



Dnia 29 stycznia 2007 roku zmarł

Andrzej Mąkowski

znakomity matematyk, niewyczerpane źródło wiedzy

aktywnie towarzyszący *Delcie* od pierwszego numeru
i wytrwale korygujący wszelkie nasze błędy

Autor i Członek Komitetu Redakcyjnego *Delty*

SPIS TREŚCI NUMERU 3 (394)

Pole magnetyczne
w galaktykach
Krzysztof T. Chyży

Wielomian Ehrharta
Marcin Hauzer

Zjawisko Aharonova–Bohma
Paweł Tomasz Pęczkowski

Dlaczego $2 \cdot 2 = 4$, czyli
twierdzenia o rekursji
Adam Kolany

∫ Wyświetlacze
ciekrokryształiczne
Mikołaj Korzyński

∫ Trójki pitagorejskie
Marcin Hauzer

Ciąg, który lubi nierówności
Jan Szejko

Zadania

Aktualności

Klub 44

Patrz w niebo

Marzec

Ω Markow rządzi (nawet
buldogami pod dywanem)
Rafał Sztencel

W następnym numerze:
Twierdzenie Lagrange'a

str. 1	„Delta” – matematyczno-fizyczno-astronomiczny miesięcznik popularny wydawany przez Uniwersytet Warszawski przy współpracy towarzystw naukowych: Polskiego Towarzystwa Matematycznego, Polskiego Towarzystwa Fizycznego i Polskiego Towarzystwa Astronomicznego.
str. 3	Komitet Redakcyjny: prof. dr hab. Andrzej Białynicki-Birula (członek rzeczywisty PAN), prof. dr hab. Bogdan Cichocki, dr Krzysztof Ciesielski – wiceprzewodniczący, dr hab. Armen Edigarian, prof. dr hab. Jan A. Gaj – przewodniczący, dr hab. Maciej Geller, prof. dr hab. Jerzy Ginter, dr Piotr Goldstein, dr Agnieszka Janiuk, prof. dr hab. Wiesław A. Kamiński, dr hab. Andrzej Majhofer, dr hab. Zbigniew Marciniak, prof. dr hab. Janusz Matkowski, <u>mgr Andrzej Mąkowski</u> , dr Adam Michalec, prof. dr hab. Ryszard J. Pawlak, dr Zdzisław Pogoda, prof. dr hab. Grzegorz Sitarski, dr Weronika Śliwa, prof. dr hab. Andrzej Woszczyk, prof. dr hab. Wiesław Żelazko.
str. 4	Redaguje kolegium w składzie: Wiktor Bartol, Ewa Czuchry, Marcin Hauzer, Krystyna Kordos – sekr. red., Marek Kordos – red. nac., Mikołaj Korzyński, Tomasz Kwast, Urszula Marciniak, Anna Rudnik, Witold Sadowski, Piotr Zalewski – z-ca red. nac. Okładki i rysunki: Emilia Bojańczyk. Rysunki techniczne: Marcin Adamski.
str. 6	Adres do korespondencji: Instytut Matematyki UW, Redakcja „Delty”, ul. Banacha 2, pokój 5450, 02-097 Warszawa, e-mail: delta@mimuw.edu.pl, tel. 022-55-44-545. Skład systemem \TeX wykonała Redakcja.
str. 8	Wydrukowano w Drukarni Naukowo-Technicznej, Oddział PAP S.A. w Warszawie, ul. Mińska 65.
str. 9	WARUNKI PRENUMERATY W FIRMIE AMOS 01-785 Warszawa, ul. Broniewskiego 8A (tel. 022-663-87-52, 022-663-11-46) Wpłaty przyjmowane są non-stop, do 10. dnia miesiąca poprzedzającego okres prenumeraty. Okres prenumeraty wynosi co najmniej trzy miesiące. Cena jednego numeru w 2007 roku wynosi 4 zł. Przy wpłacie prosimy o zaznaczenie okresu prenumeraty.
str.10	W prenumeracie zagranicznej (też przez okres co najmniej trzech miesięcy) cena numeru w 2007 r. wynosi 8 zł. W przypadku życzenia dostawy priorytetowej odpowiednią dopłatę ponosi zamawiający.
str.12	Uwaga! Dla zamawiających minimum 10 egzemplarzy każdego numeru AMOS funduje dodatkowo jeden egzemplarz pisma.
str.13	Konto AMOS-u: PKO BP SA I O/W-wa, nr 11 1020 1013 0000 0502 0004 0584
str.14	WARUNKI PRENUMERATY W RUCH-u internet: www.ruch.pol.pl, e-mail: prenumerata@okdp.ruch.com.pl
str.16	Cena prenumeraty w 2007 roku wynosi 4 zł za egzemplarz.
str.17	1. Prenumerata krajowa: wpłaty przyjmują jednostki kolportażowe „RUCH” SA właściwe dla miejsca zamieszkania. Termin przyjmowania prenumeraty: do 5 każdego miesiąca poprzedzającego okres rozpoczęcia prenumeraty. 2. Prenumerata ze zleceniem wysyłki za granicę: informacji o warunkach prenumeraty i sposobie zamawiania udziela „RUCH” SA Oddział Krajowej Dystrybucji Prasy, 01-248 Warszawa, ul. Jana Kazimierza 31/33; tel. 022-5328-731 (prenumerata płatna w walucie obcej), -816, -734, -819 (prenumerata płatna w PLN w kasie Oddziału lub na konto w banku PEKAO SA IV O/Warszawa 68 1240 1053 1111 0000 0443 0494), infolinia 0-800-1200-29. 3. Prenumerata opłacana za granicą: przelewem na nasze konto SWIFT banku: PKOPPLPWWA4; w USD PEKAO SA IV O/W-wa IBAN PL54 1240 1053 1787 0000 0443 0508; w EUR PEKAO SA IV O/W-wa IBAN PL54 1240 1053 1978 0000 0443 0511; kserokopię polecenia przelewu należy przesłać faksem pod numer +48-22-5328-731.

