Jest to streszczenie odczytu wygłoszonego na XXV Szkole Matematyki Poglądowej „Elementarne, ale niebanalne".

Co może mały zbiór?

Wiktor BARTOL, Warszawa

Cóż może być bardziej elementarnego niż mały zbiór? Zapewne tylko zbiór jeszcze mniejszy. Przypatrzmy się więc najmniejszym zbiorom.

Zbiór pusty to podstawa całej konstrukcji teorii mnogości; od niego zaczyna się hierarchia Zermelo–Fraenkla, a dochodzi...

Niemniej potężne są zbiory jednoelementowe. Można z nich zbudować, element po elemencie, dowolny zbiór. Co więcej, jednoelementowa np. grupa pochłania przez homomorfizmy wszystkie grupy, a że homomorfizmy zachowują strukturę, to w jednym elemencie skupiają się struktury wszystkich grup. Jak we Wszechświecie przed Wielkim Wybuchem.

Zbiór dwuelementowy także wiele potrafi. Na przykład, dwuelementowa algebra Boole'a jest doskonałym reprezentantem klasy wszystkich algebr Boole'a: żadna równość nie będzie w tej klasie prawdziwa, jeśli nie jest prawdziwa w algebrze dwuelementowej. Nie mówiąc już o tym, że ze zbioru dwuelementowego można zbudować każdy zbiór potęgowy...

Zbiór trzyelementowy... Taki właśnie zbiór jest jednym z głównych bohaterów opowieści, którą wypada już rozpocząć. Najpierw potrzebne jest jednak pewne przygotowanie, aby zbiór trzyelementowy wypadł na jego tle bardziej okazale.

Wiele klas algebr jest opisywanych przez równości. Na przykład, grupą jest każda algebra $(G, \cdot, ^{-1}, e)$ z jedną operacją dwuargumentową $\cdot$, jedną operacją jednoargumentową $^{-1}$ i jedną stałą (czyli operacją zeroargumentową) e, w której prawdziwe są równości

$$\begin{aligned} x \cdot (y \cdot z) &= (x \cdot y) \cdot z, \\ x \cdot e &= x, \\ e \cdot x &= x, \\ x \cdot x^{-1} &= e, \\ x^{-1} \cdot x &= e. \end{aligned}$$

Nietrudno zauważyć, że we wszystkich grupach jest prawdziwych jeszcze nieskończenie wiele innych równości, jak choćby taka:

$$(x \cdot y) \cdot ((y^{-1} \cdot z) \cdot (v^{-1} \cdot z)^{-1}) = x \cdot v.$$

Dlaczego więc, definiując grupy, nie podajemy wszystkich równości, które mają być w nich prawdziwe? Oczywiście dlatego, że w każdej algebrze, w której prawdziwe są równości występujące w tradycyjnej definicji grupy, prawdziwe są także wszystkie inne równości, takie jak równość z powyższego przykładu. „Takie", to znaczy wynikające z definiującego zbioru równości. Inaczej mówiąc, ów definiujący zbiór równości stanowi *bazę* dla zbioru wszystkich równości prawdziwych w klasie grup.

Sprecyzujmy te pojęcia. Niech Σ będzie dowolnym zbiorem równości (pewnej ustalonej sygnatury algebraicznej). Powiemy, że zbiór $\Sigma_0 \subseteq \Sigma$ jest *bazą równościową* zbioru Σ, gdy każda algebra, która jest modelem zbioru Σ_0, jest jednocześnie modelem całego zbioru Σ. Równoważnie, Σ_0 jest bazą dla zbioru Σ, gdy oba zbiory generują ten sam zbiór równości zamknięty na równościowe reguły wnioskowania. Nietrudno zauważyć, że, na przykład, zbiór $\Sigma_0 = \{x = y\}$ jest bazą równościową dla zbioru wszystkich równości ustalonej sygnatury, gdyż modelami tej równości są dokładnie algebry jednoelementowe, a każda taka algebra jest zarazem modelem każdej równości odpowiedniej sygnatury. Mówiąc inaczej, z równości $x = y$ można wyprowadzić dowolną równość $p = q$, podstawiając term p za zmienną x i term q za zmienną y (a takie podstawianie legalizuje jedna z reguł wnioskowania).

