $\mathcal{R}$ is not standard.

EXAMPLE 12

We have $e^x > 0$, $\cos(x + y) = \cos x \cos y - \sin x \sin y$ for all $x, y \in \mathbb{R}$, because these formulae are standard and true for $x, y \in {}^{st}\mathbb{R}$.

10. The special standardization principle (S_0)

The principle (I_0) informs us that we cannot create any standard infinite set without non-standard elements. The standardization principle tells us that sometimes it is possible to obtain something standard from something nonstandard. It is the third axiom of Nelson's IST. Now we formulate only some of its corollaries, namely the special standardization principle (S_0). It is as follows:

Let x be a limited real number. There exists a *unique standard* real number y such that $x \approx y$. Thus,

$$(\forall x \in \mathbb{R})(\|x\| \ll \infty \implies (\exists! {}^{st}y \in \mathbb{R})(x \approx y)). \quad (S_0)$$

The number y in (S_0) is denoted by ${}^\circ x$ and said to be *the shadow* (or *the standard part*) of x . So the shadow of $x \in \mathbb{R}$, $|x| \ll \infty$, is defined *uniquely* by

$${}^\circ x \in {}^{st}\mathbb{R} \text{ and } {}^\circ x \approx x.$$

The uniqueness of the shadow immediately implies the following proposition.

PROPOSITION 5

1° If $x \in \mathbb{R}$ is standard, then ${}^\circ x = x$.

2° The unique standard infinitesimal is zero:

$$x \in {}^{st}\mathbb{R} \implies {}^\circ x = x; \quad x \in {}^{st}\mathbb{R} \wedge x \approx 0 \implies x = 0.$$

3° If $x, y \in {}^{st}\mathbb{R}$, and $x \approx y$, then $x = y$.

Let us define

$$\mathbb{F} := \{x \in \mathbb{R} : |x| \ll \infty\}, \quad \mathbb{I} := \{x \in \mathbb{R} : x \approx 0\}.$$

Note that the set $\mathbb{F}$ of all limited real numbers and the set $\mathbb{I}$ of all real infinitesimals are *strictly external*. Indeed, $\mathbb{F}$ and $\mathbb{I}$ are non empty, $\mathbb{F}$ is bounded by every positive unlimited $\omega \in \mathbb{R}$, $\mathbb{I}$ is bounded, e.g., by 10^{-10} , but $\sup \mathbb{F}$, $\sup \mathbb{I}$, $\inf \mathbb{F}$, $\inf \mathbb{I}$ do not exist.

The immediate corollary from (S_0) is as follows:

$$\forall x \in \mathbb{F} \quad x = {}^\circ x + {}^i x, \quad {}^\circ x \in {}^{st}\mathbb{R}, \quad {}^i x \in \mathbb{I}. \quad (2)$$

The uniqueness of the shadow implies that this decomposition is unique. In other words, $\mathbb{F}$ is a *direct sum*:

$$\mathbb{F} = {}^{st}\mathbb{R} \dot{+} \mathbb{I}.$$

11. $\mathbb{R}$, ${}^{st}\mathbb{R}$, $\mathbb{F}$, and $\mathbb{I}$.

In OM, the set $\mathbb{R}$ of all real numbers together with the arithmetic operations $+$ and $\cdot$ and with the order relation $\leq$ is defined as a complete linearly ordered field. It is archimedean, i.e.

$$(\forall x \in \mathbb{R})(x \neq 0 \implies (\exists n \in \mathbb{N})(n|x| > 1)). \quad (3)$$

If x in (3) is infinitesimal, then the corresponding n must be infinite. But from (I_0) we can conclude that

$$(\forall {}^{st}x \in \mathbb{R})(x \neq 0 \implies (\exists {}^{st}n \in \mathbb{N})(n|x| > 1)).$$

An analogous *remark* also concerns other statements about $\mathbb{R}$. For instance, consider the Borel–Lebesgue lemma. Let $\mathcal{U} = \{\mathcal{U}_\alpha\}$ be an open cover of some bounded closed set $E \subset \mathbb{R}$. The lemma states that there exists a finite subcover $\mathcal{U}' = \{\mathcal{U}_{\alpha_1}, \dots, \mathcal{U}_{\alpha_n}\}$. This is true even if the exact bounds of E are infinite, and the length of a $\mathcal{U}_\alpha$ is ≈ 0 . Obviously, in this case we have $n \approx \infty$. But if E and $\mathcal{U}$ are standard, then (I_0) implies that $\mathcal{U}'$ can be chosen as standard.

The set ${}^{st}\mathbb{R} := \{x \in \mathbb{R} : st(x)\}$ is a subfield of the field $\mathbb{R}$. For instance, we have $x \in {}^{st}\mathbb{R} \wedge x \neq 0 \implies x^{-1} \in {}^{st}\mathbb{R}$, as a standard function (here $x \mapsto x^{-1}$) takes standard values at standard points. ${}^{st}\mathbb{R}$ is strictly external for the same reasons as $\mathbb{F}$.

THEOREM 1

1° The (strictly external) set $\mathbb{F}$ is a linearly ordered ring, a subring of $\mathbb{R}$. This means that

$$\begin{aligned} \mathbb{F} \subset \mathbb{R}, \quad (\forall x, y \in \mathbb{F}) (x + y \in \mathbb{F} \wedge xy \in \mathbb{F}) \\ (\forall x \in \mathbb{R}) (\forall y \in \mathbb{F}) (|x| \leq y \implies x \in \mathbb{F}). \end{aligned}$$

2° The (strictly external) set $\mathbb{I}$ is an ideal of the linearly ordered ring $\mathbb{F}$. This means that

$$\begin{aligned} \mathbb{I} \subset \mathbb{F}, \quad (\forall x, y \in \mathbb{I}) (x + y \in \mathbb{I}), \quad (\forall x \in \mathbb{I}) (\forall y \in \mathbb{F}) (xy \in \mathbb{I}), \\ (\forall x \in \mathbb{F}) (\forall y \in \mathbb{I}) (|x| \leq y \implies x \in \mathbb{I}). \end{aligned}$$

3° The map $x \mapsto {}^\circ x$ is a (strictly external) homomorphism of the ring $\mathbb{F}$ onto the field ${}^{st}\mathbb{R}$. This means that

$$\begin{aligned} (\forall x, y \in \mathbb{F}) ({}^\circ(x + y) = {}^\circ x + {}^\circ y, \quad {}^\circ(xy) = {}^\circ x \cdot {}^\circ y), \\ (x \leq y \implies {}^\circ x \leq {}^\circ y), \quad (\forall x \in \mathbb{F}) (\exists! \xi \in {}^{st}\mathbb{R}) (\xi = {}^\circ x). \end{aligned}$$

Proof is left as an exercise for the reader. For instance, let $x \in \mathbb{F}$, $y \in \mathbb{I}$. Then $|xy| < mn^{-1}$ for some $m \in {}^{st}\mathbb{N}$ and all $n \in {}^{st}\mathbb{N}$. If we take $n = km$, then we obtain $|xy| < k^{-1}$ with an arbitrary $k \in {}^{st}\mathbb{N}$, i.e. $xy \in \mathbb{I}$. Next, let $x, y \in \mathbb{F}$. According to (2), we have $xy = {}^\circ x {}^\circ y + \alpha$, where $\alpha := {}^i x {}^\circ y + {}^\circ x {}^i y + {}^i x {}^i y \approx 0$, hence ${}^\circ(xy) = {}^\circ x {}^\circ y$. Suppose that $x, y \in \mathbb{F}$, $x < y$, but ${}^\circ x > {}^\circ y$. Since ${}^\circ x + {}^i x < {}^\circ y + {}^i y$, we have ${}^\circ x - {}^\circ y < {}^i y - {}^i x$. This is a contradiction, because ${}^\circ x - {}^\circ y$ is a positive standard, and ${}^i y - {}^i x \approx 0$. And so on.

REMARK 8

The map $x \mapsto {}^\circ x$ is *idempotent*, i.e. $\forall x \in \mathbb{F} \ {}^\circ({}^\circ x) = {}^\circ x$. Its *kernel* $\{x \in \mathbb{F} : {}^\circ x = 0\}$ is the ideal $\mathbb{I}$ of infinitesimal numbers. We can say that $x \mapsto {}^\circ x$ is a *projection* of $\mathbb{F}$ onto ${}^{st}\mathbb{R}$ parallelly to $\mathbb{I}$. Clearly, $x \mapsto {}^i x$ is a *projection* of $\mathbb{F}$ onto $\mathbb{I}$, parallelly to $(\text{its kernel}) \ {}^{st}\mathbb{R} = \{x \in \mathbb{F} : {}^i x = 0\}$.

At the very end, some “exotic” examples. A finite set E is said to be *hyperfinite* if $\text{card } E \approx \infty$.

EXAMPLE 13

An external “finite” set which contains the “infinite” set ${}^{st}\mathbb{R}$.

Take an $\omega \in \mathbb{N} \setminus {}^{st}\mathbb{N}$ and form a sequence of segments $\Delta_k := [k/\omega, (k+1)/\omega[$, where $k = -\omega^2, -\omega^2 + 1, \dots, \omega^2 - 2, \omega^2 - 1$. Note that

$$\bigcup_{k=-\omega^2}^{\omega^2-1} \Delta_k = [-\omega, \omega[\supset {}^{st}\mathbb{R}.$$

None of Δ_k contains more than one standard point. Indeed, the length of Δ_k is $1/\omega \approx 0$, therefore if $x_1, x_2 \in \Delta_k$, then $x_1 \approx x_2$ and $x_1 = x_2$ whenever both x_1 and x_2 are standard. It is evident that only some Δ_k contain a standard point and never more than one. Denote by E' the set of centers of such Δ_k which do not contain a standard point, and put $E = E' \cup {}^{st}\mathbb{R}$. We see that E is a totality of $2\omega^2 \in \mathbb{N}$ points. It contains ${}^{st}\mathbb{R}$, but, unfortunately, it is defined *externally*.

The existence of an *internal* E such that ${}^{st}\mathbb{R} \subset E$ and $\text{card } E \in \mathbb{N}$ can be shown. Obviously, $\text{card } E \approx \infty$ for such E .

12. On decimal fractions

By the transfer principle (T') , each $x \in \mathbb{R}$ (standard or not) can be represented as

$$x = x_0, x_1 x_2 x_3 \dots, \quad (4)$$

where $x_0 \in \mathbb{Z}$, $\forall n \in \mathbb{N} \ x_n \in \{0, 1, \dots, 9\}$. Recall that (4) means that

$$\forall n \in \mathbb{N} \quad x_0 + \frac{x_1}{10} + \dots + \frac{x_n}{10^n} \leq x \leq x_0 + \frac{x_1}{10} + \dots + \frac{x_n}{10^n} + \frac{1}{10^n}. \quad (5)$$

By Cantor's principle of *nested segments*, (5) determines x uniquely. For this, it is necessary for the sequence $(x_n)_{n \in \mathbb{N}}$ in (4) to be internal: if $(x_n)_{n \in \mathbb{N}}$ is strictly external, then formula (4) represents no number. For instance, the fraction $0,000\dots\dots 999\dots$ (0s are at finite positions, 9s are at infinite ones) is not a real number. Obviously, x in (4) is *standard* if the sequence $(x_n)_{n \in \mathbb{N}}$ is *standard*.

PROPOSITION 6

The number $x \in \mathbb{R}$ is a positive infinitesimal if and only if its decimal expansion has the form

$$x = 0, 0\dots 0x_{\omega+1}x_{\omega+2}\dots, \quad (6)$$

where $x_{\omega+1} \neq 0$ and $\omega \approx \infty$.

Proof. From (5) and (6) we conclude that

$$\frac{x_{\omega+1}}{10^{\omega+1}} \leq x \leq \frac{x_{\omega+1} + 1}{10^{\omega+1}},$$

where $0 < x_{\omega+1} \leq 9$. Therefore, if $\omega \ll \infty$, then $x > 10^{-\omega-1} \gg 0$, and if $\omega \approx \infty$, then $x < 10^{-\omega-1} \approx 0$. ►

COROLLARY 3

Let $x \in \mathbb{R}$ be positive. According to (2), we have $x = y + z$, where $y = {}^\circ x$ is the standard part (shadow) of x , and $z = {}^i x$ is its infinitesimal part. Let $y_0, y_1 y_2 \dots$ be the decimal expansion of y . Then, for an infinite natural ω we have

$$x_0 = y_0, \quad x_1 = y_1, \dots, x_\omega = y_\omega. \quad (7)$$

Indeed, since $z \approx 0$, we have $z = 0, 0 \dots z_{\omega+1} z_{\omega+2} \dots$ for some $\omega \approx +\infty$. But $y = x - z$ and the subtraction here can only change one decimal in x before $x_{\omega+1}$. ►

REMARK 9

Let $x = x_0, x_1 x_2 \dots$. In order to find the shadow ${}^\circ x$, we only have to know x_n for $n \in {}^{st}\mathbb{N}$. Indeed, from (7) it follows that

$$\forall n \in {}^{st}\mathbb{N} \quad ({}^\circ x)_n = {}^\circ x_n.$$

On the other hand, a standard sequence $(a_n)_{n \in \mathbb{N}}$ is uniquely determined by a_n , $n \in {}^{st}\mathbb{N}$ (see Subsection 9.2, Corollary 2).

13. Some permanence principles. Robinson's lemma

An (internal) set $\tilde{N} \subseteq \mathbb{N}$ is said to be *modest* if

$$\exists m \ll \infty \quad \tilde{N} \subset \{1, 2, \dots, m\},$$

and it is said to be *greedy* if

$$\exists \omega \approx \infty \quad \{1, 2, \dots, \omega\} \subset \tilde{N}.$$

The simplest permanence principles are the following statements.

PROPOSITION 7

Let $\tilde{N} \in 2^{\mathbb{N}}$, then

$$\tilde{N} \subset {}^{st}\mathbb{N} \implies \tilde{N} \text{ is modest}; \quad (8)$$

$${}^{st}\mathbb{N} \subset \tilde{N} \implies \tilde{N} \text{ is greedy}. \quad (9)$$

Proof. Let $\tilde{N} \subset {}^{st}\mathbb{N}$, denote $N' := \{m \in \mathbb{N} : \tilde{N} \subset \{1, \dots, m\}\}$. We have $\mathbb{N} \setminus {}^{st}\mathbb{N} \subseteq N'$. But N' is internal, and $\mathbb{N} \setminus {}^{st}\mathbb{N}$ is strictly external. Therefore, $N' \neq \mathbb{N} \setminus {}^{st}\mathbb{N}$, i.e., $\exists m \ll \infty \quad m \in N'$, and (13.1) holds.

Now let ${}^{st}\mathbb{N} \subset \tilde{N}$. Put $N' = \{\omega \in \mathbb{N} : \{1, \dots, \omega\} \subset \tilde{N}\}$. By the above arguments, N' is internal. ${}^{st}\mathbb{N} \subseteq N'$, ${}^{st}\mathbb{N}$ is strictly external, therefore, N' contains some $\omega \approx \infty$. For this ω we have $\{1, \dots, \omega\} \subseteq \tilde{N}$. ►

COROLLARY 4

Let $\tilde{N} \in 2^{\mathbb{N}}$, then

$$\tilde{N} \subset \mathbb{N} \setminus {}^{st}\mathbb{N} \implies \exists \omega \approx \infty \quad \tilde{N} \subset \{\omega + 1, \omega + 2, \dots\}, \quad (10)$$

$$\mathbb{N} \setminus {}^{st}\mathbb{N} \subset \tilde{N} \implies \exists m \ll \infty \quad \{m + 1, m + 2, \dots\} \subset \tilde{N}. \quad (11)$$

Proof. If $\tilde{N} \subset \mathbb{N} \setminus {}^{st}\mathbb{N}$, then ${}^{st}\mathbb{N} \subset \mathbb{N} \setminus \tilde{N}$, therefore, $\mathbb{N} \setminus \tilde{N}$ is greedy: $\exists \omega \approx \infty$ $\{1, \dots, \omega\} \subset \mathbb{N} \setminus \tilde{N}$, hence (10) holds. (11) can be proven in an analogous way. ►

A remarkable permanence principle is the following famous

LEMMA 1 (ROBINSON'S LEMMA)

Let $x \in \mathbb{R}^{\mathbb{N}}$ and $\forall n \in {}^{st}\mathbb{N} \quad x_n \approx 0$. Then $\exists \omega \approx \infty \quad \forall n \leq \omega \quad x_n \approx 0$.

Proof. Let $\tilde{N} := \{n \in \mathbb{N} : n|x_n| < 1\}$. Since ${}^{st}\mathbb{N} \subset \tilde{N}$, $\tilde{N}$ is greedy. Let $\omega \approx \infty$ be such that $\{1, \dots, \omega\} \subset \tilde{N}$. Then $\forall n \leq \omega \quad |x_n| < n^{-1}$, therefore, $\forall n \leq \omega \quad x_n \approx 0$. ►

The other permanence principles are as follows.

PROPOSITION 8

Let $E \in 2^{\mathbb{R}}$, then

$$\begin{aligned} E \subset \mathbb{I} &\implies \exists \varepsilon \approx 0 \quad E \subset [-\varepsilon, \varepsilon] \quad (\text{modesty}), \\ \mathbb{I} \subset E &\implies \exists \varepsilon \gg 0 \quad [-\varepsilon, \varepsilon] \subset E \quad (\text{greediness}), \\ E \subset \mathbb{F} &\implies \exists a \ll \infty \quad E \subset [-a, a] \quad (\text{modesty}), \\ \mathbb{F} \subset E &\implies \exists a \approx \infty \quad [-a, a] \subset E \quad (\text{greediness}), \end{aligned}$$

where, as before, $\mathbb{I} := \{x \in \mathbb{R} : x \approx 0\}$, $\mathbb{F} := \{x \in \mathbb{R} : |x| \ll \infty\}$.

Proof. Let $\mathbb{I} \subset E$. Set $\tilde{N} := \{n \in \mathbb{N} : \forall k \leq n \quad |x| < k^{-1} \implies x \in E\}$. We have $\mathbb{N} \setminus {}^{st}\mathbb{N} \subset \tilde{N}$. By (11), $\tilde{N}$ contains some $m \ll \infty$. Therefore, $|x| \leq m^{-1} \implies x \in E$. We left the rest of the proof as an exercise for the reader. ►

14. Applications to sequences

14.1. The limit of a sequence

Let $x \in \mathbb{R}^{\mathbb{N}}$ (read “ x is a sequence of real numbers”; usually we write x_n instead of $x(n)$). We assume that the reader is familiar with the ordinary definition of convergence and the limit of x :

$$\begin{aligned} x \text{ is convergent (or Cauchy)} &\equiv (\forall \varepsilon > 0) (\exists N \in \mathbb{N}) (p, q > N \implies |x_p - x_q| < \varepsilon), \\ \lim_{n \rightarrow \infty} x_n = \ell &\equiv (\forall \varepsilon > 0) (\exists N \in \mathbb{N}) (\forall n > N) \quad |x_n - \ell| < \varepsilon. \end{aligned}$$

It turns out that, for standard sequences, this definition can be simplified, according to the aspirations of naturalists.

THEOREM 2

Let $x \in {}^{st}(\mathbb{R}^{\mathbb{N}})$, then x is convergent if and only if

$$\forall p, q \approx \infty \quad x_p \approx x_q. \quad (12)$$

Let $x \in {}^{st}(\mathbb{R}^{\mathbb{N}})$ be convergent. Then, for $n \approx \infty$ we have $|x| \ll \infty$, and therefore x_n has the shadow $\ell = {}^\circ x_n$, which is the same for all $n \approx \infty$. It is the limit of x :

$$\lim_{n \rightarrow \infty} x_n = {}^\circ(x_\omega) \quad \text{for } \omega \approx \infty.$$

Proof. Let x be a standard convergent sequence of reals x_n and $\lim_{n \rightarrow \infty} x_n = \ell$. This ℓ is standard by the transfer principle (T). Take an arbitrary $\varepsilon > 0$ and find $N \in \mathbb{N}$ such that $|x_n - \ell| < \varepsilon$ for $n > N$. Once again, by (T), this N can be chosen as standard. Let $n \in \mathbb{N} \setminus {}^{st}\mathbb{N}$. We now have $n > N$. Therefore, $|x_n - \ell| < \varepsilon$. As ε is arbitrary, we have $|x_n - \ell| \approx 0$. We see that $\forall n \approx \infty \quad {}^\circ(x_n) = \ell$. Also we know that $\forall p, q \approx \infty \quad |x_p - x_q| \leq |x_p - \ell| + |\ell - x_q| \approx 0$, that is $x_p \approx x_q$.

Conversely, assume that (12) is satisfied. Consider the set $\tilde{N} := \{n \in \mathbb{N} : |x_n - x_\omega| < 1\}$, where ω is a fixed unlimited natural number. Since the formula $p(n) \equiv |x_n - x_\omega| < 1$ is internal, the set $\tilde{N}$ is internal, too. By (12), $\mathbb{N} \setminus {}^{st}\mathbb{N} \subseteq \tilde{N}$. By (11), $\tilde{N}$ contains some *standard* n_ω . Hence (because x_{n_ω} is standard) $|x_\omega| \leq |x_{n_\omega}| + 1 \ll \infty$, so x_ω has the shadow $\ell := {}^\circ(x_\omega)$. By (12), $\forall n \approx \infty \quad x_n \approx \ell$. Now take an arbitrary standard positive ε . Note that the sentence $(\exists N \in \mathbb{N}) (\forall n > N) (|x_n - \ell| < \varepsilon)$ is true. Indeed, it is sufficient to choose $N \approx \infty$, and then $n > N \implies |x_n - \ell| < \varepsilon$. Now applying the transfer principle (T'), we get: $(\forall \varepsilon > 0) (\exists N \in \mathbb{N}) (\forall n > N) (|x_n - \ell| < \varepsilon)$. ►

EXAMPLE 14

1° The sequence $(1/n)_{n \in \mathbb{N}}$ and the number 0 are standard. For $n \approx \infty$, we have $1/n \approx 0$. Therefore, $\lim_{n \rightarrow \infty} 1/n = 0$.

2° Let $\varepsilon \neq 0$ be infinitesimal. For $n \approx \infty$, we have $1/n \approx \varepsilon$. But $\lim_{n \rightarrow \infty} 1/n \neq \varepsilon$. This fact does not contradict theorem 2, because ε is not standard.

3° Let ε be as before and $\forall n \in \mathbb{N} \quad x_n = \varepsilon$. Then $\forall n \quad x_n \approx 0$, but $\lim_{n \rightarrow \infty} x_n = \varepsilon \neq 0$. This fact does not contradict theorem 2, because the sequence $(x_n)_{n \in \mathbb{N}}$ is not standard.

4° For $n \approx \infty$ we have $n \cdot \sin 1/n \approx 1$, $(1 + 1/n)^n \approx e$, etc. Let $\mathcal{R}$ be the set of all convergent sequences $x \in \mathbb{R}^{\mathbb{N}}$. Such $\mathcal{R}$ is a partially ordered algebra on $\mathbb{R}$ relatively $+, \cdot, \leq$, defined by

$$\begin{aligned} (\alpha x + \beta y)_n &= \alpha x_n + \beta y_n, & (xy)_n &= x_n y_n, \\ x \leq y &\Leftrightarrow (\exists n \in \mathbb{N}) (\forall k > n) (x_k \leq y_k), \end{aligned}$$

here $x, y \in \mathcal{R}$, $\alpha, \beta \in \mathbb{R}$.

The main result concerning limits is the following.

THEOREM 3

The map $\lim$ is a homomorphism of the partially ordered algebra $\mathcal{R}$ onto a linearly ordered field $\mathbb{R}$, i.e.

$$\begin{aligned} \lim_{n \rightarrow \infty} (\alpha x_n + \beta y_n) &= \alpha \lim_{n \rightarrow \infty} x_n + \beta \lim_{n \rightarrow \infty} y_n, \\ \lim_{n \rightarrow \infty} (x_n y_n) &= \lim_{n \rightarrow \infty} x_n \cdot \lim_{n \rightarrow \infty} y_n, \\ x \leq y &\implies \lim_{n \rightarrow \infty} x_n \leq \lim_{n \rightarrow \infty} y_n. \end{aligned}$$

The non-standard proof of this theorem is almost trivial. At first, we note that theorem 3 is standard, therefore, by (T') , we only need to prove it for standard α, β, x, y . But for $x \in {}^{st}\mathbb{R}$ we have $\lim_{n \rightarrow \infty} x_n = {}^\circ(x_\omega)$, $\omega \approx \infty$, and according to theorem 1.3 $^\circ$ the map $r \mapsto {}^\circ r$ is a homomorphism of $\mathbb{F}$ onto ${}^{st}\mathbb{R}$. ►

Now we will obtain something better than theorem 2.

THEOREM 4

Let $x \in {}^{st}(\mathbb{R}^{\mathbb{N}})$. Suppose that for some $\omega \in \mathbb{N} \setminus {}^{st}\mathbb{N}$ and for all $n < \omega$ $n \approx \infty \implies x_n \approx 0$. Then x is convergent and $\lim_{n \rightarrow \infty} x_n = 0$.

We need the following interesting lemma to prove the theorem.

LEMMA 2

Let $k \in {}^{st}(\mathbb{N}^{\mathbb{N}})$ be strictly increasing and $\omega \in \mathbb{N} \setminus {}^{st}\mathbb{N}$. Then $(\exists n \in \mathbb{N} \setminus {}^{st}\mathbb{N}) (k_n < \omega)$. (This means that when n passes from values which are $\ll \infty$ to values which are $\approx \infty$, k cannot jump over $\omega \approx \infty$).

Proof. If $n \in {}^{st}\mathbb{N}$, then $k_n \in {}^{st}\mathbb{N}$ (the value of a standard function at any standard point is standard). Therefore, ${}^{st}\mathbb{N} \subseteq \tilde{N} := \{n \in \mathbb{N} : k_n < \omega\}$. Since the definition of $\tilde{N}$ is internal, the set $\tilde{N}$ is internal, too. By (9), there exists $n_\omega \in \tilde{N}$, $n_\omega \approx \infty$. As k is strictly increasing, $k_{n_\omega} \geq n_\omega$. Hence $k_{n_\omega} \approx \infty$ and $k_{n_\omega} < \omega$. ►

Proof of theorem 4. Suppose that x is divergent. Then, by (T') , there exist a standard $\varepsilon > 0$ and a standard, strictly increasing $k \in \mathbb{N}^{\mathbb{N}}$, such that $\forall n \in \mathbb{N} |x_{k_n}| > \varepsilon$. By lemma 2, there exists $k_n < \omega$ which is infinite, contradicting the assumption $n \approx \infty \implies x_n \approx 0$. ►

COROLLARY 5

Let $x \in {}^{st}(\mathbb{R}^{\mathbb{N}})$ and $\ell \in {}^{st}\mathbb{R}$. Suppose that for some $\omega \in \mathbb{N} \setminus {}^{st}\mathbb{N}$ and for all $n < \omega$ $n \approx \infty \implies x_n \approx \ell$. Then x is convergent and $\lim_{n \rightarrow \infty} x_n = \ell$.

Indeed, the sequence $(x_n - \ell)_{n \in \mathbb{N}}$ satisfies the condition of example 14.

We recall that $\ell \in \mathbb{R}$ is said to be a *limit point* of $x \in \mathbb{R}^{\mathbb{N}}$ if

$$(\forall \varepsilon > 0)(\forall N \in \mathbb{N})(\exists n > N)(|x_n - \ell| < \varepsilon), \quad (13)$$

or, equivalently, if x contains a subsequence which converges to ℓ .

THEOREM 5

Let $x \in {}^{st}(\mathbb{R}^{\mathbb{N}})$ and $\ell \in {}^{st}\mathbb{R}$. The number ℓ is a limit point of x if and only if $(\exists \omega \approx \infty) (x_\omega \approx \ell)$.

Proof. Let ℓ be a limit point of x and let $k \in \mathbb{N}^{\mathbb{N}}$ be a strictly increasing sequence such that $\lim_{n \rightarrow \infty} x_{k_n} = \ell$. By (T), we can choose a standard k . Then $(x_{k_n})_{n \in \mathbb{N}}$ is standard and, by (13), $\forall n \approx \infty x_{k_n} \approx \ell$.

Conversely, let $x_\omega \approx \ell$ for some $\omega \approx \infty$. Choose standard $\varepsilon > 0$ and $N \in \mathbb{N}$. Then, there exists $n > N$ such that $|x_n - \ell| < \varepsilon$ (namely, $n \approx \infty$). By (T') , we have (13). ►

EXERCISE 9

- 1° Let $x \in {}^{st}(\mathbb{R}^{\mathbb{N}})$ be decreasing and positive. Assume that $x_{\omega_0} \approx 0$ for some $\omega_0 \approx \infty$. Prove that $\forall \omega \approx \infty \ x_{\omega} \approx 0$.
- 2° Construct $x \in {}^{st}(\mathbb{R}^{\mathbb{N}})$ which is divergent, but $x_n \approx 0$ for $\omega_1 < n < \omega_2$, where $\omega_1 \approx \infty$ and $\omega_2 - \omega_1 \approx \infty$.

Acknowledgement

Thanks to the referees for many useful pieces of advice and recommendations.

References

- Albeverio, S., Fenstad, J., Høegh-Krohn, R., Lindstrøm, T.: 1986, *Nonstandard Methods in Stochastic Analysis and Mathematical Physics*, Academic Press, Orlando.
- Ballard, D.: 1994, Foundational Aspects of "Non"standard Mathematics, *Contemporary Mathematics* **176**. AMS.
- Conway, J.: 1976, *On Numbers and Games*, Academic Press, London, New York, San Francisco.
- Cutland, N. J.: 1983, Nonstandard Measure Theory and its Applications, *Bull. London Math. Soc.* **14**(6), 529-589.
- Cutland, N. J.: 1988, *Nonstandard Analysis and its Applications*, Cambridge University Press.
- Davis, M.: 1977, *Applied Nonstandard Analysis*, Wiley, New York.
- Diener, F.: 1983, *Cours d'analyse nonstandard*, Offices des Publ. Univ., Alger.
- Diener, F., Reeb, G.: 1989, *Analyse Nonstandard*, Hermann, Ed. Sc. Arts.
- Diener, F., Stroyan, K. D.: 1988, Syntactical Methods in Infinitesimal Analysis, w: N. J. Cutland (red.), *Nonstandard Analysis and its Applications*, Cambridge University Press, 258-281.
- Goldblatt, R.: 1998, *Lectures on the Hyperreals*, Springer, New York.
- Gordon, E. I., Kusraev, A. G., Kutateladze, S. S.: 2002, *Infinitesimal Analysis*, Kluwer Academic Publishers, Dordrecht.
- Hrbaček, K.: 1979, Nonstandard Set Theory, *Amer. Math. Monthly* **86**(8), 659-677.
- Hrbaček, K.: 1987, Axiomatic Foundations for Nonstandard Analysis, *Fund. Math.* **98**(1), 1-24.
- Hrbaček, K., Lessman, O., O'Donovan, R.: 2015, *Analysis with Ultrasmall numbers*, CRC Press Taylor & Francis Group.
- Hurd, A. E., Loeb, P. A.: 1985, *Introduction to Nonstandard Real Analysis*, Academic Press, New York.
- Kanovei, V., Reeken, M.: 2004, *Nonstandard Analysis, Axiomatically*, Springer.
- Kawai, T.: 1981, Axiom System of Nonstandard Set Theory, *Logic symposia, Hakone 1979, 1980*, Springer, 57-65.
- Keisler, H. J.: 1976, *Foundations of Infinitesimal Calculus*, Prindle, Weber and Schmidt, Boston, Massachusetts.

- Lindstrøm, T.: 1988, An invitation to nonstandard analysis, w: N. J. Cultland (red.), *Nonstandard analysis and its applications*, Cambridge University Press, Cambridge, 1-105.
- Lutz, R., Goze, M.: 1981, *Nonstandard Analysis: a Practical Guide with Applications*, Vol. 881, Lect. Notes in Math.
- Luxemburg, W. A. J., Robinson, A.: 1972, *Contributions to Nonstandard Analysis*, North-Holland Publishing Comp., Amsterdam, London.
- Lyantse, W., Kudryk, T.: 1997, *Introduction to Nonstandard Analysis*, VNTL Publishes, Lviv.
- Mendelson, E.: 1963, *Introduction to Mathematical Logic*, D. Van Nostrand Company, Inc., Princeton.
- Nelson, E.: 1977, Internal Set Theory: A New Approach to Nonstandard Analysis, *Bull. Amer. Math. Soc.* **83**(6), 1165-1198.
- Nelson, E.: 1987, *Radically Elementary Probability Theory*, Princeton University Press, Princeton, New Jersey.
- Nelson, E.: 1990, *Ann. Pure Appl. Logic*, The Syntax of Nonstandard Analysis.
- Robert, A.: 1985, *Analyse Non-standard*, Presse Polytechnique Romandes, Lausanne.
- Robert, A.: 1988, *Nonstandard Analysis*, Wiley, New York and Chichester.
- Robinson, A.: 1966, *Non-standard Analysis*, North-Holland Publishing Company, Amsterdam.
- Shoenfield, J.: 1967, *Mathematical Logic*, Wesley.
- Stroyan, K. D., Luxemburg, W. A. J.: 1976, *Introduction to the Theory of Infinitesimals*, Academic Press, New York, London.

Department of Mathematics
L'viv National University
L'viv, 79000, Ukraine
e-mail kudryk@mail.lviv.ua

Bożena Olik-Pawlik

O zjawisku zniekształcania obrazu matematyki*

Abstract. In this paper, we identify the phenomenon of *false conviction*. We analyze the definition of *false conviction* and provide the typology of *false convictions*, including their logical, geometric, cognitive and methodological types. Every type of *false conviction* is exemplified by empirical data.

1. Wprowadzenie

Nie można (...) wrzucać do tego samego worka pojęciowego, z jednej strony postaw i zachowań ucznia, który mówi lub pisze byle co, bo nie interesuje go rozwiązanie problemu, chwilowych pomyłek, wywołanych brakiem uwagi, błędów wynikających z luk w wiedzy, które mogą być łatwo poprawione przez samego ucznia, gdy pozostawi mu się czas na refleksję lub uzupełnienie wiadomości, wniosków opartych na lekkomyślnym zgadywaniu itp., z drugiej zaś strony świadomego rozumowania logicznie niepoprawnego, oraz błędów ujawniających głęboko zakorzenione fałszywe koncepcje, których uczeń nawet z uporem broni, jeżeli mu na to pozwalamy (Krygowska, 1989, s. 146).

Tak apelowała Z. Krygowska w przemówieniu wygłoszonym podczas inauguracji konferencji zorganizowanej przez Międzynarodową Komisję do Studiowania i Ulepszania Nauczania Matematyki¹, na której błąd w matematyce szkolnej był głównym przedmiotem rozważań. W swojej wypowiedzi Z. Krygowska zwróciła uwagę na występowanie różnorodnych typów uczniowskich błędów i wyraźnie podkreśliła, iż szczególnego traktowania wymagają te błędy, które są rezultatem głęboko zakotwiczonych fałszywych koncepcji czy nieprzypadkowego a niepoprawnego rozumowania. Błędy nazwała *istotnymi i nieuniknionymi składnikami ludzkiej działalności*. Stwierdziła też, że *zakres analiz i badań (dotyczących problemów związanych z błędami) jest tak rozległy i ich wyniki tak bogate, że nie jest już możliwe opanowanie indywidualną wiedzą tego, co już zostało wypracowane*.

*On distorting the image of mathematics

2010 Mathematics Subject Classification: Primary: 97C70; Secondary: 97D40

Key words and phrases: Research on teacher education, False conviction

¹ Konferencja odbyła się w Sherbrooke (Kanada) w dniach od 27 lipca do 1 sierpnia 1987 roku. Temat konferencji: The role errors play in the learning and teaching of mathematics (Rola błędów w uczeniu się i nauczaniu matematyki).