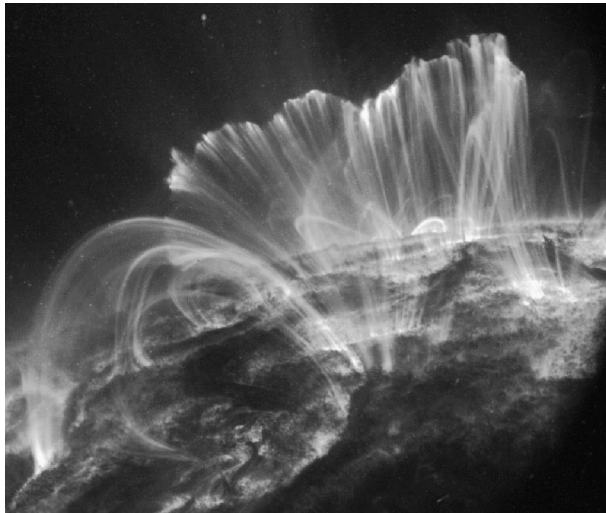
Numery archiwalne (od 1987 r.) można nabyć w Redakcji osobiście lub listownie.
Strona internetowa (streszczenia, artykuły archiwalne, linki itd.):

<http://www.mimuw.edu.pl/delta>

Wydawca: Uniwersytet Warszawski

Cena 1 egzemplarza 4 zł

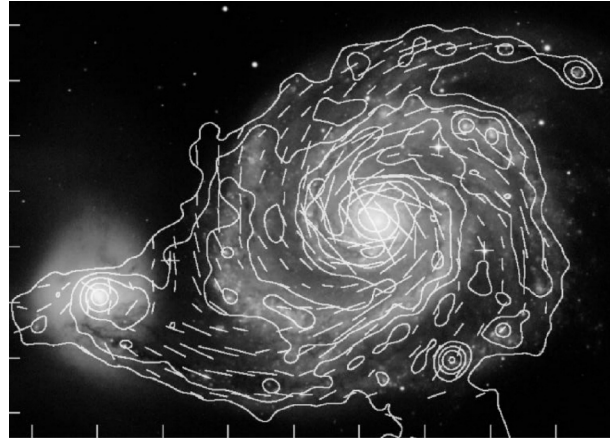
Pola magnetyczne występują powszechnie we Wszechświecie. Badanie ich struktury, ewolucji i roli w różnych kosmicznych obiektach jest jednym z najbardziej fascynujących, lecz trudnych przedsięwzięć współczesnej astrofizyki. Wydaje się, że im lepiej przyglądamy się polu, tym bardziej też doceniamy jego rolę. Otóż w przypadku Ziemi rozumiemy już dzisiaj, że pole magnetyczne było i jest niezbędnym czynnikiem umożliwiającym istnienie życia na Ziemi. A to ze względu na okalającą naszą planetę magnetosferę, która chroni nas przed zabójczym strumieniem cząstek wiatru słonecznego. Pole magnetyczne Ziemi o typowym natężeniu około 0,5 Gs (Gs (gaus) = 10^{-4} T (tesli)) umiemy już numerycznie modelować, odtwarzać jego przebiegunowania i raptowne spadki jego natężenia, z jakimi mamy do czynienia w przeciągu ostatniego stulecia. Analogiczne pola istnieją także na Jowiszu i Saturnie (okładka), obserwowane choćby w widowskich zjawiskach zórz polarnych. Znacznie silniejsze pola, osiągające natężenia kilkuset gausów, występują na Słońcu. Są one źródłem szeregu zjawisk rozgrywających się w chromosferze i koronie słonecznej i są przyczyną cyklicznej aktywności Słońca (rys. 1).



Rys. 1. Zjonizowana materia wyrzucana z powierzchni Słońca porusza się wzdłuż linii pola magnetycznego i powraca ku jego powierzchni tworząc pętle (fot. TRACE, NASA).

