O kolorowaniu map

Dr Cezary BOWSZYC

Na mapie politycznej jakiegoś kontynentu obszary państw zwykle kolorowane są w ten sposób, że państwa sąsiadujące ze sobą wzdłuż pewnej linii granicznej (lub kilku takich linii) mają różne barwy. W ubiegłym wieku wyraźnie postawiono pytanie: Jaka jest najmniejsza liczba barw wystarczająca do tego, by można było pomalować nimi dowolną mapę na płaszczyźnie (z zachowaniem powyższego warunku)?

Będziemy zajmowali się jedynie takimi mapami na całej płaszczyźnie, które zawierają skończoną liczbę obszarów o granicach regularnych, złożonych ze skończonej liczby krawędzi (być może krzywych) połączonych w wierzchołkach i tworzących zamkniętą linię bez przecięć. Za obszar będziemy również uważać nieograniczoną część płaszczyzny. Np. mapa na rysunku 1 zawiera 9 wierzchołków, 14 krawędzi i 7 obszarów. Na następnych dwóch rysunkach pokazane są przykłady map nie spełniających założonych warunków: na rys. 2 jeden z obszarów sąsiaduje ze sobą wzdłuż krawędzi, na rys. 3 brzeg jednego z obszarów składa się z dwóch linii zamkniętych. Zbiór wierzchołków i krawędzi będziemy nazywać grafem mapy. Przy poczynionych założeniach graf mapy ma tę własność, że można przejść z jednego punktu grafu do dowolnego innego punktu grafu nie wychodząc z niego. Mówimy w tej sytuacji, że graf jest spójny.

Oznaczmy przez α_0 liczbę wierzchołków, przez α_1 liczbę krawędzi, a przez α_2 liczbę obszarów rozważanej mapy. W przykładzie z rys. 1 jest $\alpha_0 - \alpha_1 + \alpha_2 = 2$. Okazuje się, że dla dowolnej mapy płaskiej o spójnym grafie zachodzi wzór Eulera:

$$\alpha_0 - \alpha_1 + \alpha_2 = 2.$$

(Można sprawdzić, że wzór ten zachodzi dla mapy z rys. 2, ale nie zachodzi dla mapy z rys. 3, bo wartość lewej strony wzoru równa się 3.)

Wykażemy przez indukcję względem liczby krawędzi α_1, że ogólniej, gdy graf mapy płaskiej rozpada się na s składowych spójnych, to

$$\alpha_0 - \alpha_1 + \alpha_2 = 1 + s.$$

Gdy $\alpha_1 = 0$, a więc gdy nie ma w ogóle krawędzi, mamy na płaszczyźnie jedynie α_0 wierzchołków tworzących graf o $s = \alpha_0$ składowych i jeden obszar: $\alpha_2 = 1$. Zatem w tym przypadku wzór zachodzi. Załóżmy, że wzór ten zachodzi dla dowolnej mapy płaskiej o liczbie krawędzi mniejszej od α_1 i rozważmy mapę o liczbie krawędzi α_1. Usuńmy jedną krawędź K z grafu. Jeżeli K leży na brzegu dwóch obszarów, to po usunięciu K obszary te połączą się, zatem α_1 i α_2 zmaleją o 1, a α_0 i s pozostaną takie same. W tym przypadku obydwie strony równości nie zmienią się. Jeżeli zaś K leży na brzegu tylko jednego obszaru (jak na rys. 2), to po usunięciu K składowa grafu zawierająca K rozpadnie się na dwie, liczba s składowych grafu wzrośnie o 1, α_1 zmaleje o 1, a α_0 i α_2 pozostaną bez zmian. W tym przypadku obydwie strony równości wzrosną o 1. Z zasady indukcji wynika więc słuszność dowodzonego wzoru, a więc również wzoru Eulera (dla $s = 1$ — por. artykuł J. A. Rempały w Delcie 1/1976).

Możemy założyć, że w każdym wierzchołku naszej mapy schodzą się co najmniej 3 krawędzie (bo gdy schodzą się tylko dwie, to można zlikwidować wierzchołek i połączyć te dwie krawędzie w jedną). Liczba par: ⟨krawędź — wierzchołek będący jej końcem⟩ jest równa $2\alpha_1 \geqslant 3\alpha_0$ na mocy powyższej uwagi. Wynika stąd, że

Mapa zawiera przynajmniej jeden obszar, którego brzeg składa się z pięciu lub mniejszej liczby krawędzi.