Problematyka rozprawy doktorskiej (Pawlik, 2004c), niepublikowanej do tej pory w całości, której fragmenty składają się na niniejszy artykuł, koncentruje się na błędach dostrzeżonych w rozumowaniach początkujących studentów matematyki – przyszłych nauczycieli. Są to takiego rodzaju błędy, których nie można, zgodnie z apelem, zlekceważyć, gdyż opierają się na niepoprawnych koncepcjach. *Falszywe przekonania*, bo tak te koncepcje nazwano, funkcjonujące w rozumowaniach uczących się matematyki, sterują ich myśleniem podczas pracy nad różnego rodzaju zadaniami, utrudniając albo wręcz uniemożliwiając ich poprawne rozwiązanie. Są one niebezpieczne dla tworzonego w umyśle uczącego się obrazu matematyki, ponieważ nie wykorzenione mogą stać się przyczyną jego zniekształcania.

2. Określenie i krótka charakterystyka fałszywych przekonań

Falszywe przekonania (w skrócie *FP*) – domniemane fundamenty niepoprawnych rozumowań; są nimi fałszywe koncepcje, błędne schematy myślowe oraz niepoprawne reguły wnioskowania. Określenie to powstało przez analogię do zaobserwowanych przez A. Bella (1992, s. 11) uczniowskich fałszywych poglądów – błędnych przekonań typu: *mnożenie zwiększa, a dzielenie zmniejsza, czy dzieli się liczbę większą przez mniejszą*.

Nazwą tą objęte są takie błędne schematy, które stosowane są przez studentów jakby automatycznie. Choć nazwane zostały przekonaniami², nie znaczy to, że są one uświadomione. Chodzi o to, że studenci postępując według tych schematów działają w dobrej intencji i wydają się być przekonani, że ich rozumowania i działania są poprawne. Warto podkreślić, że *falszywe przekonania* studenta, który nie zetknął się z wywołującym zaniepokojenie konfliktem zmuszającym go do weryfikacji swojego (błédnego) rozumowania, pozostają nieuświadomione.

Charakterystyczne dla rozwiązań opartych na *falszywych przekonaniach* jest to, że ich autorzy przystępują do pracy nad zadaniem tak, jakby już znali odpowiedź. Większość z nich nie bada, nie poszukuje, nie powołuje się na poznane wcześniej definicje czy twierdzenia. Wydaje się, że autorzy formułując rozwiązanie realizują pierwszą myśl, która im „przyszła do głowy”.

3. Przykłady i typy fałszywych przekonań

W ramach opisanych w rozprawie badań wyróżnione zostały cztery typy fałszywych przekonań. Można je krótko scharakteryzować w podany dalej sposób.

Typ I – *falszywe przekonania „logiczno-pojęciowe”*

Falszywe przekonania tego typu dotyczą rozumowań, które związane są z błédną definicją pojęcia, niepoprawną regułą wnioskowania lub ich struktura logiczna – jak to ujął Z. Powązka (2009, s. 215) – opiera się na błédnej tautologii.

² W pedagogice funkcjonuje termin **przekonanie**. W pracy (Okoń, 1984, s. 248) czytamy, iż jest to *względnie stały składnik świadomości człowieka, oparty na przeświadczeniu, że pewien stan rzeczy jest taki, a nie inny; na przekonania składają się: poznanie jakiegoś stanu rzeczy, emocjonalna aproba zdobytej wiedzy oraz skłonność do postępowania zgodnego z jej treścią; najlepszym sprawdzianem siły przekonań jest codzienna praktyka potwierdzająca zgodność wiedzy i działania*.

Typ II – fałszywe przekonania „połączenia”.

Dotyczą sklejeń³ i złożień („połączeń”) przekształceń geometrycznych. Związane są z przenoszeniem przez studentów własności pewnych obiektów matematycznych na „nowy” obiekt – „połączenie” – powstały z danych według domyślnej zasady.

Typ III – fałszywe przekonania „analogia”

Tego typu fałszywe przekonania powstały przez nieostrożne wykorzystanie zdobytej wcześniej wiedzy.

Typ IV – fałszywe przekonania „metodologiczne”

Związane są z takimi rozumowaniami, w których ujawniają się błędy natury metodologicznej – dotyczące np. roli przykładu, kontrprzykładu czy rysunku.

Zaprezentowane dalej przykłady studenckich prac służą ukazaniu zarówno samych fałszywych przekonań, jak i zwróceniu uwagi na różne ich typy.

Prezentacja każdego z zawartych tu przykładów obejmuje:

- temat zadania,
- (błędne) rozwiązanie tego zadania podane przez studenta,
- analizę przedstawionego rozwiązania,
- ogólne sformułowanie *fałszywego przekonania*, stanowiącego domyślny fundament błędnego rozumowania zawartego w omawianym rozwiązaniu.

Nie mam pewności, że wszystkie rozumowania autorów analizowanych prac przebiegały dokładnie tak, jak zostały przedstawione. Hipotetyczne rozumowania studentów zwerbalizowane zostały po to, by je opisać i rozpoznać istotę błędów popełnionych w oparciu o *fałszywe przekonania*.

Jako pierwsze przedstawiono fałszywe przekonanie *FP „odcinek”*. Choć było ono już niejednokrotnie prezentowane we wcześniejszych publikacjach (por. np. Pawlik, 2003, 2008), zamieściłam je tutaj ponownie dla przejrzystości prezentacji, ponieważ i na tym fałszywym przekonaniu oparte jest rozwiązanie zamieszczone w dalszej części tej pracy (zob. Przykład 2). Należy zaznaczyć, że inne przykłady fałszywych przekonań lub błędne prace studentów ilustrujące te fałszywe przekonania, na które zwrócono uwagę w niniejszym opracowaniu, znaleźć można w innych publikacjach B. Pawlik wskazanych w wykazie literatury.

FP „odcinek” przedstawia się następująco:

FP „odcinek”
 Obrazem odcinka AB w przekształceniu P jest odcinek $A'B'$ taki, że $P(A) = A'$,
 $P(B) = B'$
 (niezależnie od określenia przekształcenia P).

³ Sklejenia przekształceń to przekształcenia zadane dwu- lub wielonormowo. Mimo że odwzorowania tego typu nie mają dla teorii przekształceń większego znaczenia, dla celów dydaktycznych (chodzi o to, by skonstruować pojęcie złożenia przekształceń) były uwzględniane na ćwiczeniach z geometrii elementarnej.

O jego funkcjonowaniu w rozumowaniach uczących się matematyki świadczy przywołane dalej rozwiązanie zadania 1 podane przez studentkę, która nie uczestniczyła w zajęciach, w ramach których omawiane było zagadnienie obrazu odcinka w przekształceniach geometrycznych.

Przykład 1

ZADANIE 1 Oceń wartość logiczną poniższego zdania. Odpowiedź uzasadnij.

Obrazem odcinka AB w przekształceniu P jest odcinek $A'B'$ taki, że

$$P(A) = A', P(B) = B'.$$

Rozwiązanie studentki.

Obrazem odcinka AB w przekst. P jest odcinek $A'B'$ taki że $P(A)=A'$, $P(B)=B'$ jest to prawda ponieważ szukając obrazu odcinka szukamy obrazów jego końców A, B . Przekształcamy je tak że $P(A)=A'$ i $P(B)=B'$. Łączymy ze sobą A' i B' i powstaje nam obraz odcinka AB w tym przekształceniu P .

W swoim uzasadnieniu autorka pracy klarownie sformułowała błędny przepis na „wyznaczenie obrazu odcinka AB w przekształceniu P ”. Prawdopodobnie wcześniejsze doświadczenia szkolne przyczyniły się do wytworzenia opisanego schematu w umyśle rozwiązującej zadanie. Na wcześniejszych etapach edukacji uczyła się o przekształceniach zachowujących współliniowość punktów i ich uporządkowanie. Właśnie tak konstruowało się obraz odcinka w każdym z rozważanych przykładów. Wydaje się, że wielokrotne powtórzenie tego samego sposobu postępowania mogło przyczynić się do utrwalenia poglądu, iż powyższy schemat jest uniwersalny, więc można go stosować do wszystkich przekształceń geometrycznych. Wygląda na to, że dla studentki „obraz odcinka” utożsamiał się z tym, co zawsze robiła w szkole, czyli z procedurą znajdowania jego obrazu – „wyznaczyć obrazy końców odcinka i połączyć je”.

O tym, że taka procedura wielu studentom nie jest obca, przekonuje wypowiedź innej osoby, która przyznała, iż:

W szkole podstawowej, później średniej, zawsze myślałam, uczono mnie, że obrazem odcinka w przekształceniach jest odcinek. Jednak na studiach okazało się, że nie we wszystkich przekształceniach obrazem odcinka jest też odcinek.

Dla ilustracji kolejnych fałszywych przekonań – innego typu – posłuży rozwiązanie zadania 2.

Przykład 2

ZADANIE 2 Przekształcenie P płaszczyzny określono wzorami

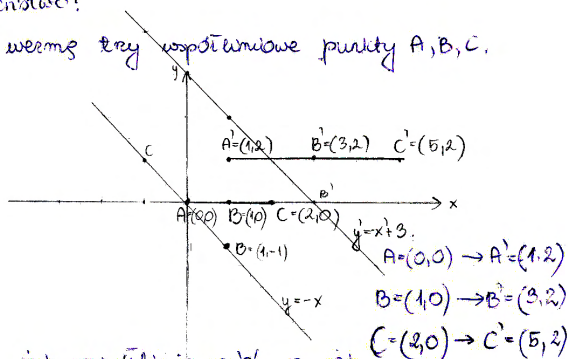
$$\begin{cases} x' = 2x + 1 \\ y' = 2y + 2 \end{cases}$$

- Czy P jest podobieństwem?
- Czy P jest przekształceniem afinicznym?

Rozwiązanie studentki.⁴ Czy to jest podobieństwo? Aby to pokazać wezmę trzy współliniowe punkty A, B, C .

Czy to jest podobieństwo?

Aby to pokazać wezmę trzy współliniowe punkty A, B, C .



Ponieważ zachowana jest współliniowość, a jak wiemy na podstawie twierdzenia, podobieństwo zachowuje współliniowość, zatem to przekształcenie jest podobieństwem.

Ponieważ zachowana jest współliniowość, a jak wiemy na podstawie twierdzenia, podobieństwo zachowuje współliniowość, zatem to przekształcenie jest podobieństwem.

Sprawdzę jeszcze dla prostej $y = -x$.

Sprawdza jeszcze dla prostej:

$$y = -x$$

$$y' = 2y + 2$$

$$x' = 2x + 1$$

$$2y = y' - 2$$

$$2x = x' - 1$$

$$y = \frac{1}{2}y' - 1$$

$$x = \frac{1}{2}x' - \frac{1}{2}$$

$$\frac{1}{2}y' - 1 = -\left(\frac{1}{2}x' - \frac{1}{2}\right)$$

$$y' - 2 = -x' + 1$$

$$y' = -x' + 3$$

Prosta $y = -x$ przeszła w tym przekształceniu na prostą $y' = -x' + 3$ zatem jest to podobieństwo.

Z tych samych powodów jest to przekształcenie afiniczne, bo jak wiemy przekształcenie afiniczne zachowuje współliniowość.

Prosta $y = -x$ przeszła w tym przekształceniu na prostą $y' = -x' + 3$, zatem jest to podobieństwo.

Z tych samych powodów jest to przekształcenie afiniczne, bo jak wiemy przekształcenie afiniczne zachowuje współliniowość.

Zaprezentowane rozwiązanie jest przykładem błędnego rozumowania kończącego się prawdziwym wnioskiem. Analizując jego strukturę logiczną można zauważyć, że autorka potraktowała występującą w nim implikację tak, jakby była równoważnością. Na wcześniej poznane i zacytowane w pracy twierdzenie: *podobieństwo zachowuje współliniowość punktów*, powołała się w ten sposób, że zachowywanie współliniowości punktów przez dane przekształcenie (warunek konieczny dla założenia twierdzenia) potraktowała jako warunek wystarczający na to, aby przekształcenie to było podobieństwem. Wygląda na to, że rozumowanie studentki przebiegało tak, jak gdyby uznawała ona fałszywą regułę dowodzenia, którą można by nazwać „regułą odrywania” postaci:

$$\frac{p \Rightarrow q, q}{p}$$

Przypuszczalne rozumowanie studentki, w „języku inkluzji”, można opisać następującym, niepoprawnym, ogólnym schematem wnioskowania:

FP „COPI”

falszywe przekonanie – „częściowo odwrócona przechodniość inkluzji”

Błędna reguła wnioskowania

$$\frac{A \subset B, C \subset B}{C \subset A}$$

Rozumowanie oparte na powyższym *falszywym przekonaniu „COPI”* studentka powtórzyła na końcu swojej pracy, stwierdzając, iż z tego samego powodu – zachowywania współliniowości punktów – przekształcenie P jest przekształceniem afinicznym.

Na uwagę zasługuje także sposób, w jaki autorka omawianego rozwiązania uzasadniła następujący fakt: *zachowana jest współliniowość* – współliniowość punktów jest niezmiennikiem danego przekształcenia. Na początku studentka wykonała rysunek, na którym zaznaczyła dwie trójki współliniowych punktów (w obu przypadkach oznaczonych jednakowo: A, B, C ; punkt A jako punkt wspólny obu trójek)

⁴ Rozwiązanie to jest przykładem pracy, którą trudno jest w sposób bezdyskusyjny zaliczyć do jednego z typów prezentowanych przykładów, ze względu na to, iż zawarte w nim rozumowanie oparte jest na więcej niż jednym *falszywym przekonaniu*.

położonych na prostych o równaniach $y = -x$ oraz $y = 0$. Wybrane punkty prostej o równaniu $y = 0$ połączyła, otrzymując odcinek AC' , dla którego punkt B jest środkiem. Następnie wyznaczyła obrazy tych punktów, obliczając odpowiednie ich współrzędne. W efekcie, łącząc obrazy punktów, utworzyła z nich odcinek $A'C'$, korzystając w tym momencie między innymi z tego faktu, który chciała wykazać (zachowywania współliniowości punktów przez przekształcenie)⁵. Wydaje się, iż rezultatem takiego rozumowania – obrazem trójki punktów odcinka jest trójka punktów odcinka – nie była usatysfakcjonowana, gdyż w dalszej części rozwiązania napisała: *Sprawdzę jeszcze dla prostej*. Równanie obrazu prostej o równaniu: $y = -x$ wyznaczyła rachunkiem analitycznym. Ostatecznie stwierdziła, że skoro wybrana przez nią prosta $y = -x$ przeszła w tym przekształceniu na prostą (tu o równaniu: $y = -x + 3$)⁶, zatem zachowana jest współliniowość.

Analiza uzasadnienia autorki pracy pozwala przypuszczać, iż dokonane przez nią uogólnienie – zaobserwowanej dla jednego obiektu własności – jest rezultatem rozumowania, które można opisać następująco:

Jeżeli A należące do Z jest W , to każdy X należący do Z jest W .

Zauważmy, iż do teoretycznego rozumowania studentki przystaje poprzednio opisane FP „COPI”:

$$\frac{W \subset Z, X \in Z}{X \in W}.$$

Wydaje się, że *falszywe przekonanie* wspomniane wyżej ma też związek z występującą w logice regułą uogólniania.

Reguła uogólniania

$$\frac{\varphi(x), x \in X}{\forall x \in X : \varphi(x)}$$

jest regułą dowodzenia, bowiem jeśli $\varphi(x), x \in X$, jest funkcją zdaniową prawdziwą w X , to dla każdego $a \in X$ prawdziwe jest zdanie $\varphi(a)$, co dowodzi, że zdanie $\forall x \in X : \varphi(x)$ jest prawdziwe” (Rasiowa, 1984, s. 235).

FP „PU”

falszywe przekonanie – „przykład – uogólnienie”

Błędna reguła uogólniania

$$\frac{\varphi(a)}{\forall x \in X : \varphi(x)},$$

gdzie a jest określonym (wiadomym, konkretnym) elementem zbioru X .

Przedmiotem rozważań autorki omawianego rozwiązania nie była dowolna prosta, lecz prosta wskazana równaniem. Być może w odczuciu studentki prosta o równaniu $y = -x$ była dowolnie wybraną prostą na płaszczyźnie. Wyznaczenie obrazu takiej prostej, a następnie stwierdzenie, że otrzymany obiekt jest również prostą, było dla autorki pracy wystarczającym argumentem potwierdzającym tezę:

⁵ Studentka ujawniła tu kolejne falszywe przekonanie – FP „odcinek”.

⁶ Studentka sporządziła błędny rysunek – nie zwróciła uwagi na to, że wyznaczony wcześniej obraz punktu A powinien należeć do obrazu prostej o równaniu $y = -x$.

Przekształcenie P zachowuje współliniowość punktów. Takie uzasadnienie, skonstruowane w oparciu o FP „ PU ”, jest „dowodem przez przykład”. Prowadzi do kolejnego fałszywego przekonania:

FP „ DPP ”
falszywe przekonanie – „dowód przez przykład”
Aby udowodnić twierdzenie, wystarczy wskazać choć jeden przykład, który je potwierdza.

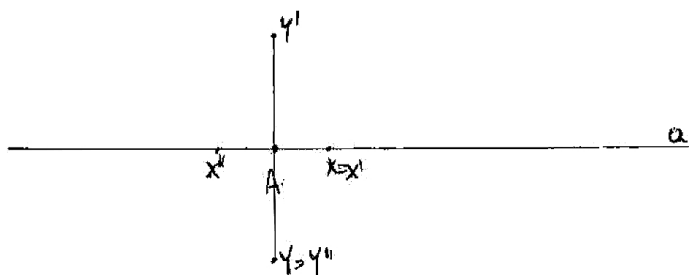
Myślę, że warto podkreślić, iż autorka ostatniej pracy w jednym rozwiązaniu trzykrotnie rozumowała podobnie w zróżnicowanych pojęciowo i metodologicznie sytuacjach. Jej praca wyraźnie ukazuje skalę zagrożenia dla obrazu matematyki wywołanego przez fałszywe przekonania „działające” w umyśle studentki.

Do kolejnej grupy fałszywych przekonań należy błędna koncepcja stanowiąca fundament kolejnego z zaprezentowanych tu rozwiązań.

Przykład 3

ZADANIE 3 Czy złożenie $S_A \circ S_a$, gdzie $A \in a$ ma punkty stałe? Jeśli tak, to ile i jakie?

Rozwiązanie studentki.



Punktem stałym jest on, ponieważ punkt stały w S_a jest punkt A w S_A

Wskazany przez autorkę rozwiązania punkt A rzeczywiście jest punktem stałym rozważanego złożenia, niestety, nie jedynym. Studentka mogła to łatwo stwierdzić, gdyby się w ogóle nad tym zastanowiła. Wybrany przez nią – w sposób szczególny, najprawdopodobniej w celu zorientowania się w sytuacji – punkt Y , jest również poszukiwanym elementem zbioru punktów stałych. Poza jednym empirycznym zabiegiem, polegającym na wyznaczeniu obrazu punktu Y , autorka nie próbowała zrobić nic więcej – nie wyznaczyła obrazów innych punktów, ani nie skorzystała z definicji punktu stałego. Rozumowała tak, jakby uważała, iż prawdziwe jest następujące twierdzenie:

FP „iloczyn”

Zbiór punktów stałych złożenia przekształceń geometrycznych składa się z wspólnych punktów stałych składowych tego złożenia.

Jako ostatnie zilustruję fałszywe przekonanie o innej od poprzedniego naturze metodologicznej.

Przykład 4

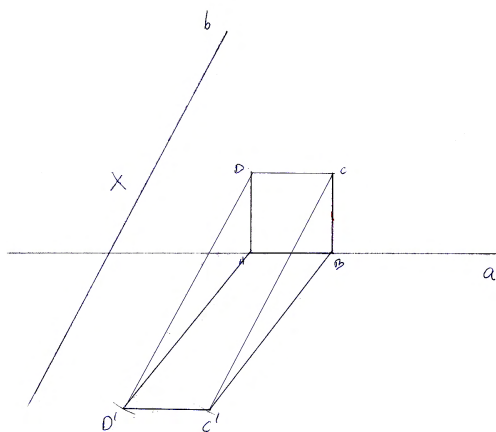
ZADANIE 4 Dane są przecinające się, nie prostopadłe proste a i b . Przekształcenie P_{-2} płaszczyzny określono w następujący sposób: Obrazem punktu X jest taki punkt X' , że istnieje $A \in a$ spełniający warunki:

$$\overrightarrow{AX} \parallel b \text{ i } \overrightarrow{AX'} = -2\overrightarrow{AX}.$$

- Czy P_{-2} jest wzajemnie jednoznaczny przekształceniem płaszczyzny na płaszczyznę?
- Wyznacz zbiór punktów stałych przekształcenia P_{-2} .
- Czy P_{-2} jest podobieństwem?
- Określ, jeśli istnieje, przekształcenie odwrotne do P_{-2} .

Odpowiedzi uzasadnij!

Rozwiązanie studenta B dotyczące punktu c).



Jak widać na przykładzie obrazem kwadratu w tym przekształceniu nie jest kwadrat

$$\frac{|AD|}{|AC|} \neq \frac{|AB|}{|BC|}$$

S. Turnau (1991) precyzując pojęcie błędu napisał, iż *błąd jest nieumyślny, tj. popełniający go działa w dobrej intencji i w swoim mniemaniu – właściwie*. Mając na uwadze przywołane stwierdzenie można być zdania, iż to, co student przedstawił jako rozwiązanie zadania (rezultat jego myślenia nad postawionym problemem) stanowiło – w jego odczuciu – właściwą odpowiedź na pytanie: „Czy P_{-2} jest podobieństwem?”. Oczywiście, problem nie tkwi w tym, że autor sporządził i powołał się w rozwiązaniu na rysunek, na którym rzeczywiście widać, że w rozważanym przekształceniu obrazem wskazanego kwadratu nie jest kwadrat⁷. Istota dostrzeżonej trudności polega na tym, że autorowi pracy wydawało się, że zrobił to, co należało zrobić, żeby odpowiedzieć na postawione w ramach egzaminu pytanie. „Sformułowana” przez studenta odpowiedź ukryta jest w rysunku; student „odpowiedział rysunkiem”. (Czytelnik pracy może spróbować tę odpowiedź „wydobyć” z przedstawionego rozwiązania.) Obiekty przedstawione na wykonanym (dokładnie) rysunku student potraktował jako prawidłowo „skonstruowany” kontrprzykład, a argumentację wzrokową, którą się posłużył, jako wystarczającą argumentację matematyczną. Rozumował tak, jakby uważał, że:

FP „DPK”
falszywe przekonanie – „dowód przez pokazanie”
Argumentacja wizualna („jak widać na rysunku”) jest rodzajem argumentacji
matematycznej.

Falszywe przekonanie *FP „COPI”*, to przykład fałszywego przekonania typu I. *FP „iloczyn”* należy do fałszywych przekonań typu „połączenia”. Typ III fałszywych przekonań zilustrowany został przez *FP „odcinek”*. Natomiast *FP „DPK”* ukazuje istotę fałszywych przekonań typu IV.

4. Zakończenie

Zaprezentowane (w rozprawie doktorskiej) rozwiązania zadań pozwalają wątpić w słuszność stwierdzenia, że *falszywe przekonania* funkcjonujące w rozumowaniach studentów zanikają samoistnie w miarę upływu czasu, czy też w wyniku rozbudowywania ich matematycznej wiedzy. Uzmysławiają, że każde nieusunięte *falszywe przekonanie* stanowi poważne zagrożenie dla wiedzy studentów. *Falszywe przekonania* mogą bowiem stać się, z jednej strony, przyczyną szeregu różnych błędów i nieporozumień trudnych do wyeliminowania bez wcześniejszego usunięcia zakotwiczonych w umysłach uczących się matematyki rozmaitych niepoprawnych schematów, koncepcji czy reguł, z drugiej zaś impulsem do formowania się w ich umysłach kolejnych *falszywych przekonań*, gdyż, jak stwierdziła Z. Krygowska:

⁷ Student B jest także autorem innego rozwiązania przemawiającego za funkcjonowaniem w jego rozumowaniu *FP „brzeg”* – *Figura to linia przedstawiona śladem przyrządu rysującego* (zob. Pawlik, 2008). W kontekście przedstawionego wcześniej fałszywego przekonania *FP „odcinek”* oraz *FP „brzeg”* pojawiają się tu pytania: Jaki zbiór punktów student miał na myśli mówiąc o kwadracie? W jaki sposób wyznaczał obraz tego kwadratu w badanym przekształceniu?

(...) matematyka jako przedmiot uczenia się zachowuje swą organizację wewnętrzną. Jest to dziedzina pojęciowo zorganizowana przez samą siebie tak, że pozornie drobna luka w rozumieniu lub wiedzy łatwo może spowodować następne nieporozumienia, narastające jedne na drugich, które uzewnętrzniają się w lawinie błędów, trudnej do zatrzymania (Krygowska, 1989, s. 142).

Literatura

- Bell, A.: 1992, Systematyczne użycie konfliktu poznawczego w nauczaniu – trzy eksperymenty, *Roczniki Polskiego Towarzystwa Matematycznego, Seria V: Dydaktyka Matematyki* **13**, 9–54.
- Krygowska, Z.: 1989, Zrozumieć błąd w matematyce, *Roczniki Polskiego Towarzystwa Matematycznego, Seria V, Dydaktyka Matematyki* **10**, 141–147.
- Okoń, W.: 1984, *Słownik pedagogiczny*, Państwowe Wydawnictwo Naukowe, Warszawa.
- Pawlik, B.: 2003, On the understanding of geometric transformations by mathematics freshmen, w: Płocki, A. (red.), *Annales Academiae Paedagogicae Cracoviensis, Folia 16, Studia Mathematica III*, Wydawnictwo Naukowe Akademii Pedagogicznej, Kraków, 155–164.
- Pawlik, B.: 2004a, *On false convictions concerning geometric transformations of the plane in mathematics students' reasoning*, referat (w postaci artykułu) przygotowany na konferencję ICME 10 Kopenhaga 2004, 4–11 lipca 2004, publikacja internetowa. (http://www.icme-organisers.dk/tsg10/articulos/Pawlik_30_updated_paper.doc).
- Pawlik, B.: 2004b, On false convictions in mathematics students' reasoning, w: *Mathematica. Proceedings of the XIth Slovak-Czech-Polish Mathematical School, Pedagogical Faculty Catholic University in Ružomberok, Ružomberok June 2nd–5th*, 138–145.
- Pawlik, B.: 2004c, *Fałszywe przekonania dotyczące przekształceń geometrycznych na płaszczyźnie w rozumowaniach studentów matematyki*. Niepublikowana rozprawa doktorska, promotor: prof. dr hab. Stefan Turnau.
- Pawlik, B.: 2005, Fałszywe przekonania dotyczące przekształceń geometrycznych na płaszczyźnie w rozumowaniach studentów matematyki, *Roczniki Polskiego Towarzystwa Matematycznego, Seria V, Dydaktyka Matematyki* **28**, 365–376.
- Pawlik, B.: 2008, Co to jest fałszywe przekonanie? Prezentacja zjawiska w oparciu o przykłady rozwiązań zadań podanych przez studentów matematyki, w: *Współczesne Problemy Nauczania Matematyki 1 (seria: Prace Monograficzne z Dydaktyki Matematyki)*, "Forum Dydaktyków Matematyki", Bielsko-Biała, 115–128.
- Powązka, Z.: 2009, O fałszywych przekonaniach obserwowanych na zajęciach z analizy matematycznej, w: Chronowski, A., Klakla, M. (red.), *Annales Academiae Paedagogicae Cracoviensis. Studia ad Didacticam Mathematicae Pertinentia II*, Wydawnictwo Naukowe Uniwersytetu Pedagogicznego, Kraków, 213–223.
- Rasiowa, H.: 1984, *Wstęp do matematyki współczesnej*, Państwowe Wydawnictwo Naukowe, Warszawa.
- Turnau, S.: 1991, O niektórych błędach dydaktycznych i ich skutkach, *Problemy dydaktyczne matematyki*, **tom V**, Wyższa Szkoła Pedagogiczna, Zielona Góra, 145–154.

*Instytut Matematyki
Uniwersytet Pedagogiczny
ul. Podchorążych 2
PL-30-084 Kraków
e-mail bpawlik@up.krakow.pl*

Katarzyna Słomczyńska

High School Identities*

Abstract. In 1969, Polish mathematician and logician, Alfred Tarski asked if all the identities true in the set of natural numbers involving the constant 1, addition, multiplication, and exponentiation can be derived from the eleven axioms that are taught at the high school level (*High School Identities*). In 1981 Alex Wilkie negatively solved this problem by constructing an identity that cannot be proved using these axioms. In this paper we survey results connected with Tarski's problem.

Wstęp

Na poziomie szkoły średniej, a nawet w nauczaniu studentów matematyki na wyższych uczelniach, niejasne pozostaje często rozróżnienie pomiędzy teorią (równościową lub pierwszego rzędu) a aksjomatami i modelami tej teorii (Baldwin, 2010). Oczywiście istnieją ważne teorie, w których rozróżnienie to jest zbędne, gdyż teoria jest wyznaczona przez jeden kanoniczny model i opisana przez skończony zbiór aksjomatów. W innych jednak przypadkach z naturalnej zdawałoby się aksjomatyki sformułowanej dla pewnego modelu nie da się wyprowadzić wszystkich twierdzeń prawdziwych w tym modelu. Co więcej, nie da się czasami takiego układu aksjomatów „poprawić” poprzez dopisanie do niego skończonej liczby twierdzeń. Najprostszego przykładu takiej sytuacji dostarcza zbiór liczb naturalnych z działaniami dodawania, mnożenia i potęgowania oraz próba aksjomatyzacji jego teorii przez tzw. *tożsamości szkoły średniej* (*High School Identities*), którą omawiam w tym artykule. Oparty jest on w dużym stopniu na przeglądowych pracach (Burris, Lee 1993); (Burris, Yeats 2004), a także innych podanych w literaturze.

1. High School Identities i problem Tarskiego

Następujące równości dla liczb naturalnych znane są każdemu, kto ukończył szkołę średnią:

*Tożsamości szkoły średniej

2010 Mathematics Subject Classification: Primary: 03C05; Secondary: 03C62

Key words and phrases: HSI-algebras, exponentiation, Wilkie's identity

$$\left. \begin{array}{l}
 (T1) \quad x + y = y + x, \quad x + (y + z) = (x + y) + z, \\
 (T2) \quad x \cdot y = y \cdot x, \quad x \cdot (y \cdot z) = (x \cdot y) \cdot z, \quad x \cdot 1 = x, \\
 (T3) \quad x \cdot (y + z) = x \cdot y + x \cdot z, \\
 (T4) \quad x^1 = x, \quad 1^x = 1, \\
 (T5) \quad x^{y+z} = x^y \cdot x^z, \quad (x \cdot y)^z = x^z \cdot y^z, \quad (x^y)^z = x^{y \cdot z}.
 \end{array} \right\} \begin{array}{l} \text{High} \\ \text{School} \\ \text{Identities} \\ \text{(HSI)} \end{array}$$

Pojawiły się one w tej postaci już w słynnej monografii Richarda Dedekinda z roku 1888: *Was sind und was sollen die Zahlen?* (Dedekind, 1888). Używając języka algebraicznego można powiedzieć, że algebra $\mathbf{N} := (\mathbb{N}, 1, +, \cdot, \uparrow)$, gdzie $\mathbb{N} := \{1, 2, \dots\}$, z jedną stałą 1 i trzema działaniami binarnymi $+, \cdot, \uparrow$, gdzie $x \uparrow y := x^y$ dla $x, y \in \mathbb{N}$, spełnia te jedenaście równości.

W roku 1969 polski matematyk i logik Alfred Tarski (1901-1983) zadał pytanie, czy z równości (T1)-(T5) można wyprowadzić wszystkie tożsamości dla liczb naturalnych, które zawierają stałą 1, dodawanie, mnożenie i potęgowanie (Doner, Tarski, 1969). Pytanie to można inaczej sformułować w następujący sposób:

PROBLEM TARSKIEGO: *Czy (T1)-(T5) tworzą zupełny układ aksjomatów teorii równościowej dla algebry $\mathbf{N}$?*

Wiadomo było, że analogiczne pytanie dla stałej 1, dodawania i mnożenia ma odpowiedź pozytywną. Dokładniej, wiadano, że identyczności (T1)-(T3) stanowią zupełny układ aksjomatów teorii równościowej dla algebry $\hat{\mathbf{N}} := (\mathbb{N}, 1, +, \cdot)$ (Henkin, 1977). Ponadto, było też wiadomo, że teoria ta jest *rozstrzygalna*, tj. istnieje algorytm pozwalający zdecydować, czy dana równość, w której występuje tylko stała 1, dodawanie i mnożenie jest prawdziwa w $\hat{\mathbf{N}}$, czy też nie, poprzez sprawdzenie wyrażen po obu jej stronach do postaci *normalnej*, którą w tym przypadku jest zwykły wielomian o współczynnikach postaci $\mathbf{k} := 1 + \dots + 1$ (k - razy), gdzie $k \in \mathbb{N}$.

Tarski dał ten problem do rozwiązania swojemu doktorantowi na uniwersytecie w Berkeley, Charlesowi Martinowi, ale temu udało się osiągnąć tylko częściowe wyniki. Pokazał on mianowicie (Martin, 1973), że problem Tarskiego ma rozwiązanie:

- *pozytywne*, gdy ograniczymy się tylko do potęgowania $(\mathbb{N}, \uparrow)$ lub do potęgowania i mnożenia $(\mathbb{N}, \cdot, \uparrow)$, oraz
- *negatywne*, gdy rozważymy strukturę z trzema działaniami $(\mathbb{N}, +, \cdot, \uparrow)$, ale bez stałej 1.

Nie rozwiązywało to jednak oryginalnego problemu. Kilka lat później wybitny amerykański logik Leon Henkin w swoim artykule (Henkin, 1977), w którym opisywał sytuację dla $\hat{\mathbf{N}}$, określił problem Tarskiego jako trudny i wciąż nierozwiązany pomimo wysiłku wielu matematyków.

Algebry spełniające równości (T1)-(T5) nazywamy *HSI-algebrami* (*HSI* jest to oczywiście akronim słów *high school identities*). Problem Tarskiego równoważny jest więc pytaniu, czy klasa algebr spełniających te same identyczności co $\mathbf{N}$ pokrywa się z klasą *HSI*-algebr. Tarski, stawiając problem, wierzył, że odpowiedź na

to pytanie jest pozytywna, ale nie umiał podać dowodu. Gdyby jednak ktoś chciał obalić przypuszczenie Tarskiego, musiałby to uzasadnić albo w sposób *syntaktyczny*, tzn. wskazać identyczność prawdziwą w $\mathbf{N}$ i niewyprowadzalną z identyczności (T1)-(T5), albo *semantyczny*, tzn. znaleźć algebrę, która należałaby do rozmaitości HSI , ale nie spełniałaby jakiegś tożsamości prawdziwej w $\mathbf{N}$.