Badania pól magnetycznych galaktyk ze względu na odległość i ograniczoną przyrządów obserwacyjnych są nieporównywalnie trudniejsze, a nasza wiedza o nich wciąż jeszcze znacznie uboższa. Najlepszą obecnie metodą wykrywania pól magnetycznych w galaktykach są obserwacje na falach radiowych. Pozwalają one na detekcję promieniowania synchrotronowego – powstałego w wyniku „wyświecania się” szybko poruszających się (relatywistycznych) elektronów w polu magnetycznym. Teoria promieniowania synchrotronowego umożliwia wyznaczenie z obserwacji wartości natężenia pola magnetycznego i jego kierunku. Tą właśnie metodą od przeszło 30 lat odkrywano pola magnetyczne w pobliskich galaktykach. Okazuje się, że ośrodek

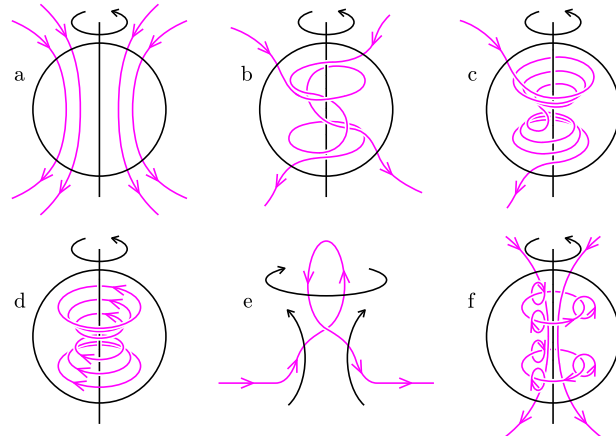
międzygwiazdowy wypełniony jest stosunkowo słabym polem magnetycznym o typowym natężeniu około 10^{-5} Gs. Możemy wyodrębnić dwie składowe tego pola. Jedna to pole chaotyczne (turbulentne), druga to pole regularne, zachowujące kierunek w skali wielu kiloparseków (galaktyka M51, rys. 2).



Rys. 2. W galaktyce spiralnej M51 odkryte pola magnetyczne biegną wzdłuż ramion spiralnych i w przestrzeni pomiędzy nimi. Pola generowane są w procesie turbulentnego dynamo, napędzanego szybką rotacją galaktyki i wybuchami supernowych. Na zdjęcie optyczne nałożone są kontury emisji radiowej, kreski reprezentują kierunek regularnego pola magnetycznego, a ich długość – jego wartość (fot. R. Beck).

Istnienie wielkoskalowych pól regularnych musi wynikać z działania jakiegoś globalnego mechanizmu, porządkującego to pole w całej galaktyce. Ponadto, gdyby w jakimś momencie „włożyć” takie pole do galaktyki, to ono samoistnie zaniknie, w przeciągu niespełna milionów lat. Zatem w galaktykach musi działać, w sposób ciągły, jakiś mechanizm podsycańca, wzmacniania pola.

Obecnie uważa się, że tym mechanizmem jest tzw. turbulentne dynamo, podobne do tego, jakie pracuje w planetach i w Słońcu (patrz rys. 3).

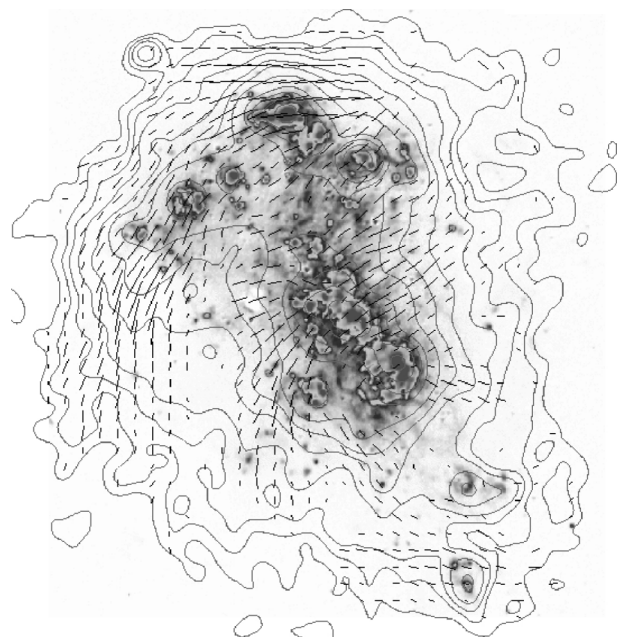


Rys. 3. Mechanizm dynamo. Niech początkowe pole magnetyczne ma kierunek południkowy (a). Na skutek rotacji różnicowej obiektu (szybszej na równiku) pole jest rozciągane i wzmacniane w kierunku równoleżnikowym (b–d). Wznoszenie lub opadanie elementów gazu pociąga za sobą pole magnetyczne. Wskutek siły Coriolisa (wynikającej z rotacji obiektu) materia wraz z polem skręca się, tworząc pętle pola (e). Zgodnie ułożone pętle (f) prowadzą do powstania wielkoskalowego, wzmocnionego pola, o początkowej, południkowej geometrii (a).

*Obserwatorium Astronomiczne Uniwersytetu Jagiellońskiego

Do jego działania potrzebna jest rotująca różnicowo przewodząca materia (zjonizowany gaz ośrodka międzygwiazdowego, czyli plazma). W ośrodku musi występować turbulencja, czyli drobnoskalowe ruchy gazu. Ponadto konieczna jest konwekcja, czyli ciągle wznoszenie i opadanie elementów gazu. Przypuszcza się, że w galaktykach za turbulencje i konwekcję odpowiedzialne są głównie wybuchy gwiazd supernowych. W tych warunkach energia kinetyczna rotacji i turbulencji jest zamieniana na energię pola magnetycznego. Jak taki proces zachodzi w galaktykach różniących się tempem wybuchów supernowych, prędkością rotacji czy wielkością? Czy jest on wystarczający do wyjaśnienia obserwowanych w galaktykach struktur pola? Czy struktura spiralna galaktyki wpływa na proces dynamy? Odpowiedzi na te i inne pytania można poszukiwać, badając zjawisko magnetyzmu w galaktykach nietypowych. Powinny one wskazać na ograniczenia naszych teoretycznych koncepcji generacji pola.