W 1937 B. H. Neumann [12] pytał o bazy dla teorii równościowych grup; wykazał, że każda grupa abelowa ma skończoną bazę równościową. W 1950 r.

R. Specht [18] przeniósł pytanie na pierścienie (ale na odpowiedź trzeba było czekać jeszcze ponad 20 lat). W latach 50. minionego wieku zaczęto zajmować się szerzej pytaniem o charakterze ogólniejszym, które możemy sformułować następująco:

Problem skończonej bazy.

Niech $\mathbb{A}$ będzie skończoną algebrą skończonej sygnatury. Czy istnieje skończona baza dla zbioru wszystkich równości prawdziwych w tej algebrze, czyli dla jej teorii równościowej?

Mówiąc krócej: *Czy każda skończona algebra skończonego typu ma skończoną bazę równościową?*

Pytanie wydaje się, na pierwszy rzut oka, nieco przesadne. Jakich to komplikacji można oczekiwać od tak prostego obiektu: skończony zbiór, skończenie wiele operacji? Co miałoby powodować, że równości prawdziwych w takiej algebrze nie da się wyprowadzić z jakiegoś skończonego podzbioru?

Przy takich oczekiwaniach pewnym rozczarowaniem mogło być pierwsze ogólne twierdzenie, dotyczące problemu skończonej bazy.

Twierdzenie (R. Lyndon [6], 1951). *Każda algebra dwuelementowa (dowolnej skończonej sygnatury) ma skończoną bazę równościową.*

Takie twierdzenie zdaje się dobrą okazją do przedstawienia dowodu, który w przypadku algebr zaledwie dwuelementowych powinien być w miarę prosty i krótki. Ale tak nie jest – i tu pojawia się ostrzeżenie, iż problem nie jest tak trywialny, jak się w pierwszej chwili wydaje. Otóż dowód Lyndona czerpie z klasyfikacji algebr dwuelementowych, przeprowadzonej w 1941 roku przez E. Posta [15], badającego dwuwartościowe funkcje logiczne. Przesuwając się klasa po klasie, Lyndon wykazuje istnienie skończonej bazy dla każdej z nich, korzystając z podanej przez Posta szczególnej postaci jej algebr.

Trzy lata później Lyndon [7] publikuje przykład 7-elementowej algebry (z jedną operacją dwuargumentową i jedną stałą) bez skończonej bazy. Okazuje się, że skończona struktura algebraiczna nie musi mieć skończonego opisu równościowego! Oto ta algebra (stałą jest 0):

$\cdot$	0	1	2	3	4	5	6
0	0	0	0	0	0	0	0
1	0	0	0	0	0	0	0
2	0	0	0	0	0	0	0
3	0	0	0	0	0	0	0
4	0	4	5	6	0	0	0
5	0	5	5	5	0	0	0
6	0	6	6	6	0	0	0

Idea dowodu, zastosowana w pracy Lyndona, przewinie się także w dalszych pracach o podobnym charakterze; przedstawimy ją za chwilę właśnie przy jednej z nich.

A więc już 7 elementów – i tylko jedna operacja dwuargumentowa plus stała – wystarczy, by zadać strukturę o nieskończenie bogatej teorii równościowej. Co się jednak dzieje między 2 (patrz twierdzenie Lyndona) a 7?