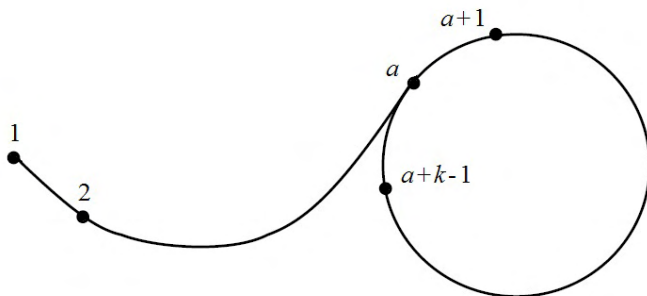
2. Klasa HSI

Klasa HSI zawiera wiele ciekawych algebr. Prosty przykład to $\mathbf{R}^+$: zbiór liczb rzeczywistych dodatnich $\mathbb{R}^+$ ze stałą 1 i naturalnymi działaniami dodawania, potęgowania i mnożenia. Inny, to zbiór potęgowy dowolnego zbioru X z tym zbiorem jako stałą 1 oraz z naturalnymi działaniami mnożeniowymi sumy i iloczynu, gdzie potęgowanie określone jest jako rzut na pierwszą współrzędną, tj. $U \uparrow V := U$ dla $U, V \subset X$. Bardziej wyrafinowany przykład (Birkhoff 1942) to algebra skończonych zbiorów częściowo uporządkowanych (rozważanych z dokładnością do izomorfizmu porządkowego) ze zbiorem jednoelementowym jako 1 oraz działaniami sumy rozłącznej, iloczynu kartezjańskiego i odwzorowań zachowujących porządek. Można ponadto pokazać, że istnieje 5 (z dokładnością do izomorfizmu) dwuelementowych HSI algebr (Asatryan, 2004), oraz 44 (z dokładnością do izomorfizmu) trójelementowe HSI algebry (Burris, Lee, 1993). Inne zasługujące na uwagę HSI -algebry to dwie nieizomorficzne algebry $\mathbf{N}_0(0)$ i $\mathbf{N}_0(1)$, otrzymane jako rozszerzenia $\mathbf{N}$ o element 0, z naturalnymi działaniami, gdzie przyjmujemy, odpowiednio, $0^0 = 0$ dla $\mathbf{N}_0(0)$ i $0^0 = 1$ dla $\mathbf{N}_0(1)$ (Asatryan, 2004).

Ciekawych przykładów skończonych HSI -algebr dostarczają nietrywialne homomorficzne obrazy $\mathbf{N}$. Aby je znaleźć, trzeba najpierw przeanalizować jak wyglądają *kongruencje*, czyli relacje równoważności zgodne z działaniami w $\hat{\mathbf{N}}$. Stosunkowo łatwo można pokazać, że muszą one mieć postać przystawiania $\equiv_{a,k}$, gdzie $a, k \in \mathbb{N}$, w którym „sklejamy” te liczby naturalne, które obie są większe lub równe a i przystają do siebie modulo k , czyli

$$m \equiv_{a,k} n \text{ wtw } m = n \text{ lub } m, n \geq a \text{ i } m \equiv_k n$$

dla $m, n \in \mathbb{N}$. Algebra ilorazowa $\hat{\mathbf{N}} / \equiv_{a,k}$ ma więc tu postać cyklu o długości k z „ogonem” o długości $a - 1$, zob. Ryc. 1.



Ryc. 1. Algebra ilorazowa $\hat{\mathbf{N}} / \equiv_{a,k}$

Powstaje pytanie: kiedy taka kongruencja w $\hat{\mathbf{N}}$ jest także kongruencją w $\mathbf{N}$, czyli kiedy zachowuje potęgowanie? Odpowiedział na nie, podając elegancką charakteryzację takich par liczb naturalnych a i k , doktorant Garretta Birkhoffa, John Dyer-Bennet (1940). Udowodnił on mianowicie, że relacja $\equiv_{a,k}$ jest kongruencją w $\mathbf{N}$ wtedy i tylko wtedy, gdy spełnione są dla każdej liczby pierwszej p i naturalnej m dwa następujące warunki:

1. jeżeli $p^m | k$, to $m \leq a$,
2. jeżeli $p | k$, to $p - 1 | k$.

Gdy rozważymy teraz szczególny przypadek $a = 1$, relacja $\equiv_{1,k}$ staje się zwykłym przystawianiem liczb naturalnych mod k . Wówczas $\hat{\mathbf{N}} / \equiv_{1,k}$ posiada element neutralny dla dodawania, którym jest klasa abstrakcji $k / \equiv_{1,k}$, i tworzy pierścień z jedynką w naturalny sposób izomorficzny z $\mathbb{Z}_k$. Stąd natychmiast wnioskujemy, że $\equiv_{1,k}$ będzie kongruencją w $\mathbf{N}$ wtedy i tylko wtedy, gdy *potęgowanie* w $\mathbb{Z}_k$ można zdefiniować w „naturalny” sposób. Jak zaraz pokażemy, działanie takie da się wprowadzić jedynie dla bardzo szczególnych wartości k . Postawienie takiej definicji nie jest możliwe na przykład w $\mathbb{Z}_3$, bo chociaż $1 \equiv 4 \pmod{3}$, to $2^1 \equiv 2 \not\equiv 2^4 \pmod{3}$.

To, dla jakich $k \in \mathbf{N}$ konstrukcja taka jest wykonalna, rozstrzyga cytowane wyżej twierdzenie Dyera-Benneta. Z pierwszego warunku wynika wtedy, że w rozkładzie k na czynniki pierwsze każda liczba może występować tylko w pierwszej potędze. Z drugiego zaś warunku wnioskujemy, że jeżeli uporządkujemy te czynniki pierwsze rosnąco, to każdy kolejny musi być równy iloczynowi pewnych mniejszych powiększonemu o 1. Teraz już łatwo można pokazać, że k musi być jedną z pięciu liczb: 1, 2, 6, $6 = 2 \cdot 3$, $42 = 2 \cdot 3 \cdot 7$, $1806 = 2 \cdot 3 \cdot 7 \cdot 43$. Więcej możliwych przypadków już nie ma, gdyż każdy iloczyn czynników pierwszych, na które rozkładamy 1806 powiększony o 1 jest albo jedną z liczb już występujących w ciągu, albo liczbą złożoną, w szczególności $1807 = 13 \times 139$. Tak więc potęgowanie jest możliwe tylko w czterech nietrywialnych pierścieniach cyklicznych: $\mathbb{Z}_2$, $\mathbb{Z}_6$, $\mathbb{Z}_{42}$ i $\mathbb{Z}_{1806}$!

Oczywiście żadna ze skończonych algebr ilorazowych algebry $\mathbf{N}$ nie może stanowić kontrprzykładu dla hipotezy Tarskiego, gdyż spełniają one wszystkie identyczności, które są prawdziwe w $\mathbf{N}$. Można jednak łatwo pokazać, że każda skończona *HSI*-algebra musi już zawierać którąś z tych algebr ilorazowych jako podalgebrę generowaną przez 1. Kontrprzykładu takiego można więc szukać rozszerzając taką algebrę ilorazową do większej, nie spełniającej już wszystkich identyczności prawdziwych w $\mathbf{N}$.

3. Kontrprzykład Wilkiego

Problem Tarskiego został rozwiązany w zaskakująco prosty sposób przez Alexa Wilkiego z Uniwersytetu w Oxfordzie (Wilkie, 1981). Podał on przykład tożsamości, która jest prawdziwa dla liczb naturalnych, a której nie da się wyprowadzić z High School Identities. Ciekawe, że preprint Wilkiego został opublikowany dopiero w roku 2000. *Tożsamość Wilkiego* ma postać równości dwóch iloczynów,

w których wielomiany zależne od jednej zmiennej x podnoszone są do potęg x i y , raz w jednej, a raz w drugiej kolejności:

$$(A^x + B^x)^y (C^y + D^y)^x = (A^y + B^y)^x (C^x + D^x)^y, \quad (W)$$

gdzie

$$\begin{aligned} A &:= x + 1, \\ B &:= x^2 + x + 1, \\ C &:= x^3 + 1, \\ D &:= x^4 + x^2 + 1. \end{aligned}$$

Tożsamość Wilkiego stanowiła sprytną modyfikację tożsamości zaproponowanej przez Martina w jego doktoracie (Martin, 1973), która jest prawdziwa w algebrze liczb naturalnych z dodawaniem, mnożeniem i potęgowaniem, ale bez jedynki, i której nie da się wyprowadzić z równości (T1)-(T3) i (T5) (ale da się ją wywieść z równości (T1)-(T5), jeżeli 1 w języku uwzględnimy!):

$$(x^y + x^y)^x (y^x + y^x)^y = (x^x + x^x)^y (y^y + y^y)^x.$$

Dlaczego tożsamość Wilkiego prawdziwa jest w $\mathbf{N}$, to znaczy $\mathbf{N} \models W$? Aby to zobaczyć, zauważmy, że w rozkładzie wielomianów C i D pojawia się ten sam czynnik $F := x^2 - x + 1$. Dokładniej, mamy $C = A \cdot F$ i $D = B \cdot F$, a stąd dla wszystkich rzeczywistych dodatnich liczb x i y (a więc w szczególności dla wszystkich liczb naturalnych) zachodzi:

$$\begin{aligned} (A^x + B^x)^y (C^y + D^y)^x &= (A^x + B^x)^y (A^y + B^y)^x F^{yx} \\ &= (A^x + B^x)^y F^{xy} (A^y + B^y)^x \\ &= (C^x + D^x)^y (A^y + B^y)^x. \end{aligned}$$

Dlaczego natomiast tożsamości Wilkiego nie da się wyprowadzić z HSI ? Istotnym powodem jest to, że w powyższym dowodzie wykorzystaliśmy wyrażenie F , które chociaż przyjmuje wyłącznie wartości naturalne dla $x \in \mathbf{N}$, to nie jest termem w języku algebry $\mathbf{N}$, czyli nie można go zapisać używając tylko stałej 1, dodawania, mnożenia i potęgowania, gdyż występuje w nim składnik ujemny. Oczywiście, nie oznacza to, że nie mogłyby *a priori* istnieć inny dowód tej tożsamości, korzystający wyłącznie z aksjomatów HSI . Wilkie podał dowód syntaktyczny, że $HSI \not\models W$, przeprowadzając go nie wprost – indukcją na długość ewentualnego takiego wyprowadzenia (Wilkie, 1981).

Pierwszy dowód semantyczny tego faktu zaproponował Yuri Gurevič (1985) konstruując 59-elementową HSI -algebrę, która nie spełnia tożsamości W . Później podawano coraz to mniejsze przykłady takich algebr. Ostatecznie Stanley Burris i Karen Yeats (2004) skonstruowali taką algebrę o 12 elementach, a rok później Jian Zhang (2005) pokazał, że najmniejsza HSI -algebra nie spełniająca W musi mieć co najmniej 11 elementów. Do dziś jednak nie wiadomo, czy taka HSI -algebra istnieje.

Otwarte pozostaje też pytanie: jaka jest najmniejsza *HSI*-algebra, która nie spełnia *jakiegś* tożsamości prawdziwej w $\mathbf{N}$ (takie tożsamości nazywa się czasem *egzotycznymi*)? Wiadomo jedynie, że musi ona liczyć co najmniej 3 elementy, gdyż, jak wspominaliśmy już, istnieje tylko 5 nieizomorficznych dwuelementowych *HSI*-algebr i okazuje się, że każda z nich spełnia wszystkie identyczności $\mathbf{N}$, więc nie mogą one stanowić kontrprzykładu dla hipotezy Tarskiego (Asatryan, 2004).

4. Skończona aksjomatyzowalność i rozstrzygalność teorii równościowej dla $\mathbf{N}$

Negatywne rozstrzygnięcie przez Wilkiego hipotezy Tarskiego sprawiło, że aktualne stało się pytanie o to, czy można zaksjomatyzować teorię równościową dla $\mathbf{N}$, dodając do jedenastu identyczności (T1)-(T5) jakiś skończony zbiór aksjomatów. Pytano więc, czy zbiór wszystkich identyczności prawdziwych w algebrze $\mathbf{N}$, który będziemy oznaczać dalej przez $Th(\mathbf{N})$, jest *skończenie aksjomatyzowalny*. Negatywnej odpowiedzi na to pytanie udzielił Gurevič (1990), który udowodnił, że teoria równościowa dla $\mathbf{N}$ nie jest skończenie aksjomatyzowalna. Wskazał on w tym celu nieskończony ciąg identyczności $(W_n)_{n \in \mathbb{N}}$ prawdziwych w $\mathbf{N}$ i podobnych w formie do W , ale zawierających tylko jedną zmienną, i pokazał, że z każdego skończonego zbioru identyczności dla $\mathbf{N}$ nie da się wyprowadzić przynajmniej jednej z nich. Identyczność W_1 uzyskuje się, wstawiając w identyczności Wilkiego 2^x za y . Ogólnie ciąg ten ma następującą postać:

$$(A_n^x + B_n^x)^{2^x} (C_n^{2^x} + D_n^{2^x})^x = (A_n^{2^x} + B_n^{2^x})^x (C_n^x + D_n^x)^{2^x}, \quad (W_n)$$

gdzie dla $n \in \mathbb{N}$

$$\begin{aligned} A_n &:= x + 1, \\ B_n &:= x^{2^n} + \dots + x + 1, \\ C_n &:= x^{2^{n+1}} + 1, \\ D_n &:= x^{4^n} + \dots + x^2 + 1. \end{aligned}$$

Chociaż algebry $\mathbf{N}$ nie da się opisać za pomocą skończonego zbioru aksjomatów, można pokazać, że teoria równościowa dla $\mathbf{N}$ jest *rozstrzygalna*, to jest istnieje algorytm (efektywna procedura), który rozstrzyga, czy dana formuła jest spełniona w $\mathbf{N}$ czy nie. Pierwszym krokiem dla pokazania tego wyniku było udowodnienie przez Martina (1973) rezultatu, stwierdzającego, że dokładnie te same tożsamości, które są prawdziwe w algebrze liczb rzeczywistych dodatnich z dodawaniem, mnożeniem, potęgowaniem i stałą 1, są też prawdziwe w $\mathbf{N}$, czyli że $Th(\mathbf{N}) = Th(\mathbf{R}^+)$.

Dowód Martina opierał się na następującym pomysłe z pracy (Hardy, 1910). Rozważamy zbiór $\mathcal{H}$ złożony z wszystkich funkcji prowadzących z $\mathbb{R}^+$ w $\mathbb{R}^+$, zdefiniowanych za pomocą stałych dodatnich, dodawania, mnożenia i potęgowania. W zbiorze tym definiujemy relację $\prec$ polegającą na tym, że jedna funkcja jest, począwszy od pewnego miejsca, silnie większa od drugiej. Okazuje się, że dla każdych dwóch różnych funkcji $f, g \in \mathcal{H}$, zachodzi albo $f \prec g$, albo $g \prec f$ (Hardy, 1910, Burris, Yeats, 2004). Innymi słowy wykresy dwóch takich funkcji nie mogą

nieskończenie wiele razy się przecinać. Każdemu termowi jednej zmiennej wyrażonemu w języku $\mathbf{N}$ da się przypisać w naturalny sposób funkcję z $\mathcal{H}$. Zatem, jeżeli równość takich dwóch termów jest nieprawdziwa w algebrze $\mathbf{R}^+$, to odpowiadające im funkcje różnią się dla dostatecznie dużych argumentów, a więc i dla pewnej liczby naturalnej. Stąd równość tych dwóch termów jest fałszywa także w $\mathbf{N}$. Sytuację, gdy termy zależą od większej liczby zmiennych, sprowadza się łatwo do przypadku jednej zmiennej za pomocą indukcji na ich liczbę.

Ostatecznie problem rozstrzygalności teorii równościowej dla $\mathbf{N}$ redukuje się więc do problemu rozstrzygalności dla $\mathbf{R}^+$, który Macintyre (1981), a kilka lat później niezależnie Gurevič (1985), rozwiązyali stosując zaawansowane metody analizy matematycznej.

Literatura

- Asatryan, G. R.: 2004, A solution to identities problem in 2-element HSI-algebras, *Math. Log. Quart.* **50**, 175-178.
- Baldwin, J. T.: 2010, Logic across the high school curriculum, *preprint*, University of Illinois at Chicago. <http://homepages.math.uic.edu/~jbaldwin/pub/logicaug20.pdf> [dostęp: 2016-01-29].
- Birkhoff, G.: 1942, Generalized arithmetic, *Duke Math. J.* **9**, 283-302.
- Burris, S., Lee, S.: 1993, Tarski's High School Identities, *Amer. Math. Monthly* **100**, 231-236.
- Burris, S., Yeats, K. A.: 2004, The saga of the High School Identities, *Algebra Universalis* **52**, 325-342.
- Dedekind, R.: 1888, *Was sind und was sollen die Zahlen?*, Verlag Friedrich Vieweg & Sohn, Braunschweig. [tłum. ang.: 1995, *What are numbers and what should they be?* Revised, edited, and translated from the German by H. Pogorzelski, W. Ryan and W. Snyder. RIM Monographs in Mathematics. Research Institute for Mathematics, Orono, ME].
- Doner, J., Tarski, A.: 1969, An extended arithmetic of ordinal numbers, *Fund. Math.* **65**, 95-127.
- Dyer-Bennet, J.: 1940, A theorem on partitions of the set of positive integers, *Amer. Math. Monthly* **47**, 152-154.
- Gurevič, R.: 1985, Equational theory of positive numbers with exponentiation, *Proc. Amer. Math. Soc.* **94**, 135-141.
- Gurevič, R.: 1990, Equational theory of positive numbers with exponentiation is not finitely axiomatizable, *Ann. Pure Appl. Logic* **49**, 1-30.
- Hardy, G. H.: 1910, *Orders of Infinity. The 'Infinitärarcalcul' of Paul Du Bois-Reymond*, Cambridge University Press, Cambridge.
- Henkin, L.: 1977, The logic of equality, *Amer. Math. Monthly* **84**, 597-612.
- Macintyre, A.: 1981, The laws of exponentiation, w: C. Berline, K. McAloon, J.-P. Res-sayre (red.), *Model Theory and Arithmetic*, Lecture Notes in Math. **890**, Springer, Berlin, 185-197.
- Martin, C. F.: 1973, *Equational Theories of Natural Numbers and Transfinite Ordinals*, Ph.D. Thesis, University of California, Berkeley, CA.

- Wilkie, A. J.: 1981, On exponentiation - a solution to Tarski's high school algebra problem, *preprint*, Oxford University, 2000, w: A. Macintyre (red.), *Connections between Model Theory and Algebraic and Analytic Geometry*, Quaderni di Matematica **6**, Naples, 107-129. <http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.15.9695&rep=rep1&type=pdf> [dostęp: 2016-01-29].
- Zhang, J.: 2005, Computer search for counterexamples to Wilkie's Identity, w: R. Nieuwenhuis (red.), *Automated Deduction – CADE-20*, Lecture Notes in Computer Science **3632**, Springer, Berlin, 441-451.

*Instytut Matematyki
Uniwersytet Pedagogiczny
ul. Podchorążych 2
PL-30-084 Kraków
e-mail ksłomcz@up.krakow.pl*

Damian Wiśniewski

Wokół liczb i szeregów harmoniczych*

Abstract. The harmonic series is one of the most celebrated infinite series of mathematics. From a pedagogical point of view, the harmonic series provides a wealth of opportunities. Applications such as *Gabriel's wedding cake* and Euler's proof of the divergence of prime numbers can lead to some very nice discussions. The main idea of this article is to survey some of unusual, insightful and inspiring divergence proofs. First of all, this article is addressed at first-year calculus students.

1. Wstęp

Liczbę $\mathcal{H}_n^{(s)} = \sum_{k=1}^n \frac{1}{k^s}$ nazywamy n -tą *liczbą harmoniczną rzędu s* , zaś szereg $\sum_{n=1}^{\infty} \frac{1}{n^s}$ *szeregiem harmonicznym rzędu s* , $s \in \mathbb{R}$. Termin *harmoniczny* wywodzi się z tego, że każdy wyraz ciągu o przepisie $a_n = \frac{1}{n}$, począwszy od drugiego, jest średnią harmoniczną wyrazów sąsiednich, to znaczy

$$a_k = \frac{2}{\frac{1}{a_{k-1}} + \frac{1}{a_{k+1}}}, \text{ dla } k \geq 2.$$

W dalszej części artykułu terminy: *liczby harmoniczne* i *szereg harmoniczny* oznaczać będą liczby harmoniczne i szereg harmoniczny rzędu pierwszego. Będziemy pisać $\mathcal{H}_n$ zamiast $\mathcal{H}_n^{(1)}$.

Jako pierwszy własności liczb harmoniczych badał Nicholas Oresme. W 1350 roku udowodnił, że szereg harmoniczny jest rozbieżny. Co więcej, pokazał, iż sumy częściowe tego szeregu spełniają nierówność $\mathcal{H}_{2^k} > \frac{k+1}{2}$, dla $k = 1, 2, \dots$

Ponad czterysta lat później oszacowaniami liczb harmoniczych zajmował się Leonhard Euler (Euler, 1734); w XIX, XX i XXI wieku to zagadnienie było badane między innymi w (Cesaro, 1885), (Lodge, 1904), (Toth, Mare, 1991), (DeTemple, Wang, 1991), (Chen, Qi, 2008) oraz (Villarino, 2004, 2007).

*Around harmonic numbers and harmonic series

2010 Mathematical Subject Classification: Primary: 97A80, 97B40, 97I30

Key words and phrases: harmonic numbers, harmonic series, divergence

Liczby harmoniczne są powszechne w nauce: w fizyce - oporność zastępcza układu oporników połączonych równolegle jest równa średniej harmonicznej poszczególnych oporności, w informatyce - w analizie algorytmów, w muzyce - przy zależnościach między tonem podstawowym a długością struny wyściowej.

Liczby harmoniczne występują w statystykach dotyczących zjawisk pogodowych. Analizując zapisy meteorologiczne w Chicago, Kifowit i Stamps (Kifowit, Stamps, 2006) zauważyli zależność między rokiem, który przyniósł rekordowe opady śniegu, a wartościami kolejnych liczb harmonicznych; generalnie po n latach obserwacji odnotowywano około $\mathcal{H}_n$ rekordowych lat.

Zadziwiająca jest sztuczka karciana oparta na pomysśle R.T. Sharpa (Sharp, 1954), która pokazuje, jak w naturalny sposób liczby harmoniczne pojawiają się w prostych sytuacjach. Przypuśćmy, że chcemy ułożyć na stole stos kart tak, aby nawis utworzony przez karty wystające poza krawędź stołu był maksymalny. Jeśli mamy k kart, d_{k+1} oznacza odległość krawędzi wierzchniej karty od krawędzi stołu, to otrzymujemy zależność $d_{k+1} = \mathcal{H}_k$ (więcej szczegółów w (Patashnik, Graham, Knuth, 2001)).

W kolejnych rozdziałach przytoczymy wybrane, różnorodne, dowody rozbieżności szeregu harmonicznego, pokażemy zastosowanie liczb harmonicznych w dowodzie twierdzenia o istnieniu nieskończenie wielu liczb pierwszych oraz wyprowadzimy podstawowe oszacowania liczb harmonicznych.

Artykuł jest adresowany przede wszystkim do studentów pierwszego roku matematyki, może stanowić przykład ciekawego wykładu w ramach kursu analizy matematycznej przy omawianiu pojęcia szeregów liczbowych i ich zbieżności. Uczniowie, którzy ukończyli szkołę średnią na profilu rozszerzonym z matematyki wiedzą, że suma nieskończonej ilości dodatnich składników może być skończona (na przykładzie sumy składników nieskończonego ciągu geometrycznego o ilorazie q , w przypadku $|q| < 1$). Z doświadczenia autora, podpartego rozmowami ze studentami pierwszego roku matematyki, wynika, że intuicja zawodzi ich w przypadku obliczenia sumy odwrotności wszystkich liczb naturalnych. Podawane są różne skończone wyniki, prawidłowa odpowiedź jest rzadkością. Główną ideą artykułu jest przedstawienie wybranych dowodów rozbieżności szeregu harmonicznego, wzbogaconych interpretacją geometryczną. Dowody te pokazują piękno i różnorodność matematyki na przystępnym poziomie.

2. Zbieżność szeregu harmonicznego

FAKT 1

Szereg harmoniczny jest szeregiem rozbieżnym.

Powyższe stwierdzenie zostało udowodnione na kilkadziesiąt (Kifowit, Stamps, 2006a) sposobów. Najpierw przytoczymy dwa z nich, najczęściej cytowane w podręcznikach analizy matematycznej (patrz (Fichtenholz, 1995), (Kaczyński, 2005)).

Dowód 1. (Nichola Oresme, 1350).

Zachodzi oczywista nierówność

$$\frac{1}{n+1} + \frac{1}{n+2} + \cdots + \frac{1}{2n} > n \frac{1}{2n} = \frac{1}{2}.$$

Jeżeli po odrzuceniu dwóch pierwszych wyrazów, pozostałe wyrazy szeregu harmonicznego rozbijemy na grupy po $2, 4, 8, \dots, 2^{k-1}, \dots$ wyrazów w każdej:

$$\underbrace{\frac{1}{3} + \frac{1}{4}}_2, \underbrace{\frac{1}{5} + \frac{1}{6} + \frac{1}{7} + \frac{1}{8}}_{2^2}, \underbrace{\frac{1}{9} + \dots + \frac{1}{16}}_{2^3}, \dots, \underbrace{\frac{1}{2^{k-1}+1} + \dots + \frac{1}{2^k}}_{2^{k-1}}, \dots,$$

to każda z tych sum z osobna będzie większa od $\frac{1}{2}$. Faktycznie,

$$\frac{1}{2^{k-1}+1} + \dots + \frac{1}{2^k} > \frac{2^{k-1}}{2^k} = \frac{1}{2} \quad \text{dla } k \geq 2.$$

Stąd

$$\mathcal{H}_{2^k} > k \cdot \frac{1}{2}.$$

Sumy częściowe nie są ograniczone z góry, więc szereg jest rozbieżny (analogicznie można pokazać, że $\mathcal{H}_{M^k} > k \cdot \frac{M-1}{M}$). $\square$

Dowód 2. Skorzystamy ze znanych nierówności $\frac{t}{t+1} \leq \ln(1+t) \leq t$ dla $t > -1$. Podstawiając w nich $t = \frac{1}{k}$, gdzie $k \in \mathbb{N}$, otrzymamy

$$\frac{1}{k+1} \leq \ln\left(1 + \frac{1}{k}\right) \leq \frac{1}{k}.$$

Sumując powyższe nierówności dla $k = 1, 2, \dots, n$ mamy

$$\frac{1}{2} + \frac{1}{3} + \dots + \frac{1}{n+1} \leq \sum_{k=1}^n \ln\left(1 + \frac{1}{k}\right) \leq 1 + \frac{1}{2} + \frac{1}{3} + \dots + \frac{1}{n}.$$

Oznaczając teraz n -tą sumę częściową szeregu harmonicznego przez s_n , otrzymujemy

$$s_{n+1} - 1 \leq \ln(n+1) \leq s_n.$$

Z tego, że $\lim_{n \rightarrow \infty} \ln(n+1) = \infty$ wynika, szereg harmoniczny jest rozbieżny. $\square$

W innej wersji powyższego dowodu korzysta się z kryterium ilorazowego dla szeregu harmonicznego i szeregu $\sum_{n=1}^{\infty} \ln\left(1 + \frac{1}{n}\right)$.

Dowód 3. (Pietro Mengoli, połowa XVII wieku) Zauważmy najpierw, że

$$\frac{1}{n-1} + \frac{1}{n+1} = \frac{2n}{n^2-1} > \frac{2}{n}, \quad \text{dla } n = 2, 3, \dots$$

Przypuśćmy, że szereg harmoniczny jest zbieżny do S . Wtedy:

$$\begin{aligned} S &= 1 + \left(\frac{1}{2} + \frac{1}{3} + \frac{1}{4}\right) + \left(\frac{1}{5} + \frac{1}{6} + \frac{1}{7}\right) + \left(\frac{1}{8} + \frac{1}{9} + \frac{1}{10}\right) + \dots \\ &> 1 + \frac{3}{3} + \frac{3}{6} + \frac{3}{9} + \dots = 1 + S. \end{aligned}$$

Otrzymaliśmy sprzeczność. $\square$

Dowód 4. (Jacob Bernoulli, 1689) Zauważmy, że jeśli $c \in \mathbb{N}$, $c > 1$, to

$$\frac{1}{c+1} + \frac{1}{c+2} + \cdots + \frac{1}{c^2} \geq (c^2 - c) \frac{1}{c^2} = 1 - \frac{1}{c}.$$

Zatem $\frac{1}{c} + \frac{1}{c+1} + \frac{1}{c+2} + \cdots + \frac{1}{c^2} \geq 1$. Stąd

$$\begin{aligned} \sum_{n=1}^{\infty} \frac{1}{n} &= 1 + \left(\frac{1}{2} + \frac{1}{3} + \frac{1}{4} \right) + \left(\frac{1}{5} + \cdots + \frac{1}{25} \right) + \left(\frac{1}{26} + \cdots + \frac{1}{676} \right) + \cdots \\ &\geq 1 + 1 + 1 + \cdots \quad \square \end{aligned}$$

Dowód 5. Przypuśćmy, że szereg harmoniczny jest zbieżny do S . Wtedy $S > 2$, ponieważ $\mathcal{H}_4 = \frac{25}{12}$. Zauważmy, że

$$\begin{aligned} S &= 1 + \left(\frac{1}{2} + \frac{1}{3} \right) + \left(\frac{1}{4} + \frac{1}{5} + \frac{1}{6} \right) + \left(\frac{1}{7} + \cdots + \frac{1}{10} \right) \\ &\quad + \left(\frac{1}{11} + \cdots + \frac{1}{15} \right) + \cdots > 1 + \frac{2}{3} + \frac{3}{6} + \frac{4}{10} + \frac{5}{15} + \cdots \\ &= \frac{2}{2} + \frac{2}{3} + \frac{2}{4} + \frac{2}{5} + \frac{2}{6} + \cdots = 2 \sum_{n=2}^{\infty} \frac{1}{n} = 2(S-1). \end{aligned}$$

Stąd $S < 2$, co jest sprzeczne z naszym wyjściowym spostrzeżeniem. $\square$

Kolejny dowód wiąże liczby harmoniczne z liczbami Fibonacciego (Kifowit, Stamps, 2006b), (Chen, Kennedy, 2012).

Dowód 6. Ciąg Fibonacciego to ciąg liczb naturalnych określany rekurencyjnie w następujący sposób:

$$f_0 = 1, \quad f_1 = 1, \quad f_{n+1} = f_n + f_{n-1}, \quad \text{dla } n = 1, 2, \dots$$

Kolejne wyrazy tego ciągu nazywane są *liczbami Fibonacciego*. Skorzystamy ze znanego faktu:

$$\lim_{n \rightarrow \infty} \frac{f_{n+1}}{f_n} = \frac{1 + \sqrt{5}}{2}.$$

Zauważmy, że

$$\lim_{n \rightarrow \infty} \frac{f_{n-1}}{f_{n+1}} = \lim_{n \rightarrow \infty} \frac{f_{n+1} - f_n}{f_{n+1}} = \lim_{n \rightarrow \infty} \left(1 - \frac{f_n}{f_{n+1}} \right) = 1 - \left(\frac{1 + \sqrt{5}}{2} \right)^{-1} \neq 0. \quad (1)$$

Mamy

$$\begin{aligned} \sum_{n=1}^{\infty} \frac{1}{n} &= 1 + \frac{1}{2} + \frac{1}{3} + \left(\frac{1}{4} + \frac{1}{5} \right) + \left(\frac{1}{6} + \frac{1}{7} + \frac{1}{8} \right) + \cdots \\ &\geq 1 + \frac{1}{2} + \frac{1}{3} + \frac{2}{5} + \frac{3}{8} + \cdots = 1 + \sum_{n=1}^{\infty} \frac{f_{n-1}}{f_{n+1}}. \end{aligned}$$

Z (1) wynika, że szereg $\sum_{n=1}^{\infty} \frac{f_{n-1}}{f_{n+1}}$ jest rozbieżny, co pociąga za sobą rozbieżność szeregu harmonicznego. $\square$

Przytoczmy jeszcze jeden z nowszych dowodów (patrz Sinha, 2013):

Dowód 7. Zauważmy, że

$$\mathcal{H}_{k+m} = \mathcal{H}_k + \frac{1}{k+1} + \frac{1}{k+2} + \cdots + \frac{1}{k+m} > \mathcal{H}_k + \frac{m}{k+m}.$$

Dla dowolnej dodatniej liczby całkowitej k mamy $\lim_{m \rightarrow \infty} \frac{m}{k+m} = 1$. Ustalając $\varepsilon \in (0, 1)$ dostajemy, że dla każdego k istnieje m takie, że $\mathcal{H}_{k+m} > \mathcal{H}_k + \varepsilon$. Stąd ciąg $(\mathcal{H}_n)$ nie jest ograniczony z góry, szereg harmoniczny jest rozbieżny. $\square$

Pomimo tego, że szereg harmoniczny jest rozbieżny, liczby harmoniczne rosną „wyjątkowo wolno”. Dla przykładu $\mathcal{H}_8 = 2\frac{201}{280}$, $\mathcal{H}_{20} \approx 3,6$, zaś $\mathcal{H}_{13000}$ ledwie przekracza 10.

W obliczu tych faktów zaskakujące może się wydawać, że

LEMAT 1

Liczba harmoniczna jest liczbą naturalną tylko dla $n = 1$.

Dowód. Oczywiście $\mathcal{H}_1 = 1 \in \mathbb{N}$. Rozważmy $\mathcal{H}_n$ dla $n > 1$ i wybierzmy k takie, że $2^k \leq n < 2^{k+1}$. W ten sposób mamy

$$\mathcal{H}_n = 1 + \frac{1}{2} + \frac{1}{3} + \cdots + \frac{1}{2^k} + \cdots + \frac{1}{n}.$$

Niech teraz M będzie najmniejszą wspólną wielokrotnością wszystkich mianowników różnych od 2^k :

$$M = \text{NWW}(1, 2, 3, \dots, 2^k - 1, 2^k + 1, \dots, n).$$

Stąd 2^{k-1} jest dzielnikiem M , zaś 2^k nim nie jest. Mnożąc $\mathcal{H}_n$ przez M otrzymamy

$$M \cdot \mathcal{H}_n = M + \frac{M}{2} + \frac{M}{3} + \cdots + \frac{M}{2^k} + \cdots + \frac{M}{n} = t + \frac{M}{2^k}, \quad \text{gdzie } t \in \mathbb{N}.$$

Z określenia M wynika, że $\frac{M}{2^k} \notin \mathbb{N}$. Stąd $M \cdot \mathcal{H}_n \notin \mathbb{N}$, więc również $\mathcal{H}_n \notin \mathbb{N}$ dla $n > 1$.

Przytoczmy teraz twierdzenie o zbieżności szeregu harmonicznego rzędu s . Z wniosku z tego twierdzenia skorzystamy w dalszej części artykułu.

FAKT 2

Szereg harmoniczny rzędu s jest zbieżny dla $s > 1$ i rozbieżny dla $s \leq 1$.

Dowód. Rozważmy trzy przypadki:

1. Niech $s = 1$. Otrzymujemy szereg harmoniczny, więc rozbieżny.

2. Niech $s < 1$. Z kryterium porównawczego mamy

$$\frac{1}{n} \leq \frac{1}{n^s} \quad \text{dla } n \in \mathbb{N}, \quad s < 1 \quad \text{i} \quad \sum_{n=1}^{\infty} \frac{1}{n} = +\infty,$$

zatem szereg harmoniczny rzędu $s < 1$ jest rozbieżny.

3. Niech $s > 1$. Łącząc wyrazy szeregu harmonicznego rzędu s w grupy, tak jak poniżej, otrzymamy sumy

$$\underbrace{\frac{1}{3^s} + \frac{1}{4^s}}_2, \quad \underbrace{\frac{1}{5^s} + \frac{1}{6^s} + \frac{1}{7^s} + \frac{1}{8^s}}_{2^2}, \quad \dots, \quad \underbrace{\frac{1}{(2^{k-1}+1)^s} + \dots + \frac{1}{(2^k)^s}}_{2^{k-1}}, \dots \quad (2)$$

Zauważmy, że

$$\frac{1}{(n+1)^s} + \frac{1}{(n+2)^s} + \dots + \frac{1}{(2n)^s} < n \frac{1}{n^s} = \frac{1}{n^{s-1}}.$$

Z powyższych zależności widać, że kolejne sumy w (2) są mniejsze od kolejnych wyrazów ciągu geometrycznego o przepisie $a_n = \frac{1}{(2^{s-1})^n}$. Zatem każda suma częściowa rozpatrywanego szeregu jest mniejsza od liczby

$$L = 1 + \frac{1}{2^s} + \frac{\frac{1}{2^{s-1}}}{1 - \frac{1}{2^{s-1}}},$$

więc szereg harmoniczny rzędu $s > 1$ jest zbieżny.