W tym sensie małe galaktyki nieregularne stanowią doskonały poligon doświadczalny dla współczesnych teorii dynamy. Po pierwsze, pozbawione są spiralnych fal gęstości. Po drugie – wolno rotują. Można się spodziewać, że efekt wzmacniania pola przez wspomniane klasyczne dynamo jest słaby i działa powoli. Tymczasem odkryte pola magnetyczne w galaktyce nieregularnej NGC4449 (rys. 4) są zadziwiająco silne, tak silne jak w typowych galaktykach spiralnych.

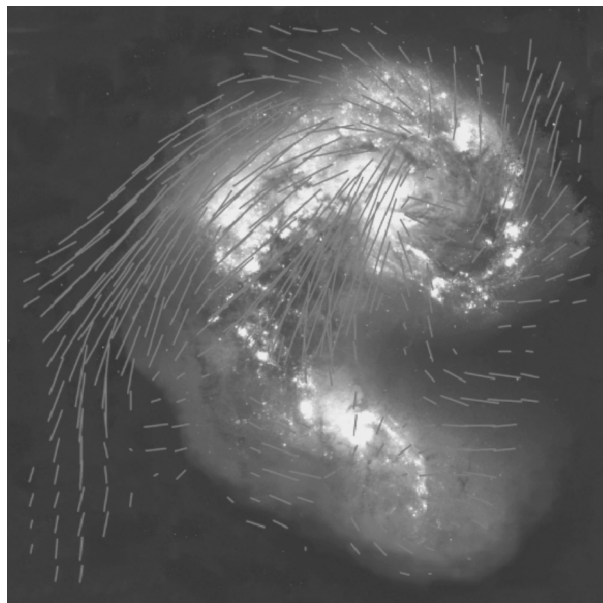


Rys. 4. Galaktyka nieregularna NGC4449 nie ma ramion spiralnych, jest stosunkowo mała i bardzo wolno rotuje. Niespodziewanie ma jednak silne pola magnetyczne wychodzące radialnie z jej centrum i okrążające ją niemal dookoła (fot. K.T. Chyży).

Musi tu zachodzić znacznie efektywniejszy proces generacji pól. Być może pasuje tutaj zmodyfikowana koncepcja tzw. szybkiego dynamy. Dodatkowo uwzględnia ona ciśnienie promieni kosmicznych (głównie relatywistycznych protonów) powstających w falach uderzeniowych po wybuchach supernowych. Prowadzi

to do szybszego unoszenia komórek gazu wypełnionych polem magnetycznym i znacznego przyspieszenia pracy dynamy.

Inną kategorią ekstremalnych galaktyk są kosmiczni piraci drogowi – galaktyki, które się zderzają. W znanym systemie „Anten” doszło do bliskiego spotkania dwóch galaktyk spiralnych i na skutek oddziaływania grawitacyjnego (pływowego) do znacznego odkształcenia ich struktur. Czy tak zdeformowany, chaotyczny ośrodek może mieć pole magnetyczne? Okazuje się, że tak (rys. 5).



Rys. 5. „Anteny” – zderzająca się para galaktyk. Na zdjęcie optyczne nałożone są kreski reprezentujące pole magnetyczne – wskazują one kierunek i regularną składową pola proporcjonalną do ich długości. Pole widoczne w środku zdjęcia zostało uporządkowane w wyniku przepływu gazu pomiędzy galaktykami, a widoczne z lewej strony zostało wyrzucone na zewnątrz układu wzdłuż „ogona pływowego”.

Jest ono nawet kilkakrotnie silniejsze niż w zwykłych galaktykach spiralnych. Z pewnością wiąże się to z dużym tempem formowania się gwiazd i częstymi wybuchami supernowych. O dziwo, silne pola magnetyczne występują również pomiędzy galaktykami. Tutaj dynamo działać nie może. Prawdopodobnie w tym przypadku pola przemieszczają się wraz z materią przepływającą pomiędzy galaktykami i porządkowane są w kierunku tego ruchu. W podobny sposób pola wyciągane są na zewnątrz układu, daleko w przestrzeń międzygalaktyczną.

Ostatnio badane galaktyki w gromadzie Panny również przejawiają niespotykane dotąd własności. Otóż pola magnetyczne prawie wszystkich tych galaktyk są asymetrycznie rozmieszczone, znacznie silniejsze po jednej stronie galaktyki. Świadczy to o działaniu jakiegoś dodatkowego procesu, poza galaktycznym dynamem. Możliwe, że związane to jest z ruchem galaktyki poprzez gęsty ośrodek gromady. Wtedy po nacierającej na ten ośrodek stronie galaktyki pole będzie ściskane, a przez to wyraźnie wzmacniane. Pole magnetyczne może też być deformowane jak w „Antenach”, podczas grawitacyjnego oddziaływania z okolicznymi galaktykami. Sprzyja temu duża

koncentracja obiektów, szczególnie w centrum gromady. Symulacje komputerowe pozwalają w takich przypadkach rozsądzić, który z tych mechanizmów zachodzi.



