Jest to rozszerzony zapis odczytu Zofii Miechowicz na XXXVI Szkole Marematyki Poglądowej, Grzegorzewice, styczeń 2006.

Kilka pomysłów na PCK

Tomasz BARTNICKI, Zofia MIECHOWICZ, Zielona Góra

Pomysł czy komputer?

Jeden z najwybitniejszych umysłów matematycznych XX wieku Paul Erdös często mawiał o Księdze, w której Bóg gromadzi doskonałe dowody twierdzeń matematycznych. Dodawał też, że nikt nie musi wierzyć w istnienie Boga, ale każdy, kto jest matematykiem, powinien wierzyć w istnienie Księgi. Wydaje się, że większość matematyków potrafi docenić piękno ukryte w najcelniejszych matematycznych dowodach, choć, z pewnością, opinie, co do tego, który dowód nadaje się do Księgi mogą się różnić. Jednak, co do pewnego dowodu, opinia wszystkich matematyków jest jednomyślna, niezależnie od tego, jaką dziedziną się zajmują i jak rozumieją pojęcie piękna. Jest nim dowód Problemu Czterech Kolorów (czyli tytułowego PCK), który, z pewnością, nie jest dowodem z Księgi. Nikt nie powie o nim, że jest elegancki, bo trudno nazwać eleganckim dowód w przeważającej części oparty na komputerowych obliczeniach, których, na dobrą sprawę, nie jesteśmy nawet w stanie sprawdzić. Trudno się też dziwić, że matematycy cały czas prześcigają się w pomysłach na znalezienie innych dróg rozwiązania tego problemu, takich, których uwieńczeniem będzie dowód nadający się do Księgi. Kilka takich pomysłów przedstawimy w niniejszym artykule.

Sam problem liczy sobie już ponad 150 lat. W roku 1852 Francis Guthrie chciał pokolorować mapę Anglii w taki sposób, żeby sąsiednie hrabstwa były rozróżnialne, tzn. żeby żadne dwa graniczące ze sobą nie miały tego samego koloru (takie kolorowanie nazywać będziemy poprawnym). Udało mu się zrobić to przy użyciu czterech barw, wysnuł więc przypuszczenie, że być może taka właśnie ich liczba wystarczy dla dowolnej płaskiej mapy. Przez jakiś czas problem pozostawał nietknięty, aż do pojawienia się pierwszego „dowodu”. W roku 1879 Alfred Bray Kempe używając prostej indukcji ze względu na liczbę obszarów udowodnił hipotezę Guthrie. Oparł się w swoim dowodzie na fakcie, wynikającym wprost ze wzoru wielościanowego Eulera, że każda mapa musi zawierać państwo, które ma co najwyżej pięciu sąsiadów. Przez 11 lat uznawano problem za rozwiązany, a dowód Kempego za poprawny. Aż w roku 1890 Percy John Heawood znalazł mapę, na której rozumowanie Kempego nie dawało pożądanych rezultatów. Dowód Kempego okazał się fałszywy, ale nie całkiem bezużyteczny, gdyż po pewnych modyfikacjach dokonanych przez Heawooda okazał się być poprawny dla nieco słabszego, ale też niebanalnego, twierdzenia o pięciu kolorach. Wtedy dopiero świat matematyczny zaczął się baczniej przyglądać hipotezie o czterech kolorach.

Historia ataków na problem jest tak bogata, że wystarczyłaby do napisania niejednego artykułu, więc przytoczymy tylko niektóre wyniki. W roku 1922 Franklin udowodnił, że każdą mapę składającą się z co najwyżej 25 obszarów można pokolorować poprawnie czterema kolorami. Po upublicznieniu tego rezultatu zaczął się prawdziwy rachunkowy wyścig. 4 lata później Reynolds pokazał że mapy co najwyżej 28-obszarowe są czterokolorowalne, w 1938 Franklin poprawił ten rezultat na 32, w 1940 Winnowi udało się udowodnić hipotezę dla map co najwyżej 36-obszarowych. Potem podniesiono tę granicę do 52 i wreszcie 92. Na tym wyścig się skończył, bo w roku 1976 Kenneth Appel i Wolfgang Haken podali pełny dowód. Jak już wspominaliśmy, nie był to dowód elegancki. Opierał się na komputerowej analizie 1936 konfiguracji. Wprawdzie w najnowszym dowodzie Robertsona, Sandersa i Seymoura z 1997 roku mamy ich już tylko 600, ale w dalszym ciągu nie obejdzie się bez pomocy komputera.