WNIOSEK 1

Ciąg $(H_n^{(s)})$ liczb harmonicznych rzędu $s > 1$ ma granicę. Nazywamy ją dztę Riemanna i zapisujemy :

$$\zeta(s) := \mathcal{H}_{\infty}^{(s)} = \sum_{k=1}^{\infty} \frac{1}{k^s}.$$

Powyższy wniosek wykorzystamy przy dowodzie Lematu 3.

Aby pobudzić wyobraźnię studentów odnośnie do rozbieżności szeregu harmonicznego, warto im przedstawić interpretację geometryczną szeregu harmonicznego, noszącą nazwę *Gabriel's wedding cake* (Fleron, 1999).

Określimy funkcję f następująco: $f(x) = \frac{1}{[x]}$ dla $x \in [1, \infty)$, gdzie symbolem $[x]$ oznaczono cechę liczby x . Obracając wykres funkcji f wokół osi OX , otrzymamy figurę przedstawioną na Ryc. 1.

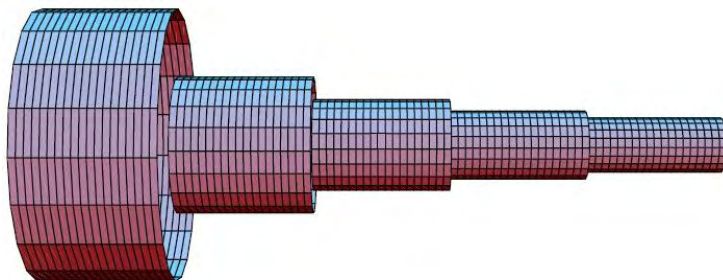
Składa się ona z nieskończonej ilości otwartych walców o wysokościach długości 1 i promieniach, kolejno, $1, \frac{1}{2}, \frac{1}{3}, \dots$. Zatem

$$V = \sum_{k=1}^{\infty} \pi \left(\frac{1}{k} \right)^2 = \pi \sum_{k=1}^{\infty} \frac{1}{k^2}.$$

Na mocy Wniosku 1 wiemy, że suma szeregu jest zbieżna do liczby $\zeta(2)$, czyli $\frac{\pi^2}{6}$ (Dunham, 1990). Stąd $V = \frac{\pi^3}{6}$. Z drugiej strony, jeśli policzymy pole powierzchni bocznej rozważanej bryły, dostaniemy

$$P = \sum_{n=1}^{\infty} 2\pi \cdot \frac{1}{n} = 2\pi \sum_{n=1}^{\infty} \frac{1}{n}.$$

Okazuje się więc, paradoksalnie, że rozważany tort „możemy zjeść, ale nie zamrozić” (Fleron, 1999).



Ryc. 1. Gabriel's wedding cake (Kifowit, Stamps, 2006b)

3. Liczby harmoniczne a liczby pierwsze

Liczby harmoniczne odgrywają zasadniczą rolę w dowodzie twierdzenia Eulera o istnieniu nieskończenie wielu liczb pierwszych. Najpierw przytoczymy twierdzenie o jednoznaczności rozkładu liczby naturalnej na iloczyn liczb pierwszych:

Twierdzenie 1

Każdą liczbę naturalną $n \geq 2$ można jednoznacznie (z dokładnością do kolejności czynników) zapisać w postaci iloczynu liczb pierwszych.

Dowód. Zobacz (Sierpiński, 1950).

Twierdzenie 2

Istnieje nieskończenie wiele liczb pierwszych.

Dowód. Przypuśćmy, że istnieje skończenie wiele liczb pierwszych, oznaczmy je przez $p_1, p_2, \dots, p_n$. Rozważmy szereg

$$1 + \frac{1}{p_k} + \frac{1}{p_k^2} + \dots, \quad 1 \leq k \leq n.$$

Oczywiste jest, że są one zbieżne, jako szeregi geometryczne o ilorazach $\frac{1}{p_k}$, gdyż $\left| \frac{1}{p_k} \right| < 1$. Zatem ich sumy S_{p_k} są równe $\frac{1}{1 - \frac{1}{p_k}}$. Stąd zaś

$$S = \prod_{k=1}^n S_{p_k} = \prod_{k=1}^n \left(\frac{1}{1 - \frac{1}{p_k}} \right).$$

Z drugiej strony, biorąc pod uwagę iloczyn Cauchy'ego rozważanych szeregów, mamy

$$S = \lim_{m_1, \dots, m_n \rightarrow \infty} \sum_{l_1=0}^{m_1} \frac{1}{p_1^{l_1}} \sum_{l_2=0}^{m_2} \frac{1}{p_2^{l_2}} \cdots \sum_{l_n=0}^{m_n} \frac{1}{p_n^{l_n}}$$

$$= 1 + \frac{1}{p_1} + \cdots + \frac{1}{p_n} + \frac{1}{p_1^2} + \frac{1}{p_1 p_2} + \cdots + \frac{1}{p_1 p_n} + \frac{1}{p_2^2} + \frac{1}{p_2 p_3} + \cdots + \frac{1}{p_n^2} + \cdots$$

Zauważmy, że każda z liczb, w rozkładzie których występują liczby $p_1, \dots, p_n$ znajduje się w mianowniku jednego z wyrazów iloczynu Cauchy'ego rozważanych szeregów $S_{p_1} \cdot S_{p_2} \cdot \dots \cdot S_{p_n}$ i to dokładnie jeden raz.

Z jednoznaczności rozkładu na liczby pierwsze wnioskujemy, że powyższe mianowniki odpowiadają wszystkim różnym liczbom naturalnym. Stąd naszą sumę S możemy zapisać jako $S = \sum_{i=1}^{\infty} \frac{1}{i}$. Stąd suma szeregu harmonicznego jest liczbą, bowiem S jest liczbą. Otrzymaliśmy sprzeczność. Zatem liczb pierwszych jest nieskończenie wiele.

4. Wybrane oszacowania liczb harmoniczných

Pokażemy teraz przykłady podstawowych twierdzeń dotyczących oszacowań liczb harmoniczných.

LEMAT 2

Dla $n \in \mathbb{N}$ zachodzą nierówności:

$$\frac{n+1}{2} \leq \mathcal{H}_{2^n} \leq n+1.$$

Dowód. (indukcyjny)

1. Dla $n = 1$ mamy $\mathcal{H}_2 = \frac{3}{2}$, więc $1 \leq \mathcal{H}_2 < 2$.
2. Załóżmy, że $\frac{k+1}{2} \leq \mathcal{H}_{2^k} \leq k+1$ dla dowolnie ustalonego $k \in \mathbb{N}$.

Wtedy:

$$\mathcal{H}_{2^{k+1}} = \mathcal{H}_{2^k} + \frac{1}{2^k+1} + \cdots + \frac{1}{2^{k+1}} \leq k+1 + \frac{2^{k+1} - 2^k}{2^k} = k+2.$$

Z drugiej strony

$$\mathcal{H}_{2^{k+1}} = \mathcal{H}_{2^k} + \frac{1}{2^k+1} + \cdots + \frac{1}{2^{k+1}} \geq \frac{k+1}{2} + \frac{2^{k+1} - 2^k}{2^{k+1}} = \frac{k+2}{2}.$$

Stąd $\frac{k+2}{2} \leq \mathcal{H}_{2^{k+1}} \leq k+2$.

Zatem, na mocy zasady indukcji matematycznej, $\frac{n+1}{2} \leq \mathcal{H}_{2^n} \leq n+1$ dla każdego $n \in \mathbb{N}$.

Przy pomocy liczb harmoniczych można zdefiniować liczbę γ zwaną stałą Eulera. Mianowicie:

LEMAT 3

Istnieje granica:

$$\lim_{n \rightarrow \infty} (\mathcal{H}_n - \ln n).$$

Dowód. Korzystając z rozwinięcia w szereg Maclaurina, mamy:

$$\ln \left(\frac{k}{k-1} \right) = -\ln \left(1 - \frac{1}{k} \right) = \frac{1}{k} + \frac{1}{2k^2} + \frac{1}{3k^3} + \frac{1}{4k^4} + \dots.$$

Szereg ten jest zbieżny dla wszystkich k takich, że $-1 < -\frac{1}{k} \leq 1$, więc w szczególności dla $k > 1$. Jeśli zsumujemy obie strony po $2 \leq k \leq n$, to otrzymamy kolejno

$$\begin{aligned} \sum_{k=2}^n [\ln k - \ln(k-1)] &= \sum_{k=2}^n \left(\frac{1}{k} + \frac{1}{2k^2} + \frac{1}{3k^3} + \dots \right), \\ \ln n - \ln 1 &= \sum_{k=2}^{\infty} \frac{1}{k} + \frac{1}{2} \sum_{k=2}^{\infty} \frac{1}{k^2} + \frac{1}{3} \sum_{k=2}^{\infty} \frac{1}{k^3} + \dots, \\ \ln n &= (\mathcal{H}_n - 1) + \frac{1}{2}(\mathcal{H}_n^{(2)} - 1) + \frac{1}{3}(\mathcal{H}_n^{(3)} - 1) + \dots. \end{aligned}$$

Stąd zaś

$$\mathcal{H}_n - \ln n = 1 - \frac{1}{2}(\mathcal{H}_n^{(2)} - 1) - \frac{1}{3}(\mathcal{H}_n^{(3)} - 1) - \frac{1}{4}(\mathcal{H}_n^{(4)} - 1) - \dots.$$

Zgodnie z Wnioskiem 1, na podstawie głębszych twierdzeń analizy matematycznej, dostaniemy

$$\lim_{n \rightarrow \infty} (\mathcal{H}_n - \ln n) = 1 - \frac{1}{2}(\zeta(2) - 1) - \frac{1}{3}(\zeta(3) - 1) - \frac{1}{4}(\zeta(4) - 1) - \dots,$$

gdzie po prawej stronie tej równości mamy sumę szeregu zbieżnego, a więc liczbę. Liczba ta nazywana jest stałą Eulera γ , wynosi ona w przybliżeniu 0,577218 (Euler, 1734).

Powyższy lemat nasuwa przypuszczenia, że istnieje zależność między liczbami harmonicznymi a logarytmem naturalnym.

Faktycznie, mamy następujące oszacowanie:

TWIERDZENIE 3

Dla każdego $n \geq 2$ zachodzą nierówności:

$$\ln(n+1) < \mathcal{H}_n < \ln n + 1.$$

Dowód. Określmy funkcję $f : [1, \infty) \rightarrow \mathbb{R}$ wzorem $f(x) = \frac{1}{x}$.
Rozważmy rodzinę prostokątów o wierzchołkach:

$$A_n = (n, 0), \quad B_n = (n+1, 0), \quad C_n = (n, f(n)), \quad D_n = (n+1, f(n))$$

leżących (częściowo) nad wykresem funkcji f . Każdy z tak utworzonych prostokątów ma bok długości 1 leżący na osi x , zaś długości sąsiednich boków tych prostokątów wynoszą kolejno $1, \frac{1}{2}, \frac{1}{3}, \dots$.

Zatem:

$$\int_1^{n+1} f(x) dx = \ln(n+1) < 1 + \frac{1}{2} + \frac{1}{3} + \dots + \frac{1}{n} = \mathcal{H}_n.$$

Rozważmy teraz rodzinę prostokątów o wierzchołkach:

$$A_n = (n, 0), \quad B_n = (n+1, 0), \quad C_n = (n, f(n+1)), \quad D_n = (n+1, f(n+1))$$

leżących pod wykresem funkcji f . Podobnie jak poprzednio każdy prostokąt ma bok długości 1 leżący na osi x , zaś długości sąsiednich boków tych prostokątów wynoszą kolejno $\frac{1}{2}, \frac{1}{3}, \frac{1}{4}, \dots$.

Zatem:

$$\int_1^n f(x) dx = \ln n > \frac{1}{2} + \frac{1}{3} + \dots + \frac{1}{n} = \mathcal{H}_n - 1.$$

Stąd $\mathcal{H}_n < \ln n + 1$. Ostatecznie $\ln(n+1) < \mathcal{H}_n < \ln n + 1$.

PRZYKŁAD 1

Oszacujmy wartość liczby $\mathcal{H}_{16} = \mathcal{H}_{2^4} = 3^{\frac{274399}{720720}}$.

Na mocy Lematu 2 mamy

$$2 \leq \mathcal{H}_{16} \leq 4,$$

zaś z twierdzenia 3, ponieważ $\ln 17 > 2,83$ i $\ln 16 < 2,77$ wynika, że

$$2,83 < \mathcal{H}_{16} < 3,77.$$

Literatura

- Cesaro, E.: 1885, Sur la serie harmonique, *Nouvelles Annales de Mathématiques* **4**.
 Chen, C.-P., Qi, F.: 2008, The best bounds of the n^{th} harmonic number, *The Global Journal of Applied Mathematics & Mathematical Sciences* **1**(1), 41-49.
 Chen, H., Kennedy, C.: 2012, Harmonic series meets Fibonacci sequence, *The College Mathematics Journal* **43**, 237-243.
 DeTemple, D., Wang, S.-H.: 1991, Half-integer approximations for the partial sums of the harmonic series, *J. Math. Anal. Appl.* **160**.
 Dunham, W.: 1990, *Journey Through Genius*, John Wiley & Sons.
 Euler, L.: 1734, De progressionibus harmonicis observationes, *Commentarii academiae scientiarum imperialis Petropolitanae* **19**, 150-161.

- Fichtenholz, G. M.: 1995, *Rachunek różniczkowy i całkowy, t. II*, PWN, Warszawa.
- Fleron, J. F.: 1999, Gabriel's wedding cake, *College Mathematics Journal*, 35-38.
- Kaczyński, A.: 2005, *Podstawy analizy matematycznej. Rachunek całkowy. Szeregi*, Oficyna Wydawnicza Politechniki Warszawskiej.
- Kifowit, S., Stamps, T.: 2006a, The harmonic series diverges again and again, *The AMATYC Review* **27**, 31-43.
- Kifowit, S., Stamps, T.: 2006b, Serious About the Harmonic Series II, *Prairie State College*.
- Lodge, A.: 1904, An approximate expression for the value of $1 + \frac{1}{2} + \frac{1}{3} + \dots + \frac{1}{r}$, *Messenger of Mathematics* **30**.
- Patashnik, O., Graham, R., Knuth, D.: 2001, *Matematyka konkretna*, PWN, Warszawa, 304-307.
- Sharp, R.: 1954, Problem 52: Overhanging dominoes, *Pi Mu Epsilon Journal* **1**, 411-412.
- Sierpiński, W.: 1950, *Teoria liczb, Monografie Matematyczne* **19**, Instytut Matematyczny Polskiej Akademii Nauk, Warszawa - Wrocław.
- Sinha, P.: 2013, An easy proof of the divergence of the harmonic series sum, *American Mathematical Monthly* **120**, 354.
- Toth, L., Mare, S.: 1991, Problem E 3432, *Amer. Math. Monthly* **98**.
- Villarino, M. B.: 2004, Ramanujan's Approximation to the n th Partial Sum of the Harmonic Series, *Depto. de Matematica, Universidad de Costa Rica*, 2-6.
- Villarino, M. B.: 2007, *Sharp Bounds for the Harmonic Numbers*, Depto. de Matematica, Universidad de Costa Rica.

Wydział Matematyki i Informatyki
Uniwersytet Warmińsko - Mazurski
ul. Słoneczna 54
PL-10-710 Olsztyn
e-mail dawi@matman.uwm.edu.pl

Lidia Zaręba

The role of visualisation in the process of generalisation*

Abstract. In my paper, I discuss the results of individual observations of pupils aged 13-14. These children undertook an attempt to solve problems which were designed in order to bring about an inductive type of generalizing. The main aim of the study was to classify typical methods of proceeding, which represent the pupils' ways of reasoning and to determine what influence a particular method of proceeding may have on the final outcome of their work. These results indicate the general pupils' strategies of generalizing. Presumably, visual thinking produces a positive effect on the pupils' process of generalizing and abilities to describe regularity with the aid of a letter symbol.

1. Introduction

In modern mathematics education, mastering the approach is valued more than mastering the knowledge. This is why great importance is attached to developing the mathematical activity of the students. The preferred methods allow students to discover or create mathematics. According to J. Filip and T. Rams: *Children should learn to communicate, articulate, and describe their ideas and doubts, be allowed attempts and mistakes, cooperate with each other.* They need to learn to listen to others, and, most importantly, learn on their own, "find their own approach to mathematics" (Filip, Rams, 2000, p. 9). In order to shape this approach, mathematical activities and behaviour should be nourished. A. Z. Krygowska outlined a number of such activities, including the perception and use of analogy, creating schemes, assimilation and processing of information, deduction, reducing problems into others, interpretation and use of definitions, creating and using algorithms (Krygowska, 1986), as well as generalizing (Krygowska, 1981).

The need of introducing elements of generalisation-based reasoning can be noticed in the recent years in teaching centres both in Poland as well as worldwide.

*Rola wizualizacji w procesie uogólniania

2010 Mathematics Subject Classification: Primary: 97C70; Secondary: 97-02, 97B50

Key words and phrases: generalisation, visualisation in mathematics

This is influenced by the international PISA research (*Programme for International Student Assessment*), the aim of which is to assess the knowledge, skills, and stance of 15 year old students in the context of the three concepts which were deemed most important in the development of teenagers. Emphasis is placed on *reading and reasoning in humanities, mathematics* as well as *reasoning in natural sciences*. In the context of mathematics, the skills taken into consideration include, among others, the ability of abstract thinking, analysis, and generalisation. In the context of the development of such activities, the need for the didactic activities of the teacher are also worth taking into consideration. The aim of the activities is to *develop the habit of observation, experimentation, seeking and attaining knowledge on one's own; this requires, on all levels, teaching the development of perception and use of analogies (similarities and differences), empirical inference, use of recurrent reasoning, and inductive generalisation* (Legutko, Stańdo, 2008, p. 32; own emphasis). The aforementioned mathematical activities can be developed on the basis of geometry and arithmetic. In the recent years, a significant amount of mathematics education literature is dedicated to the process of generalisation (Presmeg, 1999; Ciosek, 1995; Zaksis, Liljedahl, 2002; Legutko, 2010; Flores, 2012; Hitt, 2012; Malara, 2012; Vinner, 2012; Zaręba, 2012), emphasising the impact of developing the ability to perceive relations between numeric values, which are relations in arithmetic. Generalising the reasoning conducted on particular numbers leads to algebraic thinking, and, in its final form, to using symbolic notation.

Both algebraic thinking and using symbolic notation are issues for students in many stages of learning. This leads to a discussion regarding the ways of teaching algebra. The teacher faces new tasks which require him to engage in different types of activities, as well as creating or modifying educational situations in such a way as to make the acquisition of algebraic knowledge as accessible as possible to the students. In particular, it is necessary for the teachers to have knowledge of the nature of algebraic thinking. It is therefore important to note the results of research of the thought process of a student who generalises and uses symbolic notation. This is the type of research this article refers to. The research itself described in detail in another work (Zaręba, 2012).

Advancing from arithmetic to algebra, which usually takes place at primary school level, usually starts with generalizing by varying constants or by inductive generalizing. The latter begins by observing specific attributes of tasks, and noticing the rules and relations therein. It depends on the teacher whether and how such skills will be developed by the students. As stated by Legutko, teachers *will prefer the development of the skill of generalizing if they can generalise as well. (...) They should also have the ability to conduct a discussion with the students regarding them noticing the various rules, ways of writing things down with numbers, letters, algebraic expressions, and equations* (Legutko, 2010, p. 114).

In finding one's way to mathematics and discovering relations and rules, visualisation can be used. The role of this activity has been increasingly prominent in mathematics education literature (Castelnuovo, 1987; Kuřina, 1998; Presmeg, 1999; Tlustý, 2002; Rösken, Rolka, 2006; Yilmaz, Argün, Keskin, 2009; Sochański, 2011; Flores, 2012; Hitt, 2012). It is important for the teacher to notice how the

student makes use of it and uses it on the way to generalisation. This knowledge can be used to properly guide the ones who are struggling with mathematics.

2. Terminology - inductive type of generalizing and visualisation

This work was inspired by one of the articles regarding this topic (Yilmaz, Argün, Keskin, 2009). The authors conducted and described studies which allowed me to reflect upon my work (Zareba, 2012) in the context of noticing the role of visualisation in the process of generalizing. Analysing the results of both works seems interesting due to the fact that the studies described by the authors were conducted independently of one another and used a similar research tool, though aimed at different age groups. The subject of these studies is a specific type of generalizing, i.e. inductive generalizing. This is why the definition of the term as well as a description of visualisation are provided further.

In literature, while defining visualisation, different aspects of it are noted. This article follows the definition made of the authors of the synthetic depiction of visualisation (Hershkowitz, 1989; Zimmermann, Cunningham, 1991; Arcavi, 2003; Rösken, Rolka, 2006; Yilmaz, Argün, Keskin, 2009; Hitt, 2012), including the authors of the article which is referred to further in the text. They note: **Visualisation** is the ability, the process and the product of creation, interpretation, use of and reflection upon pictures, images, diagrams in our minds, on paper or with technological tools, with the purpose of depicting and communicating information, thinking about and developing previously unknown ideas and advancing understandings (Yilmaz, Argün, Keskin, 2009, p. 131; in: Arcavi, 2003; Hershkowitz, 1989; Zimmermann, Cunningham, 1991).

The purpose of visualisation is notable e.g. in situations where the method used for a specific case which does not change for a different case can allow a person to understand the rules governing the world of mathematics. Among such situations are e.g. interpreting the law of swapping the places of numbers in a sum (see: Figure 1) or notable product equations presented in a square (see: Figure 2; in: Siwek, 2005, p. 322).

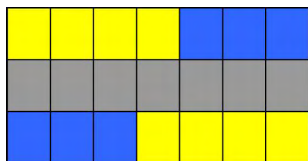


Figure 1. $4 + 3 = 3 + 4$

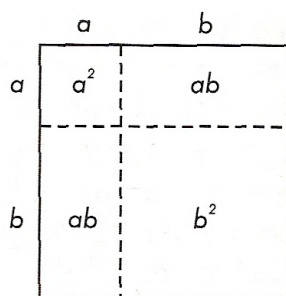


Figure 2. $(a + b)^2 = a^2 + 2ab + b^2$

Generalising of the inductive type is the process of discovering general rules on the basis of observation and comparison of particular cases (Krygowska, 1977; Gucewicz-Sawicka, 1982; Nowak, 1989).

An example of this type of generalizing is presented by Z. Krygowska:

The student is to calculate a sum

$$S_n = \frac{1}{1 \cdot 2} + \frac{1}{2 \cdot 3} + \frac{1}{3 \cdot 4} + \dots + \frac{1}{(n-1) \cdot n}.$$

He begins his reasoning by calculating the first three terms of the equation S_n . The result is:

$$S_1 = \frac{1}{2}, \quad S_2 = \frac{2}{3}, \quad S_3 = \frac{3}{4}.$$

Those are three specific theorems. The student notices that these equations can be a result of the equation $S_n = \frac{n}{n+1}$ by substituting $n = 1$, $n = 2$, and $n = 3$ respectively. The theorem $\bigwedge_{n \in N} S_n = \frac{n}{n+1}$ is a more general theorem than the three specified theorems, as all of them can be a result of the general theorem by specifying it. All of the three specific theorems are certainly true, but the general theorem is, in this stage, still a hypothesis which needs to be studied.

(Krygowska, 1977, pp. 112-113)

A way of merging visualisation with inductive generalisation can be observed in the following task regarding handshakes:

Find the number of handshakes performed when the following number of people meet: 2 people, 3 people, 4 people, 5 people, 6 people, 10 people, and 100 people. How many handshakes will there be, if n people meet?

Here, generalisation manifests in studying the consecutive specific situations of the task (see: Figure 3; Baranowska, 1992, p. 11). All of them regard a specific number of people shaking their hands. The analysis of the procedure or the results (the number of handshakes) leads to solving the task for n people. Depending on the reasoning used, the solution can be presented in multiple ways, e.g. by use of algebraic expressions: $\frac{n(n-1)}{2}$ or $1 + 2 + 3 + \dots + (n-1)$.

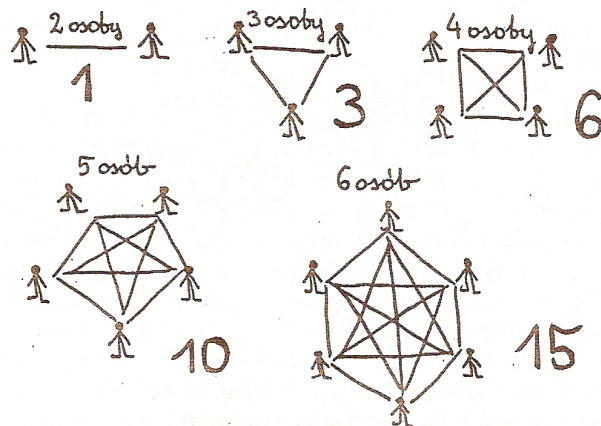


Figure 3. Handshakes

What turns out to be immensely helpful in solving the task is making drawings for the specific cases of the task and their proper interpretation. The role of such drawings in generalizing is outlined further. The description is made in the context of the aforementioned Turkish and Polish studies. The following paragraphs contain insight regarding both studies, as well as the conclusions drawn by the authors and my own thoughts on the matter.

3. Description and conclusions from Turkish research

The Turkish research involved teachers from Turkish upper secondary schools. One of the tasks they solved was this:

*How many straight lines can pass through n points
whenever three of them are non collinear?*

The problem was not formulated generally for n points at first; it started with two, then three, then four points, gradually increasing until the subject uses generalisation. Thus, the subjects were provoked to perform inductive generalisation. Furthermore, every subject was interviewed, which helped if and how the subjects visualizes and verbalizes while solving the task.

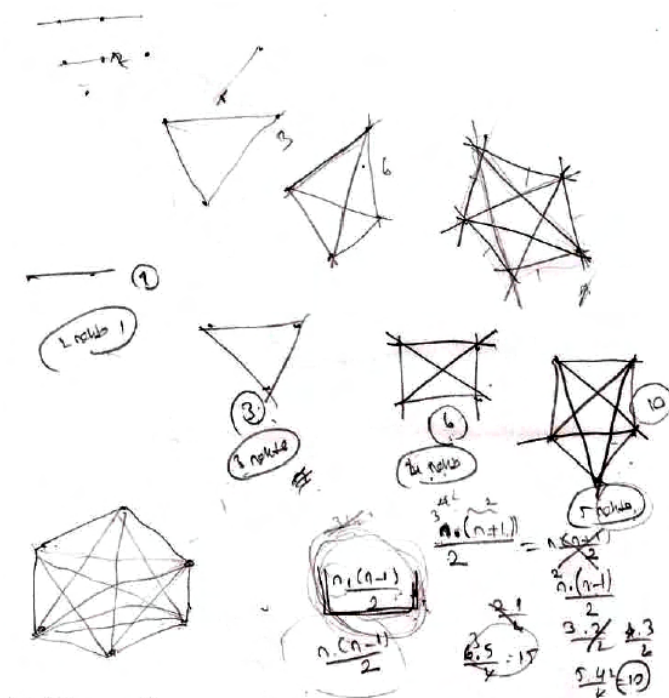


Figure 4. A teacher's solution

The authors describe two representative solutions of the subjects. I outline one of them (see: Figure 4; Yilmaz, Argün, Keskin, 2009, p. 132); provide a short

description, the thoughts and conclusions of the authors; I quote selected fragments of the interviews used in the work to justify my interpretation, based on my experience with similar studies of my own¹.

As stated by the authors, the teacher (the subject whose work is presented) used visualisation during her process of generalisation, which helped her discover the proper algebraic expression. The work process of the subject is presented below; crucial steps are supplemented with quotes from the work and own comments.

1. The subject knew that only one line goes through two different points. Despite this, she made a drawing of the situation. In the words of the authors, visual thinking was important for her to make sure her answer was right.
2. In the case of three points, at first she drew them as collinear points, which is when she noticed that they are not. This shows that the drawing helped the subject to find one of the conditions of the task, i.e. the noncollinearity of points.
3. The authors state that in the case of 4 and 5 points, the teacher used visualisation as a way of thinking of possible lines, as the subject visualized the task (which, according to the authors, is indicated by drawing points and lines going through all pairs of points) and found the solution by calculating the drawn lines:

When four points were given she directly drew two points and added other two points that are non collinear and then drew the lines by counting. She passed six lines and she was sure about the points. Because she counted and couldn't find anymore lines. So, visualisation was her way to think about the possible lines. And when five points were given the same thing happened.

(Yilmaz, Argün, Keskin, 2009, p. 134)

Taking into account the definition of visualisation provided in the introduction, I think that while the subject did make use of the drawing, I would not define making use of a drawing at this stage as visualisation; it appears as if the drawings are not yet linked to any in-depth thought regarding the solution, rather being used as a tool which helps pinpoint and calculate the proper lines. The drawing stage is, I think, the phase of understanding the task, creating the basis of observation and in-depth thought, which can help focus the subject's attention of different aspects which lead to generalisation. If the subject notices the essence of the general solution in their drawings, I will treat such a drawing as well as the reasoning behind it as a visualisation leading to generalisation.

¹ My interpretation of the work presented by the authors differs to some extent, the interpretation slightly deviating from the one presented by the authors, although I do regard the examples and the analysis presented in the work as very interesting, as well as stimulating reflection and discussion as well as exchanging points of view of research results. I would like to thank the authors of the work for their inspiration.

4. In the case of 5 points, the subject found – as a result of drawing and calculating lines – only 9 (instead of 10) lines. This seems to be a crucial element of her work. This is when the subject develops the need of generalisation, creating a formula. The following is a fragment of the conversation which proves this (S and I are remarks of the subject and observer respectively).

S: /.../ I have nine lines (she thinks). But it mustn't be nine? According to me I must have ten lines.

I: Why?

S: (she restart to count the lines that she drew).one, two, three, ..., nine (she adds the last nine) and ten. I have ten lines.

I: Why did you think that you must have ten lines?

S: I thought to reduce to a formula.

(Yilmaz, Argün, Keskin, 2009, p. 134)

The fact that the subject expected a different result than the one she received from her calculations suggests, in my opinion, that at this point she had already noticed the **arithmetical relation between the received numbers**. She probably had not noticed the relation between the way of drawing the lines and their quantity yet; if that were the case, she would have used a system of drawing lines in order to not omit any. However, in the case of 5 points, the tenth line was missing. The fact that the subject noticed the relation only in the scope of the numbers (instead of both the numbers and the drawing) is also supported by the fact that the authors, when describing the work of the subject, only pay attention to **drawing** and calculating, and not **drawing in a specific way**.

It is worth questioning in what way did the subject create the proper algebraic expression, as well as, particularly, whether visualisation played a role at this stage. Presented below is a fragment of a conversation with the subject related to the discovery of the formula as well as two interpretations (the authors' and mine) in the context of the use of visualisation at this stage of her work.

I: O.K. Can you reduce a formula?

S: Hmm ... Now ... **I passed one line from two points**. (she draws again two points and a line and write 1 near the picture). Then I passed three lines from three points. (she draws again and write 3 near it). If I think four points (she draws again four points and **draws the lines by counting**) one, two, .. and six. O.K. I have six lines (and she writes 6 near the picture). So, If I think five points (**she draws the same picture by counting**) one ... ten. I have ten lines. (she writes 10 near it and then she controls the pictures and thinks) **one from two points, three from three points, six from four points, ten from five points**. Hmm ... I must formulate it with n . (she thinks) I think to have a sequence and then to reduce to a formula. (she writes the numbers of lines and thinks). **I pass a line from two points** (writes $\frac{n(n+1)}{2}$ and **tries the numbers on this formula**). **Three lines from three points**, so, I must write something in the parenthesis that must simplify the number 2, O.K. It must have 2 in it (she thinks and writes $\frac{n(n+1)}{2}$ and thinks the numbers) no, this isn't (she crosses out that she wrote and thinks). If I write $\frac{n(n-1)}{2}$?

I: Why did you do this?

S: When I try these points; one from two, three from three .. (she tries the numbers of the points on the formula). The formula works for five points. If I try for six points (she writes for six points and has 15). I have 15 lines. ??? Is it true? Hmm. . . (she draws six points and passes lines from them by counting) one, two, ... fourteen, fifteen. Yes, I have 15 lines. It is true.

(Yilmaz, Argün, Keskin, 2009, p. 134; own emphasis)

The authors, noticing the useful role of visualisation in the generalisation, interpret the conversation as follows.

When reducing the formula, the denominator of the formula was 2. She recognized that a line passes through two points. So she put 2 in the formula. Then she thought $n(n+1)$ for the part of remain and she tried the numbers on this. The lines that she drew helped her to recognize that the formula was not true. Then she thought $n(n-1)$ and she tried again and found the formula. Also **during this time visualisation helped her to find the formula.**

(Yilmaz, Argün, Keskin, 2009, p. 134; own emphasis)

The authors emphasize that the subject made drawings and calculated lines, but my concern is whether the subject noticed anything more in the way she drew the lines, whether there was any thought process put into this, whether dividing by 2 and creating factors n , $n-1$, or $n+1$ was just a way of matching the equations to the received numbers in such a way as to make the whole formula work for the specific numbers which were the solutions in their specific cases, or whether the role of visualisation here did not only consist of making a drawing, but also on “seeing” the elements of the formula in the drawing. It is my understanding (considering the definition of visualisation outlined earlier) that **visualisation is the art of looking, but also seeing.** In this task, this would indicate seeing the general formula based on the visual representation. The visualisation could then be considered a vessel for the general solution. However, in the described work it seems that despite the drawings, the subject did not make much use of them: her comments concerned strictly the numbers received, which she compared just before attempting to create a formula: **“one for two points, three for three points, six for four points, ten for five points.** Hmm... I need to formulate this using n .” Therefore, she was analysing the numbers and trying to cover them with a formula with the n variable. It is sensible to think of visualisation in the context of one of the remarks of the subject that **the denominator has to be 2, as one line goes through 2 points.** However when taking the entire remark into account, it does not have to indicate a general observation which was used in further reasoning. The subject did not state that a line goes through **all** pairs of points; she could have only noticed the fact that the first solved case (of two points) the answer was one. This number is twice as small as the variable n , which in this case equals 2. This could suggest dividing by 2 in the sought-after formula.

The role of visualisation in the presented task could be presented in the reasoning shown for the 5-point case in Figure 5.

In conclusion, it is my point of view that in the example analysed above, visualisation does not play a significant role in the process of generalisation. While

a drawing was indeed used, as the task is of a geometric nature, the subject created the drawing in order to find the numerical results in specific cases of the task. Creating the general formula did not, however, involve the drawing. The subject found the formula by adapting algebraic expressions to the numbers received. The reasoning of the subject lacks an important element of visualisation – reflecting on the drawing. This thought process – in terms of the definition of visualisation used by me and the authors of the article – should lead to *discovering new ideas, new knowledge, and deeper reasoning*.

