ROCZNIKI POLSKIEGO TOWARZYSTWA MATEMATYCZNEGO
Seria II: WIADOMOŚCI MATEMATYCZNE XXXV (1999)

Zdzisław Skupień (Kraków)

Prosty dowód twierdzenia Cantora–Bernsteina

Motto (od S. Hartmana [8]):
Nie będziemy gadać niepotrzebnych rzeczy.
S. I. Witkiewicz (*Witkacy*), *Szewcy*

Celem niniejszych uwag jest spopularyzowanie najbardziej może elementarnego dowodu twierdzenia Cantora–Bernsteina, podającego naturalny i nietrywialny warunek konieczny i wystarczający równoliczności dwóch dowolnych zbiorów A, B.

1. Twierdzenie Cantora–Bernsteina orzeka, że słaba nierówność $\leqslant$ między mocami zbiorów jest relacją antysymetryczną. Oznacza to, że dla dowolnych zbiorów A, B, jeśli $|A| \leqslant |B|$ oraz $|B| \leqslant |A|$, to $|A| = |B|$.

Tak sformułowane twierdzenie wydaje się oczywiste i jest takie, ale tylko w przypadku skończonych mocy $|A|$, $|B|$. Równoważne sformułowanie twierdzenia odwołuje się do definicji równoliczności (w terminach istnienia bijekcji $A \to B$).

Każda funkcja utożsamiana jest ze swoim wykresem. Inaczej mówiąc, funkcja f jest zbiorem par $\langle x, f(x)\rangle$, czyli strzałek od x do $f(x)$, gdzie x przebiega dziedzinę funkcji. Elementy x i $f(x)$ nazywamy odpowiednio początkiem (bełtem) i końcem (grotem) strzałki $x \mapsto f(x)$. Obraz zbioru A poprzez funkcję f oznaczamy symbolem $f[A]$.

Twierdzenie Cantora–Bernsteina. *Jeśli odwzorowania $f : A \to B$ oraz $g : B \to A$ są iniekcjami, to istnieje bijekcja $h : A \to B$.*

D o w ó d. Załóżmy, że zbiory A, B są rozłączne (można bowiem rozważać rozłączne zbiory $A \times \{\emptyset\}$, $B \times \{\{\emptyset\}\}$). Załóżmy, że $f \cup g$ jest zbiorem wszystkich strzałek między elementami zbioru $A \cup B$. Mówimy, że elementy x i y są *połączalne*, jeśli $x = y$ lub istnieje ciąg strzałek prowadzących od jednego elementu do drugiego. „Konstruujemy” bijekcję h. Podzielmy $A \cup B$ na trzy podzbiory: $A_A \cup B_A$ składa się z wszystkich elementów połączalnych z $A \setminus g[B]$, $A_B \cup B_B$ – połączalnych z $B \setminus f[A]$, zaś $A_C \cup B_C$ – to pozostała część zbioru $A \cup B$. Oznaczenia są takie, że $A = A_A \cup A_B \cup A_C$ oraz $B = B_A \cup B_B \cup B_C$. Rozważane trzy podzbiory są parami rozłączne, bo każdy element jest końcem co najwyżej jednej strzałki. Nadto $f[A_A] = B_A$, $g^{-1}[A_B] = B_B$ oraz $g^{-1}[A_C] = B_C = f[A_C]$. Dlatego bijekcja h może być następującą unią zacieśnień odwzorowań f i g^{-1}:

$$h = f|(A_A \cup A_C) \cup g^{-1}|A_B. \qquad \square$$

Konstrukcja w dowodzie ma oczywiście charakter egzystencjalny.

Uwaga 1. Konstrukcję bijekcji h w powyższym dowodzie można zmodyfikować, jeśli podzbiór $A_C \neq \emptyset$. Mianowicie relacja „być połączalnym" jest relacją równoważności na zbiorze $A \cup B$, ponieważ każdy element jest początkiem dokładnie jednej, końcem zaś co najwyżej jednej strzałki. *Klasami połączalnych* nazwijmy odpowiadające klasy równoważności. Możemy więc przyjąć, że bijekcja h – to f na A_A, g^{-1} na A_B oraz albo f albo g^{-1} na każdym niepustym przecięciu podzbioru A_C z pojedynczymi klasami połączalnych. Najprostszą modyfikację bijekcji otrzymujemy przyjmując

$$h = f|A_A \cup g^{-1}|(A_B \cup A_C).$$

2. Impulsem do rozważań jest najnowszy artykuł Mioduszewskiego [17] przedstawiający wersję dowodu w terminach iteracji i orbit odwzorowania zbioru w siebie. Inny powód, to interesująca książka [11], w której to twierdzenie jest nazywane twierdzeniem Cantora–Schrödera–Bernsteina. Autorzy tej książki twierdzą, że 45% autorów używa nazwy „Twierdzenie Cantora–Bernsteina", a drugie 45% – „Twierdzenie Schrödera–Bernsteina". Nie ujawniają, co czyni 10% (ale sami do tej grupy należą). Z kolei autorzy artykułu [16] należą do pierwszej z tych grup, ale uzasadniają możliwość używania nazwy „Twierdzenie Cantora–Dedekinda–Bernsteina". Wszystkie cztery osoby wymienione w tych nazwach, to matematycy niemieccy.