W innym języku

Problem czterech barw można równoważnie sformułować w bardziej formalnym języku, a najodpowiedniejszy jest niewątpliwie język teorii grafów. Jeżeli w środku każdego państwa umieścimy punkt, a punkty, które leżą w państwach sąsiednich połączymy liniami, dostaniemy graf, którego wierzchołkami będą owe punkty, a krawędziami linie je łączące. Będzie to pewien szczególny graf zwany dualnym do danej mapy. Ze sposobu, w jaki go tworzymy, widzimy, że będzie on narysowany na płaszczyźnie tak, że żadne jego dwie krawędzie nie będą się przecinały. Grafy, które dają się przedstawić w ten sposób nazywamy planarnymi. Istnieje wzajemna odpowiedniość między takimi grafami, a mapami na płaszczyźnie. Każdemu grafowi planarnemu odpowiada pewna mapa, a z każdej mapy można zrobić graf. Kolorowanie mapy będzie się sprowadzało do poprawnego kolorowania wierzchołków grafu. Mamy więc następującą równoważność:

Cztery kolory wystarczają do pokolorowania dowolnej mapy na płaszczyźnie.

$$\Leftrightarrow$$

Każdy graf planarny jest czterokolorowalny.

Język teorii grafów nie jest jedynym, w jakim da się ten problem opowiedzieć, co postaramy się pokazać w dalszej części artykułu.

Pomysł pierwszy

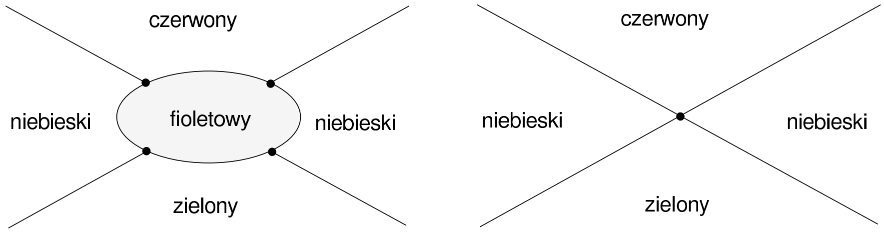
Pierwszym (po dowodzie Kempego) pomysłem na PCK było przeformułowanie problemu kolorowania mapy na problem poprawnego kolorowania krawędzi grafów sześciennych (zwanych czasem kubicznymi lub 3-regularnymi), czyli takich, w których każdy wierzchołek jest stopnia 3. Zauważmy, że rysunek mapy na płaszczyźnie możemy traktować jak płaską reprezentację grafu planarnego. Wierzchołki umieszczamy tym razem w punktach, w których spotykają się granice trzech lub więcej państw, zaś krawędziami grafu będą linie graniczne pomiędzy tymi punktami. Badając kolorowania ścian (regionów) grafów planarnych możemy ograniczyć się wyłącznie do tych, które faktycznie wyglądają jak używane w rzeczywistości mapy, co, jak się później okaże, nie wpłynie na ogólność naszych rozważań.

Zatem mamy pokolorować ściany danej płaskiej reprezentacji (tzn. takiej, że krawędzie nie przecinają się nigdzie poza wierzchołkami) grafu planarnego G, który posiada następujące własności:

(a) jest spójny,
(b) nie zawiera mostu (czyli krawędzi, której usunięcie powoduje rozspojenie grafu),
(c) jest grafem sześciennym (tzn. w każdym wierzchołku spotykają się dokładnie trzy krawędzie).

Pierwsze dwie własności są dość oczywiste dla map na płaszczyźnie. Gdyby graf nie był spójny, to kolorowanie jego różnych składowych można by traktować jak kolorowanie oddzielnych wysp na mapie. Natomiast nieistnienie mostu gwarantuje nam, że każda krawędź jest granicą dokładnie dwóch różnych regionów, gdyż w przeciwnym razie istniałoby państwo, które graniczy samo ze sobą.
Trzecia własność wymaga dokładniejszego wyjaśnienia. Jest bardzo mała szansa, że gdy weźmiemy dowolną prawdziwą mapę geograficzną, to napotkamy w niej wierzchołek stopnia większego niż 3. Oznaczałoby to, iż w jednym punkcie zbiegają się granice co najmniej 4 regionów (państw, hrabstw, województw, stanów itp.). Taki rzadki przypadek zachodzi na mapie Stanów Zjednoczonych, gdzie w jednym punkcie krzyżują się granice czterech stanów: Arizona, Utah, Kolorado i Nowy Meksyk, a miejsce to dla podkreślenia jego wyjątkowości trafnie nazwano Four Corners (Cztery Narożniki). Istnieją jednak czysto matematyczne przesłanki, które pozwolą nam ograniczyć się do rozważania jedynie map sześciennych. Załóżmy bowiem, że znamy metodę kolorowania wszystkich sześciennych map czterema kolorami, a trafiła nam się mapa, w której występują

