

# deltam

MATEMATYKA – FIZYKA – ASTRONOMIA – INFORMATYKA

$h=6,63 \cdot 10^{-34} \cdot \nu$   $\lambda = 3,141592 \dots$

NR 11 (522) 2017

CENA 4 ZŁ VAT 5%

PL ISSN 0137-3005 | NR IND 35-550 X  
MIESIĘCZNIK

[www.deltami.edu.pl](http://www.deltami.edu.pl)

**Ciemna  
materia  
w kolorze**


**Patrz str. 7**



Nakład: 3500 egz.

Gra w kości atomami  
*Krzysztof Rejmer* str. 1

 Zadania str. 3

 Mikroskop z kropli wody  
*Stanisław Bednarek* str. 4

Skąd się bierze długość Plancka?  
*Michał Bejger* str. 5

Zbyt wielkie, by zgasnąć, czyli kryzys w sferach niebieskich  
*Wojciech Hellwing* str. 7

 Niemożliwe!  
*Michał Adamaszek* str. 11

Dlaczego problem  $P \stackrel{?}{=} NP$  jest tak trudny?  
*Wojciech Czerwiński* str. 14

Migawka informatyczna:  
Paradoks Russella str. 18

Informatyczny kącik olimpijski (109):  
Gra w karty str. 19

 Owcze tajemnice  
*Magdalena Fikus* str. 20


Klub 44 str. 21

Dowolne cztery proste,  
z których żadne dwie nie są  
równoległe, a żadne trzy nie  
mają punktu wspólnego... str. 21

Prosto z nieba:  
Wielki Finał misji Cassini str. 22

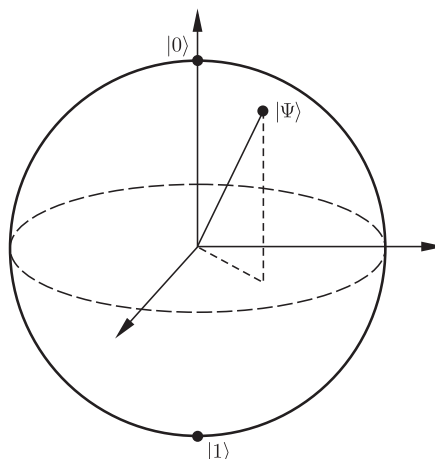
Niebo w listopadzie str. 22

Aktualności str. 24

 Krzywe i połamane  
*Joanna Jaszńska* str. 25

## W następnym numerze:

Skąd tyle szumu wokół komputerów kwantowych?



## I co tu robi ta sfera?

Miesięcznik *Delta – matematyka, fizyka, astronomia, informatyka* jest wydawany przez Uniwersytet Warszawski przy współpracy towarzystw naukowych: Polskiego Towarzystwa Matematycznego, Polskiego Towarzystwa Fizycznego, Polskiego Towarzystwa Astronomicznego i Polskiego Towarzystwa Informatycznego.

Komitet Redakcyjny: dr Waldemar Berej, dr Piotr Chrzastowski-Wachtel, dr Krzysztof Ciesielski – wiceprzewodniczący, prof. dr hab. Bożena Czerny, dr Andrzej Dąbrowski, prof. dr hab. Marek Demiański, prof. dr hab. Krzysztof Diks, dr Tomasz Greczyło, prof. dr hab. Paweł Idziak, dr hab. Agnieszka Janiuk, dr hab. Marcin Kiraga, prof. dr hab. Andrzej Majhofer, prof. dr hab. Zbigniew Marciniak, dr hab. Zygmunt Mazur, dr Adam Michalec, prof. dr hab. Michał Nawrocki – przewodniczący, dr Zdzisław Pogoda, dr Paweł Preś, prof. dr hab. Wojciech Rytter, prof. dr hab. Paweł Strzelecki.

Redaguje kolegium w składzie: Wiktor Bartol, Michał Bejger, Szymon Charzyński – z-ca red. nac., Wojciech Czerwiński, Tomasz Kazana, Piotr Kaźmierczak, Krystyna Kordos – sekr. red., Marek Kordos – red. nac., Kamila Łyczek, Katarzyna Małek, Łukasz Rajkowski, Anna Rudnik, Krzysztof Rudnik, Piotr Zalewski.

Adres do korespondencji:

Instytut Matematyki UW, Redakcja *Delta*, ul. Banacha 2, pokój 4020, 02-097 Warszawa, e-mail: [delta@mimuw.edu.pl](mailto:delta@mimuw.edu.pl) tel. 22-55-44-402.

Okładki i ilustracje: Anna Ludwicka Graphic Design & Serigrafia; rysunki techniczne: Stanisław Walczak.

Skład systemem  $\text{T}_{\text{E}}\text{X}$  wykonała Redakcja.

Wydrukowano w Drukarni Greg, ul. Górczewska 216 p. 101, 01-460 Warszawa.

PRENUMERATA

Garmond Press: [www.garmondpress.pl](http://www.garmondpress.pl)

Kolporter: [www.kolporter.com.pl](http://www.kolporter.com.pl)

RUCH S.A.: [www.ruch.com.pl](http://www.ruch.com.pl), infolinia 804-200-600

Prenumerata realizowana przez RUCH S.A.:

Cena prenumeraty w 2018 roku wynosi 4 zł za egzemplarz.

Zamówienia na prenumeratę w wersji papierowej można składać bezpośrednio na stronie [www.prenumerata.ruch.com.pl](http://www.prenumerata.ruch.com.pl)

Ewentualne pytania prosimy kierować na adres e-mail: [prenumerata@ruch.com.pl](mailto:prenumerata@ruch.com.pl) lub kontaktując się z Centrum Obsługi Klienta RUCH

pod numerem: 801 800 803 lub 22 693 70 00 – czynne w dni robocze w godzinach 7<sup>00</sup>–17<sup>00</sup>. Koszt połączenia wg taryfy operatora.

Numery archiwalne (od 1987 r.) można nabyć w Redakcji osobiście lub listownie.

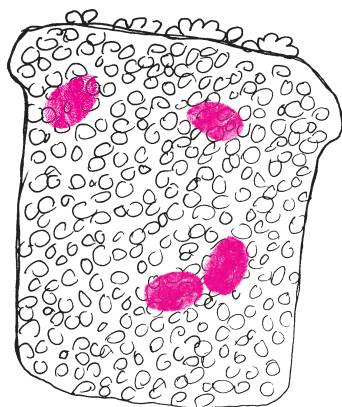
Strona internetowa (w tym artykuły archiwalne, linki itd.): [deltami.edu.pl](http://deltami.edu.pl)

Można nas też znaleźć na [facebook.com/Delta.czasopismo](https://www.facebook.com/Delta.czasopismo)

Wydawca: Uniwersytet Warszawski

# Gra w kości atomami

Krzysztof REJMER



Atomizm ma prastare korzenie, wie to przecież każdy. Jak również każdy wie, że twórcami atomistycznej hipotezy byli Leukip i Demokryt. Nie musi to być jednak prawda; niektórzy greccy filozofowie jej ojcostwo przypisywali legendarnemu fenickiemu protofilozofowi, Mochosowi z Sydonu. A już zapewne mało kto słyszał o atomizmie hinduskim czy (późniejszym) arabskim.

Grecy, z wyjątkiem epikurejczyków, atomizmu nie pokochali; obca była im idea próżni, w której miały się poruszać niezniszczalne atomy. Lukrecjusz w poemacie *De rerum natura* porzucił pierwotny demokrytejski determinizm ruchu, pisząc o pędzących w próżni atomach, że losowo odchylają swój bieg od prostoliniowego, wskutek czego się zderzają; zjawisko to nazwano parenklizą (*clinamen*). Ta myśl musiała budzić zgrozę nie mniejszą niż sama idea próżni. W IV wieku naszej ery brawurowo rozprawił się (metodą *reductio ad absurdum*) z atomami wielce uczony Laktancjusz. Otóż atomy – jak pisał – aby się łączyć, musiałyby mieć jakieś uszka i haczyki, te zaś mogą się łatwo odłamać, co przeczy niezniszczalności atomów. *Ergo*, nic takiego istnieć nie może.

Na stulecia wygnane z nauki i błakające się po jej peryferiach, uważane czasem za – wprawdzie użyteczną – ale tylko hipotezę, pełni obywatelstwa zyskały dopiero dzięki pracy Einsteina o ruchach Browna i doświadczeniu Perrina, a raczej dzięki zgodności doświadczenia z teorią. Na ruchy Browna spojrzeć można jako na spóźnione zwycięstwo lukrecjuszowskiej parenklizy, choć nieco może *au rebours*. Bo to losowe zderzenia zmieniają prostoliniowy bieg atomów, a nie na odwrót.

Badania nad zjawiskiem promieniotwórczości doprowadziły do wniosku, że atomy muszą mieć jakąś strukturę, w dodatku elektryczną, że zawierają ładunki dodatnie i ujemne. Ale skoro istnieje struktura, to jaka?

Pierwszy model zaproponowany przez Thomsona nazwano modelem ciasta z rodzynkami. Atom miał być rozciąglą obszarem naładowanym dodatnio (ciasto), w którym tkwią ujemne punktowe elektrony (rodzynki). W wyniku różnych zdarzeń elektrony mogą zostać z atomu wybite. Model ten jednak szybko odrzucono, ponieważ nie wyjaśniał poprawnie rozpraszania promieniowania alfa przez materię. Marsden i Geiger rozpraszali cząstki alfa na folii ze złota. Ich wyniki były zgodne z założeniem, że ładunek dodatni nie rozlewa się w atomie w całej jego objętości, ale jest skoncentrowany w niewielkim jądrze otoczonym przez krążące wokół elektrony. Tak narodził się model planetarny Rutherforda.

Historycznie rzecz ujmując, powiemy, że to właśnie owo klasyczne (dziś zaliczane do dziesiątki najpiękniejszych doświadczeń w fizyce) doświadczenie rozproszeniowe wyeliminowało model Thomsona. Mało kto jednak wie, że można go odrzucić również w oparciu o dane spektroskopowe. Model Thomsona można bez trudu skwantować (tak samo jak model planetarny) i znaleźć jego rozwiązanie dla najprostszego przypadku: atomu wodoru. Wtedy możemy przyjąć, że elektron (mający ładunek ujemny) tkwi w środku sferycznie symetrycznej, jednorodnie dodatnio naładowanej kuli, albo wokół tego środka się porusza. O ile tylko elektron nie wyskoczy poza kulę, atom będzie przypominał oscylator harmoniczny.

Jeżeli przez  $R$  i  $m$  oznaczymy odpowiednio promień naładowanej dodatnio kuli i masę elektronu, to rozwiązując falowe równanie Schrödingera (Henry Zatzikis, 1958 rok), możemy wyznaczyć poziomy energetyczne. Okazuje się, że są opisane takim samym wzorem (przyjmijmy układ cgs), jaki otrzymujemy dla skwantowanego modelu planetarnego. Numeruje je tylko jedna (naturalna) liczba kwantowa  $n$ :

$$E_n = -\frac{me^4}{2\hbar^2 n^2},$$

Strona techniczna: musimy rozwiązać równanie Schrödingera osobno w obszarze dodatniego ładunku i poza nim, a potem te rozwiązania zszyć. Szukamy przy tym funkcji falowej całkowalnej w kwadracie.

gdzie  $e$  i  $\hbar$  to jak zawsze ładunek elektronu i stała Plancka. Ale uwaga: o ile w modelu planetarnym  $n$  ma dowolną wartość, to w kwantowym modelu Thomsona musi spełniać warunek

$$n \geq \frac{e}{\hbar} \sqrt{\frac{mR}{3}}$$

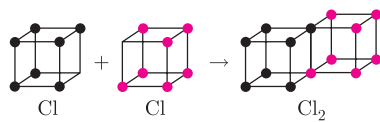
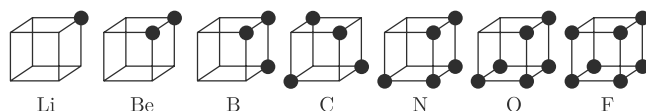
Można jednak tak dobrać wartość promienia  $R$ , żeby dopuszczalne były wszystkie liczby naturalne. Druga różnica jest już druzgocąca: brak degeneracji. Stany kwantowe atomu wodoru numerowane są przez trzy liczby kwantowe  $n, l, m$ , gdzie  $-l \leq m \leq l$ . Liczby  $l$  i  $m$  związane są z momentem pędu elektronu. Energia zależy tylko od  $n$ , tak więc jednej wartości energii odpowiada kilka poziomów energetycznych. To właśnie nazywamy degeneracją. Okazuje się, że model Thomsona nie dopuszcza rozwiązań o  $l \neq 0$ . I to falsyfikuje go w równym stopniu, co wspomniane doświadczenie rozproszeniowe.

Mechanika kwantowa, jak wiadomo, zaczęła się od prekwantowego modelu Bohra (1913 rok). Ale nie była to jedyna próba zbudowania jakiegoś prostego modelu atomu. Ta, o której teraz opowiemy, dziś zupełnie zapomniana, jest dziełem amerykańskiego fizykochemika, Gilberta Newtona Lewisa (1875–1946); był on również twórcą nazwy „foton”. Model Lewisa został opracowany dużo wcześniej niż model Bohra, jednakże opublikowany dopiero w 1916 roku.

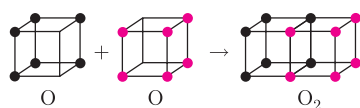
Rzecz trzeba od razu, że był on nieco dziwaczny, bo atomy miałyby być sześcianami. Czyżby jednak Bóg grał w kości, w co powątpiewał Einstein? Dokładnie rzecz biorąc, sześcianami miały być powłoki elektronowe; elektrony (w liczbie od 1 do 8) miałyby znajdować się w wierzchołkach sześcianu. W odróżnieniu od Bohra, Lewis przyjmował, że oddziaływanie elektryczne na subatomowych odległościach nie spełnia prawa Coulomba, choć trzeba od razu zaznaczyć, że jego model miał czysto jakościowy charakter i nie pozwalał na żadne obliczenia.

Z punktu widzenia chemii istotna jest tylko ta najbardziej zewnętrzna powłoka, czyli powłoka walencyjna. Te wewnętrzne, całkowicie wypełnione, wchodzą w skład jądra, rozumianego tutaj jako nadwyżka ładunku dodatniego. Jak widać, jądro atomu Lewisa nie jest dokładnie tym samym, czym jest jądro atomu Bohra.

Liczba elektronów na walencyjnej powłocy odpowiadałaby grupie układu okresowego. I tak, atomy grupy pierwszej miałyby mieć na walencyjnym sześcianie jeden elektron, atomy grupy drugiej dwa elektrony i tak dalej...



Rys. 1



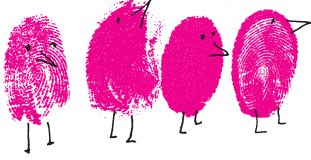
Rys. 2

Atomy, dążąc do zapełniania powłok, miałyby też współniać (uszka i haczyki!) elektrony walencyjne. Była to pierwsza, w podstawowym zarysie zachowana do dziś, teoria wiązania kowalencyjnego. Inaczej jednak wyobrażamy sobie współnianie. Rysunki 1 i 2 pokazują pojedyncze oraz podwójne wiązanie kowalencyjne.

Model Lewisa obejmował jedynie osiem grup głównych układu okresowego i błędnie przypisywał powłocy walencyjnej helowców osiem elektronów, co zasadniczo jest prawdą, ale nie dla helu, który ma tylko dwa elektrony. Lewis doskonale zdawał sobie sprawę ze słabości swego modelu, podobnie jak z trudności wyjaśnienia powstawania wiązania potrójnego, które wymagało przyjęcia dodatkowej hipotezy (psującej prostotę i elegancję) o łączeniu się elektronów w pary wierzchołkowe i wiązanie trzech sześcianów krawędziami.

Pozostaje jeszcze wyjaśnić, skąd wzięła się ta magiczna ósemka. Chodziło tu o wyjaśnienie tak zwanej reguły Abegga dotyczącej wartościowości pierwiastków. Dziś możemy sformułować ją w języku mechaniki kwantowej następująco. Najtrwalszą konfiguracją jest ta, dla której wypełnione są walencyjne powłoki:  $s$  ( $l = 0$ ) zawierająca  $2(2l + 1) = 2$  elektrony, oraz  $p$  ( $l = 1$ ) zawierająca

$2(2l + 1) = 6$  elektronów (liczba  $2l + 1$  jest degeneracją ze względu na orbitalny moment pędu, a dodatkowa dwójka uwzględnia degenerację ze względu na spin elektronu). A więc razem 8 elektronów. Wyjaśnienie reguły Abegga Lewisowi w ramach jego modelu udało się o tyle, że był on do niej po prostu pomysłowo dopasowany. Tak więc (popuśćmy wodze fantazji) w końcu Lewis musiał przyznać:



– Zwyciężyłeś, Austriaku!

Tym Austriakiem był, oczywiście, Schrödinger. W pierwszej chwili chciałem napisać: Duńczyku, ale na szczęście powstrzymałem się w porę. Wszak w odróżnieniu od naiwnego modelu Lewisa, na pozór bardziej wyrafinowany model Bohra zupełnie nie nadawał się do opisu atomów wieloelektronowych, a zatem i wiązań chemicznych!

Dodajmy może jeszcze, że urodzony w 1869 roku w Gdańsku Richard Wilhelm Heinrich Abegg w 1899 roku został profesorem Politechniki Wrocławskiej. Współpracował z Ostwaldem, Ahrreniusem i Nernstem. Prócz chemii pasjonował się fotografią i baloniarstwem, był założycielem *Schlesischen Vereins für Luftfahrt*. Zginął w kwiecie wieku, w 1910 roku, w wyniku katastrofy balonu *Schlesien* w pomorskim Cieszynie (Tessin) położonym w pobliżu Koszalina.



## Zadania

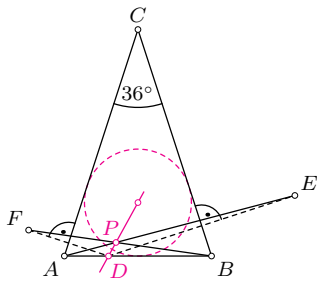
Redaguje Łukasz BOŻYK

**M 1546.** Rozważmy ciąg  $(a_n)$  zadany przez  $a_1 = 1$  oraz  $a_{n+1} = a_n^2 + 1$  dla  $n \geq 1$ . Rozstrzygnąć, czy istnieje dodatnia liczba całkowita  $n$ , dla której

$$\prod_{k=1}^n (a_k^2 + a_k + 1)$$

jest kwadratem liczby całkowitej.

Rozwiązanie na str. 15



**M 1547.** Punkt  $D$  leży na boku  $AB$  trójkąta  $ABC$ , w którym  $AC = BC$  oraz  $\sphericalangle ACB = 36^\circ$ . Punkty  $E$  i  $F$  są symetryczne do punktu  $D$  odpowiednio względem prostych  $BC$  i  $AC$ . Odcinki  $AE$  i  $BF$  przecinają się w punkcie  $P$ . Wykazać, że środek okręgu wpisanego w trójkąt  $ABC$  leży na prostej  $DP$ .

Rozwiązanie na str. 17

**M 1548.** Znaleźć najmniejszą dodatnią liczbę całkowitą  $n$  o następującej własności: dla każdego pokolorowania dokładnie  $n$  pól tablicy o wymiarach  $7 \times 7$  na czarno istnieje kwadrat  $2 \times 2$ , który zawiera co najmniej trzy czarne pola.