W przypadku przeciwnym brzeg każdego obszaru składałby się z co najmniej sześciu krawędzi i licząc pary: ⟨krawędź — obszar, do którego przylega ta krawędź⟩ uzyskalibyśmy nierówność $2\alpha_1 \geqslant 6\alpha_2$. Uwzględniając równanie Eulera i nierówności $\alpha_0 \leqslant \frac{2}{3}\alpha_1$, $\alpha_2 \leqslant \frac{1}{3}\alpha_1$ mielibyśmy $2 = \alpha_0 - \alpha_1 + \alpha_2 \leqslant \frac{2}{3}\alpha_1 - \alpha_1 + \frac{1}{3}\alpha_1 = 0$, co nie jest prawdą.

Można stąd wywnioskować przez indukcję względem liczby obszarów α_2, że

Dowolną mapę płaską można pomalować pięcioma barwami.

Załóżmy więc, że twierdzenie nasze zachodzi dla map o liczbie obszarów mniejszej niż α_2 i rozważmy mapę o α_2 obszarach. Jeżeli zawiera ona obszar O sąsiadujący z czterema (lub mniej) obszarami, to, na mocy założenia indukcyjnego, po pokolorowaniu pozostałych obszarów pozostanie nam jeszcze jedna z pięciu barw do pomalowania obszaru O. W przypadku przeciwnym istnieje obszar O, sąsiadujący z pięcioma obszarami kolejno O_1, O_2, O_3, O_4 i O_5 (rys. 4). Jeżeli O_1 nie sąsiaduje z O_3 i $O_1 \neq O_3$, to likwidując dwie krawędzie łączymy obszary O, O_1 i O_3, uzyskując mapę o mniejszej liczbie obszarów i, na mocy założenia indukcyjnego, możemy ją pokolorować. Po wprowadzeniu z powrotem usuniętych krawędzi stwierdzamy, że obszary O_1 i O_3 mają tę samą barwę. Zatem obszary sąsiadujące z O są pomalowane co najwyżej czterema barwami. Obszar O możemy więc pomalować nie wykorzystanym kolorem. Jeżeli zaś O_1 i O_3 sąsiadują lub $O_1 = O_3$, to O_2 nie może sąsiadować z O_4 i $O_2 \neq O_4$ (rys. 5). Wtedy postępujemy z obszarami O, O_2, O_4 podobnie jak poprzednio z obszarami O, O_1, O_3. Rys. 6 przedstawia mapę, w której każdy z czterech obszarów ograniczonych sąsiaduje z każdym. Takiej mapy nie można pomalować mniej niż czterema barwami, ale cztery wystarczą. Wobec tego najmniejsza liczba barw wystarczająca do pomalowania dowolnej płaskiej mapy wynosi 4 albo 5. Nasuwa się pytanie, czy wystarczą cztery barwy.

Rozważmy dowolną mapę na płaszczyźnie. Wewnątrz każdego obszaru wybierzmy punkt jako stolicę. Stolice sąsiadujących państw można połączyć krzywą przecinającą krawędź, wzdłuż której sąsiadują te obszary w ten sposób, by uzyskane krzywe nie przecinały się ze sobą poza stolicami. Otrzymany graf nazywamy grafem dualnym, a mapę przezeń wyznaczoną — mapą dualną. Rys. 7 ilustruje te pojęcia; liniami przerywanymi oznaczone są krawędzie grafu dualnego.

Gdybyśmy potrafili znaleźć mapę, w której pięć obszarów sąsiadowałoby każdy z każdym, rozwiązaniem naszego problemu byłoby: potrzeba i wystarczy użyć pięciu barw. Niestety taka mapa nie istnieje. Gdyby bowiem istniała, to w grafie dualnym znaleźlibyśmy podgraf położony na płaszczyźnie o $\alpha_0 = 5$ wierzchołkach połączonych $\alpha_1 = \binom{5}{2} = 10$ krawędziami każdy z każdym. Ze wzoru Eulera liczba obszarów mapy wyznaczonej przez ten graf wynosiłaby $\alpha_2 = 2-\alpha_0+\alpha_1 = 7$. Każdy z nich posiadałby brzeg złożony z co najmniej trzech krawędzi, więc $2\alpha_1 \geqslant 3\alpha_2$ czyli $20 \geqslant 21$, co nie jest prawdą. Na rys. 8 przedstawiono pięć punktów. Nie można ich połączyć na płaszczyźnie każdy z każdym bez dodatkowego przecięcia.