W jednej z galaktyk spiralnych gromady w Pannie, NGC4569, odkryto jeszcze inną strukturę pól magnetycznych. Otóż, symetrycznie po obu stronach galaktyki, daleko poza jej optycznym dyskiem, widnieją rozległe obszary emisji radiowej, przepełnione regularnym polem magnetycznym. Na pierwszy rzut oka ta podwójna struktura przypomina budowę odległych kwazarów. Jednak w przeciwieństwie do nich galaktyka ta nie zawiera aktywnego jądra ani masywnej czarnej dziury. Jej centrum zajmuje natomiast liczna populacja młodych, masywnych gwiazd. To one są prawdopodobnie przyczyną tego fenomenu. Są źródłem bardzo silnych wiatrów (znacznie silniejszych niż wiatr słoneczny) i wybuchów supernowych. Zdolne są wypychać materię międzygwiazdową wraz z polem magnetycznym poza granice galaktyki. „Wyrzucaniu” pól magnetycznych na tak wielkie, niespotykane dotąd odległości, pomogło prawdopodobnie przejście galaktyki blisko centrum gromady. Spowodowało ono odarcie galaktyki z pyłu i gazu w zewnętrznych jej częściach. Przez to w rzadszym ośrodku wypchnięcie materii wraz z polem magnetycznym stało się znacznie łatwiejsze. Szacunki pokazują, że oba obłoki z polem formowały się w czasie około 30 mln lat, a zmagazynowana w nich energia odpowiada wybuchom około 100 tysięcy supernowych.



Prezentowane tutaj badania ujawniły istnienie całego bogactwa struktur pola magnetycznego w galaktykach. Do ich wyjaśnienia klasyczny proces dynamiki jest dalece niewystarczający. Tam, gdzie istnienia pola w ogóle się nie spodziewano, a więc w galaktykach nieregularnych, pola okazały się bardzo silne. W zderzających się galaktykach dynamo również pracuje efektywniej niż w typowych galaktykach spiralnych. Ponadto pola generowane w procesie dynamiki są w tym przypadku silnie modyfikowane przez wielkoskalowe przepływy gazu oraz wyrzucane są daleko w przestrzeń międzygalaktyczną. Inną jeszcze ewolucję pól magnetycznych mamy w galaktykach w gromadzie Panny. Anomalie w polu magnetycznym wskazują na silną modyfikację ośrodka międzygwiazdowego spowodowanego ruchem galaktyk przez ośrodek gromady. Można podejrzewać, że obserwowane efekty będą się nasilać w odleglejszych galaktykach, pochodzących z wcześniejszych epok kosmologicznych. Ośrodek był wtedy znacznie gęstszy, a zderzenia galaktyk bardziej powszechne. Nowym wyzwaniem staje się zatem pytanie o strukturę pola magnetycznego w galaktykach wczesnego Wszechświata i możliwy wpływ pola na formowanie się pierwszych kosmologicznych struktur. Pozostaje mieć nadzieję, że już niedługo nowe instrumenty radioastronomiczne pozwolą rzucić pierwsze światło na te problemy.



Wielomian Ehrharta

Każdy, kto słyszał o twierdzeniu Picka, wie, że istnieje prosty związek między polem wielokąta o wierzchołkach w punktach kratowych, czyli takich, które mają wszystkie współrzędne całkowite, a liczbą tychże punktów w nim zawartych. Dla wielościanów podobna relacja także istnieje, choć jest bardziej zawiła. Zysk jest taki, że da się ją uogólnić do wyższych wymiarów, co pokazał E. Ehrhart w 1962 roku. Ustalmy zatem n -wymiarowy wielościan W w \mathbb{R}^n o wierzchołkach w punktach kratowych. Z takim wielościanem możemy związać w naturalny sposób ciąg liczbowy. Mianowicie, dla dowolnej liczby całkowitej dodatniej t liczymy, ile jest punktów kratowych w wielościanie W rozdętym t -krotnie, czyli w obrazie W przy jednokładności o skali t . Wynik oznaczamy przez k_t i ciąg gotowy. Ehrhart zaobserwował, że ciąg ten może być opisany przez pewien wielomian stopnia n o współczynnikach wymiernych:

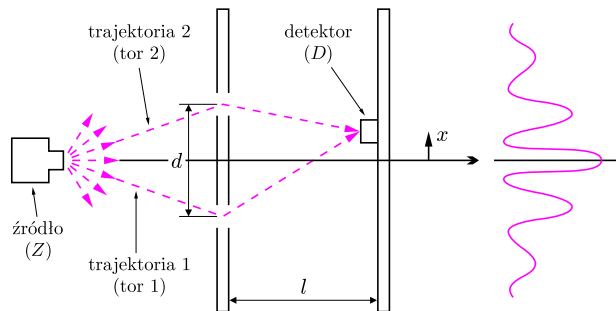
$$k_t = a_n t^n + a_{n-1} t^{n-1} + \dots + a_0.$$