W 1964 r. W. W. Wiszin [19] opublikował przykład następnej algebry z jedną operacją dwuargumentową (takie algebry nazywa się *grupoidami*), tym razem składającej się z 4 elementów, a już rok później poprzeczka opadła na najniższy możliwy poziom, gdy W. L. Murski [10] zbudował grupoid 3-elementowy bez skończonej bazy. Zatrzymajmy się chwilę przy tej algebrze. Oto ona:

$\cdot$	0	1	2
0	0	0	0
1	0	0	1
2	0	2	2

Dowód opiera się na następującym pomyśle: budujemy rodzinę równości, po jednej równości n zmiennych dla każdej liczby naturalnej n, po czym dowodzimy, że po pierwsze, każda z nich jest prawdziwa w naszej algebrze, a po drugie, równość n zmiennych z tej rodziny nie jest konsekwencją żadnego zbioru równości, zawierających mniej niż n różnych zmiennych. Wówczas teza już wynika łatwo: gdyby istniała skończona baza równościowa dla całej teorii równościowej tej algebry, to siłą rzeczy występowałoby w niej skończenie wiele zmiennych, powiedzmy, k. Ale wtedy równość ze zbudowanej wcześniej rodziny, która ma $k+1$ zmiennych, nie może być konsekwencją tej bazy, chociaż należy do teorii równościowej algebry.

W rodzinie równości zbudowanej przez Murskiego, znalazły się równości postaci następującej (pomijamy kropkę reprezentującą operację dwuargumentową):

$$(*) \quad x_1(x_2(x_3\ldots(x_{n-1}(x_nx_1))\ldots)) = (x_1x_2)(x_n(x_{n-1}(\ldots(x_4(x_3x_2))\ldots)).$$

Analiza operacji algebry wykazuje, że istotną rolę przy wyznaczaniu wartości tego rodzaju termów odgrywają elementy sąsiadujące „w szerszym sensie". Powiemy, że a sąsiaduje z b (w danym termie p), gdy w p występują obok siebie dwa nawiasy, w których skrajnie lewymi elementami są, odpowiednio, a i b (nawiasy mogą zawierać tylko jednoelementowy ciąg zmiennych i wtedy formalnie ich w zapisie nie ma), a więc gdy sąsiadują w nim dwa podtermy postaci $(a\ldots)(b\ldots)$. Otóż jeśli w każdym z dwóch termów p i q zmiennych $x_1, x_2, \ldots, x_n$ (dla pewnego n) pierwszą zmienną od lewej jest x_1, a jedynymi parami sąsiadujących (w szerszym sensie) zmiennych są pary $x_1x_2, x_2x_3, \ldots, x_{n-1}x_n, x_nx_1$, to w algebrze zachodzi równość $p = q$. Te warunki spełniają termy każdej z równości prezentowanej wyżej rodziny, zatem każda z tych równości jest prawdziwa w algebrze Murskiego.

Do wykazania drugiego warunku potrzebnego do uzyskania tezy wystarczy teraz znaleźć własność, przysługującą każdej równości n zmiennych, wyprowadzalnej ze zbioru równości o mniejszej liczbie zmiennych, i udowodnić, że równość n zmiennych ze zbudowanej wyżej rodziny tej własności nie ma. Taką własnością (ograniczoną do termów spełniających warunek opisany w poprzednim akapicie) jest własność następująca: po obu stronach równości występuje zmienna x_2, sąsiadująca (w szerszym sensie) jednocześnie z wystąpieniem zmiennej x_1 i z wystąpieniem zmiennej x_3. Jak łatwo zauważyć, równość $(*)$ tej własności nie ma.

Podobne rozumowanie występuje w niemal wszystkich pracach zawierających przykłady algebr bez skończonej bazy, w szczególności we wspomnianej wcześniej pracy R. Lyndona [7].

Wydawałoby się, że algebra Murskiego jest odosobnionym przypadkiem, zbudowanym przez autora *ad hoc* i nie związanym z żadną szerszą teorią. Jeśli nawet tak sądził sam Murski, po kilkunastu latach okazało się, że jest inaczej, niemal wręcz przeciwnie.