Tak więc nie mamy rozwiązania problemu. Hipoteza, że cztery barwy wystarczą do pomalowania dowolnej płaskiej mapy mimo wysiłków wybitnych umysłów do niedawna nie była poprawnie udowodniona ani obalona. Stopniowo wzrastała liczba n państw w twierdzeniach typu: *Jeżeli liczba państw α_2 jest nie większa od n, to mapę można pomalować czterema barwami.* Udowodniono rozmaite warunki równoważne hipotezie czterech barw. Badano własności ewentualnych map, których nie można pomalować czterema barwami. W ubiegłym roku w Stanach Zjednoczonych udało się rozwiązać pozytywnie problem czterech barw, ale trzeba było odwołać się aż do pomocy ... maszyn matematycznych; podobno życia ludzkiego nie starczyłoby do sprawdzenia bardzo wielu możliwości...

Idea w uproszczeniu była taka: Zamiast malować obszary mapy można malować wierzchołki grafu dualnego tak, by połączone krawędzią wierzchołki miały rozmaite kolory. Można dalej założyć, że obszary wyznaczone przez ten graf mają brzegi złożone z trzech krawędzi, wprowadzając dodatkowe krawędzie, jak to pokazane jest na rys. 9 (linie przerywane). Takie mapy krótko nazwiemy *triangulacjami*. W tej sytuacji $2\alpha_1 = 3\alpha_2$. Jeżeli p_i oznacza liczbę krawędzi schodzących się w i-tym wierzchołku dla $i = 1, 2, \ldots, \alpha_0$, to $2\alpha_1 = \sum_{i=1}^{\alpha_0} p_i$. Ze wzoru Eulera otrzymujemy $2 = \alpha_0 - \frac{1}{3}\alpha_1 = \alpha_0 - \frac{1}{6}\sum_{i=1}^{\alpha_0} p_i$. Uwzględniając, że α_0 jest sumą α_0 jedynek, otrzymujemy

$$\sum_{i=1}^{\alpha_0} (6-p_i) = 12.$$

Rozwiązanie zadania F 41. Gdybyśmy mieli tylko jedną elektrodę, z której wypływa prąd I i przyłożyli ją do punktu A, to ze względu na symetrię sieci prąd rozpływałby się jednakowo we wszystkich czterech kierunkach i ku punktowi B płynąłby prąd o natężeniu $I/4$. Podobnie, dla samej elektrody zbierającej prąd I przyłożonej do punktu B, gałęzią AB płynąłby prąd o natężeniu wynoszącym również $I/4$. Dla dwóch elektrod rozkład prądów jest sumą rozkładów od poszczególnych elektrod, a zatem przy podłączeniu naszej sieci do źródła w punktach A i B przez odcinek AB będzie płynął prąd o natężeniu $I/4+I/4 = I/2$. W związku z tym między punktami A i B będzie napięcie $U_{AB} = r \cdot I/2$. Z określenia oporu zastępczego R mamy $U_{AB} = R_{AB}I$, a zatem $R_{AB} = r/2$. Zadanie powyższe jest zadaniem bardzo starym i znanym, jednakże mimo to — ze względu na bardzo ciekawe i pouczające zastosowanie symetrii — zostało ono wykorzystane jako jedno z zadań stopnia wstępnego XXVI Olimpiady Fizycznej. Jest rzeczą interesującą, że pokrewne zadanie, gdy punkty A i B nie są węzłami najbliższymi, nie daje się łatwo rozwiązać. Muszę przyznać, że na znalezienie np. R_{AC}, gdzie C jest punktem pokazanym na rysunku w tekście zadania, poświęciłem sporo czasu, ale bez skutku. Możliwe jednak, że ktoś z Czytelników będzie miał więcej szczęścia. Czytelników, którzy znajdą wartość R_{AC}, prosiłbym o nadesłanie rozwiązania do Redakcji.
Najciekawsze rozwiązanie zostanie przedstawione w „Delcie". Termin nadsyłania rozwiązań: 31 lipca br.