Co więcej, wartości tych współczynników można zinterpretować geometrycznie. Objętość wielościanu to a_n , a jego charakterystyka Eulera–Poincarégo równa się a_0 . Z kolei a_{n-1} to połowa sumy odpowiednio unormowanych objętości ścian wielościanu. W przestrzeni $(n-1)$ -wymiarowej, zawierającej pewną ścianę wielościanu, objętość obliczamy przy założeniu, że najmniejszy zawarty w niej $(n-1)$ -wymiarowy prostopadłościan o wierzchołkach w punktach kratowych ma objętość 1. Znaczenie pozostałych współczynników wielomianu Ehrharta objawia się jednak dopiero przy badaniu pewnej stowarzyszonej z rozważanym wielościanem rozmaitości algebraicznej.

Zjawisko Aharonova–Bohma

Paweł Tomasz PEŹCZKOWSKI*

Jednym z bardziej spektakularnych przykładów pokazujących różnicę między mechaniką klasyczną i mechaniką kwantową jest zjawisko Aharonova–Bohma. Zjawisko to, które za chwilę przedstawimy dokładniej, zostało przewidziane już w 1949 roku przez Ehrenberga i Sidaya, a potem niezależnie przez Y. Aharonova i D. Bohma w 1959 roku. Przeanalizujemy je szczegółowo na przykładzie projektu doświadczenia przedstawionego w pracy [1]. Polega ono na badaniu interferencji elektronów w obecności pola magnetycznego.



Rys. 1. Doświadczenie interferencyjne z elektronami bez pola magnetycznego ($\mathbf{B} = 0$).

W doświadczeniu przedstawionym na rysunku 1 elektrony wychodzące ze źródła Z przepuszcza się przez dwie szczeliny, które są umieszczone w odległości d od siebie. Fale de Broglie'a tych elektronów, przechodząc przez szczeliny, uginają się, interferując ze sobą. Za przesłoną znajduje się ekran z ruchomym detektorem D zliczającym elektrony docierające doń w jednostce czasu. Badając liczbę elektronów rejestrowanych przez detektor, możemy wyznaczyć rozkład gęstości prawdopodobieństwa dotarcia elektronów do poszczególnych obszarów ekranu. Rozkład ten jest wynikiem interferencji dwóch fal prawdopodobieństwa, pochodzących z przejścia przez oba otwory w przesłonie. Wynik interferencji (wzmocnienie lub osłabienie sygnału) zależy od różnicy faz fal

$$\Delta\varphi = \frac{x}{l} \cdot \frac{d}{\lambda}.$$

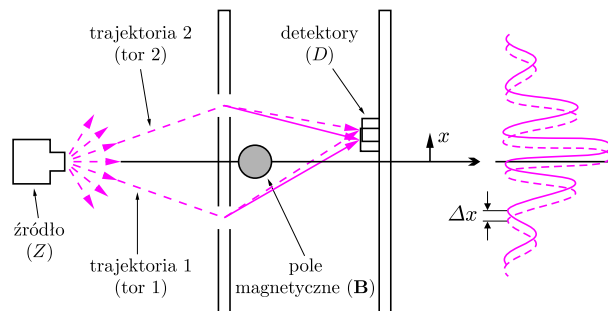
Jeżeli $x = 0$, to fale mają zgodne fazy (wtedy $\Delta\varphi = 0$) i prawdopodobieństwo znalezienia się elektronu w tym obszarze osiąga maksimum. Gdy $\Delta\varphi = \pi$, to fale mają przeciwne fazy, co odpowiada minimum prawdopodobieństwa znalezienia się elektronu w danym obszarze. W rezultacie otrzymujemy rozkład w postaci prążków interferencyjnych.

Rozważmy teraz wpływ pola magnetycznego na ten układ doświadczalny. W opisie klasycznym ruch elektronu w polu magnetycznym można opisać równaniem

$$m\ddot{\mathbf{r}} = q(\dot{\mathbf{r}} \times \mathbf{B}),$$

gdzie m , q , $\dot{\mathbf{r}}$, $\ddot{\mathbf{r}}$ jest odpowiednio masą, ładunkiem, prędkością i przyspieszeniem elektronu, \mathbf{B} zaś indukcją pola magnetycznego. W tym opisie pole magnetyczne

wpływa na ruch elektronu, gdy jego tor przechodzi przez obszar o niezerowej indukcji magnetycznej ($\mathbf{B} \neq 0$). Załóżmy jednak, że pole magnetyczne \mathbf{B} nie występuje wzdłuż klasycznych trajektorii elektronów 1 i 2. Można to osiągnąć, umieszczając solenoid za przesłoną w ten sposób, aby znajdował się pomiędzy obydwoma otworami.



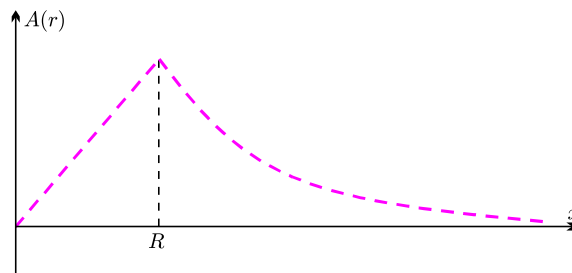
Rys. 2. Doświadczenie interferencyjne z elektronami z polem magnetycznym (solenoid).