wierzchołki stopni większych niż 3. Możemy wówczas każdy z takich wierzchołków nieco „rozdmuchać” i zastąpić go nowym fikcyjnym regionem graniczącym z tymi regionami, których granice uprzednio się w nim spotykały. Otrzymamy w ten sposób nieco bardziej skomplikowaną mapę, lecz będzie ona sześcienna, a więc możliwa do pokolorowania czterema kolorami. Łatwo już teraz zauważyć, że przywrócenie wyjściowej mapy poprzez ponowne ściągnięcie nowych regionów do pojedynczego wierzchołka nie popsuje nam poprawności pokolorowania (rysunek 1). Możemy wreszcie sformułować twierdzenie, które jest



Rys. 1. Ściągnięcie rozdmuchanego wierzchołka

historycznie pierwszym pomysłem na wyrażenie PCK w nieco innym języku. Pochodzi ono od Petera Guthrie Taita, który jednak nie potrafił podać jego formalnego dowodu (znalezionego o wiele później). Obecnie twierdzenie to należy już do klasyki teorii grafów i odgrywa w niej dużą rolę, gdyż większość późniejszych równoważnych sformułowań PCK była nim właśnie motywowana.

Twierdzenie 1. *Mapa sześcienna może być pokolorowana czterema kolorami wtedy i tylko wtedy, gdy krawędzie odpowiadającego jej grafu mogą być pokolorowane trzema kolorami.*

Dowód. ($\Rightarrow$) Załóżmy, że sześcienna mapa, której odpowiada graf G została pokolorowana czterema kolorami, które oznaczmy symbolami 00, 01, 10 i 11. Pokażemy teraz jak z takiego kolorowania otrzymać kolorowanie krawędzi trzema kolorami 01, 10 i 11. Kolor krawędzi jest wyznaczany jednoznacznie przez kolory regionów, które ta krawędź rozdziela, według reguły dodawania po współrzędnych modulo 2, co daje 6 następujących możliwości 00+01=01, 00+10=10, 00+11=11, 10+01=11, 11+01=10, 11+10=01. Widać natychmiast, że wśród kolorów krawędzi nie pojawi się kolor 00, gdyż oznaczałoby to, że regiony po obu stronach tej krawędzi są identyczne. Niewiele trudniej zauważyć (wystarczy rozpatrzyć cztery różne przypadki), że kolory krawędzi (nadane zgodnie z powyższą regułą), które spotykają się w jednym wierzchołku muszą być wszystkie różne. Inne uzasadnienie poprawności kolorowania krawędzi wynika z faktu, że gdy sporządzimy pełną tabelkę działań na zbiorze kolorów {00,01,10,11} to okaże się ona izomorficzna z tabelką działań grupy czwórkowej Kleina, a jej niektóre własności gwarantują ową poprawność.