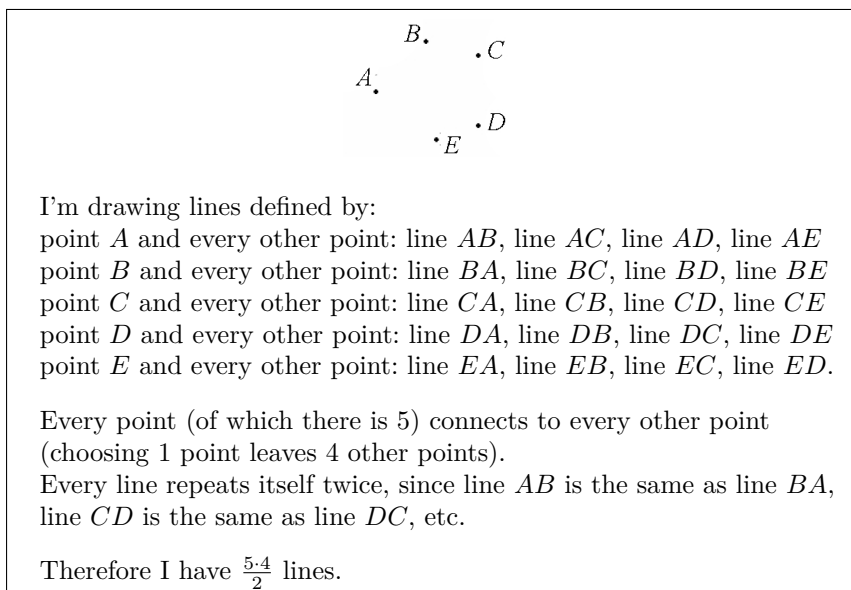


Figure 5. Solution involving the use of visualisation

In my opinion, the role of visualisation is reflected very well in the second example presented by the authors. In the cited work, the subject gets the following numbers in consecutive cases (2, 3, 4, 5, and 6 points): 1, 3, 6, 10, 15; she notices that the number of lines increases in the following case by 2, 3, 4, and 5 when compared to the previous case. Searching for a general equation, she finds the visual justification of the growth:

S: /.../ (he adds one point to the picture for four points by drawing). When I add one point, one point is added for each one. When I add one point to 4 points, 4 lines are added. I mean that 4 lines are added to the number of lines for 5 points. So, I think, $n - 1$ lines will be added to the number of line for n points. I must know $n - 1^{th}$ to know n^{th} . /.../

(Yilmaz, Argün, Keskin, 2009, p. 135)

Situations like this combine the algebraic equation with its graphic representation. The subject creates a drawing which ends up being thought-provoking. She notices more components of the desired equation while drawing.

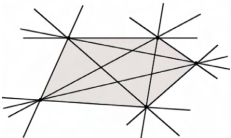
The two approaches of the subjects presented below clearly show the difference between visualisation and a drawing which is not used for generalisation. This difference

is outlined in the next paragraph. While describing the solutions of students contained therein, I detail the solutions which are related to a properly made drawing or chart. However, I emphasize the role of the solutions which contain both a drawing and a thought process related to the drawing.

4. Description and conclusions from Polish studies

The Polish studies described herein were aimed at 13-14 year old middle school students. Their main goal was the description of the process the students go through to perform inductive generalisation and write it down by using symbolic notation. By analysing these studies, the occurrence of visualisation can be pinpointed alongside its role in creating a symbolic generalisation. The students were solving various tasks during individual conversations; I will focus on a problem similar to the one related to the Turkish studies. It is clearly visible in Figure 6. The solution, in the case of n lines, leads to the same formula as in the case of the Turkish studies, i.e. $\frac{n \cdot (n-1)}{2}$, or alternately: $1 + 2 + 3 + \dots + (n - 1)$.

1. A convex pentagon is given. Every two vertices of this polygon define one straight line.
Were all the possible straight lines drawn?
What is the number of possible straight lines?



2. Draw a triangle and a convex quadrilateral.
In both cases draw all the straight lines defined by each pair of vertices of the figure.
How many such lines are defined by the vertices of a triangle?... Of a quadrilateral?...

3. Fill in the table.

Number of the angles of a polygon	Number of straight lines defined by each pair of the polygon's vertices
3	
4	
5	
6	
7	
10	
24	
345	

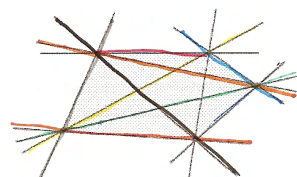
4. Let n denote a number of vertices of a polygon. How can the number of the straight lines defined by this n -gon's vertices be described?

Figure 6. Research issue of Polish research

By solving the tasks of the card, the students advanced through subsequent specific cases: for $n = 2, 3, 4$, etc. In every one of them, they created a certain situation model. Solving the problem within the framework of a particular model is connected with an interpretation based on a produced sketch or arithmetical scheme. In the work-sheets of the participants I have distinguished six mathematical models. Due to the size of the article, only three will be described in more detail – those which are relevant to the concept of visualisation. The description will be based on fragments of the work of one of the students. Other models will only be acknowledged.

1. **Geometrical model (G)** – a pupil indicates lines at the drawing and arrives at the solution by counting the previously indicated lines (see: figure 7).

1. Dany jest pięciokąt wypukły. Każde dwa wierzchołki tego wielokąta wyznaczają jedną prostą.
Czy narysowano wszystkie proste? **TAK**....
Ile ich jest?**10**.



2. Narysuj trójkąt i czworokąt wypukły.
W obu przypadkach narysuj wszystkie proste wyznaczone przez każde dwa wierzchołki figury.
Ile takich prostych wyznaczają wierzchołki trójkąta?**3**..... A czworokąta?**6**.....

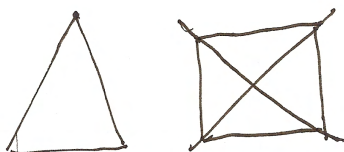


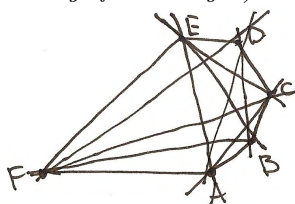
Figure 7. Excerpt of the work of the Polish student

2. **Geometrical–arithmetical model (G-A)** – in order to find the number of lines in a given particular case a pupil indicates them in a way which guarantees that no lines has been omitted. He uses this method to create the arithmetic expression which serves as the solution in the case of a particular polygon.

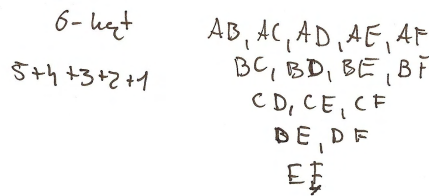
O: For the hexagon, you didn't know whether you had drawn all the straight lines. **Let's try to find a method that would allow us to state that all the straight lines are already drawn.**

S: I think I know. Let's say, for the hexagon.

(He points at the drawing of the hexagon)



S: (The student creates and makes comments on the undermentioned note). *AB, aha, with A. That would be: AB, AC, AD, AE, AF. And now for B: BC, BD, BE, BF and that's all of them. And now for C: CD, CE, CF (he finishes in silence).*



S: I've noticed that at the beginning I got 5 straight lines and this number decreased. It was 5, 4, 3, 2, 1.

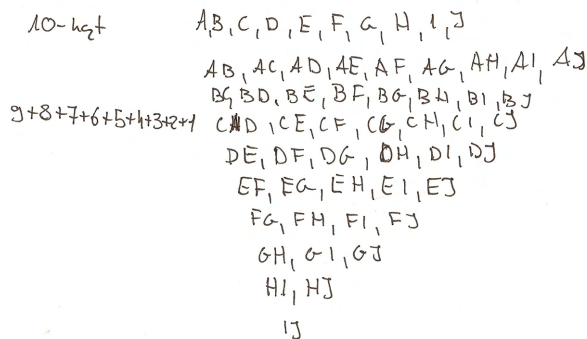
O: So, how can this number be obtained?

S: 5 + 4 + 3 makes 12...

O: I don't want to know the number, but the method.

S: 5 + 4 + 3 + 2 + 1.

For a decagon



S: On the beginning, I had 9, so 9 + 8 + 7 + 6 + 5 + 4 + 3 + 2 + 1.

Visualisation in the cited work is not just the drawing, which is used to calculate the lines; there is a general blueprint of a solution to be observed for every case of the task – this blueprint or drawing allows for the generalisation to be observed in the form of a formula. A specific layout of the lines (in the form of an “upside-down pyramid”) allows to observe a sum, the subsequent components of which present the number of lines in a particular location of the layout.

3. **Arithmetical model (A)** – the student identifies the number of lines with a respective arithmetic expression without appealing to geometric interpretation of the solution of this problem.

S: (When the number of the angles of a polygon is 24) I think I'm not going to write all the numbers down, it might be done in a shorter way, but how...

O: If you do it this in a way you have here (for the decagon) – would it be OK?

S: Yes, but in the case of a polygon that has 345 interior angles, I would need lot of sheets.

O: Maybe you could give up something; take a look: for the hexagon you have been still drawing. For the decagon you've given up drawing – and you have less work to do. And now – what could you give up to make it even shorter?

S: Aaaah! (flash) I only need to use this expression (he points at the sum of the numbers from 9 to 1). I mean, if there were nine of

these numbers, I can write: $1 + 2 + 3$, and so to 9. /.../ It will be $1 + 2$ and dots and $+ 23$.

$24 - 6 = 7$

$1 + 2 + 3 + 4 + 5 + 6 + 7 + 8 + 9 + 10 + 11 + 12 + 13 + 14 + 15 + 16 + 17 + 18 + 19 + 20 + 21 + 22 + 23$

O: And for 345-gon?

S: It will be *(he writes the sum)*.

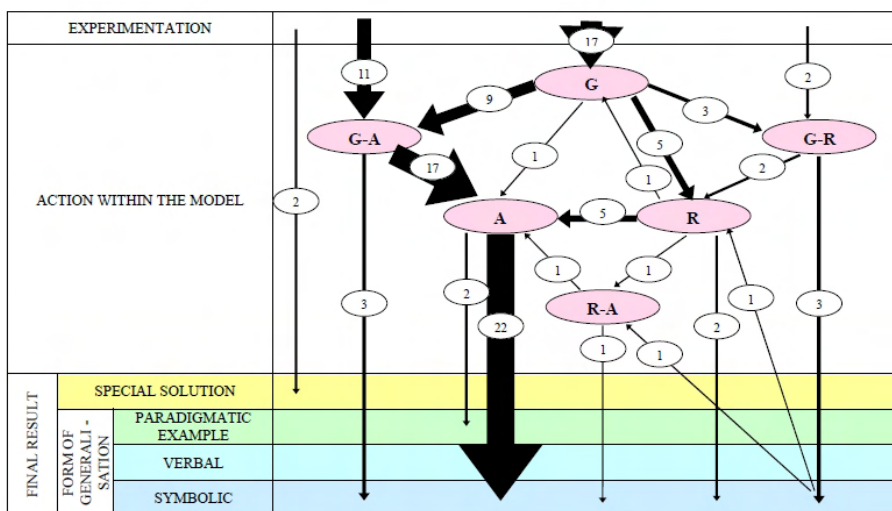
$1 + 2 + 3 + \dots + 345$

To observe the application of the models described above in the full works of the subjects, I will point out the three remaining models. They are related to the perception of a recursive relationship.

4. **Geometrical-recursive model (G-R)** – every time after drawing the next point, the student “sees” new – when compared to the previous situation – lines. This corresponds to the second solution presented in the article described above.
5. **Recursive model (R)** – while analysing (in isolation from the sketch) the number of lines for a few particular cases, a pupil notices a recursive relationship between them, notices the consecutive increase by 2, 3, 4, ... of the number of lines with respect to the number of lines for the previous polygon.
6. **Recursive-arithmetical model (R-A)** – the student notices the recursive relationship (that is, he uses the **R** model), notices how the number of lines increases subsequently by 2, 3, 4,... in comparison to the “previous polygon,” and he creates a formula based on this relationship, which allows to find the number of lines for a given polygon without having to refer to its previous case.

The analysis of the behaviour of the subjects suggests that activity related to mathematical models plays an important role in the generalisation process. In particular, before performing generalisation and writing it down using symbolic notation, the student’s reasoning makes use of a specific sequence of models used. I have noted the aforementioned models in the work of every student and indicated, by using arrows on the drawing, the process the student goes through from the moment of familiarizing himself with the task (the experimentation phase), through working with various models, to solving the task and performing generalisation. Scheme 1 contains various types of solutions (task solutions only for specific cases, solution in the form of a paradigmatic example, and generalisation in verbal or symbolic form; see: Zaręba, 2012, p. 130; own translation). In the context of this work, the arrows relevant to the task are those pointing at the solution in a symbolic form. The numbers on the arrows represent the number of people who advanced from the model where the arrow begins to the one the arrow is pointing at.

Scheme 1 shows that the activity of all subjects begins in model **G** or a related model (**G-A** or **G-R**). It is worth noting that model **G** is, in my opinion, unrelated to visualisation, as the link between the drawing and the suitable numeric expression is established by models **G-A** and **G-R**; the most common path (**G** → **G-A** → **A**) indicates that model **G** is simply the introduction to **G-A** model reasoning, in which the activities of most subjects gradually deviate from geometry and move mostly towards model **A**, which is a scheme with use of numbers. This is one step away from generalisation and creating a formula containing a variable. The students transition from a given arithmetic expression to an algebraic one.



Scheme 1. Synthesis of Paths Leading to the Generalisation

The most effective way is by using model **G-A** which links the student's drawing to the arithmetic expression. This is the stage of the student's work in which I can see a lot of potential, as this is the stage which involves the use of visualisation in generalisation. It seems that this is the path the students should be encouraged to take – so that they can “see” the proper relationships and the formula in the drawing they're looking at.

5. Conclusion

The authors of the works described in the article refer to the role of visualisation in the process of generalisation. The definition of visualisation is the same among them, but their interpretations differ in regards to solutions which use schemes or drawings. This is reflected in the analysis of the first solution presented in the article describing the Turkish research. The Turkish authors are of the opinion that the subject uses visualisation, while the author of the Polish work does not acknowledge this. It seems as though the student's drawing is not coupled with an appropriate thought process. The reflecting should be based on the drawing, not on the numerical result on which the drawing was based. In the first case the formula of the identified relationship can be seen in the drawing, whereas in the second case the drawing is not directly related to the identified relationship. Furthermore, different outlooks on the same process show how difficult it is to understand the thought process of someone who is solving a given task, and how important it is to properly interpret the verbal and non-verbal messages of the person.

In the case of the second solution described by the Turkish authors, there is no discrepancy in the interpretation of visualisation. This is why in the scope of the work where drawings were used to “see” the appropriate formula, both the Turkish (regarding adults) and Polish (regarding students) research shows that visualisation is helpful in the process leading to generalisation. The analysis of the paths taken by the students leading to generalisation seems to suggest **how to organize the didactic process aimed at developing this activity, especially if the generalisation takes on a symbolic form with use of a variable**. Taking the presented research into account, it is desirable to suggest the way described above: from a geometric view of a problem to an arithmetical one, although a smooth

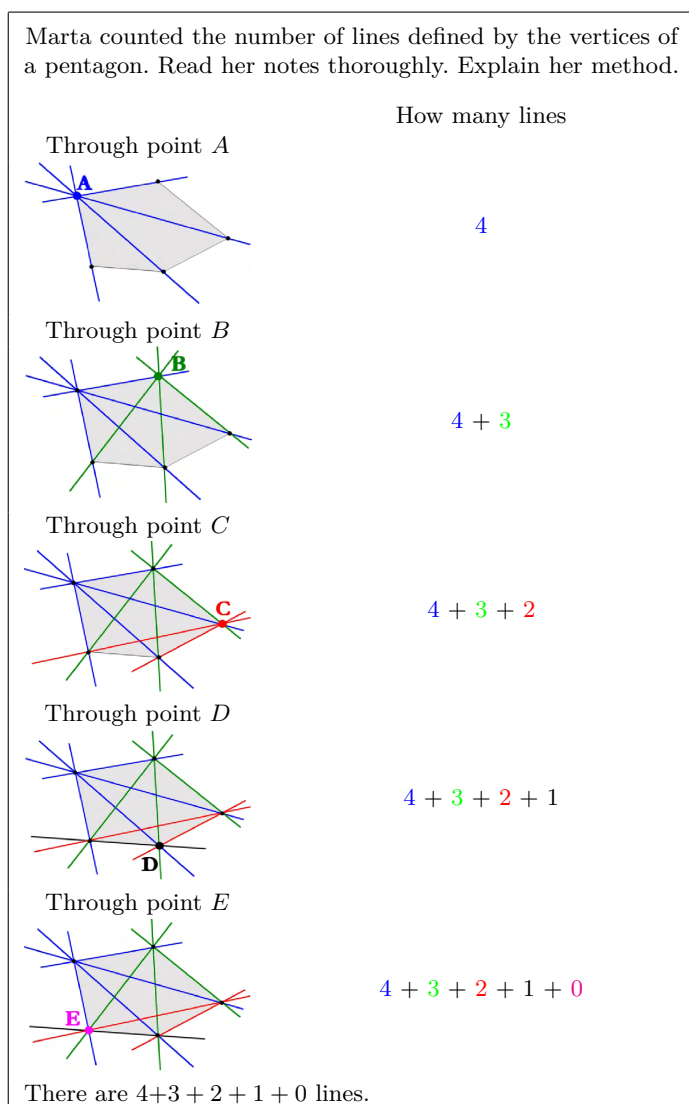


Figure 8. Visual hint

transition between them is key. Therefore, it would also be desirable for the first tasks of a tool designed to develop this type of activity to have visual representations, and for models of a geometric character to be suggested by the teacher during a student's initial steps. Suggestions which direct the student towards using a drawing to "see" the relationships also seem important. For instance, in the Polish research described above (regarding the number of lines defined by all pairs of vertices of a given polygon), the following suggestions and questions proved to be valuable:

- Let's find a method that would allow us to state that all the straight lines are already drawn.
- So, how can this number be obtained? (when the student obtains a solution by reading it from the drawing)

- I don't want to know the exact number now, but the way to obtain it.
- Geometrical directions facilitating perception of arithmetic relationships. Exemplary suggestions are presented in Figures 5 and 8 (see: Zaręba, 2012, p. 88; own translation).

The suggestions presented above are just some of the ways the students can be directed towards generalisation. I think the potential is in the students themselves, and the role of a good teacher is to listen to the student and to arrange their thinking in such a way so that they can make use of their potential.

References

- Arcavi, A.: 2003, The role of visual representations in the learning of mathematics, *Educational Studies in Mathematics* **52**, 215–241.
- Baranowska, M.: 1992, Czy ogrodnik jest pracowity?, *Nauczyciele i Matematyka. Kwartalnik Stowarzyszenia Nauczycieli Matematyki* **1**.
- Castelnuovo, E.: 1987, Umiejętność widzenia w matematyce. Kilka uwag dydaktycznych o intuicji i rozumowaniu dedukcyjnym, *Dydaktyka Matematyki* **7**, 17–25.
- Ciosek, M.: 1995, O roli przykładów w badaniu matematycznym, *Dydaktyka Matematyki* **17**, 5–85.
- Ciosek, M.: 2012, Generalisation in the process of defining a concept and exploring it by students, w: B. Maj-Tatsis, K. Tatsis (red.), *Generalisation in mathematics at all educational levels*, Wydawnictwo Uniwersytetu Rzeszowskiego, Rzeszów (Poland), 38–56.
- Filip, J., Rams, T.: 2000, *Dziecko w świecie matematyki*, Oficyna Wydawnicza „Impuls”, Kraków.
- Flores, A.: 2012, Geometric Representations in the Transition from Arithmetic to Algebra, w: F. Hitt (red.), *Representation and Mathematics Visualisation*, Working Group Representations and Mathematics Visualisation (1998–2002), Cinvestav – IPN, 35–55.
- Gucewicz-Sawicka, I.: 1982, Proces uogólniania w nauczaniu matematyki, w: Z. Krygowska (red.), *Podstawowe zagadnienia dydaktyki matematyki*, PWN, Warszawa, 107–118.
- Hershkowitz, R.: 1989, Visualisation in geometry; Two side of the coin, *Focus on Learning Problems in Mathematics* **11**(1), 61–76.
- Hitt, F.: 2012, *Representation and Mathematics Visualisation*, Working Group Representations and Mathematics Visualisation (1998–2002), Cinvestav – IPN. http://www.er.uqam.ca/nobel/r21245/varia/Book_RMV_PMENA.pdf.
- Hitt, F., Santos, M.: 1999, *Proceedings of the Twenty First Annual Meeting. North American Chapter of the International Group for the Psychology of Mathematics Education*, Vol. **1**, ERIC/CSMEE, Columbus.
- Krygowska, Z.: 1977, *Zarys dydaktyki matematyki*, part 3, WSiP, Warszawa.
- Krygowska, Z.: 1981, Główne problemy i kierunki badań współczesnej dydaktyki matematyki, *Dydaktyka Matematyki* **1**, 7–60.
- Krygowska, Z.: 1986, Elementy aktywności matematycznej, które powinny odgrywać znaczącą rolę w matematyce dla wszystkich, *Dydaktyka Matematyki* **6**, 25–41.

- Kuřina, F.: 1998, Jak myřl uczynić widzianą, *Dydaktyka Matematyki* **20**, 73-88.
- Legutko, M.: 2010, Umiejętność matematycznego uogólniania wśród nauczycieli i studentów matematyki specjalności nauczycielskiej (na przykładzie serii zadań „schodki”), *Annales Universitatis Paedagogicae Cracoviensis, Studia ad Didacticam Mathematicae Pertinentia* **III**, 79-115.
- Legutko, M., Stańdo, J.: 2008, Jakie działania powinny podjąć polskie szkoły w świetle badań PISA, *Prace monograficzne z dydaktyki matematyki. Współczesne problemy nauczania matematyki* **1**, Bielsko-Biała, 19-34.
- Maj-Tatsis, B., Tatsis, K.: 2012, *Generalisation in mathematics at all educational levels*, Wydawnictwo Uniwersytetu Rzeszowskiego, Rzeszów (Poland).
- Malara, N. A.: 2012, Generalisation processes in the teaching/learning of algebra: students behaviours and teacher role, w: B. Maj-Tatsis, K. Tatsis (red.), *Generalisation in mathematics at all educational levels*, Wydawnictwo Uniwersytetu Rzeszowskiego, Rzeszów (Poland), 57-90.
- Novotná, J., Moraová, H., Krátká, M., Stehlíková, N.: 2006, *Proceedings of the 30th Conference of the International Group for the Psychology of Mathematics Education*, Vol. 4, Prague. <http://ase.tufts.edu/education/earlyalgebra/publications/2006/PME30.pdf>.
- Nowak, W.: 1989, *Konwersatorium z dydaktyki matematyki*, PWN, Warszawa.
- Presmeg, N. C.: 1999, On Visualisation and Generalisation in Mathematics, w: F. Hitt, M. Santos (red.), *Proceedings of the Twenty First Annual Meeting. North American Chapter of the International Group for the Psychology of Mathematics Education*, Vol. 1, ERIC/CSMEE, Columbus, 151-155.
- Rösken, B., Rolka, K.: 2006, A picture is worth a 1000 words – The role of visualisation in mathematics learning, w: J. Novotná, H. Moraová, M. Krátká, N. Stehlíková (red.), *Proceedings of the 30th Conference of the International Group for the Psychology of Mathematics Education*, Vol. 4, Prague, 457-464.
- Siwek, H.: 2005, *Dydaktyka matematyki. Teoria i zastosowania w matematyce szkolnej*, WSiP, Warszawa.
- Sochański, M.: 2011, *Wizualizacje w matematyce wobec tradycji epistemologicznej*. Rozprawa doktorska (praca niepublikowana) obroniona w 2011 roku w Uniwersytecie im. Adama Mickiewicza w Poznaniu.
- Tlustý, P.: 2002, Obrázek jako názorný prostředek vizualizace matematické myřlenky, *Disputationes Scientificalae Universitatis Catholicalae in Ruřomberok*, Katolická Univerzita, Ruřomberok, roč. 2, č. 1, 84-87.
- Vinner, S.: 2012, Generalisations in everyday thought processes and in mathematical contexts, w: B. Maj-Tatsis, K. Tatsis (red.), *Generalisation in mathematics at all educational levels*, Wydawnictwo Uniwersytetu Rzeszowskiego, Rzeszów (Poland), 22-37.
- Yilmaz, R., Argün, Z., Keskin, M. Ö.: 2009, What Is the Role of Visualisation in Generalisation Processes: The Case of Preservice Secondary Mathematics Teachers, *Humanity & Social Sciences Journal* **4**(2), 130-137.
- Zaksis, R., Liljedahl, P.: 2002, Generalisation of patterns: the tension between algebraic thinking and algebraic notation, *Educational Studies in Mathematics* **49**, Kluwer Academic Publishers. Printed in the Netherlands, 379-402.

Zaręba, L.: 2012, *Matematyczne uogólniania. Możliwości uczniów i praktyka nauczania*, Wydawnictwo Naukowe Uniwersytetu Pedagogicznego, Kraków.

Zimmermann, W., Cunningham, S.: 1991, *Visualisation in Teaching and Learning Mathematics*. Washington, Mathematical Association of America Washington, DC, USA.

*Instytut Matematyki
Uniwersytet Pedagogiczny
ul. Podchorążych 2
PL-30-084 Kraków
e-mail lzareba@up.krakow.pl*

Annales Universitatis Paedagogicae Cracoviensis

Studia ad Didacticam Mathematicae Pertinentia VII (2015)

ISSN 2080-9751

Historia matematyki, Konferencje

Piotr Błaszczyk

Nota o *Lehrbuch der Algebra. Einleitung* Heinricha Webera*

W matematyce greckiej rozwinięto dwie teorie proporcji: wielkości (czyli odcinków, figur, brył, kątów) oraz liczb naturalnych. Euklides skodyfikował proporcję wielkości – w Księdze V *Elementów*, a proporcję liczb w Księdze VII. W Księdze X natomiast rozważał proporcje, w których występowały i odcinki, i liczby – nie wyjaśnił jednak, która definicja pozwala na takie rozwiązanie. Kwestia ta jest o tyle ważna, że nie wszystkie twierdzenia Księgi V stosują się do proporcji liczb i faktycznie w przeszłości kilka razy próbowano rozwiązać ten problem, by wymienić arabskie edycje *Elementów* czy słynną interpretację Księgi X przedstawioną przez Simona Stevina w książce *Arithmétique* (1585).

Do pierwszych dekad XVII wieku antyczne teorie proporcji pełniły w matematyce taką rolę, jaką liczby rzeczywiste i ułamki pełnią współcześnie. Etapem pośrednim między teorią wielkości a współczesną arytmetyką liczb rzeczywistych była arytmetyka odcinków stworzona przez Kartezjusza w *La Géométrie* (1637). Kartezjusz wprowadził mianowicie mnożenie, dzielenie i pierwiastkowanie odcinków, a odpowiednie konstrukcje oparł na proporcji wielkości oraz założeniu, że istnieje odcinek jednostkowy, 1. Posługiwał się przy tym rachunkami, w których łączył odcinki, liczby wymierne oraz pierwiastki z liczb. Można pokazać, że odcinki, $\mathbb{Q}$, wraz z kartezjańskimi działaniami tworzą ciało uporządkowane $(\mathbb{Q}, +, \cdot, 1, <)$ – dokładniej: część dodatnią ciała uporządkowanego – a ponadto, że w strukturze tej można odtworzyć arytmetykę ułamków¹.

Pierwsze nowożytnie definicje liczby były oparte na pojęciu stosunku — tak postępowali między innymi Newton i Euler². Nie były one jasne, gdyż greckie pojęcie stosunku nie było jasne. Okres niepewności co do pojęcia liczby trwał jeszcze do wczesnych lat 70. XIX wieku, kiedy to w pracach Charlesa Mèraya (1869), Eduarda Hainego (1872), Georga Cantora (1872) i Richarda Dedekinda (1872) przedstawiono pierwsze konstrukcje liczb rzeczywistych. Pojęcie liczby zyskało ostateczny kształt w roku 1900, kiedy David Hilbert w artykule *Über den Zahlbegriff*

*Note on Heinrich Weber's *Lehrbuch der Algebra. Einleitung*

¹ Zob. P. Błaszczyk, K. Mrówka, *Kartezjusz, Geometria. Tłumaczenie i komentarz*, Universitas, Kraków 2015.

² Zob. P. Błaszczyk, K. Mrówka, *Euklides, Elementy, Księgi V–VI. Tłumaczenie i komentarz*, CCPres, Kraków 2013, s. 102–104.

przedstawił pierwszą aksjomatykę liczb rzeczywistych³.

Heinrich Weber we *Wstępie* do książki *Lehrbuch der Algebra* (1895) połączył antyczne teorie proporcji i na tej podstawie skonstruował ciało liczb rzeczywistych⁴. Za pomocą stosownych definicji rozwiązał problem z Księgi X *Elementów*, podał jasną definicję stosunku, a relację proporcji zastąpił równością stosunków, stwarzając w ten sposób kolejną, obok konstrukcji Mèraya, Heinego, Cantora i Dedekinda, XIX-wieczną konstrukcję liczb rzeczywistych.

Z uwagi na odniesienia historyczne obecne w rozprawie Webera przypomnimy teraz grecką definicję stosunku (*Elementy* V.3) i proporcji (*Elementy*, V.5):

„Stosunek jest pewną relacją w odniesieniu do miary dwóch wielkości tego samego rodzaju”.

„Mówi się, że w tym samym stosunku są wielkości pierwsza $[A]$ do drugiej $[B]$ i trzecia $[C]$ do czwartej $[D]$, gdy te same wielokrotności pierwszej $[nA]$ i trzeciej $[nC]$ jednocześnie przekraczają, są jednocześnie równe lub jednocześnie mniejsze od tych samych wielokrotności drugiej $[mB]$ i czwartej $[mD]$, wziętych w odpowiedniej kolejności, zgodnie z dowolnym mnożeniem każda z dwóch każdej z dwóch”⁵.

Definicję proporcji można zapisać symbolami w następujący sposób:

$$\begin{aligned} A : B :: C : D \Leftrightarrow_{df} (\forall m, n)((nA >_1 mB \Rightarrow nC >_2 mD) \\ \wedge (nA = mB \Rightarrow nC = mD) \\ \wedge (nA <_1 mB \Rightarrow nC <_2 mD)), \end{aligned}$$

przy czym wielkości A, B należą do struktury $\mathfrak{M}_1 = (M_1, +, <_1)$, co znaczy, „są tego samego rodzaju”, są na przykład odcinkami, zaś C, D mogą należeć do innej struktury, $\mathfrak{M}_2 = (M_2, +, <_2)$, będąc na przykład trójkątami.

Słowo „stosunek” występuje często w *Elementach*, ale w rekonstrukcji teorii proporcji, gdy odtwarzany jest sens matematyczny kolejnych twierdzeń, można je pominąć. Proporcję zaś można potraktować jak relację czteroargumentową między wielkościami A, B, C, D lub też jak relację między parami wielkości (A, B) , (C, D) . Na podstawie definicji V.5 można więc wyznaczać proporcje wielkości bez ustalania, czym jest stosunek.

Niżej pokażemy, że w ujęciu Webera stosunek jest po prostu parą uporządkowaną wielkości, równość stosunków jest modyfikacją definicji V.5, a liczba jest klasą równych stosunków.

1. Liczby naturalne

Konstruując liczby rzeczywiste Mèray, Heine, Cantor i Dedekind przyjmowali arytmetykę liczb wymiernych. Współczesne konstrukcje zakładają arytmetykę liczb

³ Zob. P. Błaszczyk, *Nota o rozprawie Eduarda Heinego Elemente der Functionenlehre, Annales Universitatis Paedagogicae Cracoviensis. Studia ad Didacticam Mathematicae Pertinentia* VI (2014), 139–152 oraz P. Błaszczyk, *Nota o Über den Zahlbegriff Davida Hilberta, Annales Universitatis Paedagogicae Cracoviensis Studia ad Didacticam Mathematicae Pertinentia* IV (2012), 195–198.

⁴ Zob. H. Weber, *Lehrbuch der Algebra*, Friedrich Vieweg und Sohn, Braunschweig 1895; wyd. drugie: 1898.

⁵ P. Błaszczyk, K. Mrówka, *Euklides. Elementy. Księgi V–VI. Tłumaczenie i komentarz*, op. cit. s. 18.

naturalnych i *via* odpowiednie relacje równoważności tworzą kolejne rozszerzenia zbioru $\mathbb{N}$, czyli $\mathbb{Z}$, $\mathbb{Q}$ oraz $\mathbb{R}$. Porządek historyczny konstruowania kolejnych zbiorów był natomiast następujący: w roku 1872 znano już konstrukcje liczb rzeczywistych, ale pierwszą konstrukcję liczb wymiernych podano dopiero w 1895, w referowanej właśnie pracy Webera⁶.

Weber otwiera swój wykład przypomnieniem podstawowych praw arytmetyki liczb naturalnych. Zwracając uwagę na struktury występujące w referowanej pracy zauważmy, że liczby naturalne stanowią tu system

$$(\mathbb{N}, +, \cdot, <),$$

w którym działania mają znane własności (przemienności, łączności i rozdzielności), a porządek liniowy $<$ jest zgodny z dodawaniem i mnożeniem.

Porządek liniowy Weber definiuje jako relację spełniającą prawo trychotomii oraz przechodnią.

„Zbiór nazywa się *uporządkowanym*, gdy spośród jakichkolwiek dwóch różnych jego elementów zawsze jeden z nich uchodzi za większy, i to tak, iż z $a > b$, $b > c$ zawsze wynika $a > c$ ”.

Z uwagi na algorytm Euklidesa, trzeba ponadto przyjąć, że porządek liczb naturalnych charakteryzuje zasada indukcji w następującej wersji: nie istnieje nieskończony malejący ciąg liczb naturalnych;

„[...] ponieważ liczby $a, a_1, a_2, a_3, \dots$ stale się zmniejszają, z konieczności po skończonej liczbie kroków otrzymamy dzielenie bez reszty”.

2. Liczby wymierne

„Zbiór uporządkowany o tej własności, iż między każdymi dwoma elementami zawsze znajdują się jeszcze inne elementy nazywa się *gęstym*. Zbiór gęsty można utworzyć, gdy zbierze się liczby naturalne w pary i rozważy te pary jako elementy zbioru. Te pary nazywa się ułamkami i oznacza przez $m : n$ lub $\frac{m}{n}$, a dwa takie ułamki $m : n$ oraz $m' : n'$ uważa się za sobie równe, gdy $mn' = nm'$. Jeśli zbierze się wszystkie równe między sobą ułamki w jeden element, to otrzyma się rozmaitość, która jest uporządkowana, gdy ustali się, iż $m : n$ jest większy od $m' : n'$, gdy $mn' > nm'$ ”.

W powyższym fragmencie znajdujemy następujące definicje:

$$m : n \equiv p : q \Leftrightarrow_{df} m \cdot q = n \cdot p, \quad (1)$$

$$m : n \succ p : q \Leftrightarrow_{df} m \cdot q > n \cdot p. \quad (2)$$

Dalej Weber dowodzi, że porządek ułamków jest gęsty.

2.1. Definicję (1) można przedstawić jako interpretację twierdzeń VII.18, 19 z *Elementów*. Stanowią one, że cztery liczby są w proporcji, $m : n :: p : q$, gdy

⁶ Pierwszą konstrukcję liczb naturalnych przedstawił Johann von Neumann w artykule *Zur Einführung der transfiniten Zahlen* (1923).

równe są odpowiednie iloczyny, $m \cdot q = n \cdot p$. Weber porzucił niejasną definicję stosunku liczb i zamiast dowodzić twierdzeń VII.18 (jeżeli $m : n :: p : q$, to $m \cdot q = n \cdot p$) oraz VII.19 (jeżeli $m \cdot q = n \cdot p$, to $m : n :: p : q$) połączył ich tezy w definicję.

Definicja (2) z kolei jest interpretacją definicji V.7 *Elementów* określającej porządek stosunków wielkości. Brzmi ona:

„Przy tych samych zaś wielokrotnościach, gdy wielokrotność pierwszej $[lA]$ przekracza wielokrotność drugiej $[kB]$, a wielokrotność trzeciej $[lC]$ nie przekracza wielokrotności czwartej $[kD]$, wtedy mówi się, że pierwsza jest w większym stosunku do drugiej niż trzecia do czwartej”.

Definicję Euklidesa można tak zapisać:

$$A : B \succ C : D \Leftrightarrow_{df} (\exists k, l)(lA >_1 kB \wedge lC \leq_2 kD),$$

gdzie A, B, C, D oznaczają wielkości, zaś $lA = \underbrace{A + \dots + A}_{l\text{-razy}}$ itd. to odpowiednie wielokrotności.