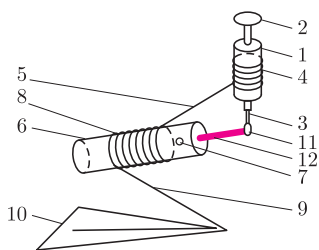
Rozwiązanie na str. 17

Przygotował Andrzej MAJHOFER

**F 939.** Przed otwarciem spadochronu skoczek często stara się wydłużyć czas swobodnego spadania, ustawiając ciało w pozycji poziomej w celu zmaksymalizowania oporu powietrza. Oszacuj, jaką maksymalną prędkość może w takim locie osiągnąć skoczek o masie  $m = 80$  kg i wzroście 1,8 m. Dla uproszczenia załóż, że podczas całego „lotu” powietrze ma stałą gęstość w przybliżeniu równą gęstości przy ziemi i temperaturę około  $20^\circ\text{C}$ . Stała gazowa  $R = 8,314$  J/K, ciśnienie atmosferyczne  $p \approx 10^5$  N/m<sup>2</sup>, średnia masa molowa cząsteczek powietrza  $\mu = 29$  g/mol, przyspieszenie ziemskie  $g \approx 10$  m/s<sup>2</sup>.  
Rozwiązanie na str. 13

**F 940.** Cienki, nieprzewodzący pierścień został jednorodnie naelektryzowany ładunkiem  $Q$ . Pierścień spoczywa na poziomej, gładkiej powierzchni, a cały układ znajduje się wewnątrz bardzo długiej cewki prostopadłej do płaszczyzny pierścienia. Ile wyniesie końcowa prędkość kątowa  $\omega$  pierścienia, gdy wewnątrz cewki indukcja pola magnetycznego wzrośnie od zera do wartości  $B_0$ ?

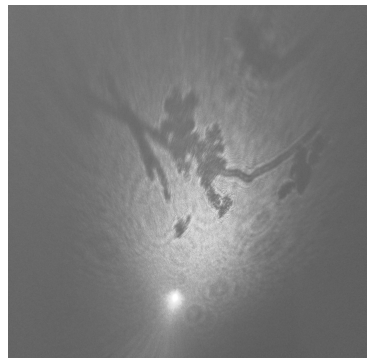
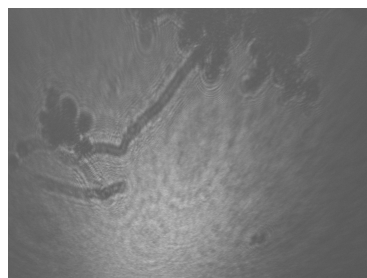
Rozwiązanie na str. 13



Rys. 1. Budowa mikroskopu:  
1 – strzykawka, 2 – tłoczek strzykawki,  
3 – obcięta igła strzykawki, 4 – górna  
spirala z drutu, 5 – łącznik, 6 – wskaźnik  
laserowy, 7 – włącznik wskaźnika,  
8 – dolna spirala z drutu, 9 – wspornik,  
10 – podstawa, 11 – kropla wody,  
12 – wiązka światła laserowego.



Fot. 1. Wygląd najprostszej wersji  
zbudowanego mikroskopu.



Fot. 2. Przykłady obrazów uzyskanych za  
pomocą mikroskopu z kropli wody.

## Mikroskop z kropli wody

Stanisław BEDNAREK

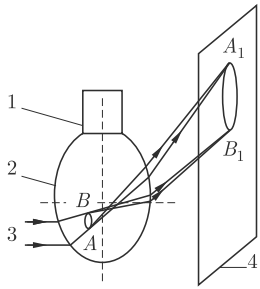
Pierwszy mikroskop optyczny zbudowali Holendrzy Hans i Zachariasz Janssenowie w 1590 r. Ich prototyp powiększał zaledwie dziesięciokrotnie, czyli tyle, ile powiększa obecnie dobrej jakości lupa. Istotnego postępu dokonał ich rodak Antoni van Leeuwenhoek, uzyskując dwustuczterdziestokrotne powiększenie. Gdy obserwował on pod tym mikroskopem kroplę wody, miał powiedzieć, że odkrył w niej życie. My wykorzystamy kroplę wody do zbudowania niezwykle prostego i interesującego mikroskopu.

W tym celu potrzebne będą: wskaźnik laserowy – najlepiej zielony, drut miedziany o średnicy około 1 mm i długości około 1 m, strzykawka jednorazowego użytku o pojemności 2 ml z grubą igłą, arkusz białego papieru, taśma klejąca, klamerka do przypinania ubrań podczas suszenia albo ściskacz do papieru, kleszcze uniwersalne (tzw. kombinerki), trójkątny pilnik do metalu, nożyczki, woda z różnych źródeł – instalacji wodociągowej, butelki z wodą mineralną, oczka wodnego itp.

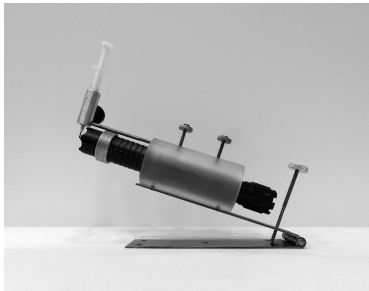
Sposób wykonania mikroskopu przedstawia rysunek 1. Drut miedziany owijamy ciasno kilka razy wokół strzykawki, rozpoczynając nawijanie od jednego z jego końców. Można wykorzystać zarówno drut w izolacji plastikowej lub emaliowej, jak i bez niej. Uzyskujemy w ten sposób górną spiralę, w którą można będzie wsunąć strzykawkę i ustalić jej położenie. Drut miedziany jest dość miękki i cienki, więc spiralę łatwo jest nawinąć ręcznie. Następnie odginamy drut kombinerkami i pozostawiamy jego prostoliniowy odcinek (łącznik) o długości kilku centymetrów. Postępując podobnie jak poprzednio, nawijamy dolną spiralę wokół wskaźnika laserowego. Znowu odginamy drut i pozostawiamy kilkucentymetrowy, prostoliniowy odcinek (wspornik). Z pozostałego kawałka drutu wyginamy podstawkę, np. w kształcie trójkąta równobocznego o boku 10–15 cm. W ten sposób wykonaliśmy statyw mikroskopu. Z igły do strzykawki odpiłowujemy ukośnie ścięty koniec, tak żeby otrzymać wylot prostopadły do jej osi. Posługujemy się w tym celu małym, trójkątnym pilnikiem o drobnych nacięciach, zwracając uwagę, żeby do igły nie dostały się opiłki i brzegi jej końca były gładkie. Do strzykawki nabieramy wody, nakładamy na nią igłę i wsuwamy strzykawkę w górną spiralę statywu. W dolną spiralę statywu wsuwamy wskaźnik laserowy. Do ściany przyklejamy arkusz białego papieru, używając w tym celu kawałków taśmy klejącej. Będzie on stanowił ekran do prowadzenia obserwacji. W ten sposób zakończyliśmy montaż mikroskopu. Widok jednego z mikroskopów, wykonanych zgodnie z podanym opisem, przedstawia fotografia 1.

Teraz przystąpimy do uruchomienia i regulacji naszego przyrządu. Powoli naciskamy tłoczek strzykawki, żeby utworzyć kroplę wody zwisającą z dolnego końca igły. Włączamy wskaźnik laserowy na pewien czas – można w tym celu posłużyć się klamerką do przypinania ubrań lub ściskaczem do papieru, które nakładamy na włącznik wskaźnika. Używając wskaźnika laserowego, należy zachować ostrożność i nie kierować wiązki światła w stronę oczu oraz lustrzanych powierzchni. Wiązka światła powinna padać na zwisającą kroplę. Żeby to osiągnąć, zwykle potrzebne jest przesunięcie wskaźnika lub strzykawki w utrzymujących je spiralach albo wygięcie znajdującego się między nimi łącznika. Mikroskop ustawiamy w odległości kilkudziesięciu centymetrów od ekranu, na którym obserwujemy obszar w kształcie zbliżonym do koła, oświetlony w kolorze światła wysyłanego przez laser. Na tym obszarze pojawiają się małe obiekty o nieregularnym kształcie i prążkowanych brzegach. Są one ciemniejsze niż oświetlony obszar i większość z nich porusza się powoli ku górze. Dwa przykłady obrazów, zaobserwowanych przy użyciu zielonego wskaźnika laserowego, pokazane są na fotografiach 2.

Spróbujmy teraz w uproszczeniu wyjaśnić, skąd biorą się te tajemnicze obiekty? W tym celu posłużymy się rysunkiem 2. Wskaźnik laserowy wysyła równoległą



Rys. 2. Uproszczone wyjaśnienie wytwarzania obrazu przez mikroskop; 1 – obcięty koniec igły, 2 – kropla wody, 3 – wiązka światła laserowego, 4 – ekran,  $AB$  – przedmiot,  $A_1B_1$  – obraz przedmiotu  $AB$ .



Fot. 3. Bardziej uniwersalna wersja mikroskopu zbudowana z zastosowaniem zawiasu i przystosowana do zamocowania wskaźników laserowych o różnych rozmiarach. Widoczny jest tutaj niebieski wskaźnik laserowy o mocy około 1 W.

wiązkę światła, która pada najpierw na powierzchnię kropli znajdującą się bliżej jego otworu wyjściowego. Powierzchnia ta ma bardzo mały promień krzywizny i działa jak soczewka o dużej zdolności skupiającej. Wewnątrz kropli mogą znajdować się kryształki nierozpuszczalnych w wodzie tlenków żelaza (pochodzących ze stalowych rur wodociągowych), glony, pyłki kurzu itp. – zależy to od źródła, z którego została pobrana kropla. Obiekty te są mniej przezroczyste niż woda, a ich rozmiary wynoszą  $10^{-3}$ – $10^{-1}$  mm. Mniejsza przezroczystość powoduje, że dają one obraz cieniowy. Na ich brzegach i niejednorodnościach zachodzi również dyfrakcja, a dalej interferencja światła. Małe rozmiary tych obiektów i spójność wiązki światła laserowego ułatwiają obserwację powstających przy tym prążków interferencyjnych. Światło dochodzące do powierzchni kropli, znajdującą się dalej od wskaźnika, niesie informację o zawartych w niej obiektach. Powierzchnia ta ma również bardzo mały promień krzywizny i działa jak lupa o dużym powiększeniu. Skutkiem tego na ekranie tworzy się cieniowo-interferencyjny obraz obiektów zawartych w kropli. Obraz ten jest powiększony i odwrócony. Z tego powodu większość widocznych na ekranie obiektów porusza się ku górze, chociaż w kropli one opadają. Przedstawiony sposób tworzenia się obrazów charakteryzuje się dużą czułością. Skutkiem tego widoczne są też prążki odpowiadające zafalowaniu powierzchni kropli wskutek wstrząsów czy ruchom konwekcyjnym w jej wnętrzu. Jeżeli zmierzmy średnicę kropli i obrazu, to możemy obliczyć powiększenie naszego mikroskopu. Im dalej kropla będzie umieszczona od ekranu, tym to powiększenie będzie większe, ale wtedy jasność obrazu zmaleje i obserwacja szczegółów oraz ich sfotografowanie staną się trudniejsze.

Mikroskop, w którym elementy połączone są kawałkiem drutu miedzianego, jest bardzo łatwy do wykonania, ale podczas dłuższej pracy wymaga ponownej regulacji i pasuje tylko do rozmiarów jednego wskaźnika laserowego. Jeżeli chcielibyśmy poeksperymentować ze wskaźnikami o różnych rozmiarach i długościach fali światła, to możemy pokusić się o zbudowanie bardziej uniwersalnej wersji mikroskopu. Można wykorzystać w nim tulejki o większych otworach i duży zawias, który świetnie nadaje się jako gotowy element mocujący i regulacyjny. Wygląd takiej wersji mikroskopu przedstawia fotografia 3. Warto też zauważyć, że zjawisko dyfrakcji światła ogranicza możliwości rozróżniania szczegółów, czyli zdolność rozdzielczą tradycyjnych mikroskopów optycznych. To właśnie zjawisko jest również istotnym czynnikiem decydującym o tworzeniu się obrazu i jego interesujących właściwościach w opisanym tu mikroskopie.

## Skąd się bierze długość Plancka?

*Michał BEJGER*

Niestety, do naszych czasów nie dotrwała żadna z prac Demokryta, a informacje o nich można jedynie czerpać z (najczęściej krytycznych i nie wprost) wzmianek w pismach m.in. Platona i Arystotelesa.

Już około 2500 lat temu Leukip i Demokryt z Abdery wiedzą, że świat składa się wyłącznie z niepodzielnych elementów materii („atomów”) oraz z próżni. Mechanika Newtona (XVII w.) operująca pojęciami absolutnego czasu i absolutnej przestrzeni umożliwia po raz pierwszy w historii jakościowy opis zachowania się cząstek materii (np. planet, a także fotonów uznawanych wówczas za korpuskuły). Faraday i Maxwell (XIX w.) dodają do listy składników świata pojęcie pola (np. pole elektromagnetyczne). Teoria względności Einsteina (XX w.) łączy czas i przestrzeń: dzięki temu na początku minionego stulecia świat składa się z czasoprzestrzeni, cząstek i pól, przy czym pole grawitacyjne zostało powiązane z nieeuklidesową geometrią czasoprzestrzeni. Przełomy związane z mechaniką kwantową (Planck, Bohr, Born, Jordan, Heisenberg, Dirac, Feynman. . .) doprowadzają do kolejnej unifikacji: cząstki są w istocie konsekwencją istnienia pól (kwantami pól). W połowie XX wieku do opisu świata potrzeba więc jedynie kwantowych pól oraz czasoprzestrzeni. Aktualnym wyzwaniem jest zrozumienie, w jaki sposób skwantować pole grawitacyjne: dotychczas było ono bowiem zawsze opisywane przy użyciu klasycznej, choć

Naturalne jednostki miary są to, słowami samego Plancka, *jednostki długości, masy i czasu, niezależne od właściwości ciał lub substancji, zachowujące swoje specjalne znaczenie niezależnie od czasów i kultur, także pozaziemskich i pozaludzkich* (Planck 1899)

Badaczem, który jako pierwszy zwrócił uwagę na fundamentalną cechę odróżniającą pole grawitacyjne od innych pól był Matvei P. Bronstein (1906–1938). Obecnie jest uważany za pioniera teorii kwantowej grawitacji i to jego argumenty z prac z roku 1936 przytaczam w tekście.

Charakterystyczną masę Plancka  $m_p$  można dostać porównując długość fali Comptona cząstki o masie  $m$ ,  $\lambda_c = h/mc$ , i promień czarnej dziury Schwarzschilda  $R = 2Gm/c^2$ . Krytyczna maksymalna masa kwantowej cząstki to  $m = \sqrt{\pi}m_p$ .



krzywoliniowej, geometrii. Jeśli grawitacja jest kwantowa, powinien istnieć charakterystyczny, minimalny rozmiar kwantu przestrzeni, przy którym czasoprzestrzeń przestaje wydawać się ciągłą.

Zanim oszacujemy ten rozmiar, przypomnijmy *naturalne jednostki miary* zaproponowane przez Maksa Plancka w 1899 roku:

$$x_p = c^\alpha G^\beta \hbar^\gamma,$$

gdzie  $c$  oznacza prędkość światła,  $G$  stałą grawitacji Newtona, a  $\hbar$  zredukowaną stałą Plancka ( $\hbar = h/2\pi$ , gdzie  $h$  to stała Plancka). Dobierając odpowiednie potęgi  $\alpha$ ,  $\beta$  i  $\gamma$  można dostać charakterystyczną długość  $l_p = (\hbar G/c^3)^{1/2} \approx 10^{-33}$  cm, masę  $m_p = (\hbar c/G)^{1/2} \approx 10^{-5}$  g oraz czas  $t_p = (\hbar G/c^5)^{1/2} \approx 10^{-43}$  s.

Wraz ze sformułowaniem zasady nieoznaczoności Heisenberga,  $\Delta x \Delta p > \hbar$ , określającej pary wielkości (np. położenie  $x$  i pęd  $p$ ), których nie da się zmierzyć jednocześnie z dowolną dokładnością, pojawiło się pytanie, czy w związku z tym nieoznaczoność przeszkadza przy pomiarze składowych jakiegoś interesującego nas kwantowego pola. W 1933 roku Bohr i Rosenfeld dostarczyli odpowiedzi dla pola elektromagnetycznego: mimo kwantowej nieoznaczoności można mierzyć składowe tego pola z dowolną dokładnością. Czy podobnie jest dla pola grawitacyjnego? Problem polega na tym, że w przeciwieństwie do przypadków kwantowych pól *znajdujących się* w czasoprzestrzeni, w której dokonujemy pomiaru, pole grawitacyjne *jest* czasoprzestrzenią.

W praktyce pomiar oznacza detekcję oddziaływania pola z jakąś materialną cząstką, która znajduje się w położeniu  $x$  z określoną dokładnością  $l$ . Stosując zasadę Heisenberga nieoznaczoność położenia musi spełniać równość  $\Delta x < l$ , czyli  $\Delta x > \hbar/\Delta p$ , a więc  $\Delta p > \hbar/l$ . Dla dużych wartości pędu średnia wartość kwadratu pędu  $p^2$  jest większa od  $(\Delta p)^2$ , co daje oszacowanie  $p^2 > (\hbar/l)^2$ . Wynik ten jest oczywistą konsekwencją zasady nieoznaczoności: duża dokładność pomiaru położenia wymaga dużych pędów. Z tego powodu w doświadczeniach akceleratorowych, w których bada się coraz mniejsze skale odległości, zderzają się bardzo szybkie cząstki. Energia relatywistycznej cząstki o masie spoczynkowej  $m_0$  i pędzie  $p$  to  $E = mc^2 = \sqrt{(m_0c^2)^2 + (cp)^2}$ . Dla dużych pędów energia związana z masą spoczynkową  $m_0c^2$  jest zaniedbywalna, więc można przyjąć, że  $E = cp$ . Dokładny pomiar położenia wymaga zatem jak największej energii w jak najmniejszej objętości.

Jak dużo energii można zgromadzić w jednym miejscu? W teorii względności energia jest równoważna masie ( $m \equiv E/c^2$ ), więc podobnie do „zwykłej” masy zakrzywia przestrzeń wokół miejsca, w którym jest zlokalizowana. Interesującym nas ograniczeniem na ilość masy-energii jest krytyczne zakrzywienie, przy którym powstaje czarna dziura. Rozmiar czarnej dziury (promień Schwarzschilda) to  $R = 2Gm/c^2$ . Zmniejszanie  $l$ , czyli coraz bardziej precyzyjna lokalizacja, prowadzi do, jak nakazuje zasada nieoznaczoności, zwiększania masy-energii, jednak jedynie do momentu, w którym  $R = l$ , ponieważ dla  $R > l$  nasz system pomiarowy zaczyna znajdować się wewnątrz czarnej dziury, do której z definicji nie mamy dostępu. Łącząc powyższe oszacowania, otrzymujemy minimalny rozmiar wyznaczony przez prawa fizyki:

$$l = \frac{mG}{c^2} = \frac{EG}{c^4} = \frac{pG}{c^3} = \frac{\hbar G}{lc^3} \quad \text{czyli} \quad l = l_p \equiv \sqrt{\frac{\hbar G}{c^3}} \approx 1.616 \cdot 10^{-33} \text{ cm},$$

a więc (niespodzianka!) minimalnym rozmiarem okazuje się długość Plancka  $l_p$ . Mimo, że przedstawione argumenty opierają się w zasadzie o klasyczną fizykę, ich znaczenie jest jasne: dla odległości porównywalnych z  $l_p$  czasoprzestrzeni nie można już uważać za ciągłą, a dla długości mniejszych od  $l_p$  pojęcie odległości w ogóle traci sens.

Ziarnistość czasoprzestrzeni w najmniejszych skalach prowokuje pytania o możliwe odstępstwa od uznanych klasycznych teorii (np. czy istnienie minimalnego rozmiaru wymaga modyfikacji szczególnej teorii względności?), ale z drugiej strony rozwiązuje paradoks Zenona z Elei, o którym na pewno wiedział Demokryt: na skwantowanej bieżni Achilles zawsze dogoni żółwia.