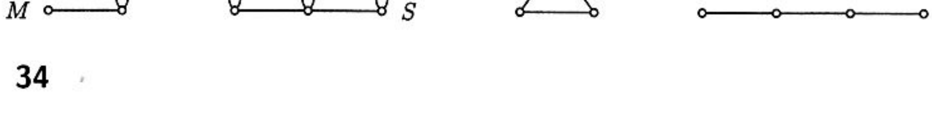
Badając problem skończonej bazy, C. Shallon [16] wprowadziła w 1979 roku pojęcie algebry grafu. Niech $G = (V, E)$ będzie grafem (być może z pętlami) ze zbiorem wierzchołków V i zbiorem krawędzi E. Określmy na zbiorze $V \cup \{0\}$, gdzie $0 \notin V$, dwuargumentową operację $\circ$ w sposób następujący:

$$a \circ b = \begin{cases} a & \text{gdy } (a,b) \in E, \\ 0 & \text{w przeciwnym przypadku.} \end{cases}$$

Algebra $(V \cup \{0\}, \circ)$ jest algebrą grafu G.

W 1987 r. ukazała się praca K. Bakera, G. McNulty'ego i H. Wernera [2], zawierająca następujące piękne twierdzenie:

Twierdzenie. *Algebra grafu G ma skończoną bazę równościową wtedy i tylko wtedy, gdy graf G nie zawiera indukowanego (pełnego) podgrafu izomorficznego z jednym z poniższych grafów:*



Spróbujmy dla przykładu zbudować algebrę grafu M. Bez trudu otrzymujemy taką tabelkę operacji:

$\cdot$	0	1	2
0	0	0	0
1	0	0	1
2	0	2	2

To oczywiście algebra Murskiego. Okazuje się, że sporadyczny przykład mieści się w szerszym kontekście jako sytuacja typowa. (Dodajmy, że C. Shallon wykazała wcześniej brak skończonej bazy równościowej dla algebry grafu S).

Po wyczerpaniu kwestii ilościowej (ile elementów może mieć najmniejsza algebra bez skończonej bazy?) nasuwają się co najmniej dwa dalsze pytania. Po pierwsze, jak często zdarza się, że skończona algebra skończonej sygnatury nie ma skończonej bazy równościowej? Po drugie, czy istnieją jakieś w miarę ogólne warunki, zapewniające istnienie takiej bazy?

Na pierwsze pytanie znów odpowiedzi udzielił (kilkanaście lat później) Murski [11], wykazując, że stosunek liczby k-elementowych grupoidów bez skończonej bazy równościowej do liczby wszystkich grupoidów k-elementowych (grupoidy rozpatrujemy, rzecz jasna, z dokładnością do izomorfizmu) jest rzędu k^{-6}. Jeśli jednak nie rozumieć określenia "ilościowe" zbyt dosłownie, należy wspomnieć o pewnym problemie, sformułowanym przez Alfreda Tarskiego w początkach lat 60. ubiegłego wieku:

Czy zbiór wszystkich algebr skończonej sygnatury określonych na skończonych podzbiorach zbioru $\mathbf{N}$*, które mają skończoną bazę równościową, jest rekurencyjny?*

W latach 80. R. McKenzie pokazał, że problem można sprowadzić do pytania o półgrupy, po czym kilkanaście lat później sprawę rozstrzygnął [9]: *zbiór wszystkich półgrup na skończonych podzbiorach zbioru* $\mathbf{N}$*, które nie mają skończonej bazy, nie jest rekurencyjny, zbiór zaś tych półgrup, które skończonej bazy równościowej nie mają, nie jest rekurencyjnie przeliczalny.* W dowodzie McKenzie w bardzo pomysłowy sposób przypisuje skończone algebry maszynom Turinga, wiążąc problem skończonej bazy z problemem stopu dla tych maszyn.