Wzór ten podał A. B. Kempe już w końcu ubiegłego wieku. Wynika z niego m.in., że każda triangulacja musi zawierać wierzchołki, w których schodzi się nie więcej niż pięć krawędzi (co w dualnej formie udowodniliśmy już wcześniej). Gdyby istniały triangulacje nie dające się pomalować czterema barwami, to istniałaby wśród nich *triangulacja minimalna*, tzn. o najmniejszej ilości wierzchołków. Kempe udowodnił, że w triangulacji minimalnej nie mogą wystąpić wierzchołki, w których schodzi się mniej niż pięć krawędzi, skąd wynika, że w triangulacji minimalnej musiałoby wystąpić co najmniej 12 wierzchołków pięciokrotnych, skąd (już nie tak prosto) wynika, że w takiej triangulacji co najmniej jeden wierzchołek pięciokrotny musiałby sąsiadować bądź z pięciokrotnym, bądź z sześciokrotnym ...
Rozważania tego rodzaju doprowadziły do stworzenia pojęcia *zbioru nieuniknionego* t.j. listy konfiguracji (grafów) o następującej własności: gdyby istniała triangulacja minimalna, to musiałaby ona zawierać przynajmniej jedną konfigurację z tej listy. Znane są przykłady różnych zbiorów nieuniknionych: najprostszy z nich to lista jednoelementowa, na której występuje konfiguracja złożona z jednego wierzchołka o krotności 5 i pięciu połączonych wierzchołków sąsiednich.

Konfiguracja nazywa się *redukowalną*, jeśli nie może ona wystąpić w żadnej triangulacji minimalnej. Najprostszy przykład konfiguracji redukowalnej to — jak wynika z tw. Kempego — konfiguracja złożona z wierzchołka czterokrotnego i jego sąsiadów. Na ogół sprawdzanie redukowalności jest zadaniem bardzo praco- i czasochłonnym.

G. D. Birkhoff zauważył w r. 1913, że twierdzenie o czterech barwach byłoby udowodnione, gdyby udało się podać przykład zbioru nieuniknionego złożonego z samych konfiguracji redukowalnych: z istnienia takiej listy wynikałoby, że nie może istnieć triangulacja minimalna, a więc tym bardziej nie może istnieć jakakolwiek triangulacja nie dająca się pomalować czterema barwami.
W znalezieniu takiej listy dopomógł fakt, że wzór Kempego można interpretować w sposób ... elektryczny: z każdym wierzchołkiem triangulacji związana jest liczba $6-p_i$, którą można traktować jako ładunek elektryczny, znajdujący się w tym wierzchołku. Rozładowanie grafu polega na przemieszczaniu ładunków dodatnich pomiędzy wierzchołkami, a graf uznaje się za rozładowany, jeśli we wszystkich wierzchołkach są ładunki niedodatnie. Wzór Kempego orzeka więc, że każda triangulacja nie da się w pełni rozładować.
K. Appel i W. Haken opisali pewien algorytm D (służący do rozładowywania grafu) oraz sporządzili listę Δ (liczącą 1936 konfiguracji), o których udowodnili, że *jeśli triangulacja minimalna nie zawiera żadnej konfiguracji z listy Δ, to daje się rozladować przy pomocy algorytmu D.*

Ponieważ jednak żadna triangulacja nie daje się rozładować, to twierdzenie to orzeka po prostu, że Δ jest zbiorem nieuniknionym. Jednocześnie zaś metoda budowania konfiguracji należących do Δ pozwalała przypuszczać, że każda z tych konfiguracji jest redukowalna. Należało to sprawdzić. W zasadzie wiadomo było, jak to zrobić i rzecz sprowadzała się do przeprowadzenia ogromnie pracochłonnych, ale mechanicznych analiz. Tu właśnie dopomogły odpowiednio zaprogramowane komputery.