# Zbyt wielkie, by zgasnąć, czyli kryzys w sferach niebieskich

Wojciech HELLWING\*

\*Instytut Kosmologii i Grawitacji,  
Uniwersytet w Portsmouth

Sam termin *too big to fail* jest sporo starszy: został spopularyzowany przez amerykańskiego kongresmena Stewarta McKinneya w 1984 roku podczas kongresowych przesłuchań w sprawie ubezpieczeń federalnych depozytów.

Model LCDM przewiduje, że Kosmos w około ćwierci swojej masy składa się z grawitującej, acz nieświecącej ciemnej materii, a 70% przypada na ciemną energię, czyli kosmiczny dziwoląg zapewniający paliwo do wciąż przyspieszającej ekspansji czasoprzestrzeni. W końcu, pozostałe 5% tego, co istnieje, to tzw. materia barionowa, czyli znane nam na co dzień atomy (głównie wodór i hel), fotony i cała menażeria standardowego modelu cząstek elementarnych.



Układ gromad galaktyk Pocisk  
(Copyright NASA, CFA).

Model LCDM głosi, że każda galaktyka – od niepozornych karłów po olbrzymie galaktyki eliptyczne – zanurzona jest w o wiele większym i masywniejszym niż ona sama halo ciemnej materii. Grawitacja tej niewidzialnej materii utrzymuje gwiazdy na orbitach i powstrzymuje gorący gaz przed ucieczką z galaktyk. Bez ciemnej materii ani gwiazdy, ani galaktyki nie mogłyby powstać.

Kiedy parę lat temu światową gospodarką wstrząsał kryzys finansowy, termin „zbyt wielkie, by upaść” (*too big to fail*) ponownie zyskał na popularności. Twierdzenie to powtarzało wielu światowych przywódców i szefów banków centralnych, a jego główne przesłanie sprowadza się do tezy, że niektóre instytucje finansowe są zbyt duże i mają zbyt wielki wpływ na rynek, by pozwolić im upaść (zbankrutować). Nikt oczywiście wtedy nie przypuszczał, że podobny termin zostanie użyty kilka lat później do opisanego poważnego kryzysu, który nastąpił w sferach raczej odległych od światowej finansjery, bo w sferach niebieskich.

Historia ta zatoczyła już kilka kręgów, a jej tematyka jest wciąż gorąco dyskutowana w międzynarodowym środowisku kosmologów. Rozpoczęła się ona od artykułu naukowego opublikowanego w 2011 roku przez Mike’a Boylan-Kolchina i współpracowników pt. *Too big to fail? The puzzling darkness of massive Milky Way subhaloes*, czyli *Zbyt wielkie, by zgasnąć? Zagadkowa ciemność masywnych subhalo Drogi Mlecznej*, w którym autorzy przekonywali, że trudno jest pogodzić liczbę i parametry obserwowanych galaktyk satelitarnych Drogi Mlecznej (DM) z przewidywaniami standardowego modelu kosmologicznego.

Standardowy model kosmologiczny, czyli model zimnej ciemnej materii ze stałą kosmologiczną  $\Lambda$  (*Lambda Cold Dark Matter*, LCDM) jest największym osiągnięciem kosmologii ostatnich trzech dekad: tłumaczy zaskakująco dużą liczbę obserwacji astronomicznych dotyczących struktury, historii i innych własności naszego Wszechświata za pomocą zaledwie sześciu parametrów. W naszej opowieści najistotniejszym elementem będzie ciemna materia, o której, co prawda, nie wiemy zbyt wiele, ale mamy wiele istotnych obserwacji potwierdzających jej istnienie pochodzących z obserwacji świecącego gazu, dynamiki zderzeń galaktyk (np. gromady galaktyk Pocisk), soczewkowania grawitacyjnego i promieniowania relikтового tła. Model LCDM postuluje, że ciemna materia jest zimna. Termin ten nie odnosi się tutaj do temperatury w zwyczajowym rozumieniu, a do tego, jak duże są prędkości resztkowe cząstek ciemnej materii pochodzące z czasów, kiedy Wszechświat był młody i gorący. Zimna ciemna materia charakteryzuje się małymi prędkościami resztkowymi, gromadzi się więc w skupiska (zgęstki) o rozmiarach i masach od planet począwszy aż na wielkich gromadach galaktyk skończywszy (różnica ponad 20 rzędów wielkości masy!). Standardowa kosmologia postuluje, że rosnące zgęstki ciemnej materii, które kosmologowie nazywają halami (od słowa *halo*, czyli aureola, nie mylić z wysokogórkimi łąkami!), utworzyły domeny, w których mogły z czasem powstać i urosnąć wszystkie obserwowane obecnie galaktyki.

LCDM postuluje, że fizyka CM jest prosta, co pozwala oszacować, ile galaktyk – i jakiej wielkości – powinno się znajdować we Wszechświecie. W szczególności wobec galaktyk tak sporych jak DM możemy przewidzieć, ile powinny one mieć satelitów, takich jak np. Obłoki Magellana, oraz jaki powinien być rozkład ich rozmiarów i mas.

Model LCDM przewiduje istnienie o wiele większej liczby satelitów Drogi Mlecznej, niż rzeczywiście obserwujemy. Deficyt ten jest znany jako problem brakujących satelitów. Jego rozwiązanie okazało się jednak dosyć proste. Otóż, o ile rzeczywiście model LCDM przewiduje, że w galaktyce o wielkości Drogi Mlecznej powinny się znajdować setki skupisk ciemnej materii w przedziale mas galaktyk karłowatych, to większość z nich jest zbyt mała, by mogły w nich kiedykolwiek powstać gwiazdy, a więc i widzialne galaktyki. W uproszczeniu jest tak dlatego, że w przeszłości Wszechświat przeszedł dosyć burzliwą fazę,

Pierwsze gwiazdy były o wiele masywniejsze niż populacja obecnie obserwowanych gwiazd, były jednocześnie bardzo gorące, i zgodnie z prawem Wiena świeciły mocno w energetycznym ultrafiolecie. Ponieważ były bardzo masywne, kończyły zazwyczaj swój żywot w potężnych eksplozjach supernowych, które dodatkowo zapelniały kosmos energetycznymi fotonami.

nazywaną epoką rejonizacji, kiedy neutralne od czasu rekombinacji wodór i hel zostały ponownie zjonizowane przez ogromną ilość fotonów UV wyemitowanych przez pierwsze pokolenie gwiazdowych pionierów. Duża energia wewnętrzna gorącego zjonizowanego gazu wymaga odpowiednio głębokiej studni potencjału grawitacyjnego, by zatrzymać go przed ucieczką. Okazuje się, że większość pierwotnych skupisk ciemnej materii była za mało masywna, by osiągnąć ten efekt. Na okładce przedstawiona jest wizualizacja symulacji z projektu APOSTOLE grupy dr Tilla Sawali. Przedstawiony obszar z grubsza obejmuje rozmiary naszej lokalnej Grupy Galaktyk. Prawa dolna połowa rysunku na okładce przedstawia gęstość ciemnej materii. Widzimy, że model LCDM przewiduje istnienie setek tysięcy małych i większych zgęstek CM. Dwa największe skupiska gwiazd i CM odpowiadają DM i galaktyce Andromedy. Widzimy też, że tylko największe zgęstki CM zawierają gwiazdy, bo tylko one były w stanie utrzymać w sobie wystarczająco dużo gazu, podstawowego budulca gwiazd. Bez gazu zaś nie można, oczywiście, utworzyć gwiazd. Co więcej, nawet jeżeli niektóre małe hala ciemnej materii utworzą galaktyki, to ich kariera nie trwa długo: potężne siły pływowe większych galaktyk są w stanie zaledwie podczas kilku okrążeń orbitalnych dosłownie rozszarpać je na strzępy. Ich resztki obserwujemy jako bardzo blade ogony i pierścienie pływowe gwiazd orbitujących wokół m.in. naszej Galaktyki.

Nie jest zatem łatwo zaistnieć jako mała galaktyka satelitarna. Jeżeli jednak już się przetrwa i epokę rozgrzewania gazu przez pierwsze pokolenie gwiazdowych olbrzymów, i grawitacyjne siły pływowe, to ma się kilka spokojnych miliardów lat na orbitowanie dookoła galaktyki centralnej. Tu dochodzimy do sedna problemu *Zbyt wielkie, by zgasnąć* (ZWBZ). Model LCDM przewiduje, że w okolicy DM powinno być sporo dużych i dosyć gęstych satelitów, bo przecież tylko tacy „twardziele” mogli przetrwać kosmiczne zapasy. Jednak jeżeli zestawimy przewidywania komputerowych symulacji LCDM z obserwowaną gromadką satelitów DM, to okaże się, że widzimy ich o wiele za mało. Wyjaśnijmy przy okazji termin gęstość satelitów. Ponieważ grawitacja ciemnej materii zdecydowanie dominuje w tych układach, to właśnie ona dyktuje zwykłej materii, po jakich orbitach się poruszać. Badając ruch gazu i gwiazd w różnych odległościach  $R$  od środka galaktyki, możemy oszacować głębokość studni potencjału grawitacyjnego oraz krzywą rotacji galaktyki, tzn. prędkość orbitalną w funkcji odległości  $V(R)$ . Niemalże wszystkie skupiska ciemnej materii bez względu na ich wielkość można opisać jednym uniwersalnym profilem gęstości  $\rho(R) = \frac{\rho_0}{\left(\frac{R}{R_s} + \frac{R}{R_s}\right)^2}$  nazwanym NFW od nazwisk jego twórców Julio Navarro, Carlosa Frenka i Simona White'a. Parametr  $\rho_0$  oznacza gęstość centralną, zaś  $R_s$  jest promieniem skalowania. Całkując ten profil względem promienia i przyrównując do całkowitej masy halo ciemnej materii, można wykazać, że  $\rho_0$  nie jest niezależnym parametrem i zależy od  $R_s$ . Dla halo o zadanej masie całkowitej  $M$  profil NFW jest zatem określony przez tylko jeden parametr: promień  $R_s$ . Odpowiada on miejscu, w którym logarytmiczne nachylenie profilu gęstości,  $\frac{d \log \rho(R)}{d \log R}$  wynosi  $-2$ . To oznacza, że dla  $R > R_s$  masa rośnie już wolniej niż  $R$ , a zatem prędkość orbitalna  $V(R)$  maleje. Wartość  $R_{\max}$  maksimum krzywej rotacji galaktyki  $V_{\max}$  jest zatem bezpośrednio związana z profilem gęstości  $\rho(R)$  i głębokością studni potencjału skupiska ciemnej materii.

Przy założeniu sferycznej symetrii i orbit kołowych  $V(R) = \sqrt{GM(<R)/R}$ , gdzie  $G$  to stała grawitacji Newtona,  $R$  to odległość od środka galaktyki, a  $M(<R)$  to całkowita masa zawarta wewnątrz kuli o promieniu  $R$ . Kształt krzywej rotacji będzie zależał jedynie od tego, jak wygląda wewnętrzny rozkład masy w galaktyce  $M(R)$ , a zatem od profilu gęstości ciemnej materii.

Z uwagi na wymagania obserwacyjne – potrzeba wielu gwiazd, by móc rzetelnie wyznaczyć uśrednioną prędkość orbitalną – pomiaru dokonuje się dla promienia kuli, z wewnątrz której dochodzi połowa całego obserwowalnego światła galaktyki ( $R_{1/2}$  to *half light radius*).

Dla większości znanych nam galaktyk satelitarnych DM zmierzono wartości prędkości orbitalnej  $V(R_{1/2})$ . Symulacje komputerowe zazwyczaj nie dostarczają dokładnych wartości  $R_{1/2}$ , używa się ich jednak do porównania przewidywań modelu z obserwacjami. Wystarczy zauważyć, że jeżeli danemu z symulacji satelicie ma odpowiadać jakakolwiek rzeczywista obserwowana galaktyka, to krzywa rotacji  $V(R)$  symulowanego halo musi gdzieś się przecinać ze zmierzoną encją  $V_{1/2}(R_{1/2})$  danej galaktyki. To spostrzeżenie doprowadziło Boylan–Kolchina i współpracowników do konkluzji, że jest niezmiernie trudno uzyskać w LCDM halo o masie odpowiadającej masie Galaktyki, które zawierałoby populację satelitów o krzywych  $V(R)$  kompatybilną z obserwacjami.

Masywne hala są po prostu zbyt wielkie, by zgasnąć – powinny zachować wystarczająco wiele gwiazd, by ich blask zarejestrowały nasze teleskopy.

Oszacowania ustalają masę halo naszej Galaktyki w przedziale  $(1 \div 3) \cdot 10^{12} M_{\odot}$ . Gdyby masa halo była mniejsza,  $(8 \div 14) \cdot 10^{11} M_{\odot}$ , to przewidywania dotyczące populacji satelitów można by łatwo pogodzić z obserwacjami.

Model ciepły przewiduje powstawanie takiej samej jak LCDM liczby dużych i średnich galaktyk oraz mniej zgęstek o masach galaktyk karłowatych; powinny one być także mniej gęste.

W astronomii widmowej *las Lymana-alfa* to seria linii absorpcyjnych w widmach odległych galaktyk i kwazarów. Linie odpowiadają przejściu wzbudzonemu elektronu w atomie wodoru. Ponieważ obłoki neutralnego wodoru są rozmieszczone na różnych odległościach wzdłuż całej linii widzenia, powodują one powstanie całego lasu linii w widmie.

Symulacje komputerowe, z których Boylan-Kolchin i inni wysnuli przewidywania LCDM, zawierały tylko CM, więc siłą rzeczy nie mogły uwzględnić zmian jej profilu przez wybuchy supernowych.

Co gorsza, model przewidywał satelity masywne, z krzywymi rotacji o sporych wartościach maksymalnych, które nigdzie nie przecinały się z żadnym punktem pomiarowym. Gdyby sytuacja ta dotyczyła tylko przypadków małych wartości  $V_{\max}$ , można by było użyć argumentów podobnych jak w przypadku problemu brakujących satelitów. Jednak model przewidywał istnienie zbyt wielu dużych satelitów o profilach gęstości niepasujących do obserwacji systemu satelitów DM. Ujmując sprawę w liczby, w otoczeniu DM powinno znajdować się co najmniej kilka satelitów o parametrach z przedziału pomiędzy Obłokami Magellana a galaktyką karłowatą w Strzelcu, jednak nigdzie ich nie widać. Gdzie się podziały?

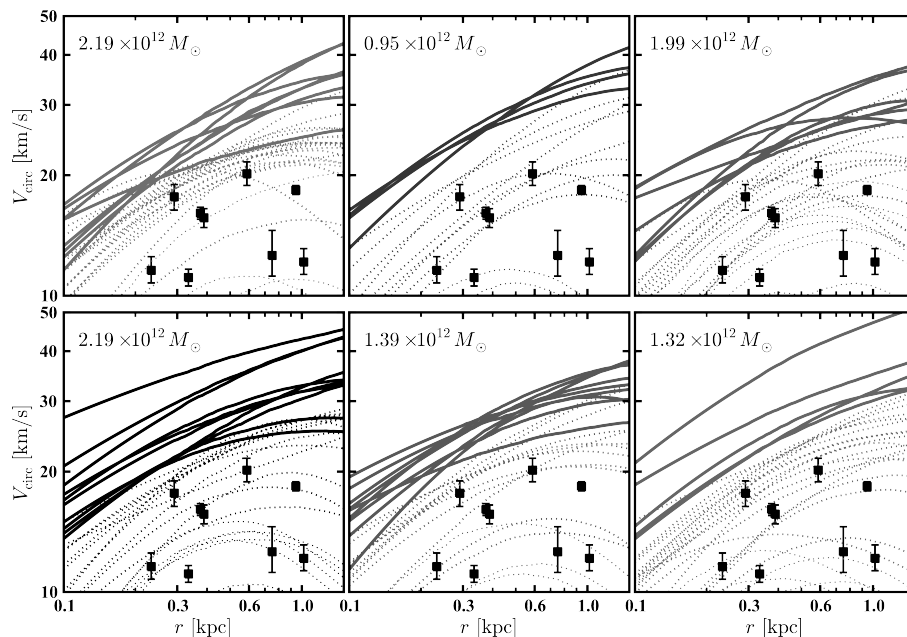
Wkrótce po publikacjach grupy Boylan-Kolchina z 2011 i 2012 roku pojawiła się cała seria artykułów dyskutująca możliwe sposoby rozwiązania zagadki. Po pierwsze, warto zauważyć, że problem ZWBZ pojawia się, gdy równocześnie próbujemy dopasować do obserwacji liczebność i gęstość satelitów. Różne fizyczne wielkości wpływają niezależnie na te dwie statystyki. Liczebność sporych satelitów silnie zależy od całkowitej masy halo centralnej galaktyki, można zatem uzgodnić model i obserwacje, postulując, że całkowita masa DM jest niższa, niż zakładaliśmy. Inne oszacowania oraz obecność Obłoków Magellana faworyzują natomiast dużą masę DM; sugerowałyby to niespotykaną wyjątkowość DM, bo bardzo trudno jest znaleźć w symulacjach halo centralne o małej masie, które równocześnie zawiera tak masywne satelity, jak oba Obłoki Magellana.

Problem ZWBZ zgłębia się również, poszukując alternatywnego mechanizmu fizycznego, który sprawiłby, że małe hala CM są mniej gęste. Popularna alternatywa modelu LCDM, w której taki efekt pojawia się naturalnie, to model ciepłej ciemnej materii (LWDM – *Lambda Warm Dark Matter*). W modelach „ciepłych” pierwotne prędkości cząstek ciemnej materii są znacząco większe, co oznacza, że ich masy spoczynkowe są odpowiednio mniejsze niż zimnej CM. Taka charakterystyka powoduje, że na skalach kiloparseków i mniejszych (czyli porównywalnych z rozmiarami galaktyk karłowatych) skupiska CM zostają niejako wygładzone. Jest tak dlatego, że większe prędkości cząstek CM pozwalają im wydostawać się spod grawitacji małych zaburzeń gęstości. Im cieplejszy model LWDM, tym mniejsze gęstości centralne i mniej małych halo. W ten sposób model LWDM w naturalny sposób rozwiązuje problem ZWBZ.

Czyżby to był więc koniec tej historii? Czy możemy uznać, że model LWDM lepiej pasuje do danych? Niestety, sprawy są bardziej skomplikowane. Istnieją mocne ograniczenia obserwacyjne na „ciepłotę” modelu LWDM. Obserwacje lasu linii Lymana- $\alpha$  informują nas dokładnie, ile w przestrzeni między galaktykami znajduje się małych obłoków neutralnego wodoru. Ponieważ takie obłoki powstają w niewielkich zagęszczeniach CM, obserwacje te dają nam wiedzę o minimalnej dopuszczalnej masie i liczbie małych zgęstek CM i sugerują, że CM może być co najwyżej letnia: dopuszczalne prędkości resztkowe są zbyt małe, by rozwiązać problem ZWBZ.

Trzecią popularną linią ataku na problem ZWBZ jest odwołanie się do fizyki galaktycznego gazu i gwiazd. Powstawaniu gwiazd towarzyszą bardzo energetyczne procesy, takie jak np. wybuchy supernowych. Gdyby w centrum małej galaktyki karłowatej wybuchło w stosunkowo krótkim czasie wiele supernowych, to rozgrzany i zjonizowany przez te wybuchy gaz uleciałby poza galaktykę. Zmniejszona w ten sposób energia potencjalna doprowadziłaby w efekcie do rozrzedzenia się profilu halo CM. Ilość energii, jaka jest potrzebna do znaczącego rozrzedzenia halo CM, jest jednak bardzo duża. Galaktyka, w której doszłoby do takiej niesłychanej kanonady, szybko utraciłaby cały gaz i nie mogąc wytworzyć nowych gwiazd, pozostałaby bardzo bleda. Tymczasem niektóre satelity DM mają jasności setek milionów słońc. Przystudiowanie innych procesów związanych z barionami, włączając w to aktywne supermasywne czarne dziury czy galaktyczne pola magnetyczne, dostarczyło dowodów, że nie

Reprodukcja rysunku 3 z pracy Boylan–Kolchin i in. 2012 (MNRAS 422, 1203–1218). Sześć paneli przedstawia krzywe rotacji  $V(r)$  satelitów zamieszkujących sześć różnych hal centralnych – analogów DM. U góry każdego panelu podana jest całkowita masa halo DM. Punkty ze słupkami błędów oznaczają pomiary dokonane dla 11 klasycznych satelitów DM. Na każdym rysunku grupa linii zaznaczona pogrubieniem odpowiada dużym i gęstym satelitom, które w żadnym punkcie swoich krzywych rotacji nie są zgodne z danymi.



są one w stanie zmniejszyć gęstości CM w galaktykach satelitarnych tak, by rozwiązać problem ZWBZ.