($\Leftarrow$) Na odwrót załóżmy, że mamy zadane pokolorowanie krawędzi grafu G trzema kolorami 01, 10 i 11 takie, że każdy z kolorów występuje dokładnie raz przy każdym wierzchołku. Rozpatrzmy podgraf grafu G otrzymany przez usunięcie z niego wszystkich krawędzi w kolorze 01. Otrzymamy graf, którego każdy wierzchołek jest stopnia 2, a więc jest sumą rozłącznych cykli. Mimo że cykle te są krawędziowo i wierzchołkowo rozłączne to ich położenie na płaszczyźnie może być bardzo różne na przykład jedne mogą leżeć wewnątrz lub na zewnątrz innych. Zbadajmy teraz położenie każdego regionu wyjściowej mapy względem wszystkich tych cykli i nadajmy na pierwszej współrzędnej wartość 0 tym regionom, które leżą wewnątrz parzystej liczby cykli, zaś 1 tym, które leżą wewnątrz nieparzystej liczby cykli. W podobny sposób ustalmy wartość na drugiej współrzędnej z tym, że teraz usuwamy z grafu wszystkie krawędzie w kolorze 10. W konsekwencji takiego postępowania każdy region otrzyma jeden z wyników ze zbioru $\{00, 01, 10, 11\}$, który możemy potraktować jako zbiór czterech różnych kolorów. Wystarczy jedynie pokazać, że dla każdej krawędzi e dwa regiony, dla których jest ona wspólną granicą, otrzymały różne kolory. Wynika to wprost ze sposobu nadawania wartości regionom na poszczególnych współrzędnych. Jeżeli bowiem krawędź e ma kolor 10, to regiony, które leżą po jej przeciwnych stronach musiały dostać różne wartości, gdy rozpatrywaliśmy cykle na krawędziach w kolorach 10 i 11 (usunęliśmy wtedy z grafu krawędzie w kolorze 01), a zatem różnią się one na pierwszej współrzędnej. Jeżeli krawędź e ma kolor 01, to z podobnego powodu regiony, które rozdziela muszą się różnić na drugiej współrzędnej. Jeśli wreszcie ma ona kolor 11, to jej przyległe regiony różnią się na obu współrzędnych, gdyż krawędzie w kolorze 11 pozostawały w grafie przy określaniu wartości zarówno pierwszej jak i drugiej współrzędnej poszczególnych regionów.

Zatem we wszystkich przypadkach dowolna krawędź rozgranicza dwa regiony o różnych kolorach, co kończy dowód. □

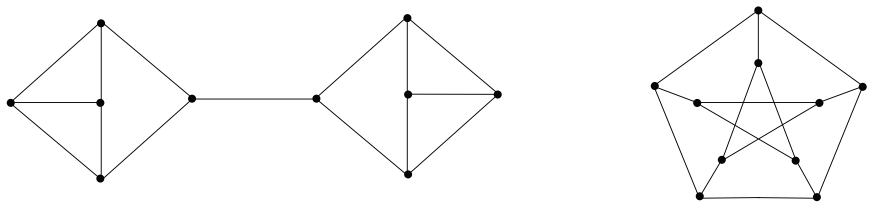
Uwzględniając uwagi poczynione wcześniej, że kolorowanie dowolnych map można sprowadzić do kolorowania map sześciennych oraz, że w ogólności kolorowanie regionów na mapach jest równoważne kolorowaniu wierzchołków grafów planarnych możemy sformułować dwa równoważne twierdzenia.

Wierzchołki dowolnego grafu planarnego można pokolorować przy użyciu 4 kolorów.

$\Leftrightarrow$

Krawędzie dowolnego sześciennego i bezmostowego grafu planarnego można pokolorować przy użyciu 3 kolorów

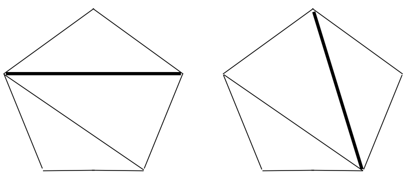
Na koniec warto nadmienić, że w drugim twierdzeniu zarówno założenie o planarności jak i nieistnieniu mostu jest istotne. Na rysunku 2 mamy przykłady dwóch grafów sześciennych, z których pierwszy zawiera most, zaś drugi (zwany grafem Petersena) nie jest planarny. Można z łatwością sprawdzić, że poprawne pokolorowanie krawędzi każdego z nich wymaga czterech kolorów.



Rys. 2. Złe grafy sześcienne

Pomysł drugi

Oderwiemy się teraz na chwilę od problemów kolorowania map i grafów i przedstawimy pewien ciekawy problem kombinatoryczno-geometryczny. Załóżmy, że mamy dany dowolny wielokąt wypukły. Dokonujemy jego triangulacji przez dorysowanie maksymalnej liczby nieprzecinających się przekątnych. Oczywiście im większa liczba boków wielokąta tym większa jest liczba sposobów dokonania takiej triangulacji (wyraża się ona liczbą Catalana). Mając ustaloną triangulację możemy z niej otrzymać inną dokonując operacji, którą nazywać będziemy *flipem*. Pojedynczy flip wykonujemy w ten sposób, że w zadanej triangulacji wybieramy dwa trójkąty, które mają wspólny bok (a zatem tworzą one czworokąt z jedną przekątną), a następnie w tym czworokącie usuwamy istniejącą przekątną i zastępujemy ją drugą (rysunek 3). Można spostrzec, że wszystkie triangulacje zadanego wielokąta są (w sensie operacji flipowania) równoważne, czyli prawdziwe jest:

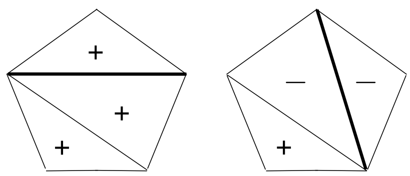


Rys. 3. Pojedynczy flip

Twierdzenie 2. *Dla dowolnych dwóch triangulacji tego samego wielokąta wypukłego istnieje skończona sekwencja flipów przeprowadzająca jedną triangulację w drugą.*

Dowód. (Indukcja ze względu na liczbę boków wielokąta)

Dla najmniejszego przypadku, czyli czworokąta, istnieją tylko dwie możliwe triangulacje, które można zamieniać jedna w drugą przy użyciu pojedynczego flipu.
Załóżmy, że teza jest prawdziwa dla wszystkich par striangulowanych wielokątów o liczbie boków niewiększej niż n. Aby pokazać, że dla dwóch dowolnych triangulacji T i T' $n+1$-kąta istnieje odpowiednia sekwencja flipów przeprowadzająca T w T', zauważmy wpierw, że każda triangulacja musi zawierać trójkąt, którego dokładnie dwa boki są jednocześnie bokami wielokąta, a co za tym idzie – z ich wspólnego wierzchołka nie wychodzą żadne przekątne. Oznaczmy przez v' taki wierzchołek w triangulacji T', odpowiadający mu zaś wierzchołek w T przez v. Jeżeli z wierzchołka $v \in T$ nie wychodzą żadne przekątne, to po usunięciu wierzchołków v i v' wraz z przyległymi bokami otrzymamy dwa nowe striangulowane wielokąty wypukłe o n bokach, a zgodnie z założeniem indukcyjnym istnieje odpowiednia sekwencja flipów. Wystarczy tylko zauważyć, że owa sekwencja przeprowadza również T na T'. W przypadku, gdy z wierzchołka $v \in T$ wychodzą jakieś przekątne, możemy sekwencję flipów zacząć od tych właśnie przekątnych, co w rezultacie doprowadzi do poprzedniej sytuacji. $\square$



Rys. 4. Pojedynczy oznakowany flip

Załóżmy następnie, że każdy z trójkątów w triangulacji wielokąta wypukłego ma przypisany znak „+” lub „−”. Dla tak oznakowanej triangulacji zdefiniujmy operację *oznakowanego flipu* w podobny, jak powyżej sposób, z jednym zastrzeżeniem, że jest ona dopuszczalna tylko dla tych par przyległych trójkątów, które posiadają te same znaki. Ponadto każdy taki flip zamienia znaki na przeciwne w nowej parze trójkątów (rysunek 4). Załóżmy, że mamy dwie różne triangulacje T i T' tego samego wielokąta i pierwsza z nich jest w jakiś sposób oznakowana. Nietrudno zauważyć, że nie dla każdego takiego oznakowania znajdziemy sekwencję flipów przeprowadzającą T w T', a dla niektórych oznakowań wręcz niemożliwe będzie wykonanie jakiegokolwiek flipu. Możemy jednak zadać pytanie, czy istnieje przynajmniej jedno oznakowanie T, które umożliwi nam przeprowadzenie T w T' w skończonej liczbie oznakowanych flipów. Ręczna analiza niewielkich przypadków lub wspomagana komputerowo analiza nieco większych pozwala przypuszczać, że jest to zawsze możliwe. Patrząc na prostotę dowodu poprzedniego twierdzenia wydawać by się mogło, że udowodnienie podobnego w wersji ze znakami będzie równie łatwe i eleganckie. Niestety, jak do tej pory żaden dowód (nawet trudny i nieelegancki) tej hipotezy nie jest znany. Co więcej w 1999 roku Shalom Eliahou pokazał [4] zaskakującą implikację.

Dla dowolnych dwóch triangulacji tego samego wielokąta wypukłego istnieje takie oznakowanie pierwszej z nich, że możliwe jest przeprowadzenie jej w drugą po skończonej sekwencji oznakowanych flipów.

$$\Rightarrow$$

Wierzchołki dowolnego grafu planarnego można pokolorować przy użyciu 4 kolorów.