*

Nasza Ziemia jest kulą i można by pytać o liczbę barw dla map na powierzchni kuli czyli sferze, albo jeszcze ogólniej na dowolnej powierzchni. Rys. 10 pokazuje, że jeżeli ze sfery usuniemy jeden punkt, np. biegun północny, to resztę powierzchni kuli będziemy mogli przez rzutowanie z tego bieguna odwzorować wzajemnie jednoznacznie na płaszczyznę. Mapie na sferze będzie odpowiadała mapa na płaszczyźnie i widać, że liczba barw dla map sferycznych jest taka sama jak dla płaskich, a więc wynosi 4.

Jest rzeczą dziwną, że analogiczne zagadnienia dla innych powierzchni zostały rozwiązane wcześniej (w 1968 r. przez G. Ringela i I. W. T.Youngsa, bez uciekania się do pomocy maszyn matematycznych), podczas gdy najprostsze powierzchnie: sfera czy płaszczyzna najdłużej stawiały opór w wyjawieniu swych tajemnic. Trudności w przypadku innych powierzchni zresztą polegały na czym innym: nie na tym, by wykazać, że każdą mapę można pomalować odpowiednią ilością kolorów (znaną już P. J. Heawoodowi w końcu ubiegłego wieku), ale na tym, by na powierzchni wskazać mapę o tylu obszarach parami sąsiadujących ze sobą. Okazuje się np., że dla map na *torusie* czyli ,,dętce'' najmniejsza liczba barw wynosi 7. Rys. 11 podaje przykład mapy na torusie złożonej z 7 obszarów, z których każde dwa sąsiadują ze sobą.

Rozwiązanie zadania M 121. Możemy przyjąć, że dla pewnej liczby α zachodzą równości: $a = \dfrac{1}{\sin\alpha}$, $b = \dfrac{1}{\cos\alpha}$ (dlaczego?), gdzie $\sin\alpha\cos\alpha \neq 0$.
Zachodzi oczywista nierówność.

$$(\cos^3\alpha - \sin^3\alpha)^2 \geqslant 0,$$

z której wynika, że
$\cos^6\alpha + \sin^6\alpha \geqslant 2\cos^3\alpha\sin^3\alpha,$
$\cos^4\alpha(1-\sin^2\alpha) + \sin^4\alpha(1-\cos^2\alpha) \geqslant$
$\geqslant 2\cos^3\alpha\sin^3\alpha,$
$\cos^4\alpha + \sin^4\alpha \geqslant \cos^4\alpha\sin^2\alpha + 2\cos^3\alpha\sin^3\alpha +$
$+\sin^4\alpha\cos^2\alpha,$

$$\frac{1}{\sin^4\alpha} + \frac{1}{\cos^4\alpha} \geqslant \frac{1}{\sin^2\alpha} + \frac{2}{\cos\alpha\sin\alpha} + \frac{1}{\cos^2\alpha},$$

$$\frac{1}{\sin^4\alpha} + \frac{1}{\cos^4\alpha} \geqslant \left(\frac{1}{\sin\alpha} + \frac{1}{\cos\alpha}\right)^2,$$

$a^4 + b^4 \geqslant (a+b)^2$, c.d.n.

Rozwiązanie zadania M 123. Zauważmy, że liczba -1 jest pierwiastkiem tego równania. Zachodzi równość
$ax^5 + bx^4 + cx^3 + cx^2 + bx + a =$
$= a(x+1)(x^4 + Ax^3 + Bx^2 + Ax + 1),$
gdzie $A = \dfrac{b-a}{a}$, $B = \dfrac{c-b+a}{a}$.
Musimy więc rozwiązać równanie

$$x^4 + Ax^3 + Bx^2 + Ax + 1 = 0.$$

Równanie to jest równoważne równaniu

$$x^2 + \frac{1}{x^2} + A\left(x + \frac{1}{x}\right) + B = 0,$$

gdyż zero nie jest jego pierwiastkiem.
Podstawmy $x + \dfrac{1}{x} = y$. Jest wówczas
$x^2 + 2 + \dfrac{1}{x^2} = y^2,$
skąd $y^2 - 2 + Ay + B = 0$. Do rozwiązania tego równania wystarczają cztery działania arytmetyczne i wyciąganie pierwiastków kwadratowych. Podobnie jest z równaniem $x + \dfrac{1}{x} = y$, równoważnym równaniu $x^2 - yx + 1 = 0$, co kończy dowód.