Wygląda więc na to, że żadne z proponowanych podejść nie oferuje w pełni satysfakcjonującego rozwiązania problemu brakujących gęstych satelitów w systemie Drogi Mlecznej. Na szczęście nie potrzebujemy wcale uciekać się do tak radykalnych postulatów, jak mała masa DM lub całkiem nowy model CM. Rozwiązanie zagadki kryje się w fizyce barionowej, a raczej w tym, że większość symulacji kosmologicznych jej nie uwzględnia. Grupa badaczy z Uniwersytetu w Durham pod kierownictwem Tilla Sawali użyła nowoczesnych kodów komputerowych projektu EAGLE do przeprowadzenia szeregu symulacji (projekt APOSTOLE) zaprojektowanych specjalnie do badania systemów, takich jak DM i Lokalna Grupa Galaktyk. Okazało się, że gdy poprawnie uwzględnimy procesy ogrzewające gaz galaktyczny, takie jak rejonizacja i supernowe, to nawet w sporych halach CM chwilowy ubytek masy na skutek „wywiania” gorącego gazu powoduje znaczące zmniejszenie się potencjału grawitacyjnego. Mimo że po jakimś czasie rozgrzany gaz ochładza się i powraca do galaktyk (co do tej pory uwzględniano teoretycznie w symulacjach bez barionów), to czasowe osłabienie grawitacji macierzystej galaktyki zauważalnie spowalnia tempo jej wzrostu. Ponieważ ostateczna masa skupiska CM równa się danemu tempu akrecji materii scałkowanemu względem kosmicznego czasu, to te same hale CM w symulacjach z pełną fizyką barionową ostatecznie odznaczały się całkowitą masą mniejszą o 10–20%. To zaś od razu przekładało się na zmianę krzywych  $V(R)$ , a zwłaszcza maksimum  $V_{\max}$ . Rysunek pokazuje, że obniżenie o 10–20% grupy krzywych niepasujących do danych w naturalny sposób sprawia, że większość z nich staje się zgodna z obserwacjami.

W ten oto sposób okazało się, że to nie model LCDM miał problemy z wyjaśnieniem obserwacji, tylko kosmologowie zmagali się z uzyskaniem rozsądnych przewidywań z założeń teoretycznych modelu. Mimo że ostatecznie wcale nie ma galaktyk zbyt wielkich, by zgasnąć, to nie jest to jeszcze koniec tej historii. Standardowy model kosmologiczny odnosi swoje główne sukcesy na największych kosmologicznych skalach rzędu setek milionów parseków. Skale znacznie mniejsze (rozmiarów galaktyk) wciąż ukrywają wiele tajemnic i nie jest pewne, czy LCDM sobie z nimi poradzi. Ciemna materia pozostaje ulotna: dotąd nie udało się zidentyfikować w naziemnych laboratoriach cząstki elementarnej za nią odpowiedzialnej. Nie wiemy zatem, czy CM jest zimna czy ciepła. Ostatnia dekada przyniosła wiele ciekawych obserwacji, które świadczą raczej na korzyść hipotezy ciepłej ciemnej materii, ale to już opowieść na inną okazję.

Projekt EAGLE (*Evolution and Assembly of Galaxies and their Environments*), czyli ewolucja i montaż galaktyk i ich środowisk, to seria nowoczesnych i dokładnych symulacji komputerowych, których celem jest zrozumienie, jak powstają i ewoluują galaktyki. Symulacje wyróżniają się implementacją (w uproszczeniu rzecz jasna) niemal wszystkich istotnych procesów fizycznych związanych z galaktykami i obejmują m.in. hydrodynamikę i ewolucję chemiczną ośrodka międzygwiazdowego, rejonizację i chłodzenie promieniste, cząsteczkowe i atomowe gazu, powstawanie gwiazd, galaktyczne wiatry powodowane przez supernowe, powstawanie oraz akrecja na supermasywne czarne dziury.



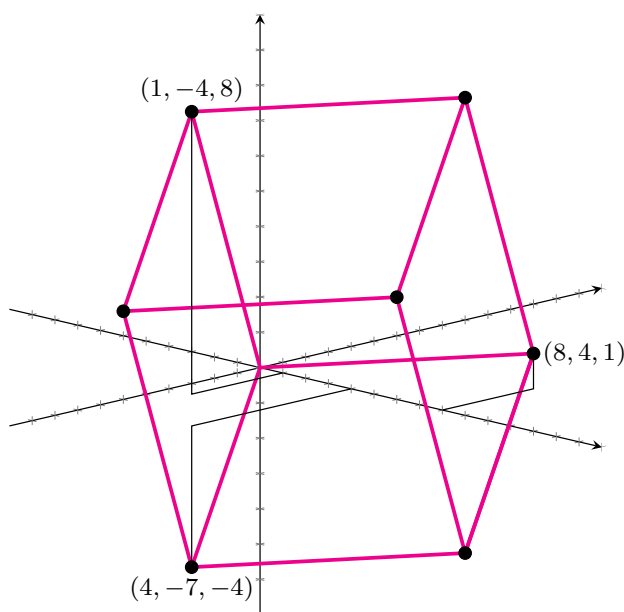


# mała delta

## Nieosiągalne!

- Lolek, chodź, pokażę Ci jedną stronę w Internecie...
- No, co tam masz? Bolek, wyłącz tego Reksia. Mama mówiła, że nie wolno nam oglądać konkurencji!
- Nie, nie to. Tutaj, w drugiej zakładce. Ktoś na tym forum napisał fajne zadanie. Czytaj.
- *Jeżeli sześcian w przestrzeni ma wszystkie wierzchołki w punktach kratowych, to jego krawędź ma całkowitą długość.*
- Zacząłbyś czytać jakieś porządne fora. Przecież to jest oczywista bzdura.
- Niby czemu?
- Bolek, nie widzisz, skąd to się wzięło? Ktoś pomyślał, że taki sześcian musi mieć krawędzie równoległe do osi współrzędnych. I wtedy to prawda. Ale przecież można wziąć obrócony sześcian, o, zaraz Ci taki jeden narysuję. Proszę! Ma wierzchołki w punktach kratowych? Ma!

- Loluś, ale on akurat ma krawędź całkowitej długości...
- No nie żartuj...  $\sqrt{8^2 + 4^2 + 1^2} = \sqrt{81} = 9$ ... rzeczywiście! Patrz, jak mi się przyfariło! Ale chyba widzisz, że można trochę poeksperymentować ze współzrędnymi i przykład wyjdzie. Poza tym *moralnie* to nie może być prawda.
- Moralnie?
- Tak, bo na płaszczyźnie to nieprawda. Zdanie *jeżeli kwadrat na płaszczyźnie ma wszystkie wierzchołki w punktach kratowych, to jego bok ma całkowitą długość* jest fałszywe: bez problemu można narysować taki kwadrat o boku  $\sqrt{5}$  albo nawet  $\sqrt{2}$ .
- I jak to się ma do sześcianu?
- Tak od razu to nijak. Ale mówiłem, że to moralny argument: skoro na płaszczyźnie analogiczne twierdzenie nie zachodzi, to co dopiero w przestrzeni, gdzie można dużo dowolniej manipulować współzrędnymi!
- Mnie te moralne argumenty jakoś nie przekonują. Na prostej wszystko gra.
- Na prostej?
- No tak, zobacz: *jeżeli odcinek na prostej ma oba końce w punktach kratowych, to jego długość jest liczbą całkowitą.*
- Teraz to ale Amerykę odkryłeś... Dobra, już dobra, skonstruuję Ci precyzyjny kontrprzykład.



- Jaką długość krawędzi byś chciał?
- To ja poproszę  $\sqrt{117} = 3\sqrt{13}$ .
- Ambitnie! Ale dobra, niech będzie. Więc tak: jeden wierzchołek będzie w punkcie  $0 = (0, 0, 0)$ , następny w jakimś punkcie odległym o  $\sqrt{117}$ , na przykład...  $u = (10, 4, 1)$ . Teraz zgadniemy sobie jakiś wektor  $v$  prostopadły do  $u$  i tej samej długości. Powiedzmy, no... czekaj... mam:  $v = (2, -7, 8)$ .
- Na pewno są prostopadłe?



– Nie wierzysz mi? Odległość między ich końcami wynosi

$$|u - v| = \sqrt{(10 - 2)^2 + (4 + 7)^2 + (1 - 8)^2} = \sqrt{234}$$

i mamy

$$|u|^2 + |v|^2 = \sqrt{117}^2 + \sqrt{117}^2 = \sqrt{234}^2 = |u - v|^2,$$

więc kąt między  $u$  i  $v$  jest prosty z twierdzenia Pitagorasa.

– Chyba z twierdzenia odwrotnego do...

– Cicho siedź! Pedant się znalazł. Poza tym łatwo sprawdzić, że Twoje pytanie sprowadza się do obliczenia *iloczynu skalarnego*

$$\langle u, v \rangle = u_1v_1 + u_2v_2 + u_3v_3 = 10 \cdot 2 + 4 \cdot (-7) + 1 \cdot 8 = 0.$$

Skoro wyszło 0, to kąt jest prosty. Uczyłem się o tym na matematyce.

Więc podsumujmy: punkty  $0, u, v$  i  $u + v = (12, -3, 9)$  wyznaczają kwadrat o boku  $\sqrt{117}$ . Zgadza się?

– Na razie dobrze.

– To będzie podstawa naszego sześcianu. Teraz wystarczy znaleźć trzeci wektor  $w$ , prostopadły do obu  $u$  i  $v$ . Ugh... to chyba potrwa. Przynies coś do picia.

– Czekaj, o tym to ja się nauczyłem na fizyce. To się nazywa *iloczyn wektorowy*  $u \times v$ . Podstawia się dwa wektory i wychodzi wektor prostopadły do nich obu, czyli w Twojej notacji  $\langle u \times v, u \rangle = \langle u \times v, v \rangle = 0$ . Chcesz wzór?

– Dawaj!

– Mam go tu w zeszycie. Uwaga:  $u \times v = (u_2v_3 - u_3v_2, u_3v_1 - u_1v_3, u_1v_2 - u_2v_1)$ .

– Trochę długi, ale to nic, podstawiamy:  $u \times v = (39, -78, -78)$ . To już, bierzemy wektory  $u, v$  i  $u \times v$  i rozpinamy na nich sześcian! A nie mówiłem!

– Za bardzo się rozpędziłeś. Ten Twój „sześcian” to tylko prostopadłościan. Trzeci wektor ma długość  $\sqrt{39^2 + (-78)^2 + (-78)^2} = \sqrt{13689} = 117$ , a nie  $\sqrt{117}$ . Zapomniałem Ci powiedzieć: długość wektora  $u \times v$  jest równa polu równoległoboku rozpiętego przez  $u$  i  $v$ .

– Czyli w naszym przypadku polu kwadratu:  $\sqrt{117} \cdot \sqrt{117} = 117$ . To by się nawet zgadzało. Ehhh... Ale zaraz, nie wystarczy przeskalować ten wektor? Kierunek jest przecież dobry, tylko długość się nie zgadza!

– Bingo! Mamy wektor o długości 117, a chcieliśmy  $\sqrt{117}$ , czyli trzeba każdą współrzędną podzielić przez  $\sqrt{117}$ ! No nie... to niedobrze. Wychodzi  $(\sqrt{13}, -2\sqrt{13}, -2\sqrt{13})$ . To nie jest punkt o współrzędnych całkowitych.

– Widzę właśnie. No trudno. Pierwsze koty za płoty. Weźmiemy jakieś inne  $u$  i  $v$ . A jak znowu się nie uda, to zmienimy ten  $\sqrt{117}$  na  $\sqrt{13}$  albo  $\sqrt{257}$  i na pewno w końcu znajdziemy dobry przykład. A tak w ogóle, to czemu ja mam się męczyć? Bolek, Ty miałeś informatykę, weź to zaprogramuj i niech komputer szuka. W jakim języku wy się tam uczycie?

– W COBOLu. Ale wiesz co, to chyba nie jest najlepszy pomysł...

– Bo co? Paru prostych pętli w COBOLu nie umiesz napisać? Nie bądź cienias!

– Nie no, umiem. Ale to chyba nic nie da, bo ja już widzę jak to się skończy.

– To weź mnie oświeć...

– Powiedzmy, że będziemy chcieli zrobić taką konstrukcję, biorąc za krawędź  $\sqrt{13}$  albo  $\sqrt{257}$ , albo w ogóle  $\sqrt{n}$  dla jakiegoś naturalnego  $n$ . Wybierzemy dwa prostopadłe wektory  $u$  i  $v$  długości  $\sqrt{n}$  i całkowitych współrzędnych tak jak Ty...

**Rozwiązanie zadania F 939.**

Siła oporu  $F_{op}$ , z jaką powietrze działa na ciało poruszające się z prędkością  $v$  wynosi

$$F_{op} = \frac{1}{2} C \rho S v^2,$$

gdzie  $\rho$  jest gęstością powietrza,  $S$  polem powierzchni przekroju ciała (prostopadłego do kierunku ruchu), a  $C$  współczynnikiem zależnym od kształtu ciała o wartości rzędu 1 (od 0,1 do 1). W dobrym przybliżeniu powietrze spełnia równanie stanu gazu doskonałego. Korzystając z tego równania dla danych zadania, otrzymujemy (dla temperatury w skali Kelvina  $T = 293\text{ K}$ ):

$$\rho = \frac{p\mu}{RT} \approx 1,2\text{ kg/m}^3.$$

Przyjmijmy, że szerokość ciała skoczka wynosi średnio 30 cm – w ramionach na pewno więcej, ale za to dla nóg mniej, co dla wzrostu 1,8 m daje  $S \approx 0,54\text{ m}^2$ . Maksymalna prędkość odpowiada sytuacji, gdy siła oporu  $F_{op} = mg$ . Po podstawieniu powyższych oszacowań do równania otrzymujemy

$$v = \sqrt{\frac{2mg}{C\rho S}} \approx 50 \frac{\text{m}}{\text{s}} = 180 \frac{\text{km}}{\text{h}}.$$

Podczas zawodów, przy skokach z wysokości około 2000 m po około 12 s skoczki osiągają stałe prędkości około 190 km/h bliskie otrzymanej w rozwiązaniu.

**Rozwiązanie zadania F 940.**

Podczas narastania indukcji pola magnetycznego w pierścieniu pojawia się siła elektromotoryczna  $\mathcal{E}$ . Zgodnie z prawem indukcji Faradaya mamy

$$\mathcal{E} = -\frac{d\Phi}{dt},$$

gdzie  $\Phi$  oznacza strumień indukcji przez powierzchnię cewki. Oznacza to, że w każdym punkcie pierścienia, stycznie do niego, na ładunki działa pole elektryczne o wartości  $E = \mathcal{E}/(2\pi r)$ , gdzie  $r$  jest promieniem pierścienia. Linowa gęstość ładunku na pierścieniu wynosi  $\rho = Q/(2\pi r)$ , a więc na odcinek  $dl$  pierścienia działa siła  $dF = \rho E dl$  i moment siły względem jego środka to  $dM = r dF = r\rho E dl$ . Całkowity moment siły „obracający” pierścień wynosi  $M = 2\pi r^2 \rho E$ . Ten moment siły nadaje pierścieniowi przyspieszenie kątowe  $d\omega/dt = M/I$ , gdzie  $I = mr^2$  jest momentem bezwładności pierścienia. Otrzymujemy więc

$$\begin{aligned} \frac{d\omega}{dt} &= -2\pi r^2 \frac{Q}{2\pi r} \frac{1}{2\pi r^3 m} \frac{d\Phi}{dt} = \\ &= -2\pi r^2 \frac{Q}{2\pi r} \frac{1}{2\pi r^3 m} \frac{\pi r^2 dB}{dt}. \end{aligned}$$

Ostatnia równość wynika z definicji strumienia jednorodnego pola  $B$  przez pole koła. Po uproszczeniu powtarzających się czynników dostajemy:

$$\frac{d\omega}{dt} = -\frac{Q}{2m} \frac{dB}{dt}$$

i ze względu na początkowe wartości  $\omega = 0$  oraz  $B = 0$ , ostatecznie otrzymujemy  $\omega = QB_0/(2m)$ .

– Wybierzmy albo i nie. Skąd wiesz, że zawsze się da?

– Nieważne, jak się nie da, to i tak nici z konstrukcji, więc powiedzmy, że się udało. Teraz obliczasz wektor  $u \times v$  i tak jak poprzednio, on ma współrzędne całkowite i długość  $\sqrt{n} \cdot \sqrt{n} = n$ . Czyli jest za długi, bo miało być  $\sqrt{n}$ . Więc musimy go skrócić  $\sqrt{n}$  razy, bo tylko tam może być trzeci wierzchołek sześciannu. . .

– Nie tylko tam, jeszcze po przeciwnej stronie!

– Teraz to Ty jesteś pedantem. Znak nic nie zmienia, bo tak czy inaczej wektor  $(u \times v)/\sqrt{n}$  nie ma współrzędnych całkowitych, tylko niewymierne, bo wszędzie się kręci ten nieszczęsny  $\sqrt{n}$ !

– Chyba że. . .

– Że co?

– . . . że  $\sqrt{n}$  jest całkowity. Na przykład  $\sqrt{25}$  albo  $\sqrt{36}$ !

– Czyli już! No bo tak jakby pokazaliśmy, że nie da się skonstruować kontrprzykładu. A to chyba to samo, co udowodnić twierdzenie?

– Tak myślę. Weź, zapiszmy to jakoś zgrabnie i możemy opublikować na tym forum.

– Nie ma sensu, tam już jest jedno rozwiązanie.

– Tam już jest. . . Co? Bolek, Ty mnie zapędzasz w kozi róg, a tam już czeka gotowe rozwiązanie?! Czeka, jak ja cię dorwę. . .

– Lolek, czekaj, a! To nie tak! Bo ja tego rozwiązania z forum tak do końca nie rozumiem! Ono jest jakieś magiczne. Twoje jest dużo lepsze.

– Dobra, przeczytajmy razem, co oni tam piszą. *Objętość sześciannu o krawędzi  $\sqrt{n}$  wynosi  $V = \sqrt{n}^3$ .* Na razie chyba wszystko jasne?

– Tak, tak. Ale patrz dalej: *Ponieważ wszystkie wierzchołki mają współrzędne całkowite, więc objętość  $V$  jest liczbą całkowitą.* Tego do końca nie rozumiem. Dalej już jest łatwo: *Z równości  $V^2 = n^3$  wynika, że  $n$  jest kwadratem liczby całkowitej, czyli krawędź  $\sqrt{n}$  jest całkowita.*

– Faktycznie magia. To jak to jest z tą objętością? Jest jakiś oczywisty powód, że sześciannu o wierzchołkach w punktach kratowych musi mieć całkowitą objętość?

– Trochę o tym myślałem i wydaje mi się, że to może być prawda nawet dla dowolnego równoległościanu. Bo, na przykład, na płaszczyźnie wszystko działa: wzór na pole równoległoboku rozpiętego przez wektory  $u = (u_1, u_2)$  i  $v = (v_1, v_2)$  to  $|u_1 v_2 - u_2 v_1|$ , więc dla współrzędnych całkowitych wynik jest liczbą całkowitą. W przestrzeni też powinien być podobny wzór na objętość. Prawda?

– Wiesz co, Bolek?

– No co?