Pomysł trzeci

Ostatni pomysł, który przedstawimy zawdzięczamy Hasslerowi Whitneyowi. Opowiedział on znany nam już dobrze problem w zupełnie innym języku. O dziwo, do opowiedzenia PCK nie trzeba używać żadnych rysunków, grafów czy geometrii. Da się to zrobić w czysto algebraiczny sposób.

Rozważmy sumę n symboli $a_1 + a_2 + \cdots + a_n$ o wartościach całkowitych. Ponieważ dodawanie jest łączne, więc wstawienie pewniej liczby nawiasów do takiej sumy nie zmieni wyniku końcowego. Może się jednak zdarzyć, że dla dwóch różnych układów nawiasów otrzymywać będziemy różne sumy częściowe. Załóżmy, ponadto, że pośród n symboli mamy umieścić dokładnie $n-1$ par nawiasów, i to w taki sposób, żeby na każdym etapie kolejność dodawania była określona jednoznacznie. Dla czterech symboli dopuszczalne są na przykład takie układy nawiasów

$$((a_1 + a_2) + (a_3 + a_4)), \quad (((a_1 + a_2) + a_3) + a_4),$$

ale już układy $(((a_1 + a_2 + a_3)) + a_4)$ lub $(((a_1 + a_2 + a_3 + a_4)))$ są niedozwolone.

Hassler Whitney wykazał równoważność:

Dla każdych dwóch układów nawiasów na sumie n symboli $a_1 + \cdots + a_n$ możemy przypisać zmiennym a_i różne od zera wartości z grupy $\mathbb{Z}_4$ tak, aby wszystkie częściowe sumy obu układów były niezerowe.

$$\Leftrightarrow$$

Wierzchołki dowolnego grafu planarnego można pokolorować przy użyciu 4 kolorów.

Można szybko sprawdzić, że dla podanych wyżej układów nawiasów na czterech symbolach, przy podstawieniu $a_1 = 1$, $a_2 = 1$, $a_3 = 2$, $a_4 = 1$ otrzymamy 0 jako jedną z sum częściowych, natomiast przyporządkowanie
$a_1 = 1$, $a_2 = 1$, $a_3 = 1$, $a_4 = 2$ będzie dobre.

W poszukiwaniu dowodu z Księgi

Przedstawiliśmy tu zaledwie kilka pomysłów na PCK, ale możemy zapewnić, że jest ich znacznie więcej ([5]). Niestety żaden z nich nie doprowadził do znalezienia alternatywnego dowodu twierdzenia o czterech kolorach. Nasuwa się więc fundamentalne pytanie, czy w ogóle istnieje jakikolwiek dowód PCK, który mógłby się znaleźć we wspomnianej na początku Księdze. Trzeba uczciwie przyznać, że matematycy są w tej kwestii podzieleni. Niektórzy nadal szukają jakiegoś przeformułowania PCK, które następnie próbują udowodnić przeróżnymi elementarnymi metodami. Są oni jednak w znacznej mniejszości, a niestety większość – dowiadując się, że jakiś ciekawy problem jest równoważny z twierdzeniem o czterech kolorach – po prostu przestaje się nim interesować zakładając z góry, że rozwiązanie go jest sprawą beznadziejną.

Postawmy więc po raz ostatni pytanie, czy istnieje dowód PCK z Księgi? Matematycy, gdy nie mogą skonstruować jakieś obiektu o zadanych własnościach, często zadawalają się niekonstruktywnym dowodem na jego istnienie. Może więc i w tym przypadku najlepszym pomysłem będzie udowodnienie, że istnieje owa Boska Księga, a w niej przepiękny, elementarny dowód twierdzenia o czterech kolorach. Wówczas fakt, że go nigdy nie poznamy, będzie dla wielu sprawą drugorzędną.

Literatura

[1] M. Aigner, G. Ziegler, *Dowody z księgi*, Wydawnictwo Naukowe PWN, Warszawa 2002

[2] T. Bartnicki, J. Grytczuk, *O dwóch takich co kolorowali mapę*, Delta (9/2005)

[3] V. Bryant, *Aspekty kombinatoryki*, WNT, Warszawa 1997

[4] S. Eliahou, *Signed Diagonal Flips and the Four Color Theorem*, Erop. J. Combinatorics (1999) **20**, 641–647

[5] P. Kainen, T. Saaty, *The four-color problem. Assault and conquest*, McGraw-Hill International Book Co., New York-Bogota-Auckland, 1977

[6] R. Wilson, *Cztery barwy wystarczą, czyli o kolorowaniu map*, Delta (6/2004)