Stosując tę definicję do liczb m, n, p, q , otrzymujemy

$$m : n \succ p : q \Leftrightarrow (\exists k, l)(lm > kn \wedge kq \geq lp).$$

Przyjmując dalej $k = p, l = q$ dostajemy

$$m : n \succ p : q \Leftrightarrow (qm > pn \wedge pq \geq qp),$$

lub krótko

$$m : n \succ p : q \Leftrightarrow qm > pn.$$

Zbiór ułamków dodatnich będziemy oznaczać dalej tak, jak Weber, literą $\mathfrak{R}$. Definiując dodawaniem przekrojów (zobacz niżej §3) wykorzystuje Weber dodawanie ułamków, liczby wymierne tworzą więc strukturę

$$(\mathfrak{R}, +, <).$$

3. Ciągłość

W rozważaniach o ciągłości Weber powtarza tezy rozprawy Richarda Dedekinda *Stetigkeit und irrationale Zahlen*⁷. Ciągłość jest zatem cechą porządku liniowego wyrażoną za pomocą przekrojów.

„Podział zbioru uporządkowanego $\mathfrak{M}$ na dwie części A, B tego rodzaju, że każdy element a z A jest mniejszy od każdego elementu b z B będzie nazywany *przekrojem* w $\mathfrak{M}$ i oznaczany, odpowiednio, przez (A, B) ”.

Definicja ciągłości jest zaś następująca:

„Gdy każdy przekrój w zbiorze gęstym $\mathfrak{M}$ jest wytworzony przez określony element μ , to zbiór ten nazywa się ciągłym”.

⁷ R. Dedekind, *Stetigkeit und irrationale Zahlen*, Friedrich Vieweg & Sohn, Braunschweig 1872.

Porządek zbioru $\mathfrak{M}$ jest więc ciągły, gdy dla każdego przekroju (A, B) zbioru $(\mathfrak{M}, <)$ zachodzi

$$(\exists \mu \in \mathfrak{M})(\forall x \in A)(\forall y \in B)(x \leq \mu \leq y).$$

Następnie Weber definiuje zbiór wszystkich przekrojów liczb wymiernych:

„Rozmaitość $\mathfrak{R}$ może jednak służyć jako punkt wyjścia do konstrukcji zbioru ciągłego. Ogół wszystkich przekrojów w $\mathfrak{R}$ jest z pewnością rozmaitością, którą można oznaczyć przez $\mathfrak{S}$. Jeśli rozważymy dwa jej różne elementy $\alpha = (A, B)$ oraz $\alpha' = (A', B')$, to albo A będzie częścią A' , albo A' częścią A . Jeśli bowiem jakkolwiek element a należy do A , to także każdy mniejszy od niego element z $\mathfrak{R}$ należy do A . Jeśli A jest częścią A' , to chcemy nazywać α mniejszym od α' , przez co zbiór $\mathfrak{S}$ staje się *uporządkowany*”.

Zatem

$$\mathfrak{S} = \{(A, B) : (A, B) \text{ jest przekrojem zbioru } (\mathfrak{R}, <)\}.$$

Niech $\alpha, \alpha' \in \mathfrak{S}$, gdzie $\alpha = (A, B), \alpha' = (A', B')$. Porządek w $\mathfrak{S}$ jest więc zadany definicją

$$\alpha < \alpha' \Leftrightarrow_{df} A' \subset A.$$

„Jeśli uznamy przekroje wytworzone przez ułamki wymierne za równe samym tym ułamkom wymiernym i nazwiemy je krótko przekrojami wymiernymi, to zbiór $\mathfrak{S}$ zawiera zbiór $\mathfrak{R}$ i zbiór $\mathfrak{S}$ jest w każdym bądź razie zbiorem gęstym. Zbiór $\mathfrak{S}$ jest jednak także *ciągły*”.

Przyjmując utożsamienie liczby wymiernej μ z przekrojem $((-\infty, \mu], (\mu, \infty))$ dostajemy, że $\mathfrak{R} \subset \mathfrak{S}$.⁸ Weber pokazuje, że zbiór $\mathfrak{R}$ jest gęsty w $\mathfrak{S}$ i dowodzi, że porządek przekrojów jest ciągły. Dalej, definiuje dodawania przekrojów:

„Rozumiemy teraz przez sumę $(A, B) + (A', B') = (A'', B'')$ lub $\alpha + \alpha' = \alpha''$ ten przekrój w $\mathfrak{R}$, który otrzymuje się, gdy do A'' skieruje się ułamek wymierny α'' tylko wtedy, gdy istnieje α w A oraz α' w A' o tej własności, że $\alpha'' \leq \alpha + \alpha'$. [...] Przekroje wytworzone przez ułamki wymierne μ, μ' dają w wyniku dodawania przekrój wytworzony przez $\mu + \mu'$ ”.

Przekroje dodatniej części osi liczb wymiernych tworzą więc strukturę

$$(\mathfrak{S}, +, <),$$

a przy tym porządek przekrojów jest ciągły.

4. Zbiór mierzalny

Pojęcie *zbiór mierzalny* odnosi Weber do struktury algebraiczno-porządkowej $\mathfrak{M} = (M, +, <)$ scharakteryzowanej następującymi aksjomatami: dodawanie jest działaniem wewnętrznym, łącznym i przemienne, porządek jest liniowy, a ponadto spełnione są następujące zależności:

⁸ W istocie liczbie μ odpowiadają dwa przekroje: wyżej wskazany oraz $((-\infty, \mu), [\mu, \infty))$.

$$W1 \ (\exists n \in \mathbb{N})(na > b),$$

$$W2 \ (\forall a, b)(\exists c)(a > b \Rightarrow b + c = a),$$

$$W3 \ (\forall a, b)(a + b > a), \quad \text{gdzie term } na \text{ oznacza } \underbrace{a + \dots + a}_{n\text{-razy}}.$$

Można pokazać, że aksjomaty te są równoważne aksjomatom

$$E1 \ (\forall x, y)(\exists n \in \mathbb{N})(nx > y),$$

$$E2 \ (\forall x, y)(\exists z)(x < y \Rightarrow x + z = y),$$

$$E3 \ (\forall x, y, z)(x < y \Rightarrow x + z < y + z).$$

W innym miejscu pokazaliśmy, że aksjomaty E1–E3 charakteryzują Euklidesa pojęcie wielkości⁹.

Na podstawie aksjomatów W1–W3 Weber dowodzi kilku istotnych twierdzeń o proporcjach *wielkościach mierzalnych*. W następnym paragrafie zaprezentuję Webera definicje stosunku i proporcji — to najciekawszy fragment referowanej rozprawy. Wcześniej jednak przedstawimy uwagę dotyczącą metodologii.

4.1. Konstruując liczby rzeczywiste Weber rozważa dwa zbiory mierzalne: przekroje $(\mathfrak{S}, +, <)$ oraz stosunki elementów z $\mathfrak{S}$, który oznaczamy symbolem $(\Lambda, \oplus, <)$. Obydwa zbiory spełniają warunki W1–W3, dlatego definicje stawiane w dowolnym zbiorze mierzalnym mogą być powtórzone i w $(\mathfrak{S}, +, <)$, i w $(\Lambda, \oplus, <)$. Podobnie własności W4–W6 (patrz niżej §5) wyprowadzane z aksjomatów W1–W3 przysługują i $(\mathfrak{S}, +, <)$, i $(\Lambda, \oplus, <)$. Ponadto, porządek obydwu struktur jest ciągły, dlatego własność W7 (patrz niżej §5) przysługuje i $(\mathfrak{S}, +, <)$, i $(\Lambda, \oplus, <)$.

Aksjomatyczna charakterystyka zbioru mierzalnego w istotny sposób skraca więc cały wywód Webera. Wskazany zabieg metodologiczny nie został jednak jasno przedstawiony w rozprawie. Przypomnijmy zatem, że sama idea metody aksjomatycznej została wprowadzona dopiero przez Hilberta w *Über den Zahlbegriff* i wcześniej w pierwszym wydaniu *Grundlagen der Geometrie*¹⁰.

5. Liczba

5. Prezentowane niżej definicje stosunku i proporcji są stawiane w dowolnym zbiorze mierzalnym $(\mathfrak{M}, +, <)$.

„Przechodzimy teraz do definicji *stosunku* (*Verhältnisse*), który od Starożytności były uważany za podstawę nauki o liczbach, a który najpierw występuje u Euklidesa”.

W istocie, za pomocą pojęcia stosunku, Weber podaje nową definicję proporcji.

(a) Stosunki wymierne.

„Założmy najpierw, że istnieją dwie liczby całkowite m, n takie, że $na = mb$, jak to np. zawsze ma miejsce, gdy $\mathfrak{M}$ jest systemem liczb naturalnych, lub gdy

⁹ Zob. P. Błaszczyk, K. Mrówka, *Euklides, Elementy ...*, op. cit, rozdz. 3.

¹⁰ Zob. D. Hilbert, *Über den Zahlbegriff*, op. cit. oraz D. Hilbert, *Grundlagen der Geometrie*, w: *Festschrift zur Enthüllung des GAUSS-WEBER Denkmals in Göttingen*, Leipzig 1899.

a, b są dwoma współmiernymi odcinkami; gdy wtedy p, q są innymi liczbami całkowitymi, to $qa = pb$ wtedy i tylko wtedy, gdy $mq = np$. Para liczb p, q jest całkowicie określona poprzez to żądanie, o ile dojdzie jeszcze warunek, iż p, q mają być liczbami względnie pierwszymi. Może być wtedy $m = hp, n = hq$, gdzie h jest dowolną liczbą całkowitą. W tym przypadku nazywamy proporcję $a : b$ wymierną i uważamy ją za równą ułamkowi wymiernemu $m : n$ lub $p : q$. Te ułamki wymierne mogą odtąd być uważane za proporcje liczb całkowitych”.

Stosunek wymierny $a : b$ jest więc tak definiowany:

$$a : b = m : n \Leftrightarrow_{df} na = mb. \quad (3)$$

Przy okazji tej i następnych definicji, Weber powołuje się na Księgę V *Elementów*. W istocie definicja (3) to teza twierdzenia X.5.

(b) Porównanie stosunku z ułamkiem.

„Wracamy teraz do jakiejkolwiek rozmaitości mierzalnej $\mathfrak{M}$ i bierzemy z niej jakiejkolwiek elementy a oraz b . Jeśli wybierze się, co zawsze jest możliwe, dwie liczby naturalne m, n tak, że $na > mb$, to proporcja $a : b$ jest nazywana większą od proporcji wymiernej $m : n$, lub $\frac{a}{b} > \frac{m}{n}$ [...] Podobnie wynika, że jeśli $n'a < m'b$, to $\frac{a}{b} < \frac{m'}{n'}$ ”.

Mamy zatem:

$$a : b \succ m : n \Leftrightarrow_{df} na > mb, \quad a : b \prec m : n \Leftrightarrow_{df} na < mb. \quad (4)$$

(c) Porządek i równość stosunków.

„Jeśli teraz $a : b$ oraz $\alpha : \beta$ są dwiema jakimikolwiek proporcjami, które dla zwięzłości chcemy oznaczać przez e, ε , i których elementy należą do tych samych lub różnych rozmaitości, to możliwe są dwa przypadki: 1) Między e oraz ε nie leży żadna proporcja wymierna μ lub 2) istnieje proporcja wymierna między e oraz ε . W przypadku 1) obie proporcje e oraz ε nazywają się *wzajem równymi* [...] W przypadku 2) proporcje e, ε nazywają się *nierównymi*. Może zatem być albo $\alpha) e < \mu < \varepsilon$, albo $\beta) e > \mu' > \varepsilon$ i oba te przypadki nawzajem się wykluczają, ponieważ z $e < \mu < \varepsilon, \varepsilon < \mu'$ wynika, że $\mu < \mu'$, a w konsekwencji dostaje się $e < \mu'$ ”.

Mamy zatem

$$a : b = \alpha : \beta \Leftrightarrow_{df} \neg(\exists \mu)(a : b \succ \mu \succ \alpha : \beta) \wedge \neg(\exists \nu)(a : b \prec \nu \prec \alpha : \beta),$$

gdzie μ, ν są liczbami wymiernymi.

Definicję nierówności stosunków tak zapiszemy

$$a : b \succ \alpha : \beta \Leftrightarrow_{df} (\exists m, n \in \mathbb{N})(na > mb \wedge m\beta > n\alpha). \quad (5)$$

Podobnie,

$$a : b \prec \alpha : \beta \Leftrightarrow_{df} (\exists m, n \in \mathbb{N})(na < mb \wedge m\beta < n\alpha). \quad (6)$$

Teraz można pokazać, że

$$a : b = \alpha : \beta \Leftrightarrow \neg(a : b \succ \alpha : \beta) \wedge \neg(a : b \prec \alpha : \beta). \quad (7)$$

Następnie, na mocy definicji jest:

$$(a : b \prec \alpha : \beta) \vee (a : b = \alpha : \beta) \vee (a : b \succ \alpha : \beta), \quad (8)$$

a ponadto zachodzi dokładnie jeden z tych warunków.

Przechodniość relacji $\prec$ wynika wprost z zależności

$$(m : n \prec a : b \prec p : q) \Rightarrow mq < np. \quad (9)$$

Pojęcie liczby:

„Gdy teraz zbierzemy razem wszystkie równe między sobą stosunki, to otrzymamy pojęcie gatunkowe, które oznaczamy jako *liczbę* w sensie ogólnym. Liczba jest zatem nazwą lub znakiem dla pewnej rozmaitości, której elementami są właśnie te proporcje, której równe są jednej spośród nich. W tym pojęciu liczby są zawarte proporcje wymierne, a w konsekwencji również liczby naturalne, jako proporcje $m : 1$; owe proporcje wymierne tworzą *liczby wymierne*”.

Stosując współczesny język powiedzielibyśmy, że liczba jest klasą abstrakcji.

Jakkolwiek Weber relatywizuje pojęcie liczby do „pewnej rozmaitości”, to dalej przyjmujemy, że liczby są stosunkami przekrojów. Przyjmujemy zatem, że liczby tworzą zbiór

$$\Lambda = \{a : b \mid a, b \in \mathfrak{G}\}.$$

Na podstawie definicji (5), (6) oraz własności (8), (9) liczby tworzą zbiór uporządkowany liniowo

$$(\Lambda, \prec).$$

Działania na liczbach będą zdefiniowane w oparciu o własności *wielkości mierzalnych*.

6. Dodawanie stosunków

Definicje działań poprzedzają cztery twierdzenia: dla dowolnych $a, b \in \mathfrak{M}$ zachodzą następujące zależności:

$$W4 \quad (\forall a, b, c)(a > b \Rightarrow a : c \succ b : c),$$

$$W5 \quad (\forall a, b, c)(b > c \Rightarrow a : c \succ a : b),$$

$$W6 \quad \forall a, b, c, d)(a : b = c : d \Rightarrow a : c = b : d),$$

$$W7 \quad (\forall a, b, c)(\exists d)(a : b = c : d).$$

Dwa pierwsze odpowiadają twierdzeniu V.8 z *Elementów*, trzecia – twierdzeniu V.16, czwarta natomiast było znana w matematyce greckiej jako założenie o istnieniu czwartej proporcjonalnej, współcześnie zaś jest przyjmowana jako aksjomat teorii wielkości¹¹. Spójrzmy na dowód W7:

„Twierdzenie jest bezpośrednią konsekwencją założonej ciągłości. Albowiem gdy weźmie się np. element x w $\mathfrak{M}$ z A lub B , odpowiednio do tego, czy $x : b$ jest

¹¹ Zob. P. Błaszczyk, K. Mrówka, *Euklides, Elementy. Księgi V-VI. Tłumaczenie i komentarz*, op. cit., s. 115–117.

mniejsze czy większe od $c : d$, to otrzyma się przekrój, który jest wytworzony przez element a taki, że spełniony jest warunek $a : b = c : d$. Twierdzenie to zachodzi jednak również dla pewnych zbiorów nieciągłych, np. dla ułamków wymiernych”.

Tak więc, gdy dane są trzy przekroje b, c, d należące do $\mathfrak{S}$, to istnieje dokładnie jeden przekrój a , taki, że zachodzi $a : b = c : d$. Element a jest wyznaczony przez przekrój (A, B) , gdzie

$$A = \{x \in \mathfrak{S} \mid c : d \succ x : b\}, \quad B = \{x \in \mathfrak{S} \mid x : b \succeq c : d\}.$$

6.1. Przy ustalonym c , na podstawie twierdzenia W7, dla dowolnego stosunku $e : f$ można znaleźć równy mu stosunek $b : c$. Na tej podstawie definiowane jest dodawanie:

$$\frac{a}{c} \oplus \frac{e}{f} =_{df} \frac{a+b}{c}, \quad \text{gdzie} \quad \frac{e}{f} = \frac{b}{c}. \quad (10)$$

„Ta reguła obejmuje, jako szczególny przypadek, dodawanie ułamków wymiernych, a żeby uzasadnić ją w ogólności trzeba potem jeszcze tylko pokazać, że jeśli $a : c = a' : c'$ oraz $b : c = b' : c'$, to musi być również $(a+b) : c = (a'+b')c'$ ”.

Regułę, o której pisze Weber, można tak zapisać

$$(W8) \quad (a : c = a' : c', \quad b : c = b' : c') \Rightarrow (a+b) : c = (a'+b') : c',$$

W *Elementach* występuje ona jako twierdzenie V.24.

Dzięki definicji (10) liczby tworzą zbiór mierzalny

$$(\Lambda, \oplus, \prec).$$

„Tym samym dowiedziono zatem, że także liczby, tak, jak je zdefiniowaliśmy, tworzą zbiór mierzalny. Jeśli a oraz c są wzięte z rozmaitości ciągłej, to stosunki $a : c$ tworzą zbiór ciągły, także przy ustalonym c , a ponieważ istnieją w ogóle zbiory ciągłe, więc również liczby tworzą zbiór ciągły”.

Ostatnia uwaga oznacza, że skoro porządek w strukturze przekrojów, czyli w $(\mathfrak{S}, +, \prec)$, jest ciągły, to porządek w strukturze $(\Lambda, \oplus, \prec)$ także jest ciągły.

Do struktury liczb należą stosunki $na : ma$, te zaś są równe ułamkom $m : n$. Mając to na uwadze mówimy, że do struktury $(\Lambda, \oplus, \prec)$ należą ułamki, w szczególności liczby naturalne, które są utożsamiane ze stosunkami $n : 1$.

7. Mnożenie i dzielenie stosunków

Zacznijmy od cytatu:

„Jeśli teraz $\alpha, \beta, \gamma, \delta$ są liczbami, to z proporcji $\alpha : \beta = \gamma : \delta$ można ustalić dowolną z tych czterech liczb poprzez trzy pozostałe. Jeśli przyjmie się $\delta = 1$, to otrzyma się *mnożenie* $\alpha = \beta\gamma$, a przemienność czynników jest konsekwencją twierdzenia, że $\alpha : \gamma = \beta : \delta$ wynika z $\alpha : \beta = \gamma : \delta$. Jeśli szuka się γ , to otrzymuje się *dzielenie*; a z podanej wyżej definicji dodawania wynika podstawowa formuła $\alpha(\beta + \gamma) = \alpha\beta + \alpha\gamma$. Cztery sposoby rachowania są zatem wykonalne w dziedzinie liczb, z jedynym ograniczeniem, iż przy odejmowaniu odjemnik musi być mniejszy od odjemnej”.

Mamy zatem

$$\alpha = \beta \otimes \gamma \Leftrightarrow_{df} \alpha : \beta = \gamma : 1. \quad (11)$$

Przemienność mnożenia liczb wynika z twierdzenia W6, zachodzi bowiem

$$\alpha : \beta = \gamma : 1 \Leftrightarrow \alpha : \gamma = \beta : 1,$$

stąd

$$\alpha : \gamma = \beta : 1 \Leftrightarrow \alpha = \gamma \otimes \beta.$$

Można pokazać, że zachodzi także rozdzielność mnożenia względem dodawania. Istotnie, na mocy definicji (11) jest

$$(\alpha \otimes \beta) : \alpha = \beta : 1, \quad (\alpha \otimes \gamma) : \alpha = \gamma : 1,$$

a stąd, na podstawie twierdzenia W8,

$$((\alpha \otimes \beta) \oplus (\alpha \otimes \gamma)) : \alpha = (\beta \oplus \gamma) : 1,$$

co, znów na mocy definicji (11), daje

$$(\alpha \otimes \beta) \oplus (\alpha \otimes \gamma) = \alpha \otimes (\beta \oplus \gamma).$$

Można też pokazać, że iloczyn $\otimes$ jest zgodny z porządkiem, ale tego prawa Weber w ogóle nie wymienia w swojej rozprawie.

7.1. Przyjmijmy, że zachodzi proporcja $\alpha : \beta = \gamma : 1$.

„Jeżeli szukamy liczby γ , dochodzimy do dzielenia”:

$$\alpha \div \beta = \gamma \Leftrightarrow_{df} \alpha : \beta = \gamma : 1.$$

Mamy zatem $\alpha : \beta = (\alpha \div \beta) : 1$, czyli $\alpha = \beta \otimes (\alpha \div \beta)$, w szczególności $1 = \beta \otimes (1 \div \beta)$ ¹².

8. Liczby ujemne

Niech x oznacza każdy element dotąd zdefiniowanego systemu liczbowego, który teraz chcemy określać jako system *liczb dodatnich*. Bierzymy ten system liczbowy po raz drugi i dla odróżnienia, w tym drugim systemie, który określamy ma być jako system *liczb ujemnych*, każdy element oznaczamy przez $-x$. Drugi system porządkujemy teraz dokładnie odwrotnie niż pierwszy, tak, że wszędzie, gdzie w pierwszym systemie x występuje jako „większa”, w tym drugim ustalamy $-x$ jako „mniejszą”, i na odwrót. Dodawanie oraz odejmowanie dla $-x$ są objaśnione podobnie jak dla x , tak, że ma być $(-x) + (-y) = -(x + y)$; $(-x) - (-y) = -(x - y)$.

Pozwalając sobie na uproszczenia, bo w tej części wykład Webera jest zupełnie jasny, przyjmijmy $\mathfrak{L} = -\Lambda \cup \Lambda$, gdzie $-\Lambda = \{-\alpha : \alpha \in \Lambda\}$. Zobaczmy teraz, jak

¹² Czytelnik łatwo znajdzie podobieństwo między Webera definicjami mnożenia i dzielenia stosunków, a definicjami mnożenia i dzielenia odcinków podanymi przez Kartezjusza; zob. P. Błaszczyk, K. Mrówka, *Kartezjusz, Geometria...*, op. cit., s. 157–160.

Weber definiuje liczbę 0. Otóż jeden jedyny przekrój zbioru $(\mathfrak{L}, \prec)$ tworzy lukę, mianowicie $(-\Lambda, \Lambda)$ ¹³.

„Aby zapewnić tu ciągłość, musimy zatem dodać odpowiadającą przekrojowi $(-x, x)$ jeszcze jedną liczbę o znaku 0, która zdefiniowana jest właśnie przez ów przekrój. Mamy wtedy uporządkowany zbiór ciągły, z obu stron nieograniczony, tj. pełny ciąg liczb rzeczywistych”.

W zbiorze $\mathfrak{L}$, za pomocą reguł na znakach, jak np. $x(-y) = (-x)y = -(xy)$, definiuje Weber dodawanie i mnożenie.

„Poglądowo przedstawia się ciąg liczbowy przez punkty, wychodząc od ustalonego i oznaczonego przez 0 punktu prostej i nanosząc liczby dodatnie jako odcinki z jednej, powiedzmy prawej, strony, a liczby ujemne z drugiej (lewej) strony. Obraz sumy dwóch odcinków $z_1 + z_2$ otrzymuje się, gdy poczynawszy od punktu z_1 naniesie się w prawą lub lewą stronę odcinek o długości $\pm z_2$, w zależności od tego, czy z_2 jest dodatnia czy ujemna”.

* * *

Przedkładane tłumaczenie *Lehrbuch der Algebra. Einleitung* zostało przygotowane przez Profesora Jerzego Pogonowskiego na podstawie pierwszego wydania. W roku 1925 Wydawnictwo Kasy im. Mianowskiego wydało *Podręcznik algebry wyższej* w przekładzie Samuela Dicksteina. *Wstęp* z tej edycji różni się znacząco od tłumaczenia zamieszczonego w niniejszym tomie.

¹³ Przyjmując np. przyporządkowanie $\alpha \mapsto (\Lambda, \alpha) =_{df} -\alpha$, można uznać, że $-\Lambda$ jest zbiorem, który powstaje ze zbioru Λ na mocy aksjomatu zastępowania. Porządek ze zbioru $(\Lambda, \prec)$ w „naturalny” sposób przedłużany jest na $\mathfrak{L}$.

Heinrich Weber

Podręcznik algebry. Wprowadzenie*

W naszych rozważaniach przyjmujemy, że *liczby naturalne* $1, 2, 3, \dots$ oraz reguły, wedle których rachuje się na tych liczbach, są znane i wprzód podane. Podstawowymi sposobami rachowania, tak zwanymi *czterema podstawowymi działaniami arytmetycznymi*, są: *dodawanie*, *mnożenie*, do którego przynależy też potęgowanie traktowane jako powtarzanie mnożenia, *odejmowanie* oraz *dzielenie*. Pierwsze dwa nazywają się bezpośrednimi sposobami rachowania; wyróżniają się one tym, że mogą być bez ograniczeń wykonywane w świecie liczb naturalnych. Pierwsza z niebezpośrednich lub odwrotnych operacji, odejmowanie, jest wykonalna jedynie wtedy, gdy odjemna jest większa od odjemnika.

Cel dzielenia można pojmować na dwa sposoby. W pierwszym elementarnym ujęciu pyta się, ile razy dzielnik *zawarty* jest w dzielnej. Liczba nie jest zawarta w liczbie od niej mniejszej. Jeśli jednak dzielna jest równa lub większa od dzielnika, to odpowiedź na pytanie o wynik dzielenia daje *iloraz* oraz w większości przypadków *resztę*, która jest mniejsza od dzielnika. Jeśli nie pozostaje żadna reszta, to mówi się o *dzieleniu bez reszty*, lub że *dzielna jest podzielna przez dzielnik*, lub że *dzielnik jest czynnikiem* liczby, która przedstawia dzielną.

Ten cel, który stawia się już na pierwszych poziomach sztuki rachowania, prowadzi do głębiej leżącego rozróżnienia liczb, będącego podstawą całej teorii liczb.

Ponieważ czynnik nigdy nie może być większy od liczby, której jest czynnikiem, każda liczba ma tylko skończoną liczbę czynników. Każda liczba jest podzielna przez 1 oraz przez siebie samą, a liczba, która nie ma poza tym żadnych dzielników, nazywa się *liczbą pierwszą*. Samej liczby 1 z określonych powodów nie zalicza się do liczb pierwszych. Jeżeli dwie liczby są podzielne przez trzecią, to również suma oraz różnica obu pierwszych jest podzielna przez trzecią; a jeśli jakaś liczba jest podzielna przez drugą, a ta druga przez trzecią, to także pierwsza jest podzielna przez trzecią.

Dwie liczby zawsze mają wspólny dzielnik 1. Gdy nie mają one żadnego innego wspólnego dzielnika, jak np. pary liczb 5 i 7 lub 21 i 38, to nazywają się te dwie liczby *względnie pierwszymi*, lub *liczbami bez wspólnych dzielników*.

Wśród wspólnych dzielników jakichkolwiek dwóch danych liczb jeden jest największy i jest bardzo ważnym zadaniem znalezienie tego największego wspólnego

*Textbook of Algebra. Introduction

dzielnika. Prowadzi do tego postępowanie, znane pod nazwą algorytmu wyszukiwania największego wspólnego dzielnika i znajdujące się już u Euklidesa¹.

Jeśli a , a_1 są dwiema liczbami, to bierze się większą z nich, niech będzie to a , jako dzielną, a mniejszą a_1 jako dzielnik oraz ustala iloraz q_1 ; a gdy dzielenie daje resztę, także resztę a_2 , a zatem $a = q_1 a_1 + a_2$, przy czym $a_2 < a_1$. Każdy wspólny dzielnik a i a_1 jest wtedy wspólnym dzielnikiem a_1 i a_2 oraz na odwrót. Jeśli postąpi się z a_1 i a_2 tak samo jak z a oraz a_1 i otrzyma, gdy jest to dzielenie z resztą, $a_1 = q_2 a_2 + a_3$, to znowu mamy $a_3 < a_2$ i każdy wspólny dzielnik a_1 i a_2 jest wtedy wspólnym dzielnikiem a_2 i a_3 oraz na odwrót. Jeśli prowadzi się dalej dzielenie w ten sposób, to, ponieważ liczby $a, a_1, a_2, a_3, \dots$ stale się zmniejszają, z konieczności po skończonej liczbie kroków otrzymamy dzielenie bez reszty, a więc na koniec wystąpić musi para równań $a_{\nu-2} = q_{\nu-1} a_{\nu-1} + a_\nu$, $a_{\nu-1} = q_\nu a_\nu$. Wtedy a_ν jest wspólnym dzielnikiem wszystkich poprzednich a , a więc także a oraz a_1 i każdy wspólny dzielnik a oraz a_1 jest dzielnikiem a_ν . Tak więc, a_ν jest *największym wspólnym dzielnikiem* a oraz a_1 i otrzymaliśmy tym samym twierdzenie, że każdy wspólny dzielnik dwóch liczb musi dzielić bez reszty ich największy wspólny dzielnik.

Jeśli a oraz a_1 są liczbami względnie pierwszymi, to ostatni dzielnik $a_\nu = 1$. Jeżeli przy tym założeniu pomnożymy otrzymane równania przez jakąkolwiek liczbę b , to wynika z tego, że b jest największym wspólnym dzielnikiem ab oraz $a_1 b$ i że każdy wspólny dzielnik ab oraz a_1 lub a oraz $a_1 b$ musi być dzielnikiem b . Jeśli zatem a_1 jest względnie pierwsza z a oraz z b , to jest też względnie pierwsza z ab , a ze szczególnego założenia, iż a_1 jest liczbą pierwszą otrzymuje się, że iloczyn tylko wtedy może być podzielny przez liczbę pierwszą, gdy co najmniej jeden z czynników jest przez nią podzielny.

Jeśli zatem iloczyn ab jest podzielny przez a_1 oraz a_1 jest względnie pierwsza z a , to b musi być podzielna przez a_1 .

Jeśli a , b są jakimikolwiek liczbami o największym wspólnym dzielniku d oraz $a = da'$, $b = db'$, to a' oraz b' są wzajem względnie pierwsze. Każda liczba m , która jest jednocześnie wielokrotnością a oraz b ma postać $m = am' = da'm'$, a stąd $a'm'$ musi być wielokrotnością b' , a zatem także m' musi być wielokrotnością b' , tj. każda liczba, która jest wielokrotnością jednocześnie a oraz b , jest podzielna przez $a'b'd$. Ta liczba $a'b'd$, która sama jest wspólną wielokrotnością a oraz b nazywa się więc *najmniejszą wspólną wielokrotnością* a oraz b .

Gdy jakaś liczba jest podzielna przez dwie liczby względnie pierwsze, to jest też podzielna przez ich iloczyn, a zatem gdy liczba m jest podzielna przez więcej liczb, z których każde dwie są względnie pierwsze, to jest ona podzielna również przez ich iloczyn.

Na tym opiera się dowód ważnego twierdzenia, że liczba m może zostać przedstawiona zawsze i tylko w jeden sposób jako iloczyn liczb pierwszych. Ponieważ wtedy dzielnik nie może być większy od dzielnej, więc z pewnością m może być podzielna tylko przez skończoną liczbę liczb pierwszych. Jeśli a jest jedną z tych liczb pierwszych, a a^α najwyższą potęgą a , przez którą m dzieli się bez reszty, to α także jest określoną liczbą. Jeśli teraz $b^\beta, c^\gamma, \dots$ będą, podobnie, najwyższymi potęgami pozostałych mieszczących się w m liczb pierwszych $b, c, \dots$, przez które

¹ *Elementy*, księga VII, 2, tom drugi w wydaniu Heiberga.

m daje się dzielić, to wtedy, ponieważ każde dwie z liczb $a^\alpha, b^\beta, c^\gamma, \dots$ są względnie pierwsze, m również musi być podzielna przez iloczyn $a^\alpha b^\beta c^\gamma \dots$, oraz m musi być równa temu iloczynowi, gdyż w przeciwnym razie jakaś inna jeszcze liczba pierwsza lub wyższa potęga liczb pierwszych $a, b, c, \dots$ mieściłaby się w m .

* * *

Podamy teraz przegląd niezbędnych w matematyce i wprowadzanych stopniowo rozszerzeń pojęcia liczby.

Przez *rozmaitość* lub *zbiór*, oznaczany w skrócie symbolem $\mathfrak{M}$, rozumiemy system obiektów lub elementów jakiegokolwiek rodzaju, który jest sam w sobie w ten sposób ograniczony oraz ukończony, że dla dowolnego podanego obiektu jest określone, czy należy on do tego systemu czy nie, wszystko jedno, czy jesteśmy w stanie w każdym poszczególnym przypadku rzeczywiście rozstrzygnąć tę sprawę czy nie.

Zbiór nazywa się *uporządkowanym*, gdy spośród jakichkolwiek dwóch różnych jego elementów zawsze jeden z nich uchodzi za większy, i to tak, iż z $a > b, b > c$ zawsze wynika $a > c$. Jeśli $a > b$ oraz $b > c$, lub, równoważnie, $a > b > c$, to mówimy, że b leży między a oraz c .

Liczby naturalne tworzą zbiór uporządkowany; między dwoma bezpośrednio po sobie następującymi jego elementami nie leży żaden inny element. Taką rozmaitość nazywa się *dyskretną*. Zbiór uporządkowany o tej własności, iż między każdymi dwoma elementami zawsze znajdują się jeszcze inne elementy nazywa się *gęstym*. Zbiór gęsty można utworzyć, gdy zbierze się liczby naturalne w pary i rozważy te pary jako elementy zbioru. Te pary nazywa się ułamkami i oznacza przez $m : n$ lub $\frac{m}{n}$, a dwa takie ułamki $m : n$ oraz $m' : n'$ uważa się za sobie równe, gdy $mn' = nm'$. Jeśli zbierze się wszystkie równe między sobą ułamki w jeden element, to otrzyma się rozmaitość, która jest uporządkowana, gdy ustali się, iż $m : n$ jest większy od $m' : n'$, gdy $mn' > nm'$. To, że ta rozmaitość jest gęsta, okazuje się tak oto: jeśli $\mu = m : n, \mu' = m' : n'$ są dwoma ułamkami oraz $\mu > \mu'$, to można przyjąć, dla dowolnej liczby h

$$\mu = \frac{hmn'}{hnn'}, \quad \mu' = \frac{hm'n}{hnn'},$$

a stąd $hmn' > hm'n$. Zawsze można dobrać h tak, że między hmn' a $hm'n$ leżą jeszcze liczby, a gdy p jest taką liczbą, to $p : hnn'$ leży między μ a μ' .

Punkty linii prostej także można uważać za zbiór uporządkowany, gdy rozumie się przez większy i mniejszy jakiekolwiek oznaczenie miejsca, np. dalej na prawo i dalej na lewo lub wyżej i niżej.

* * *

Podział zbioru uporządkowanego $\mathfrak{M}$ na dwie części A, B tego rodzaju, że każdy element a z A jest mniejszy od każdego elementu b z B będzie nazywany *przekrojem* w $\mathfrak{M}$ i oznaczany, odpowiednio, przez (A, B) . Przekrój taki powstaje, gdy wybierze się jakikolwiek element μ z $\mathfrak{M}$ i zaliczy wszystkie elementy od niego mniejsze do A , wszystkie większe do B , a sam element μ dowolnie, do A lub B . Powstają, dokładnie mówiąc, w zależności od tego, czy uczyni się to pierwsze czy drugie, dwa

przekroje, które chcemy jednak zawsze uważać za równe. Jeśli w przekroju (A, B) albo A zawiera element największy, albo B element najmniejszy, μ , to mówimy, że μ wytwarza przekrój (A, B) . Może jednak wystąpić też przypadek, że ani A nie posiada elementu największego, ani B elementu najmniejszego.