Sformułowania twierdzenia opublikował Georg Cantor [3, 4] w latach 1883 i 1895, poprawny dowód znalazł dziewiętnastoletni wówczas Felix Bernstein i przedstawił na seminarium Cantora w Halle wiosną r. 1897, po czym dowód ten opublikował E. Borel w swojej książce [2] w 1898 r. O historii tego twierdzenia, o różnych jego dowodach i o początkach teorii mnogości można poczytać w artykule Mańki i Wojciechowskiej [16], p. też Cantor [5]. Dotychczas opublikowane dowody albo bywają dość skomplikowane albo są wadliwe lub wręcz błędne. Przykładem błędnego dowodu jest dowód E. Schrödera [19], opublikowany w 1898 r. Błąd poprawiony jest przez Korselta [13]. Poprawne dowody w pierwszej dekadzie XX wieku opublikowali J. König [15], Peano [18] i Zermelo [21]. Dowody Bernsteina i Zermela przytoczone są w monografii Hausdorffa [9]. Swój oryginalny dowód przedstawił Dedekind jedynie w liście do Cantora w sierpniu 1899 r., por. publikacje w [6] oraz [5, str. 449]. W komentarzu Zermelo stwierdza, por. [5, str. 451], że jego dowód jedynie nieistotnie różni się od Dedekindowego.

Mniej lub bardziej oryginalne wersje dowodów można znaleźć w podręcznikach teorii mnogości, topologii lub analizy. Współcześnie bywa też, że dowód jest pomijany nawet w podręcznikach dla studentów matematyki (jako „stosunkowo trudny", por. [20]).

3. W przytoczonym wyżej dowodzie można rozważać parę zbiorów $D := \langle A \cup B, f \cup g\rangle$ i można ją nazwać *digrafem.* Wtedy elementy zbioru $A \cup B$ nazywane są wierzchołkami digrafu D. Zauważmy, że każdy wierzchołek digrafu D jest początkiem jednej strzałki i końcem co najwyżej jednej strzałki. Zatem digraf D jest rozłączną unią swoich składowych spójności, z których każda jest ścieżką (nieprzedłużalną i nieskończoną) lub konturem (czyli skończoną ścieżką zamkniętą). Właśnie każda klasa połączalnych, zdefiniowana wyżej w Uwadze 1, jest zbiorem wierzchołków jednej składowej digrafu. W szczególności zbiory $A_A \cup B_A$ i $A_B \cup B_B$ są pokryte przez składowe mające początki, przy czym początki te są odpowiednio w $A \setminus g[B]$ i $B \setminus f[A]$. Zbiór $A_C \cup B_C$ zaś jest pokryty przez składowe będące konturami lub obustronnie nieskończonymi ścieżkami.

Autorowi wiadomo, że od pewnego czasu popularyzowany jest grafowy dowód matematyka węgierskiego D. Königa [14] wykorzystujący nieskierowane pokojarzenia, por. Horák [10]. Warto zaznaczyć, że książka [14] jest chronologicznie pierwszą monografią teorii grafów, jej autor zaś jest synem autora pracy [15].

4. Szczególnie proste dowody twierdzenia Cantora–Bernsteina oparte są na porządku poprzedzania (lub następowania) zdefiniowanym w rozłącznej unii $A \cup B$, por. Kołmogorow i Fomin [12] oraz Halmos [7]. Dowód w [12] nie jest jednak poprawny, bo zastosowana definicja poprzedzania jest zbyt restrykcyjna. Halmos (który użył nazwy twierdzenie Schrödera–Bernsteina) przedstawia dowód na prawie stronę bardzo zbliżony do podanego wyżej (ale dłuższy). W szczególności rozważa identyczne jak wyżej podziały dwóch danych zbiorów na trzy podzbiory. Ponieważ jednak nie definiuje *explicite* równoważnościowej relacji takiej jak połączalność, więc nie uzyskuje subtelniejszego opisu struktury wytworzonej przez dane odwzorowania f i g.

5. Również Banach jest autorem wersji dowodu twierdzenia Cantora–Bernsteina. Następujący wynik, będący uogólnieniem idei Banacha, jest prostym wnioskiem z rozważań przytoczonych wyżej (a także przez Halmosa).