– Myślałem, że Ciebie nie ruszają takie *moralne* argumenty.

1. Zapisz zgrabnie rozwiązanie Bolka i Lolka.
2. Czy Bolek mógłby wybrać takie  $\sqrt{n}$ , żeby Lolek potknął się już przy próbie znalezienia wektora  $u$ ?
3. Czy Bolek mógłby wybrać takie  $\sqrt{n}$ , żeby Lolek nie mógł dobrać do  $u$  żadnego wektora  $v$ ?
4. Znajdź wzór na objętość równoległościanu rozpiętego przez trzy wektory. Podpowiedź: wyznacznik. Korzystając z niego, uzasadnij ostatnią obserwację Bolka.
5. Czy rozwiązanie Bolka i Lolka i rozwiązanie znalezione na forum są istotnie różne, czy może to tylko dwa wcielenia tego samego pomysłu?

*Małą Deltę przygotował Michał ADAMASZEK*

# Dlaczego problem $P \stackrel{?}{=} NP$ jest tak trudny?

Wojciech CZERWIŃSKI

24 maja 2000 roku Instytut Matematyczny Claya ogłosił listę siedmiu Problemów Milenijnych, czyli zagadnień, które zostały uznane za najważniejsze otwarte problemy matematyczne opierające się rozwiązaniom od lat. Wśród nich był jeden problem zaliczany do informatyki teoretycznej, o którym wielu Czytelników zapewne słyszało. Chodzi oczywiście o tytułowy problem: „Czy  $P = NP$ ”? Jest on powszechnie uznawany za najważniejsze pytanie informatyki teoretycznej.

Problem nazywamy *decyzyjnym*, jeśli odpowiedź to „tak” lub „nie” (a nie np. liczba). Klasa  $P$  (od *Polynomial time*) zawiera te problemy, które dadzą się rozwiązać w czasie wielomianowym. Definicja klasy  $NP$  (od *Nondeterministic Polynomial time*) jest nieco bardziej złożona. Powiemy, że problem  $L$  należy do  $NP$ , o ile dla każdego możliwego wejścia  $w$ , na które odpowiedź jest „tak”, istnieje *świadek* tej pozytywnej odpowiedzi  $v$ , co najwyżej wielomianowo większy niż  $w$ , taki, że gdy dostaniemy parę  $(w, v)$ , to możemy w czasie wielomianowym potwierdzić, że istotnie odpowiedź na  $w$  to „tak”. Intuicyjnie więc  $P \stackrel{?}{=} NP$  pyta mniej więcej, czy każdy problem, dla którego można sprawdzić poprawność odpowiedzi szybko (w czasie wielomianowym), można również rozwiązać szybko (w czasie wielomianowym). Już w latach 50. problem ten pojawiał się w dyskusjach, ale precyzyjnie zdefiniował go w 1971 roku Stephen Cook. Od tej pory jest otwarty, mimo że naprawdę wiele osób próbowało go rozwiązać.

Dlaczego jednak uważamy ten problem za aż tak ważny i ciekawy? Przecież jest wiele starych problemów, których nikt jeszcze nie rozwiązał. Można na to pytanie odpowiedzieć na kilka sposobów. Po pierwsze, jest bardzo wiele praktycznych problemów, o których wiemy, że należą do klasy  $NP$ , jednak nie znamy dla nich żadnego algorytmu wielomianowego. Często rozważanym przykładem jest tzw. *problem komiwojażera*, w którym mamy dany graf o  $n$  wierzchołkach (reprezentujących miasta), a na każdej krawędzi liczbę, która oznacza, ile czasu potrzebuje komiwojażer, aby przejechać między tymi właśnie miastami. Pytanie brzmi, czy może on objechać wszystkie miasta i wrócić do domu w zadanym czasie (lub szybciej). Łatwo zauważyć, że istotnie należy on do  $NP$ , świadkiem jest tutaj samo rozwiązanie, czyli cykl o wystarczająco małym czasie przejazdu. Innym przykładem, ważnym w dalszej części artykułu, jest problem *spełnialności formuł* (znany jako *SAT* od angielskiego *satisfiability* – spełnialność). Rozważamy w tym przypadku formuły logiczne zawierające  $n$  zmiennych:  $x_1, \dots, x_n$ . Możemy założyć, że formuła jest zawsze koniunkcją klauzul  $C_i$ , czyli postaci  $C_1 \wedge \dots \wedge C_k$ . Każda klauzula natomiast jest alternatywą zmiennych  $x_i$  lub ich negacji  $\neg x_i$ . Przykładowa formuła wygląda więc tak:  $(\neg x_1 \vee x_2 \vee x_3 \vee x_4) \wedge (x_2 \vee \neg x_3 \vee \neg x_4) \wedge (x_1 \vee \neg x_2) \wedge (\neg x_1 \vee x_3 \vee x_4)$ . Pytanie brzmi, czy istnieją takie wartości  $x_1, \dots, x_n \in \{0, 1\}$ , dla których formuła jest prawdziwa. W naszym przykładzie istnieją: są to, na przykład,  $x_1 = x_3 = 1$  oraz  $x_2 = x_4 = 0$ . *SAT* należy do  $NP$ , bo świadkiem spełnialności jest tu po prostu takie wartościowanie zmiennych, które spełnia formułę. A więc, tak jak i wcześniej, świadkiem jest obiekt, o którego istnieniu pytamy. Jest to częste zjawisko, ale wcale nie zawsze świadek musi być takiej postaci. Wiele praktycznych problemów da się przepisać jako problem spełnialności formuły powyższej postaci, więc bardzo przydatne byłoby umieć rozwiązywać je szybko, najlepiej w czasie wielomianowym. Niestety, nie znamy takich metod, a najlepsze znane algorytmy działają w czasie wykładniczym. Znaczący to, że nie są istotnie lepsze od bezmyślnego podstawiania wszystkich możliwych  $2^n$  wartości na zmienne  $x_i$ . Co ciekawe, oba powyższe problemy są *NP-zupełne*, co oznacza, że każdy inny problem z klasy  $NP$  da się przeformułować na *SAT* albo na problem komiwojażera niewiele (co najwyżej wielomianowo) większej wielkości. Oznacza to, że gdybyśmy umieli rozwiązać jakiś problem *NP-zupełny* (np. *SAT*), to wszystkie inne problemy z klasy  $NP$  rozwiązalibyśmy również

Równoważnie definiuje się  $P$  jako zbiór problemów rozpoznawalnych przez deterministyczne maszyny Turinga w czasie wielomianowym, a  $NP$  jako zbiór problemów rozpoznawalnych przez niedeterministyczne maszyny Turinga w czasie wielomianowym.





w czasie wielomianowym (poprzez sformułowanie w postaci SATa, a potem rozwiązanie go). A to oznaczałoby, że  $P = NP$ . Ta właśnie cecha problemów NP-zupełnych świadczy o ich znaczeniu: wystarczy rozwiązać jeden z nich, a wiemy, że  $P = NP$ , z drugiej strony, oczywiście, jeśli tylko jeden z nich nie należy do P, to  $P \neq NP$ .

Gdybyśmy umieli rozwiązywać SAT w czasie wielomianowym, to wiele problemów rozwiązalibyśmy dużo szybciej. To dobrze. Jednak byłyby też złe strony tej sytuacji. Większość szyfrów opiera się na tym, że pewne obliczenia wymagają dużo czasu i nie są znane metody istotnie lepsze od sprawdzania wszystkich możliwych wartości (jak wyżej dla wartości zmiennych  $x_i$ ). Gdyby jednak  $P = NP$ , to te obliczenia już nie potrzebowałyby tyle czasu i większość szyfrów przestałaby być bezpieczna. To źle. To kolejna motywacja do zrozumienia, czy  $P = NP$ , czy też nie. I to trochę innego rodzaju, bo w tym przypadku, jeśli zrozumiemy, że  $P \neq NP$ , to też mamy zysk: wiemy, że niektóre szyfry są chociaż trochę bezpieczne. Te dwie motywacje nie są jedynymi powodami, dla których chcemy znać odpowiedź. To tylko poszlaki, że problem jest naprawdę ważny i fundamentalny. Być może najważniejszą przyczyną jest przekonanie, które stoi u źródeł nauk podstawowych, że zrozumienie tak istotnego problemu, niezależnie od wyniku, może skutkować nieprzewidywalnymi wręcz korzyściami. Tak jak to było z odkryciem przez Watsona i Cricka, że DNA ma strukturę podwójnej helisy.

Dochodzimy teraz do pytania, o którym tak naprawdę chcę opowiedzieć, czyli: skoro uważamy ten problem za tak ważny i tylu ludzi bardzo się stara, to dlaczego wciąż jest nierozwiązany? Można by uznać, że jest to pytanie głupie: jak to dlaczego – jest po prostu trudny. Nie pozostawiamy jednak na tym poziomie odpowiedzi, bo historia jest dużo bardziej subtelna. Już przez ponad 45 lat ludzie próbowali rozwiązać problem, pewne metody rozwijały się, ale nie udawało się nimi rozwiązać  $P \stackrel{?}{=} NP$ . Więc zadawali sobie pytanie: „dlaczego nam się nie udaje?”. I często tym sposobem dochodzili do wniosku, że pewne metody po prostu nie mogą działać. Historia tych badań jest fascynująca i może nam bardzo wiele powiedzieć o teorii złożoności, a także chociaż trochę wytłumaczyć, dlaczego  $P \stackrel{?}{=} NP$  jest aż tak trudny i jakie bariery stoją na drodze.

Wszystko zaczęło się w latach 70. Po tym, jak Cook zadał pytanie, matematycy próbowali je rozwiązać. Naturalną techniką była wtedy tak zwana metoda *diagonalizacji*. Wiele pierwszych rezultatów z teorii złożoności (czyli dziedziny zajmującej się klasyfikacją problemów na bardziej złożone i mniej złożone) używa właśnie tej metody. Użyta jest np. w dowodach nierozstrzygalności problemu stopu, twierdzenia (dużo starszego) o tym, że liczb rzeczywistych jest więcej niż naturalnych, albo mniej znanych twierdzeń o hierarchii czasowej i pamięciowej. Niech *język* będzie zbiorem słów nad pewnym ustalonym alfabetem, czyli ciągów liter z tego alfabetu. Rozważmy maszyny, które dostają słowo, wykonują jakieś obliczenie i odpowiadają „tak” lub „nie”, typowy przykład to maszyny Turinga, być może z jakimiś ograniczeniami (jak działanie w czasie wielomianowym). Zbiór słów, dla których maszyna  $M$  odpowie „tak”, nazywamy językiem tej maszyny, oznaczamy go  $L(M)$ . Rozważmy teraz dwie klasy maszyn  $C_1$  i  $C_2$ . Aby wykazać, że maszyny z  $C_2$  opisują pewne języki, które nie są opisywane przez maszyny z  $C_1$ , wystarczy wskazać taką maszynę  $M \in C_2$ , że dla dowolnej maszyny  $M' \in C_1$  istnieje takie słowo  $w$ , iż maszyny  $M$  i  $M'$  potraktują je inaczej (jedna z nich powie „tak”, a druga „nie”). Tę właśnie metodę nazywamy diagonalizacją. Okazuje się, że większość dowodów, które używają tej metody, podlega *relatywizacji*. Aby powiedzieć, co to oznacza, opowiemy najpierw o wyroczni. Rozważmy jakąś wyrocznię  $O$ , która odpowiada „tak” lub „nie” na pewne pytania poprawnie i za darmo. Możemy dodać maszynie możliwość zadawania pytań do takiej wyroczni. Przykładowo wyrocznia mogłaby odpowiadać na pytanie, czy dana liczba jest potęgą liczby pierwszej. Widać wtedy, że maszyna, która miałaby za zadanie stwierdzić, czy dana liczba jest pierwsza, miałaby, być może, ułatwione zadanie, gdyby mogła

Więcej o związkach problemu  $P = NP$  z kryptologią można przeczytać w artykule Tomasza Kazany, w poprzednim numerze *Delty*.



#### Rozwiązanie zadania M 1546.

Udowodnimy, że taka liczba  $n$  nie istnieje.

Przyjmijmy oznaczenie  $f(x) = x^2 + x + 1$ . Bezpośrednio sprawdzamy, że dla dowolnego  $x$  zachodzi równość

$$f(x)f(x-1) = f(x^2).$$

Ponadto, jeżeli  $x$  jest liczbą całkowitą różną od 0, to

$$x^2 < f(x) < (x+1)^2,$$

więc  $f(x)$  nie jest kwadratem liczby całkowitej.

Udowodnimy, że dla każdego  $n \geq 1$  zachodzi równość

$$(*) \quad \prod_{k=1}^n f(a_k) = f(a_n^2),$$

co w połączeniu z obserwacją z poprzedniego akapitu zakończy dowód.

Przeprowadzimy dowód indukcyjny. Dla  $n = 1$  równość (\*) jest spełniona.

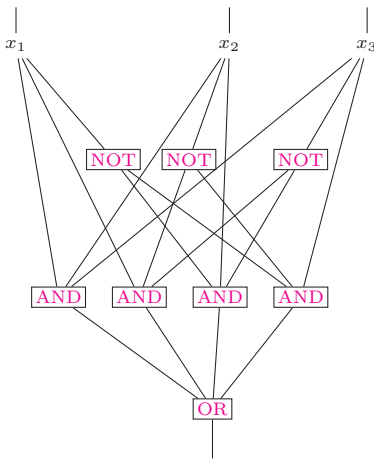
Przypuśćmy, że zachodzi ona dla pewnego  $n \geq 1$ ; wówczas

$$\begin{aligned} \prod_{k=1}^{n+1} f(a_k) &= f(a_{n+1}) \prod_{k=1}^n f(a_k) = \\ &= f(a_{n+1})f(a_n^2) = \\ &= f(a_{n+1})f(a_{n+1} - 1) = \\ &= f(a_{n+1}^2), \end{aligned}$$

co oznacza, że (\*) zachodzi również dla  $n + 1$ . To kończy rozwiązanie.

pytać taką wyrocznię. Klasę maszyn  $\mathcal{C}$  z możliwością korzystania z wyroczni  $O$  oznaczmy przez  $\mathcal{C}^O$ . Powiemy, że dowód, iż maszyny z  $\mathcal{C}_2$  rozpoznają więcej języków niż maszyny z  $\mathcal{C}_1$ , się relatywizuje, jeśli dla dowolnej wyroczni  $O$  ten dowód pokazuje również, że maszyny z  $\mathcal{C}_2^O$  rozpoznają więcej języków niż maszyny z  $\mathcal{C}_1^O$ . Jak powiedzieliśmy, zasadniczo dowody używające diagonalizacji się relatywizują. I dlatego przełomem było twierdzenie, które w roku 1975 udowodnili Baker, Gill i Solovay. Pokazali oni, że istnieje taka wyrocznia  $A$ , że  $P^A = NP^A$  oraz taka wyrocznia  $B$ , że  $P^B \neq NP^B$ . Oznacza to, że żaden dowód, który się relatywizuje, na pewno nie pokaże, że  $P \neq NP$  (bo jest wyrocznia  $A$ ) ani że  $P = NP$  (bo jest wyrocznia  $B$ ). Jeśli chcemy rozwiązać  $P \stackrel{?}{=} NP$ , to musimy iść inną drogą.

Ten wynik spowodował, że społeczność badaczy teorii złożoności zarzuciła metody diagonalizacji przy rozwiązywaniu problemu  $P \stackrel{?}{=} NP$ . W zasadzie słusznie, ale ruch ten był, być może, zbyt radykalny. Można bowiem wyobrazić sobie, że diagonalizacja jest tylko jednym z argumentów w dowodzie i wówczas cały dowód się nie relatywizuje. W każdym razie mniej więcej w tamtym czasie popularne zaczęły być obwody logiczne, to z nimi wiązano teraz nadzieję na



postęp w kwestii problemu  $P \stackrel{?}{=} NP$ . Obwód logiczny o  $n$  wejściach i jednym wyjściu oblicza funkcję  $f : \{0, 1\}^n \rightarrow \{0, 1\}$ . Wszystko można zakodować jako ciąg zer i jedynek, natomiast jedynek na wyjściu traktować jako odpowiedź „tak”, a zero jako odpowiedź „nie”. A zatem obwód logiczny jest dobrą alternatywą dla maszyn i można też używać go do opisu języków. Pomiedzy  $n$  wejściami a wyjściem obwód ma bramki jednego z trzech typów: AND, OR oraz NOT. Bramki AND i OR mogą mieć wiele wejść, a wyjście mają jedno, na którym zwracają koniunkcję (AND) lub alternatywę (OR) wejść. Bramka NOT ma jedno wejście i jedno wyjście, na którym wypuszcza zanegowane wejście. Można składać z tych bramek dowolnie skomplikowane funkcje, na rysunku przedstawiono obwód dla funkcji  $\text{xor} : \{0, 1\}^3 \rightarrow \{0, 1\}$ , która zwraca sumę wejść modulo dwa. Co prawda, każdą funkcję można opisać pewnym obwodem (Ambitnego Czytelnika zachęcamy do wykazania tego faktu), ale już gdy ograniczymy pewne parametry obwodów, to nie wszystkie funkcje dadzą się opisać i sytuacja staje się znacznie ciekawsza. Przykładowo jeden z ważniejszych wyników tamtego okresu (1983 rok) to dowód, że ciąg funkcji  $\text{xor} : \{0, 1\}^n \rightarrow \{0, 1\}$ , które zwracają sumę wejść modulo dwa, nie może być opisany ciągiem obwodów (dla coraz większej liczby zmiennych) o stałej głębokości, czyli długości najdłuższej ścieżki od wejścia do wyjścia (czyli  $\text{xor}$  nie należy do klasy  $AC^0$ ). Tego typu rezultatów niemożliwości dowodzone wtedy więcej i wierzone, że postęp w tych technikach może doprowadzić do wykazania, że  $P \neq NP$ . Oczekiwany dowód miałby podążać mniej więcej po następującej linii: 1) definiujemy jakąś sprytną miarę skomplikowania funkcji boolowskiej, 2) wykazujemy, że miara funkcji SAT (czyli przypisującej kodowaniom zero-jedynkowym spełnialnych formuł jedynek, a reszcie zero) lub jakiejś innej funkcji z klasy NP jest duża, 3) pokazujemy, że wszystkie obwody wielomianowej wielkości opisują funkcje o małej mierze. Ponieważ wiemy, że każdą funkcję z P można zapisać wielomianowym obwodem (w takim obwodzie można zakodować algorytm wielomianowy, który jest wykonywany przez bramki), to punkty 1)–3) pokazywałyby, że  $P \neq NP$ .

Niestety, (a może raczej na szczęście) i tu pojawiła się nieoczekiwana trudność. W roku 1994 Razborov i Rudich przedstawili swoją pracę na temat „dowodów naturalnych”, w której pokazywali, że takiej miary nie można zdefiniować, o ile tylko prawdziwa jest powszechnie uznawana za prawdziwą hipoteza na temat istnienia pseudolosowych funkcji. Główne idee ich dowodu, mimo, że przełomowe i bardzo pomysłowe, dadzą się pokrótce naszkicować. Autorzy pokazali też, że wszystkie wcześniejsze dowody różności klas złożoności korzystały z pewnych własności funkcji, które to własności nazwali naturalnymi. Powiemy, że własność  $\phi$  funkcji jest *naturalna*, jeśli spełnia trzy warunki: i) co najmniej  $1/2^n$  funkcji  $f : \{0, 1\}^n \rightarrow \{0, 1\}$  spełnia  $\phi$ , ii) jeśli  $f$  spełnia  $\phi$ , to nie ma obwodu wielomianowej wielkości, który ją opisuje, iii) spełnianie  $\phi$



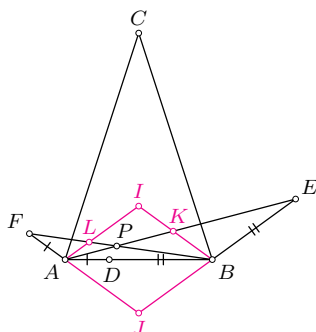
### Rozwiązanie zadania M 1547.