Gdy każdy przekrój w zbiorze gęstym $\mathfrak{M}$ jest wytworzony przez określony element μ , to zbiór ten nazywa się ciągłym.

Ciągłość, tak jak gęstość, są własnościami, które z natury rzeczy są niedostępne naszemu postrzeganiu zmysłowemu; nie można ich zatem ściśle przypisać rzeczom świata zewnętrznego, wielkościom przestrzennym, okresom czasu, masom, jakkolwiek głęboko leżą one w istocie naszego oglądu. Można jednak swobodnie konstruować czyste systemy pojęć, którym przysługuje gęstość bez ciągłości, albo też gęstość oraz ciągłość².

Przykładu zbioru gęstego dostarczają ułamki wymierne. Ta rozmaitość, którą chcemy oznaczać przez $\mathfrak{R}$, nie jest ciągła. Jeśli bowiem weźmiemy jakikolwiek ułamek wymierny $\mu = m : n$, gdzie m oraz n nie mają żadnego wspólnego dzielnika i nie są oba kwadratami liczb, to μ nie jest kwadratem ułamka wymiernego. Gdyby bowiem było $\mu = p^2 : q^2$, to z tego wynikałoby $mq^2 = np^2$, a stąd $m = p^2$, $n = q^2$, co wykluczone jest przez założenie. Gdy więc tworzymy w $\mathfrak{R}$ przekrój (A, B) , w którym do A zaliczamy każdy element a z $\mathfrak{R}$, którego kwadrat jest mniejszy od μ , a do B każdy element b , którego kwadrat jest większy od μ , to ani A nie zawiera elementu największego, ani B elementu najmniejszego, i przekrój (A, B) nie będzie wytworzony przez żaden element z $\mathfrak{R}$.

Jeżeli bowiem założymy, że $a = p : q$ jest jakimkolwiek elementem w A , a zatem $p^2 : q^2 < m : n$, lub $np^2 < mq^2$, to weźmy wtedy dowolną liczbę naturalną y oraz wybierzmy inną liczbę naturalną x tak, że $x > y$ oraz $x(mq^2 - np^2) > ny(2p + 1)$, z czego wynika:

$$x(mq^2 - np^2) > nxy(2p + 1) > n(2pxy + y^2),$$

a więc również

$$mq^2x^2 > n(px + y)^2.$$

Jeśli więc przyjmiemy $a' = (px + y) : qx$, to $a' > a$ oraz $(a')^2 < \mu$, czyli a' także jest zawarty w A ; podobnie można pokazać, że w B nie ma elementu najmniejszego.

Rozmaitość $\mathfrak{R}$ może jednak służyć jako punkt wyjścia do konstrukcji zbioru ciągłego. Ogół wszystkich przekrojów w $\mathfrak{R}$ jest z pewnością rozmaitością, którą można oznaczyć przez $\mathfrak{S}$. Jeśli rozważymy dwa jej różne elementy $\alpha = (A, B)$ oraz $\alpha' = (A', B')$, to albo A będzie częścią A' , albo A' częścią A . Jeśli bowiem jakikolwiek element a należy do A , to także każdy mniejszy od niego element z $\mathfrak{R}$

² Pokazał to DEDEKIND, któremu w ogóle zawdzięczamy podaną wyżej definicję ciągłości. Por. pisma DEDEKINDA, *Stetigkeit und irrationale Zahlen*, Braunschweig 1872, 1892, *Was sind und was sollen die Zahlen?*, Braunschweig 1888, 1893. Inne rozmaitości, którym przysługuje ciągłość, są stworzone przez WEIERSTRASSA oraz CANTORA.

Definicja ciągłości, którą za DEDEKINDEM bierzemy tu za podstawę, jest tak dalece wyczerpująca, że zbiór ciągły w tym sensie, któremu przysługuje jeszcze za chwilę podana własność mierzalności, nie może być częścią bogatszego zbioru ciągłego. Nie wiem, czy własność ta jest gdziekolwiek dowiedziona i mam nadzieję wrócić do tego przy innej sposobności. Zauważę jednak, że taka własność dowodliwa jest jedynie dla zbiorów mierzalnych. Zbiór tylko uporządkowany może być zawsze, jakkolwiek by był gęsty, pojmowany jako część zbioru jeszcze bardziej gęstego.

należy do A . Jeśli A jest częścią A' , to chcemy nazywać α mniejszym od α' , przez co zbiór $\mathfrak{S}$ staje się *uporządkowany*.

Jeśli uznamy przekroje wytworzone przez ułamki wymierne za równe samym tym ułamkom wymiernym i nazwiemy je krótko przekrojami wymiernymi, to zbiór $\mathfrak{S}$ zawiera zbiór $\mathfrak{R}$ i zbiór $\mathfrak{S}$ jest w każdym bądź razie zbiorem gęstym. Zbiór $\mathfrak{S}$ jest jednak także *ciągły*; oznaczmy bowiem przekroje przez duże litery gotyckie $\mathfrak{A}, \mathfrak{B}, \dots$ i rozważmy jakikolwiek przekrój $(\mathfrak{A}, \mathfrak{B})$ w rozmaitości przekrojów, wtedy możemy ustalić w $\mathfrak{S}$ element $\alpha = (A, B)$, przy czym przyjmujemy do A każdy ułamek wymierny, który wytwarza przekrój z $\mathfrak{A}$, a wrzucamy do B wszystkie inne ułamki wymierne, które wytwarzają przekroje w $\mathfrak{B}$. Ten przekrój α w $\mathfrak{R}$ *wytwarza* przekrój $(\mathfrak{A}, \mathfrak{B})$ w $\mathfrak{S}$. Zostanie to udowodnione, gdy pokaże się, że każdy element α' w $\mathfrak{S}$, który jest mniejszy od α , należy do $\mathfrak{A}$ oraz każdy element β' w $\mathfrak{S}$, który jest większy od α , należy do $\mathfrak{B}$.

Niech zatem $\alpha' = (A', B')$ oraz $\alpha' < \alpha$, wtedy istnieją ułamki wymierne w A , które nie są zawarte w A' ; istnieje więc wymierny μ taki, że $\alpha' < \mu < \alpha$ i ten μ wytwarza przekrój, który zawarty jest w $\mathfrak{A}$, a więc sam należy do $\mathfrak{A}$; ponieważ $\alpha' < \mu$, więc również α' należy do $\mathfrak{A}$. Całkiem podobnie pokazuje się, że każdy β' , który jest większy od α , należy do $\mathfrak{B}$, a tym samym ciągłość $\mathfrak{S}$ została udowodniona.

Ten bardzo abstrakcyjny tok rozważań daje nam pewność, że przyjęcie istnienia zbioru ciągłego nie zawiera żadnej sprzeczności, że takie zbiory istnieją co najmniej w królestwie myśli. Geometria, jak też analiza, które zawsze chętnie nawiązują do oglądu geometrycznego, od dawna milcząco przyjmowały jako swego rodzaju aksjomat istnienie zbiorów ciągłych, np. punktów linii prostej lub jakiegokolwiek innej linii ciągłej. Także rozróżnienie pomiędzy zbiorami gęstymi a ciągłymi, które leży u podstaw rozróżnienia odcinków współmiernych i niewspółmiernych nie pojawiło się u starożytnych.³

My również nie chcemy rezygnować w dalszym ciągu z pomocy oglądu geometrycznego i traktować np. punktów linii prostej bez zastrzeżeń jako zbioru ciągłego.

* * *

Zbiór uporządkowany $\mathfrak{M}$ nazywa się *mierzalnym* przy następujących założeniach: dodawanie oraz zwielokrotnienie są w ogólności wykonalne w $\mathfrak{M}$, podobnie jak odejmowanie mniejszego od większego elementu, tj. z jakichkolwiek dwóch elementów a, b (które mogą też być sobie równe) można utworzyć, wedle określonego przepisu, nowy element $a+b$ z $\mathfrak{M}$ tak, że $a+b$ jest większy od a i od b i że zachodzą znane reguły, wyrażone formułami $a+b = b+a$, $(a+b)+c = a+(b+c)$; a dla dwóch elementów a, c , z których drugi jest większy, można znaleźć trzeci element b tak, że $a+b = c$, co będzie także wyrażane znakiem odejmowania $b = c - a$. Dwa nierówne elementy $\mathfrak{M}$ mają zatem zawsze określoną *różnicę*. Z tych założeń wynika, że zwiększa się suma, gdy zwiększa się jeden z jej składników. Powtórzone, powiedzmy m -razy dodawanie tego samego elementu nazywa się zwielokrotnieniem, a jego wynik będzie oznaczany przez ma .

Dochodzi w końcu jeszcze jedno założenie, a mianowicie to, że dla każdego danego a wystarczająco duże zwielokrotnienie ma jest większe od dowolnie danego

³ Euklides, *Elementy*, księga X.

innego elementu b . Wśród elementów zbioru mierzalnego nie ma zatem elementu największego.

W gęstym zbiorze mierzalnym nie ma też elementu najmniejszego; gdyby bowiem a był najmniejszym, a b dowolnym elementem, to między b oraz $b + a$ nie mogłyby leżeć żadne elementy, ponieważ, gdy $b < c < b + a$, to z definicji mierzalności wynikałoby, że $a' = c - b$ byłby mniejszy od a . Zachodzi też wynikanie odwrotne: zbiór mierzalny, w którym nie występuje element najmniejszy, jest gęsty. Jeśli bowiem a oraz $a + b$ są jakimikolwiek dwoma elementami, to wystarczy przecież dodać do a tylko jeden element, który jest mniejszy od b , aby otrzymać element między a oraz $a + b$.

Jeśli zbiór jest ciągły, to założenia dotyczące mierzalności mogą zostać jeszcze uproszczone, ponieważ wtedy odejmowanie jest konsekwencją dodawania. Jeśli mianowicie a oraz c są dwoma elementami ciągłego zbioru uporządkowanego $\mathfrak{M}$, w którym wykonalne jest dodawanie, oraz $c > a$, to otrzymujemy przekrój (A, B) w $\mathfrak{M}$, gdy skierujemy do A wszystkie elementy x , dla których $a + x \leq c$, a do B te, dla których $a + x > c$. Ten przekrój będzie wytworzony przez element b , dla którego musi być $a + b = c$.

Liczby naturalne wedle naszej definicji tworzą zbiór mierzalny, w którym występuje najmniejszy element, a mianowicie 1. Rozmaitość ułamków wymiernych także staje się mierzalna, gdy określi się dodawanie oraz odejmowanie wedle znanych reguł rachowania na ułamkach. Szczególnie ważna, a przy tym typowa dla zbiorów mierzalnych jest rozmaitość odcinków prostoliniowych lub długości linii, które są dodawane po prostu przez leżenie przy sobie. Również zbiory materiałów, porównywanych przez ważenie oraz okresy czasu, mierzone zegarem dostarczają przykładów zbiorów mierzalnych. Sposób mierzenia nie leży tu w naturze samej rozmaitości, lecz jest wprowadzony przez myślącego obserwatora; tak więc, dla przykładu, byłoby równie dobrze rozumieć przez sumę $a + b$ dwóch odcinków a oraz b przeciwprostokątną trójkąta prostokątnego o przyprostokątnych a oraz b zamiast, jak to zwykle bywa przyjęte, odcinek złożony z przyłożonych do siebie odcinków a i b .

Aby ciągły zbiór $\mathfrak{S}$ przekrojów w rozmaitości $\mathfrak{R}$ ułamków wymiernych uczynić mierzalnym, należy najpierw zauważyć co następuje.

Jeśli $\alpha = (A, B)$ jest elementem $\mathfrak{S}$, a μ dowolnym danym ułamkiem wymiernym, to zawsze można określić w A element a' tak, że $a' + \mu = b'$ jest zawarty w B . Wybierzmy bowiem dwa dowolne elementy a, b , a wtedy można tak określić liczbę naturalną m , że $m\mu > b - a$, a więc $a + m\mu$ jest zawarty w B . Jeśli potem h jest najmniejszą liczbą całkowitą, dla której $a + h\mu$ jest zawarty w B , to $a + (h - 1)\mu = a'$ jest w A , a więc $a' + \mu = b'$ jest w B .

Rozumiemy teraz przez sumę $(A, B) + (A', B') = (A'', B'')$ lub $\alpha + \alpha' = \alpha''$ ten przekrój w $\mathfrak{R}$, który otrzymuje się, gdy do A'' skieruje się ułamek wymierny α'' tylko wtedy, gdy istnieje α w A oraz α' w A' o tej własności, że $\alpha'' \leq \alpha + \alpha'$. W istocie, (A'', B'') jest przekrojem; jeśli bowiem α'' jest zawarty w A'' , to to samo zachodzi również dla każdego mniejszego ułamka oraz istnieją ułamki, które są zawarte w A'' i ułamki, które nie są zawarte w A'' , a mianowicie ułamki o postaci $a + a'$ oraz $b + b'$. Dalej, α'' jest większy od α oraz od α' . Albowiem A'' zawiera najpierw wszystkie A , a gdy a' jest dowolnym ułamkiem w A' , to można wybrać a

w A tak, iż element $a + a'$ z A'' jest zawarty w B ; a zatem A'' jest obszerniejszy od A . Przekroje wytworzone przez ułamki wymierne μ, μ' dają w wyniku dodawania przekrój wytworzony przez $\mu + \mu'$.

* * *

Przechodzimy teraz do definicji *stosunków*, które od starożytności były uważane za podstawę nauki o liczbach, a które najpierw występują u Euklidesa⁴.

Gdy połączy się w pary elementy zbioru mierzalnego $\mathfrak{M}$ i same te pary te postrzeka jako elementy, to powstaje nowa rozmaitość; oznaczamy taką parę przez $a : b$ lub też przez $\frac{a}{b}$, odróżniamy jednak $a : b$ od $b : a$, gdy a i b są różnymi elementami i nazywamy a licznikiem, a b mianownikiem $a : b$. Chcemy nazywać te pary stosunkami i chcemy teraz ów nowy zbiór uporządkować i uczynić mierzalnym.

Załóżmy najpierw, że istnieją dwie liczby całkowite m, n takie, że $na = mb$, jak to np. zawsze ma miejsce, gdy $\mathfrak{M}$ jest systemem liczb naturalnych, lub gdy a, b są dwoma współmiernymi odcinkami; gdy wtedy p, q są innymi liczbami całkowitymi, to $qa = pb$ wtedy i tylko wtedy, gdy $mq = np$. Para liczb p, q jest całkowicie określona poprzez to żądanie, o ile dojdzie jeszcze warunek, iż p, q mają być liczbami względnie pierwszymi. Może być wtedy $m = hp, n = hq$, gdzie h jest dowolną liczbą całkowitą. W tym przypadku nazywamy stosunek $a : b$ wymiernym i uważamy go za równy ułamkowi wymiernemu $m : n$ lub $p : q$. Te ułamki wymierne mogą odtąd być uważane za stosunki liczb całkowitych.

Wszystkie równe między sobą stosunki wymierne tworzą *liczbę wymierną*, a liczby wymierne tworzą, podobnie jak ułamki wymierne, rozmaitość uporządkowaną, gęstą i mierzalną.

* * *

W rozmaitości liczb wymiernych liczby naturalne same się porządkują, gdy przez liczbę naturalną m rozumie się stosunek $m : 1$.

Wracamy teraz do jakiegokolwiek rozmaitości mierzalnej $\mathfrak{M}$ i bierzemy z niej jakiegokolwiek elementy a oraz b . Jeśli wybierze się, co zawsze jest możliwe, dwie liczby naturalne m, n tak, że $na > mb$, to stosunek $a : b$ jest nazywany większym od stosunku wymiernego $m : n$, lub

$$\frac{a}{b} > \frac{m}{n},$$

a gdy $m : n > p : q$, to również $a : b > p : q$.

Podobnie wynika, że jeśli $n'a < m'b$, to

$$\frac{a}{b} < \frac{m'}{n'}.$$

Jeśli $a : b > m : n$, to można znaleźć jedną, a w konsekwencji również dowolnie wiele liczb wymiernych $m_1; n_1$ takich, że

$$\frac{a}{b} > \frac{m_1}{n_1} > \frac{m}{n},$$

⁴ *Elementy*, księga V.

tj. można między $a : b$ oraz $m : n$ wstawić dowolnie wiele stosunków wymiernych. Aby to pokazać, wybiera się dowolną liczbę całkowitą k oraz ustala liczbę całkowitą h tak, że $h(na - mb) > kb$, co zawsze jest możliwe; wtedy

$$\frac{a}{b} > \frac{hm + k}{hn} > \frac{m}{n}.$$

Podobnie wynika, że gdy $n'a < m'b$, $h'(m'b - n'a) > k'a$, to

$$\frac{a}{b} < \frac{h'm'}{h'n' + k'} < \frac{m'}{n'}.$$

Jeśli teraz $a : b$ oraz $\alpha : \beta$ są dwoma jakimikolwiek stosunkami, które dla zwięzłości chcemy oznaczać przez e , ε , i których elementy należą do tych samych lub różnych rozmaitości, to możliwe są dwa przypadki: 1) między e oraz ε nie leży żaden stosunek wymierny μ lub 2) stosunek wymierny leży między e oraz ε .

W przypadku 1) oba stosunki e oraz ε nazywają się *wzajem równymi* i widać, że dwa stosunki, które równe są trzeciemu, są także wzajem równe; jeśli bowiem $e < \mu < \varepsilon$, to każdy inny stosunek e' jest albo mniejszy, albo równy, albo większy od μ . Jeśli e' jest równy μ , to zarówno między e i e' , jak i między e' i ε leżą stosunki wymierne. Jeśli jednak e' jest mniejszy od μ , to μ leży między e' oraz ε , a jeśli e' jest większy od μ , to μ leży między e oraz e' ; a więc e' nie może być jednocześnie równy e oraz ε .

W przypadku 2) stosunki e , ε nazywają się *nierównymi*. Może zatem być albo $\alphae < \mu < \varepsilon$, albo β) $e > \mu' > \varepsilon$ i oba te przypadki nawzajem się wykluczają, ponieważż z $e < \mu < \varepsilon$, $\varepsilon < \mu'$ wynika, że $\mu < \mu'$, a w konsekwencji dostaje się $e < \mu'$.

Zależność między wielkościami 2 α) lub 2 β) zachodzi także, gdy e lub ε zostaną zastąpione równym im elementem. Jeśli bowiem $\mu < \varepsilon$ oraz $\mu \geq \varepsilon'$, to między ε oraz ε' leży stosunek wymierny i ε , ε' nie są równe.

W przypadku 2 α) e nazywa się *mniejszym od* ε , w przypadku 2 β) e nazywa się *większym od* ε .

Między dwa nierówne stosunki można wstawić dowolną liczbę stosunków wymiernych.

Gdy teraz zbierzemy razem wszystkie równe między sobą stosunki, to otrzymamy pojęcie gatunkowe, które oznaczamy jako *liczbę* w sensie ogólnym. Liczba jest zatem nazwą lub znakiem dla pewnej rozmaitości, której elementami są właśnie te stosunki, które równe są jednemu spośród nich⁵. W tym pojęciu liczby są zawarte stosunki wymierne, a w konsekwencji również liczby naturalne, jako stosunki $m : 1$; owe stosunki wymierne tworzą *liczby wymierne*.

Liczby, które nie wywodzą się ze stosunków wymiernych, nazywają się *liczbami niewymiernymi*.

Na mocy tego, co dotąd powiedziano, liczby tworzą *zbiór uporządkowany*, a ich porządek można ustalić, gdy dla każdej liczby wybierze się jako reprezentanta jakikolwiek zawarty w niej stosunek.

⁵ Przez to pojęcie gatunkowe można też wprowadzić w prostszy i bardziej konsekwentny sposób liczby naturalne.

* * *

Spośród dwóch stosunków o tym samym mianowniku oraz różnych licznikach ten jest większy, którego licznik jest większy, a spośród dwóch stosunków o tym samym liczniku oraz różnych mianownikach ten jest mniejszy, którego mianownik jest większy.

Jeśli mianowicie a, a', b są dowolnymi elementami zbioru mierzalnego oraz $a' > a$, to najpierw wybiera się liczbę całkowitą n tak, że $na > b$ oraz $n(a' - a) > b$. Potem bierze się *najmniejszą* liczbę całkowitą m , która spełnia warunek $mb > na$; wtedy $na < mb$, ale $mb < na'$. Gdyby bowiem $mb \geq na' \geq na + n(a' - a)$, to byłoby $mb > na + b$; a więc byłoby $(m - 1)b > na$, wbrew założeniu. Dalej,

$$\frac{a}{b} < \frac{m}{n} < \frac{a'}{b},$$

a więc $a : b < a' : b$; całkiem podobnie można udowodnić, że gdy $b' > b$, to $a : b > a : b'$.

Można to twierdzenie wyrazić również tak, że stosunek rośnie wraz ze wzrostem licznika, a maleje wraz z rosnącym mianownikiem.

Łatwo otrzymuje się z tego także następujące twierdzenie: Jeśli a, b, c, d są elementami tego samego zbioru mierzalnego oraz $a : b = c : d$, to również $a : c = b : d$.

Założmy bowiem, że $a : b = c : d$, wtedy muszą istnieć dwie liczby całkowite m, n takie, że

$$na < mc$$

$$nb > md.$$

Wtedy jednak byłoby, na mocy właśnie udowodnionego twierdzenia, $na : nb < mc : md$, a więc także $a : b < c : d$, wbrew założeniu.

Wynika z tego teraz następujące *główne twierdzenie*. Jeśli spośród czterech wielkości a, b, c, d jakiegokolwiek trzy z nich są dane w ciągłym zbiorze mierzalnym, to również czwartą ustalić można w tym samym zbiorze tak, że $a : b = c : d$.

Twierdzenie jest bezpośrednią konsekwencją założonej ciągłości. Albowiem gdy weźmie się np. element x w $\mathfrak{M}$ z A lub B , odpowiednio do tego, czy $x : b$ jest mniejszy czy większy od $c : d$, to otrzyma się przekrój, który jest wytworzony przez element a taki, że spełniony jest warunek $a : b = c : d$. Twierdzenie to zachodzi jednak również dla pewnych zbiorów nieciągłych, np. dla ułamków wymiernych.

Wynika z tego, że za reprezentantów dwóch liczb zawsze można wybrać dwa stosunki, których elementy należą do tej samej rozmierności oraz mają ten sam dowolnie wybrany mianownik. Dodawanie jest wtedy tak zdefiniowane, że zachodzi

$$\frac{a}{c} + \frac{b}{c} = \frac{a + b}{c}.$$

Ta reguła obejmuje, jako szczególny przypadek, dodawanie ułamków wymiernych, a żeby uzasadnić ją w ogólności trzeba potem jeszcze tylko pokazać, że jeśli $a : c = a' : c'$ oraz $b : c = b' : c'$, to musi być również $(a + b) : c = (a' + b')c'$. Dowodzimy

tego, pokazując, że gdy $a : c = a' : c'$ oraz $(a + b) : c > (a' + b')c'$, to musi być też $b : c > b' : c'$. Niech zatem, gdy m oraz n są dwiema liczbami całkowitymi, będzie

$$\frac{a+b}{c} > \frac{m}{n} > \frac{a'+b'}{c'},$$

wtedy $n(a+b) > mc$, a więc

$$\frac{b}{c} > \frac{mc - na}{nc}, \quad \frac{mc' - na'}{nc'} > \frac{b'}{c'}.$$

Z drugiej strony, łatwo jednak wynika z założenia $a : c = a' : c'$, że także

$$\frac{mc - na}{nc} = \frac{mc' - na'}{nc'},$$

a więc $b : c > b' : c'$, c.b.d.u.

Tym samym dowiedziono, że także liczby, tak, jak je zdefiniowaliśmy, tworzą zbiór mierzalny. Jeśli a oraz c są wzięte z rozmaitości ciągłej, to stosunki $a : c$ tworzą zbiór ciągły, także przy ustalonym c , a ponieważ istnieją w ogóle zbiory ciągłe, więc również liczby tworzą zbiór ciągły.

Jeśli teraz $\alpha, \beta, \gamma, \delta$ są liczbami, to z proporcji $\alpha : \beta = \gamma : \delta$ można ustalić dowolną z tych czterech liczb poprzez trzy pozostałe. Jeśli przyjmie się $\delta = 1$, to otrzyma się *mnożenie* $\alpha = \beta\gamma$, a przemienność czynników jest konsekwencją twierdzenia, że $\alpha : \gamma = \beta : \delta$ wynika z $\alpha : \beta = \gamma : \delta$. Jeśli szuka się γ , to otrzymuje się dzielenie; a z podanej wyżej definicji dodawania wynika podstawowa formuła $\alpha(\beta + \gamma) = \alpha\beta + \alpha\gamma$. Cztery sposoby rachowania są zatem wykonalne w dziedzinie liczb, z jedynym ograniczeniem, iż przy odejmowaniu odjemnik musi być mniejszy od odjemnej.

* * *

Konstrukcja przekroju w ciągu liczb dostarcza zawsze dowodu istnienia liczby, która czyni zadość określonym żądaniom. Otrzymuje się więc przekrój (A, B) , gdy przyjmie się do A wszystkie i tylko te liczby, których kwadrat jest mniejszy od ustalonej liczby α ; temu przekrojowi odpowiada określona liczba, której kwadrat jest równy α i która będzie oznaczana przez $\sqrt{\alpha}$ i przez to udowodnione zostaje istnienie pierwiastków kwadratowych.

Do przekrojów sprowadzić można także ciągi liczbowe wprowadzone przez G. Cantora w jego definicji liczb niewymiernych⁶.

Wedle CANTORA przez ciąg liczbowy należy rozumieć jakikolwiek nieograniczony i w określony sposób uporządkowany system liczb:

$$S = x_1, x_2, x_3, x_4, \dots$$

o tej własności, że istnieje określona liczba g , poniżej której nie opada żadna liczba z S i taka, że gdy δ jest dowolnie wybraną liczbą, a x_n, x_m różnymi elementami z S ,

⁶ Cantor, Über die Ausdehnung eines Satzes aus der Theorie der trigonometrischen Reihen, *Mathematische Annalen* **5** (1872); por. też Heine, *Elemente der Functionenlehre*, Journal für Mathematik **74** (1872).

to $x_n - x_m$ lub $x_m - x_n$ pozostaje zawsze mniejsza od δ , gdy m oraz n przekroczą wystarczająco dużą liczbę.

Zawsze istnieją liczby, których nie przekraczają liczby takiego ciągu; jeśli bowiem ustalona została δ oraz odpowiednia n , to x_m , którego wartość może być też równa m , nigdy nie przekracza największej z liczb $x_1, x_2, \dots, x_n, x_n + \delta$. Otrzymuje się teraz przekrój (A, B) , gdy do B wrzuci się te liczby, których nie przekracza żadna x_n , dla wystarczająco dużych wartości n , a wszystkie pozostałe liczby (które zatem przekracza nieskończenie wiele x_n) do A . Jeśli ten przekrój będzie wytworzony przez liczbę α , to jakkolwiek mała jest ε , zawsze istnieje nieskończenie wiele liczb x_n między $\alpha - \varepsilon$ oraz α i można powiedzieć, że ta całkowicie określona przez S liczba α jest wytworzona przez ciąg liczbowy S . Wedle CANTORA, ciąg liczbowy S jest po prostu definicją liczby α . Oczywiście jedna i ta sama liczba może zostać wytworzona przez bardzo różne ciągi liczbowe. Te ciągi liczbowe należy jednak traktować jako między sobą równe. Między innymi, można liczby w S brać wszystkie jako ułamki wymierne. Również na odwrót, łatwo pokazać, że dla każdej podanej liczby α zawsze można podać ciągi liczbowe S , przez które α jest wytworzona tak, że ogół tych ciągów liczbowych S również tworzy zbiór ciągły.

* * *

Przy objaśnianiu podstawowych sposobów rachowania wynikło niewygodne ograniczenie odejmowania, od którego się teraz uwolnimy poprzez wprowadzenie zera oraz liczb ujemnych.

Niech x oznacza każdy element dotąd zdefiniowanego systemu liczbowego, który teraz chcemy określać jako system liczb dodatnich. Bierzemy ten system liczbowy po raz drugi i dla odróżnienia, w tym drugim systemie, który określamy ma być jako system liczb ujemnych, każdy element oznaczamy przez $-x$. Drugi system porządkujemy teraz dokładnie odwrotnie niż pierwszy, tak, że wszędzie, gdzie w pierwszym systemie x występuje jako „większa”, w tym drugim ustalamy $-x$ jako „mniejszą”, i na odwrót. Dodawanie oraz odejmowanie dla $-x$ są objaśnione podobnie jak dla x , tak, że ma być $(-x) + (-y) = -(x+y)$; $(-x) - (-y) = -(x-y)$.

Chcemy jednak oba te systemy uporządkować razem w taki sposób, aby każda $-x$ była mniejsza od każdej x . Otrzymujemy wtedy zbiór uporządkowany, w którym nie ma największego ani najmniejszego elementu. System ten jest też w ogólności ciągły, tylko jedyny przekrój $(-x, x)$ nie będzie wytworzony przez żaden element i tu, gdzie oba systemy się stykają występuje jeszcze naruszenie ciągłości. Aby zapewnić tu ciągłość, musimy zatem dodać odpowiadającą przekrojowi $(-x, x)$ jeszcze jedną liczbę zero lub 0, która zdefiniowana jest właśnie przez ów przekrój. Mamy wtedy uporządkowany zbiór ciągły, z obu stron nieograniczony, pełny ciąg liczb rzeczywistych.

W tak rozszerzonej dziedzinie liczbowej objaśniamy teraz ogólnie dodawanie, ustalając na drodze definicji:

$$x + (-x) = 0, \quad x + 0 = x,$$

$$\begin{aligned} x + (-y) &= x - y, \text{ gdy } x > y, \\ x + (-y) &= -(y - x), \text{ gdy } x < y. \end{aligned}$$

Przy tak określonym dodawaniu zachodzą następujące prawa, gdy z_1, z_2, z_3 są jakimikolwiek liczbami z całej dziedziny liczbowej:

$$z_1 + z_2 = z_2 + z_1, \quad (z_1 + z_2) + z_3 = z_1 + (z_2 + z_3),$$

które nazywa się prawami przemienności oraz łączności. Sumę dowolnej liczby składników tworzy się, łącząc w sumy w dowolnej kolejności po dwa składniki. O odejmowanie nie trzeba się już szczególnie troszczyć, gdy określi się $z_1 - z_2$ jako $z_1 + (-z_2)$ oraz ustali, że $-(-z) = z$.

Poglądowo przedstawia się ciąg liczbowy przez punkty, wychodząc od ustalonego i oznaczonego przez 0 punktu prostej i nanosząc liczby dodatnie jako odcinki z jednej, powiedzmy prawej, strony, a liczby ujemne z drugiej (lewej) strony. Obraz sumy dwóch odcinków $z_1 + z_2$ otrzymuje się, gdy poczynawszy od punktu z_1 naniesie się w prawą lub lewą stronę odcinek o długości $\pm z_2$, w zależności od tego, czy z_2 jest dodatnia czy ujemna.

Mnożenie oraz dzielenie w rozszerzonej dziedzinie liczbowej zostaje określone przez równania

$$\begin{aligned} x(-y) &= (-x)y = -xy \\ (-x)(-y) &= xy, \quad 0x = 0. \end{aligned}$$

Dzielenie jako odwrotność mnożenia jest zawsze możliwe, poza przypadkiem, gdy dzielnik jest zerem.

* * *

Dalsze rozszerzenie pojęcia liczby polega na wprowadzeniu *wielkości zespolonych*. Łączymy każde dwie liczby całego ciągu liczbowego w pary (x, y) i traktujemy dwie takie pary (x, y) oraz (a, b) jako równe tylko wtedy, gdy $x = a, y = b$. Pary te tworzą rozmaitość, której elementy nie są co prawda uporządkowane, ale dla których operacje rachowania: dodawanie, odejmowanie, mnożenie oraz dzielenie mają być określone wedle następujących reguł. Niech

$$(x, y) + (a, b) = (x + a, y + b),$$

$$(x, y)(a, b) = (xa - yb, xb + ya),$$

i ustalamy poza tym, że ma być $(x, 0) = x$, co nie przeczy tym równaniom. Tylko wtedy jest $(x, y) = 0$, gdy x oraz y obie są równe zero. Dalej, dla zwięzłości oznaczamy $(0, 1)$ przez i . Wtedy powyższe równości mają następujące konsekwencje:

$$(x, 0)(0, 1) = (0, x) \quad \text{lub} \quad = ix,$$

$$(x, 0) + (0, y) = (x, y) = x + yi,$$

$$x + yi + a + bi = (x + a) + (y + b)i,$$

a dla odwrotności dodawania:

$$x + yi - (a + bi) = (x - a) + i(y - b),$$

a dalej mnożenie:

$$(x + yi)(a + bi) = xa - yb + i(xb + ya), \quad i^2 = -1,$$

$$(x + yi)(x - yi) = x^2 + y^2,$$

$$x + yi = (a + bi) \frac{(a - bi)(x + yi)}{(a^2 + b^2)}$$

lub

$$\frac{x + yi}{a + bi} = \frac{ax + by + i(ay - bx)}{a^2 + b^2},$$

przez co określone jest dzielenie, oprócz przypadku, gdy $a + bi = 0$. Liczby o postaci ix nazywają się *liczbami czysto urojonymi*, a i urojoną jednostką. Liczby $a + bi$ nazywają się *urojonymi* lub *zespolonymi*. Liczby *rzeczywiste* oraz *czysto urojone* są w tym zawarte jako przypadki szczególne.

Podstawowe sposoby rachowania są zatem w dziedzinie *liczb zespolonych* $x + yi$ wykonalne bez ograniczeń (z wyjątkiem dzielenia przez zero), a rachowanie na liczbach rzeczywistych jest ich szczególnym przypadkiem.

Za GAUSSEM przedstawia się liczby zespolone $z = x + yi$ jako punkty płaszczyzny, przyjmując u podstaw prostokątny układ współrzędnych i traktując punkt o współrzędnych x, y jako obraz wartości liczbowej z . Punkty osi x przedstawiają w wyżej opisany sposób liczby rzeczywiste. Punkty osi y są obrazami liczb czysto urojonych yi . Początek układu współrzędnych jest obrazem liczby 0. Wektor promienia z punktu zerowego do punktu z ma wartość liczbową $\rho = \sqrt{x^2 + y^2}$, która jest nazywana *wartością bezwzględną*, wartością lub, wedle starszego sposobu wyrażania się, modulem liczby zespolonej z .

Jedynie liczba 0 ma wartość bezwzględną 0. Każda liczba dodatnia występuje jako wartość bezwzględna nieskończenie wielu liczb zespolonych, a obrazy punktowe wszystkich liczb o tej samej wartości bezwzględnej leżą na okręgu, którego środek leży w początku układu współrzędnych.

Dwie liczby urojone, które różnią się tylko znakiem przy i , a zatem $x + yi$ oraz $x - yi$ nazywają się urojonymi sprzężzonymi. Ich iloczyn jest kwadratem wartości bezwzględnej każdej z nich.

Jeśli jedna z dwóch sprzężzonych liczb urojonych jest równa zero, to także pozostała jest równa zero, a więc można w każdym prawdziwym równaniu liczbowym zastąpić i przez $-i$, bez naruszenia prawdziwości.