Pytanie natury jakościowej (które algebry mają skończoną bazę równościową?) doprowadziło do długiej serii prac, zawierających wyniki zarówno pozytywne (istnieje skończona baza), jak i negatywne. Początek tej serii wyznacza praca S. Oates i M. B. Powella [13], w której autorzy udowodnili (metodami specyficznymi dla teorii grup), że każda grupa skończona ma skończoną bazę równościową. Praca ta znalazła interesującą puentę niemal dwadzieścia lat później: R. Bryant [3] znalazł w 1982 roku przykład skończonej grupy, która – traktowana jako algebra z dodatkową stałą, różną od elementu neutralnego grupy – takiej skończonej bazy nie ma.

W 1969 roku P. Perkins [14] podał pewne ogólne warunki, których spełnienie przez teorię równościową półgrupy wyklucza istnienie skończonej bazy równościowej, i wykazał, że warunki te spełnia teoria 6-elementowej półgrupy macierzy 2×2 (z mnożeniem macierzy), składającej się z macierzy:

$$\begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix} \quad \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \quad \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \quad \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \quad \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix} \quad \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}$$

Perkins także wskazuje na przykład skończonej półgrupy ze skończoną bazą, która traci tę własność po dodaniu do sygnatury symbolu elementu neutralnego. Okazało się później, że brak skończonej bazy dla tej 6-elementowej półgrupy macierzy jest „zaraźliwy", implikuje bowiem brak takiej bazy równościowej dla półgrupy wszystkich macierzy 2×2, choć taka baza istnieje dla *pierścienia* tych macierzy. Ta ostatnia uwaga to konsekwencja rozstrzygnięcia problemu skończonej bazy dla pierścieni w 1973 roku, co uczyniło – niezależnie od siebie – dwóch autorów: R. Kruse [4] i I. W. Lwow [5], wykazując, że dla każdego skończonego pierścienia odpowiedź jest pozytywna. Warto dodać, że w 1976 r.

znaleziono przykład 64-elementowego pierścienia niełącznego bez skończonej bazy równościowej.

Sytuacja dla półgrup jest bardziej skomplikowana, o czym świadczy także praca Perkinsa i problem Tarskiego. Wiadomo, że skończoną bazę równościową ma każda półgrupa przemienna, każda półgrupa idempotentna, a także każda półgrupa co najwyżej 4-elementowa.

W 1977 roku K. Baker publikuje (anonsowaną już pięć lat wcześniej) pracę [1], która określa odmienny kierunek badań. Dotąd pytania formułowano wobec klas algebr pewnego ustalonego typu: grupoidy, półgrupy, grupy, pierścienie itp. Baker określa warunek niezależny od sygnatury i od charakteru algebry. Przypomnijmy, że *rozmaitość* algebr tej samej sygnatury to klasa algebr, zamknięta na obrazy homomorficzne, podalgebry i produkty proste. Biorąc wszystkie obrazy homomorficzne podalgebr produktów prostych ustalonej algebry $\mathbb{A}$, otrzymujemy rozmaitość *generowaną* przez $\mathbb{A}$. Rozmaitość jest *kongruencyjnie rozdzielna*, gdy krata kongruencji (relacji równoważności zgodnych z operacjami) dowolnej algebry z tej rozmaitości jest rozdzielna.

Twierdzenie. *Każda skończona algebra skończonej sygnatury, która generuje rozmaitość kongruencyjnie rozdzielną, ma skończoną bazę równościową.*

W literaturze można znaleźć kilka dowodów tego twierdzenia (różnych autorów).

Każda krata ma rozdzielną kratę kongruencji, zatem twierdzenie Bakera pozwala wnioskować, że każda skończona krata ma skończoną bazę równościową (choć to udowodniono wcześniej bezpośrednio).

Z kolei kraty kongruencji grup i pierścieni (tożsame w tych przypadkach odpowiednio z kratami dzielników normalnych i ideałów) są modularne, nasuwa się więc naturalne pytanie o uogólnienie pozytywnych wyników dla tych algebr na wszystkie algebry o modularnych kratach kongruencji. Najdalej idące twierdzenie R. McKenziego [8] dokłada do kongruencyjnej modularności pewien dodatkowy warunek; nie wiadomo, w jakim stopniu można jego twierdzenie osłabić.