Oznaczmy środek okręgu wpisanego w trójkąt  $ABC$  przez  $I$ , a punkt symetryczny do niego względem prostej  $AB$  przez  $J$  – wówczas czworokąt  $AIBJ$  jest rombem. Niech ponadto  $K = AE \cap BI$  oraz  $L = BF \cap AI$ .

Zauważmy, że

$$\sphericalangle JBD + \sphericalangle DBE = \sphericalangle ABI + 2 \cdot \sphericalangle ABC = 36^\circ + 2 \cdot 72^\circ = 180^\circ,$$

wobec czego punkty  $J, B, E$  są współliniowe. Analogicznie dochodzimy do wniosku, że punkty  $J, A, F$  są współliniowe.



Z równoległości par prostych  $AI \parallel BE$  oraz  $BI \parallel AF$  wynika, że

$$\frac{IL}{AL} = \frac{IB}{AF} \quad \text{oraz} \quad \frac{BK}{IK} = \frac{BE}{IA}.$$

W połączeniu z  $AD = AF$ ,  $BD = BE$  oraz  $IA = IB$  otrzymujemy więc

$$\frac{AD}{BD} \cdot \frac{BK}{IK} \cdot \frac{IL}{AL} = 1,$$

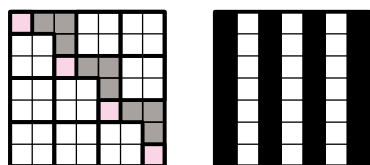
a zatem z twierdzenia odwrotnego do twierdzenia Cevy dla trójkąta  $ABI$  wynika, że proste  $AK, BL, ID$  przecinają się w jednym punkcie. Tym samym punkt  $P$  leży na prostej  $ID$ , co jest równoważne tezie zadania.



### Rozwiązanie zadania M 1548.

Odpowiedź:  $n = 29$ .

Przypuśćmy, że pola tablicy pokolorowano tak, że każdy kwadrat  $2 \times 2$  zawiera co najwyżej dwa czarne pola. Wówczas każdy z 9 białych kwadratów  $2 \times 2$  oraz każda z 3 szarych figur o polu 3 na rysunku z lewej zawiera co najwyżej 2 czarne pola, a ponadto każdy z 4 kolorowych kwadratów jednostkowych może być czarny. To oznacza, że łącznie w tablicy jest co najwyżej  $9 \cdot 2 + 3 \cdot 2 + 4 = 28$  czarnych pól.



Z kolei kolorowanie 28 pól przedstawione na drugim rysunku ma tę własność, że każdy kwadrat  $2 \times 2$  zawiera dokładnie dwa czarne pola. To oznacza, że szukana najmniejsza wartość  $n$  jest równa 29.

daje się obliczyć w czasie  $2^{O(n)}$ . Gwoli ścisłości, definicja naturalności zależy jeszcze od tego, jakie klasy złożoności chcemy rozróżniać, ale my tu skupimy się na powyższej definicji. Gdybyśmy chcieli realizować plan naszkicowany w poprzednim akapicie, to funkcja spełniałaby  $\phi$ , gdyby miała dużą miarę skomplikowania – autorzy przedstawili przekonujące argumenty, że dla każdej rozsądnej miary rzeczywiście spełnione byłyby punkty i), ii), iii). Niemniej jednak, co było głównym wynikiem pracy, o ile tylko istnieją funkcje pseudolosowe, to nie istnieje żadna naturalna własność  $\phi$  (a więc nie użyjemy jej w dowodzie). Funkcje pseudolosowe to takie, które wyglądają jak losowe z punktu widzenia obserwatora o ograniczonej mocy obliczeniowej. Idea dowodu jest następująca. Weźmy dowolną funkcję, kandydata na funkcję pseudolosową. Losowa funkcja spełnia własność  $\phi$  z prawdopodobieństwem co najwyżej  $1/2^n$ , co wynika z i). Natomiast funkcja pseudolosowa musi być prosta w opisie, więc ma obwód wielomianowej wielkości (z ii)), czyli spełnia  $\phi$  z prawdopodobieństwem 0. A zatem z prawdopodobieństwem co najmniej  $1/2^n$  własność  $\phi$  rozróżnia funkcję losową od naszej, a to, czy funkcja spełnia  $\phi$ , można obliczyć w czasie  $2^{O(n)}$  (z iii)). Więc nasza funkcja nie jest pseudolosowa, bo można ją odróżnić od losowej z sensownym prawdopodobieństwem i w odpowiednio ograniczonym czasie. Nie ma więc żadnej takiej – sprzeczność z powszechnie uznawaną hipotezą. Razborov i Rudich w 2007 roku otrzymali nagrodę Gödla za ten rezultat, przez wielu uznawany za najważniejszy wynik dotyczący problemu  $P \stackrel{?}{=} NP$ .

Również i ten wynik istotnie zmienił społeczność informatyków. Ludzie odwrócili się od obwodów logicznych, przestali dowodzić dolne ograniczenia dla nich, bo jak pokazuje powyższy wynik – nie tędy droga. Być może znów zrobiono to zbyt pochopnie, a barierę należy raczej traktować jako drogowskaz, którego nie iść, tzn. żeby nie definiować naturalnych własności w swoich dowodach. Niedawno zaczęto stosować pewne techniki omijające powyższe bariery i używające algebry. Żeby zrozumieć ideę, przyjrzyjmy się szkicowi dowodu wspomnianego wyżej faktu, że xor nie należy do klasy  $AC^0$ . Dowód bada, jak dobrze można przybliżyć funkcję  $f : \{0, 1\}^n \rightarrow \{0, 1\}$  wielomianem wielu zmiennych nad ciałem  $F_3$  o stopniu co najwyżej  $\sqrt{n}$ . Przez przybliżenie mamy na myśli zwrócenie takiego samego wyniku w większości przypadków. Okazuje się, że każdą funkcję należącą do klasy  $AC^0$  można przybliżyć całkiem niezłe, natomiast funkcji xor nie da się przybliżyć dobrze. Akurat ten dowód nie omija bariery dowodów naturalnych, ale stosując podobne techniki używające wielomianów niskich stopni nad skończonymi ciałami lub pierścieniami, można ominąć obie wspomniane bariery. Ale i ta technika ma ograniczenia; niedawna praca Aaronsona i Wigdersona z roku 2009 wprowadza nową barierę zwaną algebraizacją. Jest to uogólnienie relatywizacji. Inkluzja  $C_2 \subseteq C_1$  się algebraizuje, jeśli  $C_2^A \subseteq C_1^A$ , dla pewnej wyroczni  $A$  oraz wyroczni  $\bar{A}$ , która powstała z  $A$  poprzez rozszerzenie z funkcji boolowskich do wielomianów niskich stopni nad skończonym ciałem (dla pierścieni to też działa). Jeśli się algebraizuje, to pewnymi technikami bazującymi na wielomianach niskich stopni (tzw. algebraizującymi się) nie da się jej obalić, podobnie jak fakt, iż  $NP^A = P^A$  pokazuje, że technikami, które się relatywizują, nie da się obalić inkluzji  $NP \subseteq P$ . Autorzy wykazali też, że inkluzja  $NP \subseteq P$  się algebraizuje, czyli techniki, które się algebraizują, nie pozwolą na jej obalenie. Oznacza to, że piłeczka poszukiwaczy dowodu  $P \neq NP$  po raz kolejny została odbita.

Jaka jest przyszłość problemu  $P \stackrel{?}{=} NP$ ? Nie wiadomo. Większość badaczy twierdzi, że  $P \neq NP$ , ale są też poważani naukowcy, którzy twierdzą, że wręcz przeciwnie,  $P = NP$ . Najprawdopodobniej nieprędko się to wyjaśni, bo, jak widać, nie bez przyczyny problem ten trzyma się długo, kolejne bariery rzucają nam kłody pod nogi. Musimy uzbroić się w cierpliwość. Albo sami ruszyć do starcia z niedostępnym do tej pory problemem. Tylko pamiętajmy, że jeśli zdecydujemy się na to ostatecznie, to należy wcześniej mądrze i gruntownie przygotować się do ataku!

## Paradoks Russella

W miejscowości  $M$  jest fryzjer, nazwijmy go *superfryzjerem*, który strzyże tych i tylko tych mieszkańców miejscowości, którzy nie strzygą siebie samych. Czy superfryzjer strzyże siebie samego? Chwila namysłu pokazuje, że obie możliwości są wykluczone: nie może on strzyć siebie samego, bo strzyże tylko tych, którzy siebie sami nie strzygą; gdyby zaś sam się nie strzygł, to musiałby się strzyć, bo strzyże wszystkich tych, którzy sami się nie strzygą. A zatem, superfryzjer nie może istnieć! Pokażemy jak z powyższego faktu otrzymać różne twierdzenia matematyczne, odpowiednio definiując mieszkańców miejscowości  $M$  oraz to, kto kogo strzyże. Część tych wyników była opisana w artykule Wojciecha Czerwińskiego (*Delta* 10/2016) poświęconym metodzie przekątniowej, bardzo blisko związanej z paradoksem Russella.

**Paradoks Russella.** Bertrand Russell używał historii o fryzjerze, by ilustrować następujący problem dotyczący podstaw matematyki, odkryty przez niego w 1901 roku, i mający wielki wpływ na rozwój tej dziedziny. Czym jest zbiór? Chcielibyśmy, by elementami zbiorów mogły być inne zbiory; na przykład, okrąg to zbiór punktów na płaszczyźnie, i można rozważać zbiór wszystkich tych zbiorów, które są na płaszczyźnie okręgami o promieniu 1. Ogólniej, można by oczekiwać, że dowolne określenie postaci „zbiór wszystkich zbiorów o własności  $P$ ”, gdzie  $P$  jest precyzyjnie określoną własnością zbiorów, definiuje pewien zbiór. Pokażemy, że tak być nie może. Mieszkańcami  $M$  niech będą wszystkie zbiory, i powiemy, że zbiór  $x$  strzyże zbiór  $y$ , jeżeli  $y$  należy do  $x$ . Brak superfryzjera oznacza tyle, że określenie „zbiór wszystkich zbiorów, które nie należą do siebie samych” nie może określać żadnego zbioru.

**Twierdzenie Cantora.** Paradoks Russella był zainspirowany dziesięć lat starszą metodą przekątniową Georga Cantora. Cantor zastanawiał się, które zbiory są *przeliczalne*, tzn. wszystkie ich elementy można wypisać w nieskończonym ciągu  $x_1, x_2, x_3, \dots$ . Oczywiście, zbiór liczb naturalnych  $\{1, 2, 3, \dots\}$  jest przeliczalny, ale też zbiór liczb wymiernych w przedziale  $(0, 1)$  też jest przeliczalny:  $\frac{1}{2}, \frac{1}{3}, \frac{2}{4}, \frac{1}{4}, \frac{3}{5}, \frac{2}{5}, \frac{4}{6}, \frac{1}{6}, \frac{5}{7}, \frac{2}{7}, \frac{3}{7}, \frac{4}{7}, \frac{5}{8}, \frac{6}{8}, \frac{1}{8}, \frac{3}{8}, \frac{5}{8}, \frac{7}{8}, \dots$ , również zbiór wszystkich liczb wymiernych jest przeliczalny (ćwiczenie dla Czytelnika). Z kolei, zbiór wszystkich liczb rzeczywistych, nawet tych w przedziale  $(0, 1)$ , nie jest przeliczalny, co właśnie wykazał Cantor, i my teraz znowu wykazemy. Każda liczba rzeczywista z przedziału  $(0, 1)$  jest określona przez nieskończony ciąg cyfr, np. 0,14451323... (przy czym nie rozważamy zapisów, które od pewnego momentu mają same cyfry 9).

Niech  $c_1, c_2, c_3, \dots$  będzie ciągiem liczb rzeczywistych z przedziału  $(0, 1)$ . Zdefiniujmy liczbę  $x \in (0, 1)$  następująco: jej  $n$ -ta cyfra to 1, jeśli  $n$ -ta cyfra liczby  $c_n$  jest równa 0, i 0 w przeciwnym przypadku. Wykażemy, że  $x$  nie pojawia się w ciągu: gdyby  $x = c_k$  dla pewnego  $k$ , to  $k$  byłoby superfryzjerem w miejscowości  $M$  zamieszkałej przez liczby naturalne, w której  $m$  strzyże  $n$  wtedy, i tylko wtedy, gdy  $n$ -ta cyfra liczby  $c_m$  to 1. Sprzeczność.

**Twierdzenie Turinga.** Wybierzmy swój ulubiony język programowania, np. Java, C++, Pascal czy Python (w oryginalnym dowodzie z 1936 r. Alan Turing użył bardzo prostego „języka”, zwanego dziś *maszyną Turinga*). Kod źródłowy programu jest to napis (używający znaków dostępnych na klawiaturze komputera), który jest zgodny

ze składnią wybranego języka. Rozważmy programy, których kod źródłowy jest szczególnej postaci (\*): pierwszą wykonywaną instrukcją jest „wczytaj napis  $N$  wprowadzony przez użytkownika”, potem następuje ciąg instrukcji bez interakcji z użytkownikiem, a ostatnią instrukcją programu jest wypisanie słowa „koniec”. Z punktu widzenia nieśmiertelnego użytkownika, który uruchamia taki program i wprowadza napis  $N$ , są dwa możliwe scenariusze: albo program po pewnym czasie napisze „koniec” – mówimy wtedy, że program *akceptuje* napis  $N$  – albo nigdy nic nie napisze, tzn. „zawiesi się”. Dla wielu programów, wynik działania na danym napisie  $N$  bardzo łatwo przewidzieć. Można by pokusić się o napisanie programu  $Q$ , który dostaje na wejściu kod źródłowy dowolnego programu  $P$  postaci (\*) oraz napis  $N$ , po czym napisze „wiesza się”, jeżeli program  $P$  się wiesza dostawszy na wejściu  $N$ , oraz napisze „akceptuje” w przeciwnym przypadku. Twierdzenie Turinga, które teraz udowodnimy, mówi, że taki program  $Q$  nie może istnieć. Gdyby istniał, to skonstruowalibyśmy program  $F$  który, dostawszy kod dowolnego programu  $P$ , akceptuje go wtedy, i tylko wtedy, gdy program  $P$  uruchomiony na wejściu  $P$  się wiesza, co można stwierdzić, uruchamiając program  $Q$  na parze  $P, P$ . Wtedy  $F$  byłby superfryzjerem w miejscowości  $M$  zamieszkałej przez kody źródłowe postaci (\*), w której  $P$  strzyże  $K$  jeśli program  $P$  akceptuje  $K$ . Sprzeczność.

**Twierdzenie Gödla o niezupełności.** W uproszczeniu, ten wynik z 1931 r. mówi, że istnieje zdanie, którego ani nie da się dowieść, ani dowieść jego zaprzeczenia. Wywnioskujemy to z twierdzenia Turinga. Przez *zdanie* rozumiemy napis spełniający pewne proste wymagania składniowe. O *dowodach* zakładamy jedynie tyle, że istnieje program komputerowy, który dostawszy napis  $K$  oraz napis  $Z$  odpowiada w skończonym czasie „poprawny” bądź „niepoprawny”, w zależności od tego, czy  $K$  jest poprawnym dowodem zdania  $Z$ . Dodatkowo, każde zdanie, które ma dowód, jest prawdziwe.

Przypuśćmy, że dla każdego zdania  $Z$  albo istnieje jego dowód, albo dowód zdania „nieprawda, że  $Z$ ”, oznaczanego  $\neg Z$ . Napiszemy program  $Q$ , który działa następująco. Rozważmy zdanie  $Z$  „program  $P$  akceptuje napis  $N$ ” oraz jego negację  $\neg Z$ . Wczytaj na wejściu kod źródłowy programu  $P$  oraz napis  $N$ . Jeżeli  $Z$  ma dowód, to napisz „akceptuje”, a jeżeli  $\neg Z$  ma dowód, to napisz „wiesza się”. Żeby stwierdzić, który przypadek zachodzi, program  $Q$  przeszukuje wszystkie (coraz dłuższe) napisy, i dla każdego z nich sprawdza, czy jest dowodem  $Z$  bądź  $\neg Z$ . Z naszych założeń wynika, że w skończonym czasie znajdzie dowód albo  $Z$ , albo  $\neg Z$ , i że opisany program  $Q$  poprawnie przewiduje zachowanie programu  $P$  na wejściu  $N$ , przecząc twierdzeniu Turinga. Musi więc istnieć zdanie, które nie ma dowodu, ani którego zaprzeczenie nie posiada dowodu. Z kolei, z twierdzenia Gödla o *pełności* z 1929 roku, takie zdanie nie jest ani prawdziwe, ani nieprawdziwe – jest *niezależne* od przyjętych aksjomatów. Więcej o tym można znaleźć w moim artykule w *Delcie* 1/2017.

**Inne zastosowanie w informatyce.** W *złożoności obliczeniowej* mierzy się, jak „skomplikowana” jest własność liczb naturalnych, tj. ile kroków musi wykonać program, by stwierdzić, czy dana liczba ma tę własność. Przykładowo, własność „liczba parzysta” jest mniej skomplikowana niż własność „liczba pierwsza”. Metoda przekątniowa pozwala wszystko znacznie bardziej skomplikować.

Szymon TORUŃCZYK

## Informatyczny kącik olimpijski (109): Gra w karty

W tym odcinku IKO omówimy zadanie *Gra w karty*, które pojawiło się podczas pierwszej rundy konkursu „Potyczki Algorytmiczne 2016”.

Zadanie ma bardzo atrakcyjną warstwę fabularną, gdyż pojawia się w niej nie tylko popularny Bajtek, ale i inny amator gry w karty: Bitek. Obaj panowie pragną rozstrzygnąć, kto jest lepszy w karcianą grę *Magiczne Stwory*. Co ciekawe, dla ostatecznego wyniku ważniejsze od samej rozgrywki będzie to, co dzieje się jeszcze przed jej rozpoczęciem, a mianowicie wybór talii kart.

Zakładamy, że Bajtek dysponuje zestawem  $n$  talii kart do gry w *Magiczne Stwory*, konkretnie taliami  $A_1, \dots, A_n$ . Podobnie Bitek ma talie  $B_1, \dots, B_n$ . Do gry potrzebne są dwie talie, a więc umówiono się, że jedna z talii pochodzić będzie od Bajtka, a druga – od Bitka. Gdy już ustalimy wybór kart, to sama rozgrywka będzie deterministyczna i łatwo potrafimy z góry powiedzieć (tzn. te informacje podane są jako dane wejściowe), kto wygra, ewentualnie, że będzie remis.

Metoda wyboru talii do gry jest następująca. Naprzemiennie, poczynając od Bajtka, każdy z graczy odrzuca jedną z talii przeciwnika. Prawdziwą rozgrywkę zaczynamy, gdy każdy z graczy ma już tylko po jednej talii. Celem zadania jest stwierdzenie, czy Bajtek ma strategię wygrywającą. Jeśli nie – to, czy chociaż potrafi zapewnić sobie remis.

Zanim przejdziemy do rozwiązania tego zadania, spróbujemy się pokusić o jego *redukcję* do problemu prostszego. Twierdzę, że jeśli wybór talii byłby inny (prostszy), to nie wpłynęłoby to na szanse Bajtka. Rozważmy więc taką wersję wyboru talii, gdzie najpierw Bajtek po prostu wybiera pewną talię  $B_j$  z zestawu Bitka, po czym Bitek dowolnie dobiera do niej jakąś talię  $A_i$  z zestawu Bajtka i obaj zaczynają grać kartami  $A_i \cup B_j$ .

Dlaczego taki wybór talii nie jest istotnie różny od oryginalnego?

Po pierwsze: w obu wersjach to Bajtek decyduje, która z talii Bitka (oznaczymy ją przez  $\bar{B}_j$ ) weźmie udział w rozgrywce. Istotnie, w wersji uproszczonej jest to wprost powiedziane, natomiast w oryginalnej: to Bajtek kolejno odrzuca talie Bitka, a więc to właśnie on kontroluje, która na końcu zostanie.