Chcemy jeszcze udowodnić często stosowane twierdzenie, że wartość bezwzględna sumy dwóch różnych od zera liczb zespolonych nigdy nie jest większa od sumy bezwzględnych wartości obu składników, a równa tylko wtedy, gdy stosunek (iloraz) obu składników jest rzeczywisty i dodatni. Niech mianowicie

$$\begin{array}{lll} z = x + yi & c = a + bi & Z = (x + a) + (y + b)i \\ \rho^2 = x^2 + y^2 & r^2 = a^2 + b^2 & R^2 = (x + a)^2 + (y + b)^2, \end{array}$$

wtedy

$$(r + \rho - R)(r + \rho + R) = (r + \rho)^2 - R^2 = 2(r\rho - ax - by),$$

co z pewnością jest dodatnie, gdy $ax + by \leq 0$. Jeśli jednak $ax + by > 0$, to z

$$r^2 \rho^2 - (ax + by)^2 = (ay - bx)^2$$

wynika, że $r\rho \geq ax + by$, a równość zachodzi tylko wtedy, gdy $ay - bx = 0$. Wynika z tego zatem, że $r + \rho > R$ i tylko w szczególnym przypadku, gdy $ay - bx = 0$, $ax + by > 0$ mamy $r + \rho = R$, z czego wynika, co powiedziano.

W przedstawieniu geometrycznym twierdzenie to jest wyrazem tego, że w trójkącie jeden bok jest mniejszy od sumy obu pozostałych. Twierdzenie ma inną jeszcze konsekwencję, że wartość bezwzględna sumy nie może być mniejsza od różnicy wartości bezwzględnych składników. Jeśli bowiem zastosujemy poprzednie twierdzenie do sumy $c = Z - z$, to wynika z tego, że $r \leq R + \rho$ lub

$$R \geq r - \rho.$$

Równość ma miejsce tylko wtedy, gdy iloraz $z : c$ jest rzeczywisty oraz ujemny.

* * *

Trzeba jeszcze powiedzieć słowo o najważniejszym środku pomocniczym algebry, rachowaniu na literach. Stosowanie tego środka pomocniczego jest tak ogólne, że niekiedy terminu rachowanie na literach używa się jako synonimu dla terminu algebra. Reguły, mówiące jak rachować na takich wyrażeniach literowych przyjmujemy za wprzódki znane. Równania między wyrażeniami literowymi mogą być dwóch rodzajów; albo, po pierwsze, są to tak zwane identyczności, tj. dwa wyrażenia przyjmowane za równe mogą zostać tak przekształcone za pomocą reguł rachowania, że oba wyrażenia są dokładnie zgodne. Otrzymuje się potem z takich równań literowych prawdziwe równania liczbowe, gdy litery zostaną zastąpione przez jakiekolwiek liczby, powiedzmy rzeczywiste bądź zespolone, przy założeniu, że nie występuje przy tym żądanie dzielenia przez zero. Litery w takich równaniach są często oznaczane też jako zmienne, ponieważ można sobie wyobrazić, bez popadnięcia w sprzeczność, że za litery można podstawiać coraz to inne i inne wartości liczbowe.

Równania drugiego rodzaju między wyrażeniami literowymi nie mają tego charakteru identyczności. O wiele bardziej zawierają one żądanie stawiane liczbom, które można bez popełnienia błędu podstawić za litery. Zadanie wykrycia wartości liczbowych, które czynią zadość takim żądaniom algebra nazywa *rozwiązaniem równania*. W tych równaniach litery oznaczane są także jako „niewiadome”. Bardzo często w jednym i tym samym równaniu występują litery dwóch rodzajów, takie, za które można podstawiać dowolne wartości liczbowe oraz inne, których wartość liczbową ma dopiero zostać wykryta.

* * *

Podstawa tłumaczenia: Weber, H.: 1895, *Lehrbuch der Algebra. Einleitung*, Friedrich Vieweg und Sohn, Braunschweig, 1-20.

Tłumaczenie: Jerzy Pogonowski
22 października 2010

Annales Universitatis Paedagogicae Cracoviensis

Studia ad Didacticam Mathematicae Pertinentia VII (2015)

ISSN 2080-9751

Piotr Błaszczuk

XXX Konferencja z Historii Matematyki, Będlewo, 5–8 V 2016*

1. W maju 2016 roku, w Ośrodku Badawczo-Konferencyjnym Instytutu Matematycznego PAN w Będlewie, odbyła się XXX Konferencja z Historii Matematyki. Jej organizatorem był Witold Więśław, człowiek-instytucja, osoba, której zasługi dla historii matematyki w Polsce trudno przecenić. W. Więśław uczestniczył we wszystkich 30. Konferencjach z Historii Matematyki¹, wniósł istotny wkład w organizację, prace programowe oraz wydanie materiałów pokonferencyjnych wielu edycji, a poczynając od XXV organizował konferencje niemal samodzielnie, sporadycznie zyskując wsparcie instytucjonalne, czy to Instytutu Matematycznego Uniwersytetu Wrocławskiego, czy PTM-u, czy Instytutu Matematycznego PAN.

Okrągły jubileusz skłania do przypomnienia historii Konferencji². Oto one:

- I Ogólnopolska Szkoła Historii Matematyki, Pokrzywna, woj. opolskie, 1–6 VI 1986.
- II Ogólnopolska Szkoła Historii Matematyki: Matematyka XIX wieku, Lubin, k. Międzyzdrojów, 25–29 V 1987.
- III Ogólnopolska Szkoła Historii Matematyki: Matematyka przełomu XIX i XX wieku. Nurt mnogościowy, Jaworze, k. Bielska-Białej, 25–29 V 1988.
- IV Ogólnopolska Szkoła Historii Matematyki: Matematyka przełomu XIX i XX wieku, Pogorzelica, 8–12 V 1989.
- V Ogólnopolska Szkoła Historii Matematyki: Probabilistyka i mechanika w szkicach historycznych, Dziwnów, 9–13 V 1990.
- VII Ogólnopolska Szkoła Historii Matematyki: Problemy Hilberta, Międzyzdroje, 10–14 V 1993.
- VIII Ogólnopolska Szkoła Historii Matematyki: Prawdy i mity w historii matematyki, Rudy Raciborskie, 6–10 V 1994.
- IX Ogólnopolska Szkoła Historii Matematyki: Matematyka w Polsce 1851–1950, Międzyzdroje, 4–9 VI 1995.

*XXX Conference on History of Mathematics, Będlewo, 5–8 V 2016

¹ W. Więśław nie był obecny na pierwszej konferencji, ale jego referat został tam odczytany i zamieszczony w materiałach pokonferencyjnych.

² Nazwa konferencji zmieniała się wraz ze zmianą organizatorów: od Ogólnopolskiej Szkoły Historii Matematyki, przez Konferencję Naukową PTM z Historii Matematyki, po Konferencję z Historii Matematyki. Pamiętając o tych różnicach, dalej będzie stosować jedną nazwę Konferencja z Historii Matematyki.

- XIII Ogólnopolska Szkoła Historii Matematyki: Matematyka XVIII stulecia, Kołobrzeg, 17–21 V 1999.
- XIV Ogólnopolska Szkoła Historii Matematyki: Matematyka czasów Gaussa, Zielona Góra, 8–14 V 2000.
- XV Ogólnopolska Szkoła Historii Matematyki: Matematyka czasów Weierstrassa, Kołobrzeg, 28 V–2 VI 2001.
- XVI Ogólnopolska Szkoła Historii Matematyki: Algorytmy w dziejach matematyki, Turawa, k. Opola, 14–18 V 2002.
- XVII Ogólnopolska Szkoła Historii Matematyki: Matematyka abelowa, Nowy Sącz, 9–13 VI 2003.
- XVIII Ogólnopolska Szkoła Historii Matematyki: Słynne dzieła matematyczne i rocznice, Białystok–Supraśl, 31 V–4 VI 2004.
- XIX Ogólnopolska Szkoła Historii Matematyki: Wokół Bernoullich, Zamość, 6–10 VI 2005.
- XX Ogólnopolska Szkoła Historii Matematyki: Historia matematyki polskiej, Ustroń, k. Bielska-Białej, 21–26 V 2006.
- XXI Konferencja Naukowa PTM z Historii Matematyki: Leonhard Euler (1707–1783) – w trzechsetlecie urodzin, Iwonicz-Zdrój, 21–25 V 2007.
- XXII Konferencja Naukowa PTM z Historii Matematyki: Matematyka i matematycy polscy okresu zaborów (1795–1918), Iwonicz-Zdrój, 26–30 V 2008.
- XXIII Konferencja Naukowa PTM z Historii Matematyki: Matematyka i matematycy polscy okresu zaborów (1795–1918), Iwonicz-Zdrój, 25–29 V 2009.
- XXIV Konferencja Naukowej PTM z Historii Matematyki: Matematyka polska przełomu XIX i XX wieku, Iwonicz-Zdrój, 24–28 V 2010.
- XXV Konferencja Naukowa PTM z Historii Matematyki: Matematyka polska I połowy XX wieku, Będlewo, 23–27 V 2011.
- XXVI Konferencja Naukowa z Historii Matematyki: Matematyka polska w latach 1851–1950, Iwonicz-Zdrój, 21–25 V 2012.
- XXVII Konferencja z Historii Matematyki, Będlewo, 20–24 V 2013.
- XXVIII Konferencja z Historii Matematyki, Będlewo, 9–12 VI 2014.
- XIX Konferencja z Historii Matematyki, Będlewo, 25–28 V 2015.
- XXX Konferencja z Historii Matematyki, Będlewo, 5–8 V 2016.

Podsumowanie dziesięciu pierwszych edycji Konferencji zawierają artykuły: S. Fudali, *Drukowane pokłosie Szkół Historii Matematyki* oraz Z. Pawlikowska-Brożek, *Geneza i tematyka Szkół Historii Matematyki*³. Tytuły referatów z dwiętnastu pierwszych Konferencji oraz dane dotyczące wydawnictw pokonferencyjnych zebrał W. Więśław w nocie zatytułowanej *Poprzednie tomy z Historii Matematyki*⁴. Dodajmy jeszcze, bo to symptomatyczne, że najwięcej uczestników, blisko stu, zgromadziła VII Konferencja. Jej dorobek wydano w tomie *Problemy Hilberta w pięćdziesięciolecie śmierci ich twórcy*⁵.

³ S. Fudali, *Drukowane pokłosie Szkół Historii Matematyki*, [w:] *X Szkoła Historii Matematyki*, Uniwersytet Opolski. Zeszyty Naukowe. Matematyka 30. Opole 1997, 85–90; Z. Pawlikowska-Brożek, *Geneza i tematyka Szkół Historii Matematyki*, tamże, 115–124.

⁴ W. Więśław, *Poprzednie tomy z Historii Matematyki*, *Antiquitates Mathematicae* 1, 2007, 12–23.

⁵ W. Więśław (red.), *Problemy Hilberta w pięćdziesięciolecie śmierci ich twórcy*, Instytut Historii Nauki PAN, Warszawa 1997.

Pierwsze konferencje dedykowano wyraźnie zakreślonym tematom z historii matematyki powszechnej. Kolejne, poczynając od XX, były zorientowane na matematykę polską, skupiając się przy tym raczej na biografiach niż kwestiach matematycznych. Tę zmianę tematyki dobrze ilustruje zestawienie programów I i XXX konferencji. Oto tytuły referatów zamieszczonych w wydawnictwie pokonferencyjnym I Szkoły Historii Matematyki⁶:

- R. Duda, *Zarys rozwoju koncepcji przestrzeni.*
- J. Folta, *Rozwój geometrii w XIX wieku, niektóre rysy rozwoju społecznego, a specyfika powstawania szkół geometrii.*
- J. Konarski, *Historia rozwoju geometrii różniczkowej.*
- M. Kordos, *Początki metody dedukcyjnej.*
- M. Kordos, *Zarys historii geometrii rzutowej.*
- J. Mioduszewski, *Rozwój pojęcia continuum.*
- M. Moszyńska, *Z historii geometrii analitycznej.*
- L. Szczerba, *Historia geometrii nieeuklidesowej.*
- L. Szczerba, *Historia podstaw geometrii euklidesowej.*
- W. Wiśław, *Historia konstrukcji geometrycznych.*

A oto program XXX Konferencji:

- R. Duda, *Matematycy polscy na tle dziejów kraju w wiekach XIX i XX.*
- L. Maligranda, *Wydział Matematyczno-Fizyczny Uniwersytetu Lwowskiego w latach 1939–1941.*
- J. Prytuła, *Wykłady z matematyki na Uniwersytecie Lwowskim do 1939 roku.*
- W. Wilczyński, *Zbiory, których konstrukcje zawdzięczamy polskim matematykom.*
- A. Pałka, *Starodruki anamorficzne w zbiorach Biblioteki Jagiellońskiej.*
- W. Wiśław, *Michał Pelka Poliński (1785–1848).*
- K. Wuczyńska, *Sylwetka nauczyciela matematyki na tle rozwoju szkolnictwa wg Adama Wachulki.*
- W. Piotrowski, *Matematyka w Wolnej Wszechnicy Polskiej w okresie międzywojennym.*
- L. Maligranda, *Aleksander Pelczyński (1932–2012).*
- R. Murawski, *Warszawska Szkoła Logiczna o przedmiocie matematyki i logiki.*
- D. Ciesielska, *O wykładach Władysława Kretkowskiego (1882/83) oraz Wacława Sierpińskiego (1908) z teorii kwaternionów.*
- Z. Pogoda, *Prace Antoniego Hoborskiego (1879–1940) dotyczące nauczania geometrii.*
- S. Domoradzki, *Andrzej Pelczar (1937–2010) – życie i twórczość.*
- Z. Pogoda, *Kazimierz Żorawski (1866–1953) – w 150 rocznicę urodzin.*
- P. Błaszczyk, *Heinricha Webera konstrukcja liczb rzeczywistych.*
- S. Domoradzki, *Krzysztof Tatarkiewicz (1923–2011) – życie i twórczość.*
- R. Duda, *Stan i potrzeby historii matematyki w Polsce.*
- W. Wójcik, *Jan Śleszyński, polski logik i matematyk. Analiza osiągnięć okresu krakowskiego.*
- I. Józwik, M. Terepeta, *Prace Stefana Kempistego z teorii funkcji rzeczywistych.*

⁶ Zob. J. J. Charatonik (red.), *Zeszyty Naukowe Wyższej Szkoły Inżynierskiej w Opolu* Nr 125, *Matematyka*, Z. 12, Opole 1987.

- J. Koroński, *Matematyka w siedemdziesięcioleciu Politechniki Krakowskiej*.
- A. Dawidowicz, *Alternatywa Fredholma wczoraj i dziś*.
- D. Ciesielska, *Wybrane publikacje z czasopisma Izwiestija Cesarskiego Uniwersytetu Warszawskiego (1870–1917)*.
- J. Strelcyn, *Matematyka i matematycy na Cesarskim Uniwersytecie Warszawskim (1870–1915)*.

2. Podczas gdy w Niemczech, Anglii czy USA już dawno wytworzyły się grupy zawodowych historyków matematyki, to w Polsce historię matematyki uprawiają tylko zawodowi matematycy, właściwie jako hobby, a nie jako podstawową aktywność naukową. W rezultacie polskie środowisko wydało szereg opracowań wybranych zagadnień matematyki najnowszej, przede wszystkim XX-wiecznej. Są to rzetelne prace, pisane przez specjalistów w dziedzinach szczegółowych, oparte na materiałach źródłowych, czyli artykułach badawczych i przeglądowych oraz monografiach.

Jednakże problemy i techniki matematyki współczesnej ledwie wykraczają poza granice XX wieku i im głębiej w historię, tym bardziej bezradny staje dzisiejszy matematyk wobec dawnych, a zarzuconych już technik, konwencji, sposobów argumentacji czy wreszcie pojęć, by wymienić równość figur, wielkość, proporcję, nieskończenie małe, stosunek $\frac{0}{0}$, itp. W rezultacie matematyka XVIII, XVII wieku i dawniejsza jest dla współczesnego matematyka zrozumiała o tyle, o ile umie on powiązać ją z matematyką XX wieku, a teksty źródłowe są wówczas z konieczności interpretowane.

W polskich opracowaniach tych okresów źródła są owszem cytowane, ale ich interpretacje nie są oryginalne – w istocie powtarzają tezy podstawowych monografii z historii matematyki autorów takich jak Carl Boyer, Moris Kline czy Victor Katz⁷. Tym sposobem w polskiej literaturze ciągle obowiązuje na przykład pogląd, że Leibniz, Newton czy Euler rozwijali rachunek różniczkowy dysponując przybliżoną wersją współczesnej analizy matematycznej i dopiero w ostatnich dekadach XIX wieku, wraz z powstaniem klasycznej analizy matematycznej, sformułowano „rygorystyczne” podstawy analizy. Taki obraz nie wytrzymuje konfrontacji z tekstami źródłowymi (autorstwa Lebniza, Newtona czy Eulera, jak i tymi pochodzącymi od Cantora, Dedekinda czy Weierstrassa). I chociaż w badaniach historycznych interpretacja tekstu jest często niezbędna, to akurat w przypadku pism ojców założycieli rachunku różniczkowego, perspektywę lepszą niż klasyczna analiza matematyczna stwarza analiza niestandardowa, dziedzina, która w Polsce jest słabo reprezentowana⁸.

⁷ Zob. C. Boyer, *Historia rachunku różniczkowego i całkowego i rozwój jego pojęć*, PWN, Warszawa 1964; M. Kline, *Mathematical Thought from Ancient to Modern Times*, Oxford University Press, New York 1972; V. Katz, *History of Mathematics*, Addison-Wesley, Reading, Massachusetts 1998.

⁸ Zob. P. Błaszczyk, M. Katz, D. Sherry, Ten misconceptions from the history of analysis and their debunking, *Foundations of Science*, 18(1), 2013, 43–74; P. Błaszczyk, V. Kanovei, M. Katz, D. Sherry, Controversies in the Foundations of Analysis: Comments on Schubring’s Conflicts, *Foundations of Science* 2016 (online first); T. Bascelli et al., Leibniz versus Ishiguro: closing a quarter century of syncategoremia, *HOPOS* 6(1), 2016, 117–147; P. Błaszczyk, V. Kanovei, M. Katz, S. Kutateladze, D. Sherry, Toward a history of mathematics focused on procedures, *Foundations of Science* 2016 (przyjęte do druku); J. Bair et al., Interpreting the infinitesimal

Gdy sięgniemy jeszcze głębiej, do matematyki greckiej, to polska historia matematyki jest prowadzona bez kontaktu z materiałami źródłowymi, polegając z konieczności na opracowaniach lub spekulacjach. Co więcej, autorzy podstawowych syntez historii matematyki, jak chociażby ci wyżej wskazani, także nie pracują na tekstach źródłowych i powielają stare schematy. W rezultacie o matematyce greckiej ciągle opowiada się za pomocą liczb rzeczywistych, a w najlepszym razie za pomocą jakiegoś bliżej niesprecyzowanego ciała archimedesowego⁹, z pominięciem faktu, że grecka teoria proporcji była podstawowym narzędziem matematycznym do końca XVII wieku. Tym sposobem całościowe spojrzenia na historię matematyki, jakie wydało polskie środowisko matematyczne, ma owszem wiele walorów, ale głównie popularnonaukowych¹⁰.

3. XXX Konferencję z Historii Matematyki w sposób szczególny nazwały dwa referaty: Ady Pałki *Starodruki anamorfczne w zbiorach Biblioteki Jagiellońskiej* oraz Romana Dudy *Uwagi o sytuacji historii matematyki w Polsce*. Magister Pałka jest słuchaczką Studiów Doktoranckich przy Instytucie Matematyki Uniwersytetu Jagiellońskiego, profesor Duda to wybitny matematyk, a zarazem nestor polskiej historii matematyki. Być może spotkanie na małej konferencji tych prelegentów jest znakiem, że do polskiej nauki wkraczają już zawodowi historycy matematyki.

Słynny obraz Hansa Holbeina *Ambasadorowie* (1533, National Gallery, Londyn), z zagadkowym przedstawieniem ludzkiej czaszki w dolnej partii płótna, to przykład anamorfozy, czyli techniki zniekształcania przedstawień, które ukazują się w naturalnych proporcjach dopiero wtedy, gdy są oglądane przy spełnieniu odpowiednich warunków. Rozwinięto trzy odmiany anamorfozy: (1) *perspektywiczną*, w której przedstawienie ukazuje się w naturalnych proporcjach, gdy jest oglądane z wybranego miejsca (czyli tak jak w obrazie Holbeina), (2) *lustrzaną*, w której przedstawienie ukazuje się w naturalnych proporcjach, dopiero jako odbite na lustrzanej powierzchni cylindrycznej lub stożkowej, (3) *zwinianą*, w której przedstawienie ukazuje się w naturalnych proporcjach, gdy papier, na którym położono rysunek zostanie odpowiednio zwinięty, np. w stożek. Okazuje się, że polski grafik i rysownik Jan Ziarnko w traktacie *Perspectivae Stereo Graphicae Pars Specialis* (1619) jako pierwszy opisał matematyczne reguły *anamorfozy zwinianej*. Odkrycie to zawdzięczamy młodej badaczce z Krakowa, Adzie Palce. Rzecz została ogłoszona w prestiżowym periodyku *Print Quarterly* i jest świetnym przykładem propagowania polskiego wkładu w historię matematyki¹¹.

Uwagi o sytuacji historii matematyki w Polsce Romana Dudy wpisują się w tradycję refleksji nad stanem polskiej historii matematyki, sięgającą takich prac, jak Aleksandra i Ludwika Birkenmajerów *Najważniejsze dezyderaty nauki*

mathematics of Leibniz and Euler, *Journal for General Philosophy of Science* 2016 (przyjęte do druku); zob. także literaturę cytowaną w tych artykułach.

⁹Zob. P. Błaszczyk, K. Mrówka, *Euklides. Elementy. Księgi V–VI*. Tłumaczenie i komentarz, CCPress, Kraków 2013, 176–177.

¹⁰Zob. J. Mioduszewski, *Ciągłość. Szkice z historii matematyki*, WSiP, Warszawa 1996; W. Wiśniewski, *Matematyka i jej historia*, Wydawnictwo NOWIK, Opole 1997; M. Kordos, *Wykłady z historii matematyki*, SCRIPT, Warszawa 2005; L. Gruszecki, *Zarys dziejów matematyki*, Wydawnictwo KUL, Lublin 2009.

¹¹Zob. A. Pałka, Jan Ziarnko's Anamorphic Print 'A Pair of Lovers Embracing', *Print Quarterly* 32(1), 2015, 3–13.

polskiej w zakresie historii nauk matematycznych (1918) czy Zdzisława Opiała *Stan i potrzeby historii matematyki w Polsce* (1968).

Nawiązując do odróżnienia *history* i *heritage*, zaproponowanego niegdyś przez Ivora Grattan-Guinnessa¹², R. Duda wyróżnił dwa zasadnicze sposoby uprawiania historii matematyki: *historyczny* (skupiony na biografiach) i *matematyczny* (zajmujący się rozwojem pojęć, czy po prostu historią idei). Następnie, zwracając uwagę na specyfikę polskiego środowiska, odróżnił historię matematyki powszechnej i narodowej, a rysując program na przyszłość apelował o skupienie badań na *matematycznej* historii matematyki polskiej.

Diagnostując stan polskiej historii matematyki Profesor zwrócił uwagę na minimalne zainteresowanie historią ze strony zawodowych matematyków, wynikiem czego jest chociażby likwidacja Komisji Historii Matematyki PTM i trudności w zdobywaniu awansów naukowych w zakresie historii matematyki.

Trzeba jednak przyznać – dodajmy trochę polemicznie – że dla uprawiania matematyki znajomość historii nie jest konieczna. Zainteresowanie historią matematyki wśród matematyków, podobnie jak na przykład historią powszechną wśród polityków, jest pochodną wykształcenia ogólnego. Jeżeli historia matematyki nie jest obecna ani w nauczaniu szkolnym, ani akademickim, to zainteresowanie tą dziedziną może przyjść jedynie z zewnątrz: z referatu wysłuchanego przypadkowo gdzieś na kongresie matematycznym, czy z książki o intrygującym tytule, zauważonej u kolegi z USA. Z drugiej strony, znane są przykłady zagadnień wyrosłych na pograniczu historii i filozofii matematyki, które angażowały wybitnych, acz refleksyjnie nastrojonych matematyków. Wymieńmy dla przykładu różnice w interpretacji Księgi II *Elementów* Euklidesa – zagadnienie znane pod nazwą *algebraizacja geometrii*. Dyskusja wokół tego problemu zaangażowała z jednej strony matematyków takich jak B. Van der Waerden, H. Freudenthal, A. Weil, z drugiej, wybitnych historyków, takich jak O. Neugebauer, S. Unguru, I. Grattan-Guinness¹³.

O ile skłonność do refleksji jest przypadłością, na którą nie sposób wpłynąć, to zainteresowanie historią matematyki można rozbudzić, ukazując ją jako źródło aktualnych problemów czysto matematycznych. Kilka przykładów. Aksjomaty geometrii euklidesowej podane przez Hilberta w *Grundlagen der Geometrie* można traktować jako interpretację Ksiąg I–IV *Elementów* Euklidesa. Przyjęto w nich jednak rozwiązanie, które historyk winien zakwestionować. Otóż w systemie Hilberta porządek liniowy jest definiowany (*via* relacja *leżenia między*); tak samo jest w ujęciu geometrii euklidesowej autorstwa Karola Borsuka i Wandy Szmielew, gdzie porządki liniowe w zbiorach odcinków i kątów są definiowane¹⁴. O ile więc we współczesnych aksjomatykach porządek liniowy jest definiowany, o tyle w *Elementach* jest on pojęciem pierwotnym — fakt ten wynika z analizy Księgi V¹⁵. Zadanie matematyczne z tym związane można więc tak przedstawić: zbudować aksjomatykę

¹² Zob. I. Grattan-Guinness, *The mathematics of the past: distinguishing its history from our heritage*, *Historia Mathematica* 31, 2004, 163–185.

¹³ Przegląd stanowisk w tym trwającym już ponad wiek sporze podaje V. Blåsjö w artykule *In defence of geometrical algebra*, *Archive for History of Exact Science*, 2016 (online first).

¹⁴ Zob. K. Borsuk, W. Szmielew, *Podstawy geometrii*, PWN, Warszawa 1972; zob. także P. Błaszczyk, *Analiza filozoficzna rozprawy Richarda Dedekinda 'Stetigkeit und irrationale Zahlen'*, Wydawnictwo Naukowe AP, Kraków 2007, rozdz. II, 77–102.

¹⁵ Zob. P. Błaszczyk, K. Mrówka, *Euklides. Elementy. Księgi V–VI. Tłumaczenie i komentarz*, CCPress, Kraków 2013, 98–122, 256–269.

geometrii, w której porządek liniowy (relacja *wiekszy-mniejszy* wśród odcinków i *wiekszy-mniejszy* wśród kątów) będzie pojęciem pierwotnym. Taka aksjomatyka byłaby bliżej oryginalnego systemu Euklidesa, niż ta podana przez Hilberta.

Stworzenie nowej aksjomatyki geometrii euklidesowej jest poważnym wyzwaniem, dlatego wskażemy jeszcze dwa inne problemy wyrastające wprost z badań historycznych. Otóż wbrew powszechnemu mniemaniu wzorem metody dedukcyjnej jest nie tyle geometria Euklidesa (braki systemu Euklidesa są często powtarzane przy okazji prezentacji systemu Hilberta), ile teoria proporcji rozwinięta w Księdze V. Znane są dwa systemy aksjomatów tej teorii¹⁶. Zadanie, które czeka na rozwiązanie, to wykazanie niezależności tych aksjomatów (dla każdego systemu, musi być ono oczywiście rozważane odrębnie) – jest więc to takie samo zadanie jak to, które postawił sobie Hilbert prezentując aksjomaty geometrii euklidesowej.

I jeszcze jedno zadanie. Kartezjusz naszkicował w *Geometrii* sposób rozwiązania zagadnienia Pappusa dla dowolnej liczby prostych¹⁷. Zamysł swój doprowadził do przypadku 13 linii. Zagadnienie Pappusa ciągle czeka na ogólne rozwiązanie.

Z problemem, który hasłowo nazwiemy *Jak zainteresować zawodowych matematyków historią matematyki*, wiążą się kwestie metodologiczne, które w wystąpieniu Romana Dudy były zaznaczone tylko drobnym muśnięciem. A pytanie jest następujące: jak pisać historię matematyki, by zachować wierność materiałom źródłowym, trafić do współczesnego czytelnika, a przy tym nie popaść w ton popularnonaukowy. W książce *Euklides. Elementy. Księgi V–VI. Tłumaczenie i komentarz* pokazaliśmy, jak zamieniamy tekst Księgi V w formuły algebraiczne zrozumiałe dla współczesnego matematyka. W wyniku wskazanych zabiegów hermetyczny tekst sprzed 2 000 lat, po uprzednim przerobieniu kilku ćwiczeń i przestudiowaniu objaśnień, może być czytany z taką łatwością, jak podręcznik algebry elementarnej. Przyjętej metodzie transkrypcji (formalizacji) towarzyszą objaśnienia, takie na przykład jak to, że w greckiej matematyce w ogóle (nie tylko w Księdze V) nie ma pojęcia odpowiadającego symbolowi $+$, a formuła, powiedzmy $A + B$ w tekście greckim będzie zapisana jako „A, B”. Czytelnik otrzymuje więc reguły, które nie tylko pozwalają mu przekładać prozę matematyczną sprzed 2 000 lat na formuły algebraiczne, ale i odwrotnie, formuły algebraiczne przekładać na prozę matematyczną. Jako ciekawostkę dodajmy, że w jednej z recenzji artykułu skierowanego do pisma *Foundations of Science* otrzymaliśmy taką oto uwagę, napisaną, co wynika, z wcześniejszych partii, przez zawodowego matematyka: „An ordinary mathematician would quite appreciate a translation of Elements V.5, not just into English words but symbols that a modern person might understand”. Polski czytelnik dysponuje taką transkrypcją nie tylko definicji V.5, ale i V.4, V.7, i wszystkich 25 twierdzeń Księgi V wraz z dowodami.

Inną próbą tego rodzaju jest formalizacja arytmetyki odcinków z *Geometrii* Kartezjusza¹⁸. Pozwala ona pokazać, że chociaż Kartezjusz nie dysponował ogólnym pojęciem liczby, to mógł posługiwać się odcinkami, tak jak współczesny

¹⁶ Zob. F. Beckmann, *Neue Gesichtspunkte zum 5. Buch Euklids*, *Archive for History of Exact Sciences* IV, 1967, s. 1–144; P. Błaszczyk, K. Mrówka, *Euklides. Elementy. Księgi V–VI. Tłumaczenie i komentarz*, op. cit., 115–117.

¹⁷ Zob. P. Błaszczyk, K. Mrówka, *Kartezjusz, Geometria. Tłumaczenie i komentarz*, Universitas, Kraków 2015.

¹⁸ Ibidem, 146–166.

matematyk posługuje się elementami ciała uporządkowanego. Jeszcze inną próbę formalizacji przedstawiliśmy w serii artykułów towarzyszących tłumaczeniom klasycznych tekstów z historii liczb rzeczywistych¹⁹.

Na podobne formalizacje czeka jeszcze wiele klasycznych dzieł i zagadnień, by wymienić Archimedesesa i technikę wyczerpywania czy analizę matematyczną w ujęciu Eulera. O dziwo istnieje tylko jedna praca poświęcona transpozycji dorobku Archimedesesa na współczesny język matematyczny: E. J. Diksterhuis, *Archimedes*²⁰. Próba ta jest niezadowolająca z dwóch względów: (1) autor tworzy całą gamę własnych pojęć i oznaczeń, przez co czytelnik musi uczyć się nowego języka symbolicznego, (2) posługuje się pojęciem granicy, które jest całkowicie obce matematyce greckiej i nie przystaje do metody wyczerpywania. Jest to więc próba formalizacji, która zakłamuje tekst źródłowy. Gdy zaś idzie o Eulera, to pozytywnym przykładem formalizacji jest praca M. McKinzie, C. Tuckey, *Hidden lemmas in Euler's summation of the reciprocals of the squares*²¹. W tym przypadku, autorzy objaśnili pojęcia, które dla większości historyków wydają się niezrozumiałe oraz uzupełnili rozumowanie Eulera o twierdzenia (dowodzone w ramach współczesnej analizy niestandardowej), które pozwalają traktować jego dowody z zachowaniem współczesnych standardów poprawności²².

Na zakończenie uwag do wystąpienia Romana Dudy postawmy najważniejsze pytanie: Po co historia matematyki? Profesor nie podjął tego wątku, bo to prawie pytanie o sens życia. Spróbujemy podejść do tej kwestii pragmatycznie. Otóż Grattan-Guinness widzi oczywistą rolę historii w nauczaniu matematyki na poziomie akademickim. I faktycznie, co jakiś czas pojawiają się próby ziszczenia takiego zamysłu²³. Generalny pomysł jest taki, aby wprowadzać studenta w daną dziedzinę matematyki w porządku historycznym, a nie tak, jak to jest najczęściej, w oparciu o jakąś monografię powstałą w drugiej połowie XX wieku.

Jeszcze inną rolę może odegrać historia matematyki w filozofii. Sytuacja jest tutaj taka, że filozofowie matematyki, otoczeni murem teorii mnogości i pogrążeni w dyskusjach nad aksjomatami, już dawno stracili kontakt z rzeczywistością, to jest bieżącą matematyką (pod tym względem są w dużo gorszej sytuacji niż historycy matematyki, acz nie boją się nad tym, bo funkcjonują w ramach innych insty-

¹⁹ Zob. P. Błaszczyk, Nota o rozprawie Otto Höldera *Die Axiome der Quantität und die Lehre vom Mass*, *Annales Universitatis Paedagogicae Cracoviensis. Studia ad Didacticam Mathematicae Pertinentia* V, 2013, 129–142; P. Błaszczyk, Nota o rozprawie Eduarda Heinego *Elemente der Functionenlehre*, *Annales Universitatis Paedagogicae Cracoviensis. Studia ad Didacticam Mathematicae Pertinentia* VI, 2014, 129–144; P. Błaszczyk, Nota o *Lehrbuch der Algebra. Einleitung* Heinricha Webers, *Annales Universitatis Paedagogicae Cracoviensis. Studia ad Didacticam Mathematicae Pertinentia* VII, 2015, 129–139 (niniejszy tom)

²⁰ Zob. E. J. Diksterhuis, *Archimedes*, Princeton University Press, Princeton, New Jersey 1989 (wyd. pierwsze 1938).

²¹ Zob. M. McKinzie, C. Tuckey, *Hidden lemmas in Euler's summation of the reciprocals of the squares*, *Archive for History of Exact Science* 51, 1997, 29–57.

²² Jako przykład niezrozumienia Eulera można wskazać najnowszą pracę Jeremy'ego Graya, *The real and the complex: a history of analysis in the 19th century*, Springer, Heidelberg 2015.

²³ Zob. J. Fauvel, J., van Mannen, (Eds.), *History in Mathematics Education. The ICME Study*. Kluwer Academic, Dordrecht 2000; A. Knoebel, R. Laubenbacher, J. Lodder, D. Pengally, *Mathematical Masterpieces*, Springer, New York 2006

tucji)²⁴. Z historycznego punktu widzenia teoria mnogości jest tylko jedną z wersji całościowego ujęcia matematyki, rolą historyków byłoby natomiast opisanie innych wersji matematyki z ukazaniem ich swoistości. (Dotychczasowa praktyka jest taka, że historia matematyki jest przedstawiana jako dążenie do teorii z XX wieku.) Wówczas w odpowiedzi na pytanie, czym jest matematyka, można ukazać jej różne wersje historyczne, tak jak w odpowiedzi na pytanie, czym jest człowiek ukazuje się historycznie różne przejawy jego aktywności. Istnieją książki z historii matematyki sięgające takich fundamentalnych pytań²⁵.

4. XXXI Konferencja z Historii Matematyki odbędzie się w Będlewie, w dniach 22–26 V 2017.

*Instytut Matematyki
Uniwersytet Pedagogiczny
ul. Podchorążych 2
PL-30-084 Kraków
e-mail pb@up.krakow.pl*

²⁴ Chwalebnym przykładem filozofa sięgającego w głąb historii matematyki jest Ian Hacking; zob. jego *Why Is There Philosophy of Mathematics At All?*, Cambridge University Press, Cambridge 2014.

²⁵ Zob. J. Klein, *Greek mathematical thought and the origin of algebra*, Dover, New York 1968.