Do twierdzeń o charakterze ogólnym dodajmy jeszcze twierdzenie J. Shapiro, odwołujące się tym razem do kraty podalgebr (stosunkowo rzadko wykorzystywanej):

Twierdzenie ([17]). *Każda skończona algebra skończonej sygnatury, która generuje rozmaitość z rozdzielnymi kratami podalgebr, ma skończoną bazę równościową.*

Problem skończonej bazy, przedstawiony tu w dużym skrócie, wciąż budzi ogromne zainteresowanie matematyków i wciąż jest źródłem wielu bardzo interesujących prac. Jak słychać, w ostatnich latach R. Willard zbudował ciąg f_n (dla $n \in \mathbb{N}$) operacji na 9-elementowym zbiorze A o następującej własności: *jeśli* $\mathbb{A}_n = (A, (f_i)_{i<n})$, *to dla każdego* $n \in \mathbb{N}$ *algebra* $\mathbb{A}_{2n}$ *ma skończoną bazę równościową, natomiast algebra* $\mathbb{A}_{2n+1}$ *takiej bazy nie ma.*

Bibliografia

[1] K. A. Baker, *Finite equational bases for finite algebras in congruence distributive varieties*, Advances in Math. **24** (1977), 207–243.

[2] K. A. Baker, G. McNulty, H. Werner, *The finitely based varieties of graph algebras*, Acta Scient. Math. (Szeged) **51** (1987), 3–15.

[3] R. Bryant, *The laws of finite pointed groups*, Bull. London Math. Soc. **14** (1982), 119–123.

[4] R. Kruse, *Identities satisfied by a finite ring*, Journ. of Algebra **26** (1973), 298–318.

[5] I. W. Lwow, *Rozmaitości pierścieni łącznych, I* (ros.), Algebra i Logika, **12** (1973), 269–297; *II*, ibidem, **12** (1973), 667–688.

[6] R. Lyndon, *Identities in two-valued calculi*, Trans. Amer. Math. Soc. **71** (1951), 457–465.

[7] R. Lyndon, *Identities in finite algebras*, Proc. Amer. Math. Soc. **5** (1954), 8–9.

[8] R. McKenzie, *Finite equational bases for congruence modular varieties*, Alg. Univ. **24** (1987), 251–266.

[9] R. McKenzie, *Tarski's finite basis problem is undecidable*, Intern. Journ. Alg. Comput. **6** (1996), 49–104.

[10] W. L. Murski, *Istnienie w logice trójwartościowej zamkniętej klasy bez skończonego pełnego układu równości* (ros.), Dokl. Akad. Nauk SSSR **163** (1965), 815–818.

[11] W. L. Murski, *Liczba k-elementowych algebr z operacją binarną, które nie mają skończonej bazy równości* (ros.), Problemy Kibernetiki **35** (1979), 5–27.

[12] B. H. Neumann, *Identical relations in groups*, Math. Ann. **114** (1937), 506–525.

[13] S. Oates, M. B. Powell, *Identical relations in finite groups*, Journ. of Algebra **1** (1964), 11–39.

[14] P. Perkins, *Bases for equational theories of semigroups*, Journ. of Algebra **11** (1968), 293–314.

[15] E. Post, *Two-valued iterative systems of mathematical logic*, Ann. Math. Studies 5, Princeton Univ. Press, Princeton N.J. 1941.

[16] C. Shallon, *Nonfinitely Based Finite Algebras Derived from Lattices*, Ph. D. dissertation, UCLA 1979.

[17] J. Shapiro, *Finite equational bases for subalgebra distributive varieties*, Alg. Univ. **24** (1987), 36–40.

[18] W. Specht, *Gesetze in Ringen, I*, Math. Zeitsch. **52** (1950), 557–589.

[19] W. W. Wiszin, *Przekształcenia identycznościowe w logice czterowartościowej* (ros.), Dokl. Akad. Nauk, **150** (1963), 719.