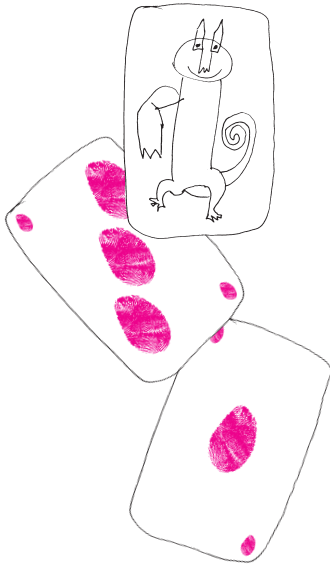
Nieco trudniej dowieść drugiej własności, to znaczy tej, że Bitek może dowolnie *dobrać* talię  $\bar{A}_i$  do  $\bar{B}_j$ . W wersji uproszczonej dostajemy to wprost z definicji, ale w oryginalnej nie jest to wcale jasne. Oczywiście, to Bitek odrzuca kolejno talie Bajtka, ale nie wie przecież z góry, jakie będzie  $\bar{B}_j$ , co niesie ryzyko odrzucenia  $\bar{A}_i$  na jakimś wczesnym etapie procesu wyboru. Wykażemy przez indukcję, że Bitek jest w stanie – mimo wszystko – zachować należytą ostrożność.

Dla  $n = 1$  nie mamy żadnego wyboru, więc baza indukcji jest trywialna. Załóżmy więc, że  $n > 1$  oraz że Bajtek zgłasza odrzucenie talii  $B_k$ . Dodatkowo oznaczymy jako  $A_m$  tę talię, którą dobraćby Bitek do  $B_k$ . Mamy teraz dwa przypadki. Pierwszy jest wtedy, gdy  $A_m$  nie byłaby dobrana dla żadnego innego  $B_j$  poza  $B_k$ . Wówczas bez żalu Bitek może odrzucić  $A_m$ . W przeciwnym przypadku zakładamy, że  $A_m$  jest dobrana do więcej niż jednego  $B_j$ . Oznacza to, że *funkcja dobierająca talie  $A_i$  do  $B_j$*  nie jest różnowartościowa, a więc istnieje talia  $A_z$ , która nigdy nie byłaby dobrana. Tym razem to z tą talią Bitek może się pożegnać i przejść indukcyjnie do problemu rozmiaru  $(n - 1)$ .

Skoro już uzasadniliśmy równoważność zadania dla znacznie prostszego modelu, możemy bardzo łatwo je rozwiązać. Po pierwsze: jeśli dla każdej talii  $B_j$  istnieje taka talia  $A_i$ , że karty  $A_i \cup B_j$  dają zwycięstwo Bitkowi, to Bitek ma strategię wygrywającą. Po drugie: jeśli dla każdej talii  $B_j$  istnieje taka talia  $A_i$ , że karty  $A_i \cup B_j$  oznaczają remis, to Bitek może zapewnić sobie remis. Okazuje się więc, że Bajtek wygrywa tylko wtedy, gdy istnieje taka talia  $B_j$ , która w połączeniu z dowolnym  $A_i$  daje zwycięstwo Bajtkowi.

Sprawdzenie, w którym z powyższych przypadków się znajdujemy, jesteśmy w stanie wykonać bardzo prosto, po prostu analizując bezpośrednio (w czasie liniowym) wczytywane dane.

Tomasz KAZANA



Owce rasy:



Texel



Scottish Blackface



skrzyżowana

## Owce tajemnice

Dawno, dawno temu – w języku współczesnych genetyków oznacza to nieomal 30 lat temu, kiedy realnie myślano o sekwencjonowaniu genomu człowieka – powstała w umyśle naukowca z National Institutes of Health, Craiga Ventera, myśl, żeby szukać produktów ekspresji genów, zamiast oznaczać po kolei nukleotydy w DNA. Ciekawiej i może nawet prościej byłoby zabrać się za oznaczanie cząsteczek powstających na matrycy DNA, stanowiących produkt pośredni w syntezie białek, czyli za RNA. Proces syntezy RNA w komórce nazywa się *transkrypcją*, zestaw wszystkich RNA transkrybowanych w danym gatunku – *transkryptomem*. Transkryptom to kopie DNA w działaniu, dynamiczny i ciekawy, bo choć w każdej komórce ciała danego organizmu znajduje się taki sam DNA, to jego transkrypcja, aktywność, zależy od tkanki, etapu życia, płci.

Genom człowieka opisały w 2000 roku dwie grupy badaczy. Pomysł badania transkryptomu, wielokrotnie potem zastosowany dla wielu innych gatunków, pozostał. Bardzo współcześnie: dla owcy! Genetycy molekularni często wymawiają to słowo wraz z imieniem – Dolly, pierwszego ssaka z klonu. Owce interesują też bardziej realistycznie rolników na całym świecie. Dają mięso, mleko, wełnę, są ważnym modelem w badaniach biomedycznych. Wiedza o ich genomie jest oczywiście przydatna, gdy myśli się o poprawie cech produkcyjnych i o charakterystyce odporności na choroby. Ogólny szkic genomu owcy powstał w 2014 r., jednak ponad połowie genów (jest ich w sumie ponad 25 tysięcy) nie umiano przypisać funkcji i rodzaju aktywności. Takich danych dostarczyć może jedynie badanie transkryptomu w różnych warunkach, wieku i płci dawcy. Jeżeli np. nieznaną z funkcji gen transkrybowany jest w makrofagach i znajduje się na genomie w sąsiedztwie znanych genów makrofagów, jest wielce prawdopodobne, że i ten gen związany jest z funkcjami immunologicznymi. Transkryptomem owiec zajęli się naukowcy szkoccy i australijscy (nietrudno zgadnąć, dlaczego w tych krajach owce interesują naukowców i... rolników) [\*]. Wcześniej poznano genomy kury, świni, kozy, bydła, ale ich transkryptomy – tylko wrywkowo, podobnie jak transkrypcyjne atlasy myszy, ludzi i koni.

Opublikowany obecnie atlas RNA dla owcy jest najbardziej szczegółowy z uzyskiwanych dla zwierząt gospodarskich, zawiera dane dla głównych narządów dorosłej owcy, zarodka i młodych osobników. Tysiące dotychczas nierozpoznanych genów zgrupowano i porównano między tkankami (wspólne procesy komórkowe, trawienie, odporność, reprodukcja, rozwój zarodka, zróżnicowanie płciowe, porównanie z innymi ssakami, w tym z człowiekiem).

Uważa się też, że te dane są bardzo ważne z medycznego punktu widzenia (dotychczas za najważniejszy model medyczny uznaje się myszy), ze względu na znaczne podobieństwa dużych ssaków do człowieka w rozmiarach i regulacji fizjologicznej i genetycznej.

Nawet na niespecjaliście atlas robi ogromne wrażenie liczbą wykonanych pomiarów i ich dogłębną charakterystyką. Korzystając z geograficznej separacji między badawczymi laboratoriami (Edynburg i Queensland), wykonano porównanie transkryptomów dwu bardzo odległych i różnych ras – Texel i szkocka Blackface i ich krzyżówki, co pozwoliło na ocenę przydatności krzyżowania ras w poszukiwaniu żądanych cech produkcyjnych.

Cóż, można tylko powiedzieć z uznaniem: meeee.

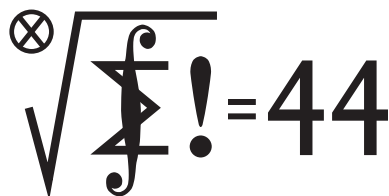
[\*] E.L. Clark, S.J. Bush, M.E.B. McCulloch, I.L. Farquhar, R. Young, L. Lefevre, et al. (2017) *A high resolution atlas of gene expression in the domestic sheep (Ovis aries)*. PLoS Genet 13(9): e1006997. <https://doi.org/10.1371/journal.pgen.1006997>

# Klub 44

Liga zadaniowa Wydziału Matematyki, Informatyki i Mechaniki,  
Wydziału Fizyki Uniwersytetu Warszawskiego i Redakcji *Delty*

## Skrót regulaminu

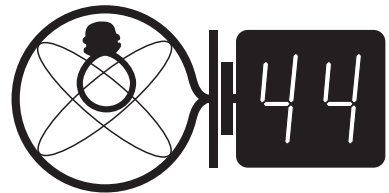
Każdy może nadsyłać rozwiązania zadań z numeru  $n$  w terminie do końca miesiąca  $n + 2$ . Szkice rozwiązań zamieszczamy w numerze  $n + 4$ . Można nadsyłać rozwiązania czterech, trzech, dwóch lub jednego zadania (każde na oddzielnej kartce), można to robić co miesiąc lub z dowolnymi przerwami. Rozwiązania zadań z matematyki i z fizyki należy przysyłać w oddzielnych kopertach, umieszczając na kopercie dopisek: **Klub 44 M** lub **Klub 44 F**. Oceniamy zadania w skali od 0 do 1 z dokładnością do 0,1. Ocenę mnożymy przez współczynnik trudności danego zadania:  $WT = 4 - 3S/N$ , gdzie  $S$  oznacza sumę ocen za rozwiązania tego zadania, a  $N$  – liczbę osób, które nadesłały rozwiązanie choćby jednego zadania z danego numeru w danej konkurencji (**M** lub **F**) – i tyle punktów otrzymuje nadsyłający. Po zgromadzeniu **44** punktów, w dowolnym czasie i w którejkolwiek z dwóch konkurencji (**M** lub **F**), zostaje on członkiem **Klubu 44**, a nadwyżka punktów jest zaliczana do ponownego udziału. Trzykrotne członkostwo – to tytuł **Weterana**. Szczegółowy regulamin został wydrukowany w numerze 2/2002 oraz znajduje się na stronie [deltami.edu.pl](http://deltami.edu.pl)



Termin nadsyłania rozwiązań: 31 I 2018

Czołówka ligi zadaniowej **Klub 44 M** po uwzględnieniu ocen rozwiązań zadań 739 ( $WT = 1,55$ ) i 740 ( $WT = 2,72$ ) z numeru 4/2017

Jerzy Cisło	Wrocław	43,94
Janusz Olszewski	Warszawa	43,42
Adam Dzedzej	Gdańsk	43,22
Roksana Słowik	Knurów	41,91
Patryk Jaśniewski	Gdańsk	40,94
Marcin Małogrosz	Warszawa	40,86
Franciszek S. Sikorski	Warszawa	39,71
Marcin Kasperski	Warszawa	39,22
Krzysztof Maziarz	Kraków	37,45



## Zadania z matematyki nr 749, 750

Redaguje Marcin E. KUCZMA

**749.** Trójkąt  $ABC$  jest opisany na okręgu o środku  $I$ , stycznym do boków  $AB$  i  $AC$  w punktach  $S$  i  $T$ . Na boku  $AC$  leży taki punkt  $P$ , że  $IP \parallel ST$ . Proste  $ST$  i  $BP$  przecinają się w punkcie  $Q$ . Dowieść, że  $QC \parallel AB$ .

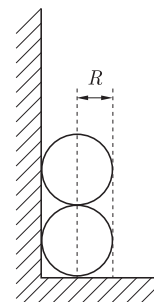
**750.** Znaleźć wszystkie pary liczb pierwszych  $p, q$  ( $p > q$ ), dla których także liczby  $(p^2 + q^2)/2$  oraz  $(p^2 - q^2)/24$  są pierwsze.

Zadanie 750 zaproponował pan Tomasz Ordowski.

## Zadania z fizyki nr 646, 647

Redaguje Elżbieta ZAWISTOWSKA

**646.** Dwie kulki o jednakowych masach i promieniach  $R$  leżą jedna na drugiej na poziomej powierzchni stykając się ze ścianą. Po zakłóceniu równowagi kulka górna ślizga się wzdłuż ściany, kulka dolna ślizga się po poziomej powierzchni, a ich prędkości początkowe są zerowe. Nie ma tarcia. Znaleźć prędkość kulki dolnej po utracie kontaktu między kulkami.



**647.** Nienaładowany, metalowy walec obraca się z prędkością kątową  $\omega$  wokół swojej osi. Walec umieszczony jest w jednorodnym polu magnetycznym, którego wektor indukcji  $\vec{B}$  jest równoległy do osi walca. Znaleźć gęstość ładunku wewnątrz walca.

## Dowolne cztery proste, z których żadne dwie nie są równoległe,

a żadne trzy nie mają punktu wspólnego...

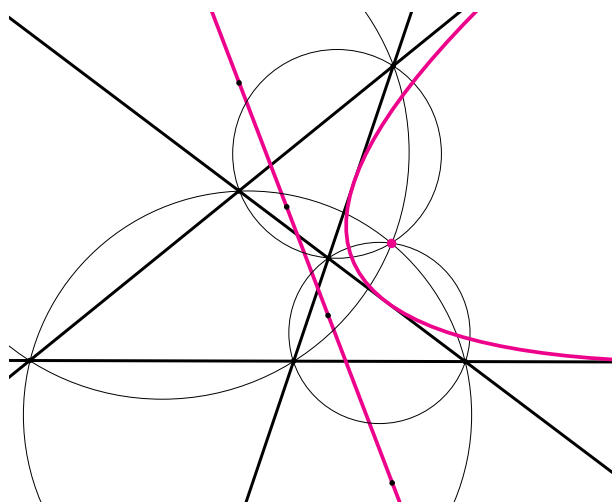
... tworzą cztery trójkąty – to każdy widzi i nikt się nie dziwi.

Ale gdy narysujemy ortocentra tych trójkątów (czyli punkty przecięcia wysokości każdego z nich – to te czarne kropki na rysunku), to okaże się, że punkty te leżą na jednej prostej – kto umie to udowodnić?

A gdy z kolei narysujemy okręgi opisane na tych trójkątach, okaże się, że wszystkie one mają punkt wspólny – czy to ma związek z poprzednim spostrzeżeniem? No i jak to udowodnić?

To teraz coś jeszcze. Gdy narysujemy kilka (lepiej kilkanaście) punktów jednakowo odległych od prostej łączącej ortocentra i od wspólnego punktu okręgów, to może nam przyjść do głowy, że punkty te leżą na paraboli, dla której kolorowa prosta jest kierownicą, a kolorowy punkt ogniskiem.

To wygląda już na kompletne szaleństwo, ale jest prawdą. Co więcej, parabola ta jest styczna do każdej z wyjściowych czterech prostych.



Cóż – trzeba przemyśleć, jak by i to udowodnić. Taka rozrywka na jesienne dżdżyste wieczory.

M. K.

Misja nosi nazwiska wielkich astronomów XVII w., Giovanniego Cassiniego (1625–1712) i Christiaana Huygensa (1629–1695). Pierwszy z nich był odkrywcą wielu księżyców Saturna oraz badaczem struktury pierścieni (jego imię nosi największa przerwa w dysku), drugi jest odkrywcą największego księżycza w układzie, Tytana.

## Prosto z nieba: Wielki Finał misji Cassini

Od chwili wystrzelenia w 1997 roku do spektakularnego Wielkiego Finału 15 września 2017 roku misja *Cassini-Huygens* (NASA/ESA/ASI) dostarczyła przełomowych informacji o Saturnie i jego księżycach. Przed misją *Cassini* bliskie, lecz pobieżne obserwacje Saturna wykonały jedynie sondy *Pioneer 11* oraz *Voyager 1 i 2* w trakcie ich podróży poza granice Układu Słonecznego. *Cassini* został specjalnie zaprojektowany do przestudiowania układu Saturna, a w szczególności, do zrzucenia lądownika *Huygens* na powierzchnię Tytana, co nastąpiło w 2005 roku. *Huygens* wykonał pierwsze w historii zdjęcia metanowych jezior i powierzchni Tytana, ukrytych do tej pory pod gęstymi (1,45 atmosfery!) chmurami. Niebagatelnym odkryciem była aktywność Enceladusa, którego lodowe fontanny zasilają jeden z zewnętrznych pierścieni (pierścień E). Pod powierzchnią Enceladusa znajduje się rozległy ocean, w którym mogą występować warunki sprzyjające życiu. *Cassini* wykonał wiele zdjęć Saturna m.in. tajemniczego heksagonu znajdującego się na jednym z biegunów (więcej informacji w [saturn.jpl.nasa.gov](http://saturn.jpl.nasa.gov)).

Misja zakończyła się spektakularnie. Zestaw coraz bardziej karkołomnych trajektorii *Cassiniego* – przelotów przez pierścienie prowadzących do ostatecznego upadku na Saturna – pozwoliły na zgromadzenie unikalnych danych niemożliwych do zebrania wcześniej ze względu na ryzyko związane z tego typu manewrami. Detektory satelity bardzo dokładnie zmierzyły pola magnetyczne i grawitacyjne Saturna co, być może, pomoże ujawnić strukturę wewnętrzną planety i umożliwi pomiar tempa rotacji jej jądra. Przeloty przez lodowe pierścienie ułatwią z kolei oszacowanie ilości tworzącej je materii. W trakcie przelotów rejestratory cząstek na pokładzie *Cassiniego* zbierały informacje o składzie pierścieni i atmosfery planety, a kamery do ostatniej chwili wykonywały bardzo dokładne zdjęcia pierścieni i chmur.

Podstawowa misja *Cassini* trwała od 2004 do 2008 roku. Sonda dotarła do Saturna w momencie przesilenia zimowego na północnej półkuli (nachylenie osi obrotu planety jest podobne do ziemskiego, 26,73°, a rok trwa prawie 29,5 ziemskich lat). W czasie 13 lat misji *Cassini* krążył wokół Saturna obserwując go w różnym oświetleniu w trakcie równonocy (sierpień 2009, misja *Equinox* w latach 2008–2010), oraz obecnie w trakcie przesilenia letniego (misja *Solstice*, 2010–2017).

Dlaczego podjęto decyzję o zakończeniu misji? Do 2017 roku, po siedmiu latach podróży z Ziemi *Cassini* spędził 13 lat na orbicie wokół Saturna. Poziom paliwa, który mu pozostał jest niski, więc operatorom misji pozostawało do dyspozycji coraz mniejsze pole manewru. Aby uniknąć prawdopodobnej przyszej kolizji niekontrolowanego już z Ziemi *Cassiniego* z jednym z interesujących księżyców Saturna, Enceladusa lub Tytana, na których być może znajdują się prymitywne formy życia, NASA zdecydowała się na bezpieczne zniszczenie sondy w atmosferze Saturna gwarantując, że *Cassini* nie zanieczyści atmosfer i powierzchni księżyców. Aktualnie nie wiadomo, jaka misja będzie wysłana jako kolejna, ale z pewnością po sukcesach *Cassiniego* nie będziemy długo czekać na kontynuację badań układu Saturna.

Michał BEJGER

## Niebo w listopadzie



W listopadzie noce są już bardzo długie, zwłaszcza pod koniec miesiąca, gdy Słońce przekroczy  $-20^\circ$  deklinacji w drodze na południe, co dzieje się corocznie około 20 listopada, czyli miesiąc przed przesileniem zimowym (w tym roku 21 grudnia). Od tego momentu położenie Słońca w południe zmieni się już tylko o 3,5 stopnia, osiągając minimum pierwszego dnia zimy. W trakcie miesiąca ubywa dnia o ponad 1,5 godziny, jednocześnie zwiększa się coraz bardziej kontrast między długością dnia w Polsce północnej i południowej. 30 listopada w miejscowości Wołosate w Bieszczadach od wschodu do zachodu Słońca mija 8 godzin i 34 minuty, natomiast na Przylądku Rozewie nad Bałtykiem – 7 godzin i 40 minut. Do pierwszego dnia zimy różnica ta zwiększy się do ponad godziny.