Potencjał wektorowy, oznaczany \mathbf{A} , to pomocnicza wielkość fizyczna, którą możemy wprowadzić zamiast indukcji magnetycznej do opisu pola magnetycznego. W przeciwieństwie do indukcji nie jest wyznaczona jednoznacznie, gdyż dla danej konfiguracji pola magnetycznego istnieje nieskończenie wiele różnych pól potencjału jej odpowiadających. Z tego powodu nie jest to wielkość bezpośrednio mierzalna.

W celu dalszych rozważań wprowadźmy potencjał wektorowy \mathbf{A} odpowiadający nieskończenie długiemu solenoidowi z rysunku 2:

$$\mathbf{A}(\mathbf{r}) = \begin{cases} \frac{\mathbf{B} \times \mathbf{r}}{2} & \text{dla } \rho \leq R, \\ \frac{R^2 \cdot \mathbf{B} \times \mathbf{r}}{2\rho^2} & \text{dla } \rho > R, \end{cases}$$

gdzie R jest promieniem cewki, ρ zaś odległością punktu od jej osi.



Rys. 3. Pole wektorowe $\mathbf{A}(\mathbf{r})$ dla nieskończenie długiego solenoidu.

Z rysunku 3 widać, że potencjał wektorowy nie znika na zewnątrz solenoidu, mimo że pole magnetyczne \mathbf{B} jest w tym obszarze równe zero. Fakt ten odgrywa istotną rolę w dalszych rozważaniach.

W ramach mechaniki kwantowej ruch elektronu w polu magnetycznym opisywany jest równaniem Schrödingera zawierającym potencjał wektorowy \mathbf{A} , a nie bezpośrednio wektor indukcji magnetycznej. Z rozwiązania tego równania wynika, że po umieszczeniu solenoidu w układzie doświadczalnym, tak jak to pokazano na rysunku 2, wystąpi przesunięcie prążków

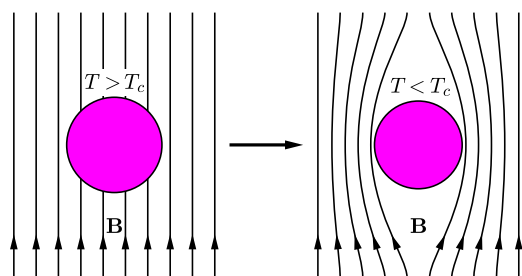
*Instytut Fizyki Doświadczalnej, Uniwersytet Warszawski

interferencyjnych na ekranie o $\frac{q}{h} \frac{1}{d} \lambda \phi_B$, gdzie ϕ_B oznacza strumień pola magnetycznego \mathbf{B} przez obszar między drogami 1 i 2. Wielkość ta nie jest równa zero, nawet gdy obie drogi przebiegają w miejscu, gdzie samo $\mathbf{B} = 0$. Wystarczy bowiem, że pole magnetyczne obecne jest gdziekolwiek pomiędzy drogami, np. tylko wewnątrz solenoidu, aby jego strumień nie zniknął. Opisywane przesunięcie prążków było obserwowane przez kilka niezależnych zespołów badawczych, np. R. G. Chambers (1960), H. Boersch i in. (1960), A. Tonomura i in. (1982).

Z punktu widzenia fizyki klasycznej zjawisko Aharonova–Bohma daje wynik zaskakujący, gdyż widoczny jest w nim wpływ pola magnetycznego na ruch elektronów w sytuacji, gdy występuje ono jedynie w obszarze, do którego elektrony nie docierają! Z tego powodu możliwość pojawienia się podobnego efektu została zakwestionowana przez niektórych fizyków: Bocchieri, Loinger (1978) i Roy, Singh (1983) zanegowali istnienie zjawiska Aharonova–Bohma, przypisując obserwowane przesunięcie prążków interferencyjnych przenikaniu pola magnetycznego \mathbf{B} na zewnątrz solenoidu.

Całkowity kres zarzutom związanym z przenikaniem pola magnetycznego położyło doświadczenie przeprowadzone przez A. Tonomurę i in. w 1986 roku. Magnes w kształcie torusa zbudowany ze stopu 80% Ni i 20% Fe (taki stop nazywa się permalloyem) został całkowicie pokryty materiałem nadprzewodzącym (w tym przypadku niobem, Nb). Grubość osłony niobowej była większa niż głębokość, na jaką pole magnetyczne wnika do nadprzewodnika (dla niobu: $\lambda_{Nb} \sim 10^{-7}$ m).

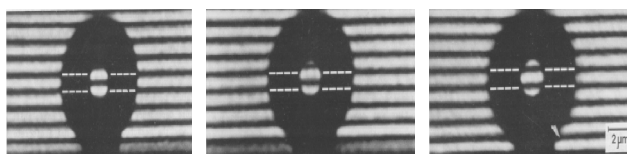
Gdy próbka nadprzewodnika, umieszczona w polu magnetycznym w temperaturze zbyt wysokiej na nadprzewodnictwo, zostanie schłodzona do temperatury niższej niż temperatura krytyczna (czyli temperatura przejścia do stanu nadprzewodzącego), to pole magnetyczne przenikające próbkę zostanie z niej wypchnięte (rys. 4). Ściślej mówiąc, pole wnika do wewnątrz nadprzewodnika tylko na niewielką głębokość, zwaną londonowską głębokością wnikania. Zjawisko to nosi nazwę efektu Meissnera–Ochsenfelda.