Twierdzenie 1. *Przy założeniach powyższego twierdzenia Cantora–Bernsteina każdy ze zbiorów A i B można przedstawić jako unię*

$$A = A_1 \cup A_2 \cup A_3, \qquad B = B_1 \cup B_2 \cup B_3$$

trzech rozłącznych podzbiorów takich, że $f[A_1] = B_1$, $g[B_2] = A_2$ *oraz* $f[A_3] = B_3$ *i* $g[B_3] = A_3$. *Nadto podzbiory* A_1, A_2 *oraz* B_1, B_2 *są niepuste, jeśli niepuste są oba podzbiory* $A \setminus g[B]$ *i* $B \setminus f[A]$.

Możliwie najmniejsze podzbiory A_1 i A_2, to odpowiednio A_A i A_B. □

Wniosek (Banach [1]). *Oba zbiory* A, B *są rozłącznymi uniami*

$$A = P \cup Q \quad \textit{oraz} \quad B = P' \cup Q'$$

odpowiednio podzbiorów P, Q *oraz* P', Q' *takich, że* $f[P] = P'$ *i* $g[Q'] = Q$.

Dowód. Można przyjąć, że P i Q są maksymalnymi zbiorami, na których ustalona bijekcja h pokrywa się odpowiednio z f i g^{-1}. Wtedy P' i Q' są obrazami P i Q poprzez odpowiednio f i g^{-1}. □

Uwaga 2. Autor przedstawił dowód na konferencji Seventh Workshop GRAPHS '3 in 1' w Krynicy 27 listopada 1998 roku, w dniu swoich urodzin i w przeddzień imienin.

Cytowane prace

[1] Stefan Banach, *Un théorème sur les transformations biunivoques*, Fund. Math. 6 (1924), 236–239.
[2] Emil Borel, *Leçons sur la théorie des fonctions*, Gauthier-Villars, Paris, 1898.
[3] G. Cantor, *Grundlagen einer allgemeinen Mannigfaltigkeitslehre*, Leipzig, 1883.
[4] G. Cantor, *Beiträge zur Begründung der transfiniten Mengenlehre*, Math. Ann. 46 (1895), 481–512 (wyd. ang.: *Contributions to the Founding of the Theory of Transfinite Numbers*, Dover, New York, 1915).
[5] Georg Cantor, *Gesammelte Abhandlungen*, red. E. Zermelo, Springer, Berlin, 1932.
[6] Richard Dedekind, *Ähnliche (deutliche) Abbildung und ähnliche Systeme*, w: *Gesammelte Math. Werke*, vol. III, Braunschweig, 1932, 447–448.
[7] P. R. Halmos, *Naive Set Theory*, Van Nostrand Reinhold Co., New York *et al.*, 1960.
[8] S. Hartman, *Wstęp do analizy harmonicznej*, PWN, Warszawa, 1969.
[9] Felix Hausdorff, *Grundzüge der Mengenlehre*, Veit u. Co., Leipzig, 1914.
[10] P. Horák, Referat na seminarium w Instytucie Matematyki AGH, Kraków, styczeń 1995.
[11] W. Just, M. Weese, *Discovering Modern Set Theory. I: The Basics*, Amer. Math. Soc., 1996.
[12] A. N. Kołmogorow, S. W. Fomin, *Elemienty tieorii funkcyj i funkcjonalnogo analiza*, Nauka, Moskwa 1972 (wyd. I: 1954).
[13] A. Korselt, *Über einen Beweis des Äquivalenzsatzes*, Math. Ann. 70 (1911), 294–296.
[14] Dénes König, *Theorie der endlichen und unendlichen Graphen*, Akad. Verlag., Leipzig, 1936 (przedruk: Teubner Verlag., Leipzig, 1986).
[15] Julius König, *Sur la théorie des ensembles*, C. R. Acad. Sci. Paris 143 (1906), 110–112.

[16] R. M a ń k a, A. W o j c i e c h o w s k a, *O dwóch twierdzeniach Cantora*, Wiadom. Mat. 25 (1984), 191–198.

[17] J. M i o d u s z e w s k i, *Twierdzenie Cantora–Bernsteina – znany dowód zapisany inaczej*, Matematyka 4'98 (272), rok 51 (1998), 207–211.

[18] Giuseppe P e a n o, *Super theorema de Cantor–Bernstein*, Rend. Circ. Mat. Palermo 21 (1906), 360–366.

[19] Ernst S c h r ö d e r, *Über zwei Definitionen der Endlichkeit und G. Cantorsche Sätze*, Nova Acta Leop. 71 (1898), 303–366.

[20] J. S ł u p e c k i, K. H a ł k o w s k a, K. P i r ó g - R z e p e c k a, *Logika i teoria mnogości*, Wyd. Nauk. PWN, Warszawa 1994, 178.

[21] Ernst Z e r m e l o, *Untersuchungen über die Grundlagen der Mengenlehre I*, Math. Ann. 65 (1908), 261–281.