Z jasnych planet Układu Słonecznego w listopadzie tylko Mars jest widoczny dobrze, zwłaszcza pod koniec miesiąca, gdy na początku świtu żeglarskiego – około godz. 5 – zajmie on pozycję na wysokości  $17^\circ$  nad wschodnim



widnokregiem. Do końca listopada jasność Czerwonej Planety urośnie z +1,8 do +1,7 magnitudo, jednak przez cały czas średnica marsjańskiej tarczy będzie miała wartość 4", zaś faza spadnie z 97 do 96%. Mars wędruje przez gwiazdozbiór Panny. 1 listopada planeta znajdzie się niecały 1° od gwiazdy 4. wielkości Zaniah ( $\beta$  Vir), 8 listopada – 100' na południe od jaśniejszej o 0,5 magnitudo Porrimy ( $\gamma$  Vir), 15 listopada w odległości niecałych 3° Marsa minie Księżyc w fazie 10%, natomiast 30 listopada planeta przejdzie 3° na północ od Spiki, czyli najjaśniejszej gwiazdy Panny.

Bardzo dobrze widoczne są planety Neptun i Uran, jednak do ich obserwacji potrzebna jest przynajmniej lornetka. Neptun przebywa mniej niż 40' od świecącej blaskiem +3,7 magnitudo gwiazdy  $\lambda$  Aquarii. 22 listopada planeta zmieni kierunek ruchu z wstecznego na prosty, co oznacza koniec jej najlepszej widoczności w tym sezonie obserwacyjnym. Neptun świeci blaskiem +7,8 wielkości gwiazdowej. 27 listopada planetę ponownie zakryje Księżyc w fazie 56%. Będzie to ostatnie z serii zakryć Neptuna przez Księżyc w tym roku, następna wystąpi w latach 2023–24. Tym razem zjawisko ominie obszary zamieszkałe przez ludzi. Będzie widoczne tylko z Antarktydy. Planeta Uran wędruje na tle gwiazdozbioru Ryb, poruszając się ruchem wstecznym. W ciągu miesiąca oddali się ona od gwiazdy  $\sigma$  Psc na 3°. Jasność planety to +5,7 magnitudo. 3 listopada rano 4,5 stopnia na południe od Urana przejdzie Księżyc w fazie 98%.

Pozostałe jasne planety widoczne są słabo, choć np. Merkury 24 listopada osiągnie swoją maksymalną elongację wschodnią i to całkiem dużą, jak na tę planetę, bo prawie 22°. Jednak niekorzystne nachylenie ekliptyki do zachodniego wieczornego widnokregu i dodatkowo położenie planety pod ekliptyką spowoduje, że zniknie ona z niebosłonu niecałą godzinę po Słońcu i z dużych szerokości geograficznych będzie niewidoczna. Znajdującą się również na wieczornym niebie planetę Saturn jeszcze w pierwszej połowie listopada można próbować dostrzec nisko nad południowo-zachodnim widnokregiem tuż po zmierzchu, lecz do końca miesiąca zniknie ona w zorzy wieczornej. Jednak w tym przypadku nie można liczyć na nic poza wzrokowym stwierdzeniem obecności planety na niebie, ewentualnie sfotografowanie jej, jako punktu na niebie wraz z otaczającymi ją gwiazdami i/lub Księżycem. W teleskopie obraz planety będzie zaburzany silnym falowaniem atmosfery. Bardzo cienki sierp Księżyca spotka się z Saturnem w dniach 20 i 21 listopada. 20 listopada Srebrny Glob będzie 2 dni po nowiu, w fazie 4%. Godzinę po zachodzie Słońca oba ciała niebieskie znajdą się na wysokości 4° nad południowo-zachodnim widnokregiem, w odległości 5°. Dobę później Księżyc zwiększy fazę do 9% i o tej samej porze Saturn znajdzie się na godzinie 5. względem Księżyca, 7° od niego.

Nad ranem w pierwszej części miesiąca oprócz Marsa świeci coraz bardziej zbliżająca się do Słońca planeta, Wenus, a w drugiej połowie listopada zastąpi ją

powracający na poranne niebo po październikowej koniunkcji ze Słońcem Jowisz. Początkowo godzinę przed świtem Wenus wzniesie się na wysokość około 5° nad wschodnim widnokregiem, lecz każdej kolejnej doby jej pozycja będzie coraz niższa: w połowie miesiąca o tej samej porze już tylko 2°, a w trzeciej dekadzie miesiąca – pod horyzontem. Wenus dąży do koniunkcji górnej ze Słońcem 9 stycznia przyszłego roku, a na niebie wieczornym pojawi się dopiero pod koniec lutego. Obecnie jasność planety wynosi –3,9 magnitudo, przy tarczy o średnicy 10" i fazie 98%. Zupełnie inaczej sprawa ma się z Jowiszem. 26 października spotkał się on ze Słońcem i dzięki wciąż korzystnemu nachyleniu ekliptyki do porannego widnokregu już w drugiej dekadzie listopada znacznie się on pojawiać nad wschodnim widnokregiem tuż przed świtem. Przez cały miesiąc planeta świeci blaskiem –1,7 magnitudo, z tarczą o średnicy 31". 13 listopada Jowisz spotka się w bliskiej koniunkcji z Wenus. Tego ranka planety oddzieli dystans zaledwie 18', lecz obie planety godzinę przed świtem zajmą pozycję na wysokości jedynie 2° nad widnokregiem, stąd obserwacja tego spotkania nie będzie łatwa. Kilka dni później, 16 i 17 listopada z obiema planetami spotka się Księżyc tuż przed nowiem. 16 listopada Srebrny Glob pokaże tarczę w fazie 5% i o godzinie 6:00 znajdzie się na wysokości 13°. Planety będzie można dostrzec na godzinie 7 względem Księżyca, w odległości 9° (Jowisz) i 12° (Wenus). Do tego czasu dystans między planetami urośnie do 3°. Dobę później Jowisza i Wenus oddzieli już 4°, zaś Księżyc w fazie 2% znajdzie się na lewo od pary planet, odpowiednio 4,5 stopnia od pierwszej i 3° od drugiej z nich. Tego ranka Księżycowi do nowiu zabraknie jedyne 30 godzin.

Oczywiście, oprócz planet Księżyc minie również kilka jasnych gwiazd. Dla nas szczególnie warte polecenia jest przejście Srebrnego Globu przez Hiady w nocy z 5 na 6 listopada. Z Polski będzie można zobaczyć m.in. zakrycie gwiazdy  $\gamma$  Tauri (od około 19:57 do 20:49) oraz Aldebarana (od około 4:00 do 4:46). Niestety, obserwacje znacznie utrudni faza Księżyca, wynosząca 95%. Drugim ciekawym zakryciem, lecz widocznym tylko w północnej części Polski, jest zakrycie gwiazdy 4. wielkości  $\sigma$  Leonis 13 listopada. Faza Księżyca będzie znacznie mniejsza, 27%, stąd choć gwiazda słabsza, to łatwiej zjawisko zaobserwować, zwłaszcza odkrycie przy ciemnym brzegu po godz. 2:30. Gwiazda zniknie za Księżycem na północ od linii mniej więcej Zgorzelec–Radom–Włodawa. Z tym, że w pierwszych dwóch miastach do zakrycia nie dojdzie.

W listopadzie co roku przez prawie cały miesiąc promieniają meteory ze słynnego roju Leonidów, z maksimum aktywności w okolicach 17 listopada. W tym roku, niestety, nie prognozuje się deszczu meteorów. Radiant roju wschodzi około godz. 23, a nad ranem można spodziewać się około 10 meteorów na godzinę. Za to Księżyc będzie wtedy bliski nowiu, zatem – o ile tylko pozwoli na to kapryśna polska pogoda – warunki obserwacyjne Leonidów będą w tym roku znakomite.

*Ariel MAJCHER*

## Nieparzystość kombinowana neutrin?

Naruszenie parzystości kombinowanej CP, czyli niezmienniczości procesów przy jednoczesnej zmianie parzystości P (zamiana orientacji układu odniesienia na przeciwną) i sprzężeniu ładunkowym C (od ang. *charge*; zamiana cząstek na ich antycząstki) w rozpadach neutralnych kaonów, odkryte w 1964 roku, było wielkim zaskoczeniem. Trzy lata później, czyli równo pół wieku temu, Andriej Sacharow zauważył, że jest to jeden z koniecznych warunków wygenerowania przewagi materii nad antymaterią w ewolucji Wszechświata (w wyniku tzw. bariogenezy), czyli warunkiem istnienia nas jako bytów materialnych.

Okazało się, że takie naruszenie jest możliwe, o ile istnieją co najmniej trzy generacje cząstek materii, bo wtedy przekształcenia między generacjami muszą być opisywane nie tylko za pomocą trzech tzw. kątów mieszania (oddających to, jak trudne jest przejście między poszczególnymi parami generacji), ale również tzw. zespolonej fazy (wynika to z warunku unitarności macierzy  $3 \times 3$ ). Naruszenie występuje, o ile faza ta nie jest wielokrotnością  $\pi$  (czyli jej część urojona jest niezerowa).

W latach sześćdziesiątych nie wiedzano nawet, że istnieją dwie generacje (np. wiedzano tylko o trzech z sześciu kwarków). Skompletowanie trzech, rozpoczęte w latach siedemdziesiątych, a zakończone tuż przed przełomem tysiącleci, wraz z uzyskaniem dowodu, że nie ma ich więcej (przynajmniej takich, w których neutrina są prawie bezmasowe) jest bardzo ciekawe.

Okazało się jednak, że poziom naruszenia CP w sektorze kwarkowym jest niewystarczający do wyjaśnienia procesu bariogenezy. Dlatego powinno istnieć dodatkowe źródło tego naruszenia w sektorze leptonowym lub w ogóle poza Modelem Standardowym.

Jeżeli chodzi o sektor leptonowy, to faza jest umieszczana w macierzy mieszania neutrin, bo wynika to z historii odkrywania Modelu Standardowego (niektórzy upierają się, że samo mieszanie, możliwe do zdefiniowania dzięki niezerowym i różnym masom neutrin, jest już wyjściem poza Model Standardowy), pojęciowo jest przystępniejsze, rachunkowo prostsze oraz, co chyba najważniejsze, pomiar jest możliwy tylko poprzez precyzyjne badanie oscylacji neutrin.

Najczulszą procedurą doświadczalną jest poszukiwanie neutrin (antyneutrin) elektronowych w wiązce powstającej z produkcji par antymion-neutrino (mion-antyneutrino) w rozpadach dodatnich (ujemnych) pionów i kaonów produkowanych w wyniku bombardowania tzw. tarczy protonami. Wiązki takie początkowo składają się prawie wyłącznie z neutrin (antyneutrin) mionowych, ale zanim doleżą do oddalonego o setki kilometrów detektora, ich skład zmienia się w wyniku oscylacji. Ponieważ tempo oscylacji zależy od parametrów macierzy mieszania

i różnic mas oraz stosunku odległości do energii neutrin, to te ostatnie dwie wielkości można dobrać tak, żeby maksimum interesującej nas oscylacji przypadało w miejscu, gdzie umieszczony jest tzw. daleki detektor. Wtedy zliczanie pojawiających się neutrin (antyneutrin) elektronowych pozwala na wyznaczenie tych parametrów sektora neutrin, od których tempo danej oscylacji zależy.

Jeżeli CP jest naruszone, to prawdopodobieństwa oscylacji dla neutrin i antyneutrin będą się różnić.

W praktyce przeprowadza się dopasowanie do wszystkich dostępnych danych, a fazę naruszenia CP można wyłuskać nawet z oscylacji samych neutrin, których produkuje się więcej (bo protony wiązki i tarczy są z materii, czyli są dodatnie – produkuje się więcej mezonów dodatnich, a w konsekwencji antymionów). Dane jednak zbiera się na zmianę, raz dla neutrin, raz dla antyneutrin.

Obecnie działają dwa takie eksperymenty, T2K w Japonii oraz NOvA w Stanach Zjednoczonych. Pierwszy z nich ogłosił latem [1], że hipotezę o braku naruszenia CP można odrzucić na poziomie istotności statystycznej 0,05. Kolokwialnie mówiąc, oznacza to, że wyniki na tyle niezgodne z taką hipotezą, co te obserwowane, mają szansę pojawić się w wyniku fluktuacji statystycznej nie większą niż raz na dwadzieścia. Dane wskazują jednocześnie na maksymalne naruszenie, co (po potwierdzeniu i przy sprzyjających wiatrach) wystarcza do wyjaśnienia bariogenezy.

Wyniki drugiego z eksperymentów [2] też sugerują, że CP jest w sektorze leptonowym naruszone, choć odnajdywane najbardziej prawdopodobne wartości parametrów nie są takie same. Z wiarygodnego źródła (w eksperymencie T2K bierze udział silna grupa naukowców z Polski) wiem, że przedstawiciele obu eksperymentów właśnie zamierzają rozmawiać o przeprowadzeniu wspólnej statystycznej analizy danych.

Zaplanowane naświetlenia mogą w ciągu kilku lat doprowadzić do polepszenia istotności statystycznej do poziomu  $3\sigma$  (prawdopodobieństwo fluktuacji mniejsze od dwóch na tysiąc), odpowiadającej tzw. ewidencji obserwowania czegoś nowego w fizyce cząstek. Oczywiście przy założeniu, że CP jest rzeczywiście naruszone w sposób maksymalny. Szans na uzyskanie poziomu „odkrycia” ( $5\sigma$ : 3 na 10 milionów) na razie nie ma. A gra toczy się o kolejną, już szóstą Nagrodę Nobla z Fizyki „za neutrina”.

Piotr ZALEWSKI

[1] T2K prezentuje dane wskazujące na możliwość łamania symetrii CP przez neutrina, <http://neutrino.fuw.edu.pl/pl/node/754>

[2] The NOvA Collaboration, *Constraints on Oscillation Parameters from  $\nu_e$  Appearance and  $\nu_\mu$  Disappearance in NOvA*, arXiv:1703.03328v2



## Krzywe i połamane

Joanna JASZUŃSKA

Oto kilka zadań związanych z istnieniem i własnościami pewnych krzywych lub łamanych na płaszczyźnie i w przestrzeni trójwymiarowej.

1. Mamy dwa ziemniaki. Wykaż, że istnieje taka krzywa zamknięta w przestrzeni trójwymiarowej, którą da się narysować na powierzchni każdego z tych ziemniaków.

2. Danych jest 9 punktów rozmieszczonych w wierzchołkach, środkach boków i środku kwadratu  $2 \times 2$ , jak na rysunku 1. Połącz wszystkie te punkty za pomocą łamanej złożonej z czterech odcinków.

3. Czy istnieje w przestrzeni taka łamana zwyczajna zamknięta, której żaden z rzutów na płaszczyznę w ustalonych trzech wzajemnie prostopadłych kierunkach nie zawiera łamanej zamkniętej?

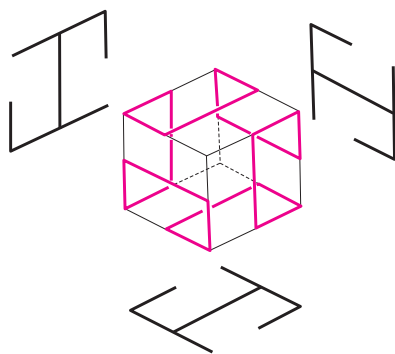
4. Pewnego poranka o godzinie  $8^{00}$  turysta wyruszył z domu u podnóża góry i o  $20^{00}$  dotarł do schroniska na szczycie. O  $8^{00}$  następnego dnia wyruszył ze szczytu tą samą trasą i o  $20^{00}$  wrócił do domu. Udowodnij, że istnieje taki punkt, w którym turysta był w oba dni dokładnie o tej samej godzinie.

5. W przestrzeni dana jest łamana zwyczajna zamknięta złożona z sześciu równej długości odcinków, przy czym odcinki pierwszy i czwarty, drugi i piąty oraz trzeci i szósty są parami równoległe. Ponadto każde dwie proste zawierające kolejne odcinki łamanej tworzą kąty płaskie równe  $120^\circ$ . Czy łamana ta musi być sześciokątem foremnym?

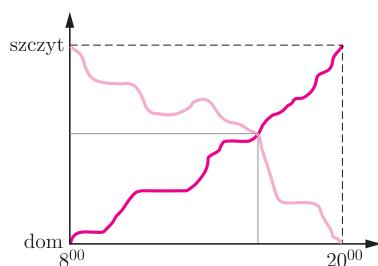
6. Robaczek wgrzył się w idealnie kuliste jabłko o średnicy 1, wydrążył w nim cieniutki tunelik o długości 0,9 i o sobie tylko znanym kształcie i wyszedł na powierzchnię w innym punkcie. Wykaż, że można to jabłko tak przeciąć na pół, by w jednej z połówek nie było śladów bytności robaczka.



Rys. 1



Rys. 2



Rys. 3. Turysta mógł zmieniać tempo marszu, robić postoje, a nawet się cofać.

Elipsoida obrotowa o ogniskach  $F, G$  i stałej  $d$  to zbiór takich punktów  $X$  przestrzeni, dla których  $XF + XG = d$ . Punkty  $X$ , dla których  $XF + XG < d$ , tworzą wnętrze elipsoidy. Elipsoida jest figurą wypukłą.

Literatura:  
Peter Winkler, *Mathematical Mind-Benders*, A K Peters 2007

### Rozwiązania

**R1.** Zamiast ziemniaków rozważmy ich duchy o dokładnie tym samym kształcie i rozmiarze. Wystarczy teraz, by jeden duch częściowo przeniknął przez drugiego. Przecięcie ich powierzchni wyznacza szukaną krzywą.  $\square$

**R2.** Startując ze środka lewego boku rysujemy w górę odcinek o długości 2 (wystaje on więc poza kwadrat); następnie w prawo w dół odcinek przechodzący przez środki górnego i prawego boku i kończący się na poziomie dołu kwadratu; potem w lewo poziomy odcinek o długości 3, kończący się w lewym dolnym rogu kwadratu; i wreszcie czwarty odcinek łamanej – przekątną kwadratu (w prawo w górę).  $\square$

**R3.** Tak, przykład takiej łamanej zaprezentowano na rysunku 2. Znajduje się ona na powierzchni pewnego sześciianu, a przedstawione jej rzuty są w kierunkach zgodnych z krawędziami tego sześciianu.  $\square$

**R4.** Narysujmy wykresy obu wędrówek turysty (rys. 3). Jedna z nich to krzywa łącząca dwa przeciwległe wierzchołki prostokąta, druga zaś łączy pozostałe dwa wierzchołki. Takie dwie krzywe muszą się gdzieś przeciąć, co kończy dowód.  $\square$

Aby nieco ściślej uzasadnić przecinanie się krzywych, można wykorzystać np. własność Darboux i zadanie 1 z *deltoidu* 12/2010. Inne rozwiązanie opisano w *deltoidzie* 11/2014.

**R5.** Nie musi. Rozważmy ośmiościan foremny o parach przeciwległych wierzchołków  $A$  i  $A'$ ,  $B$  i  $B'$ ,  $C$  i  $C'$ . Łamana  $ABCA'B'C'A$  spełnia warunki zadania.

Odpowiednie odcinki są równe, bo ośmiościan jest foremny i równoległy, bo  $ABA'B'$ ,  $ACA'C'$ ,  $BCB'C'$  są kwadratami. Ściany tego ośmiościanu są trójkątami równobocznymi, więc wszystkie pary kolejnych prostych tworzą kąty  $120^\circ$ .  $\square$

**R6.** Rozważmy elipsoidę obrotową o ogniskach w punktach wejścia i wyjścia robaczka oraz o stałej równej 0,9. Każdy punkt tuneliku odległy jest od jego końców w sumie o najwyżej 0,9, więc cały tunelik mieści się w elipsoidzie. Z kolei środek jabłka jest poza nią, gdyż suma jego odległości od ognisk równa jest  $0,5 + 0,5 = 1 > 0,9$ . Istnieje więc taka płaszczyzna przechodząca przez środek, że cała elipsoida jest po jednej jej stronie. Wtedy po drugiej stronie otrzymujemy nietkniętą przez robaczka połówkę jabłka.  $\square$