Rys. 4. Zjawisko Meissnera–Ochsenfelda. Wypychanie pola magnetycznego z wnętrza nadprzewodnika.

Niob ekranował więc pole magnetyczne od solenoidu, nie pozwalając mu przeniknąć do obszaru na zewnątrz. Przeprowadzono też badanie zależności przesunięcia obrazu interferencyjnego od temperatury. Jeżeli temperatura była niższa od temperatury krytycznej nadprzewodnika (9,2 K dla niobu), to strumień magnetyczny nie przenikał na zewnątrz nadprzewodnika.

Zauważono, że układ prążków interferencyjnych gwałtownie zmienił się w momencie przekroczenia tej temperatury (rys. 5).



Rys. 5. Obraz interferencyjny uwidaczniający przesunięcie funkcji falowej dla toroidalnego ferromagnetyka odpowiednio w temperaturze: $T = 4,5$ K, $T = 15$ K, $T = 300$ K.

Analogiczne doświadczenie wykonano również z wykorzystaniem interferencji elektronów w nadprzewodniku. Interferencję taką można obserwować przy użyciu tzw. *złącza Josephsona*. Złącze Josephsona to zestaw dwóch nadprzewodników oddzielonych cienką warstwą izolatora. W części nadprzewodzącej prąd przepływa bez oporu, czyli bez strat energii. Jeżeli warstwa izolująca jest dostatecznie cienka, to nośnik prądu elektrycznego, czyli para elektronów, może przeniknąć przez nią dzięki zjawisku tunelowania i w układzie popłynie prąd nadprzewodzący. Ponadto, gdy prąd płynie przez dwa złącza dwiema różnymi drogami, różnica faz powstała w wyniku przepływu prądu powoduje interferencję pomiędzy złączami. Jeśli w układzie zostanie umieszczony solenoid, przesunięcie prążków interferencyjnych będzie zależało od wielkości strumienia magnetycznego wewnątrz solenoidu. Podobnie jak poprzednio, pole nie jest obecne wewnątrz nadprzewodników, gdzie tym razem odbywa się ruch elektronów.

Zjawisko Aharonova–Bohma uwidacznia nielokalny charakter zjawisk opisywanych przez mechanikę kwantową. Według jej zasad pole magnetyczne może wpływać na elektrony, nawet gdy nie „dotyka” ich bezpośrednio, czyli występuje tylko w miejscach odległych od trajektorii ich ruchu.

Literatura:

- [1] Y. Aharonov, D. Bohm, *Significance of Electromagnetic Potentials in the Quantum Theory*, Phys. Rev. 115, 485, 1959.
- [2] R.G. Chambers, *Shift of an Electron Interference Pattern by Enclosed Magnetic Flux*, Phys. Rev. Lett. 5, 3, 1960.
- [3] H. Boersch, H. Hamisch, D. Wohlleben, K. Grohman, *Antiparallele Weissche Bereiche als Biprisma für Elektroneninterferenzen*, Z. Phys. 159, 397, 1960.
- [4] M. Peshin, A. Tonomura, *The Aharonov–Bohm Effect*, Springer-Verlag, Berlin, Heidelberg, 1989.
- [5] A. Tonomura i inni, *Evidence for Aharonov–Bohm effect with magnetic field completely shielded from electron*, Phys. Rev. Lett. 56, 792, 1986.
- [6] W. Meissner, R. Ochsenfeld, *Ein neuer Effect bei Eintritt der Supraleitfähigkeit*, Naturwiss. 21, 787, 1933.
- [7] R.C. Jaklevic, J.J. Lambe, A.H. Silver, J.E. Mercereau, *Quantum Interference from a Static Vector Potential in a Field-Free Region*, Phys. Rev. Lett. 12, 274, 1964.

Dlaczego $2 \cdot 2 = 4$, czyli twierdzenia o rekursji

Adam KOLANY*

Matematyk nie musi wiedzieć, ile jest $2 \cdot 2$ — on musi wiedzieć, **czemu** to jest 4. Z kolei logik zagadnięty w tej materii, zanim odpowie, najpierw zada cztery pytania: co to jest „dwa”, co to jest „ile”, co to jest „razy” i co to jest „jest”. Tę ostatnią kwestię pozostawmy filozofom, przyjmując, że wiemy, co to znaczy. Co to znaczy „ile”, nadaje się na całkiem osobny artykuł, więc i tym nie będziemy sobie teraz zaprzętać głowy (szczególnie że chodzi tylko o „cztery”). Co to jest „dwa”, wyjaśni się niebawem. Tym, czym zajmiemy się w tym artykule, jest to, co znaczy „razy”.

We współczesnej teorii mnogości liczby naturalne definiuje się jako elementy najmniejszego zbioru **induktywnego** (oznaczanego zazwyczaj symbolem ω). Najmniejszego, czyli zawartego w każdym innym zbiorze induktywnym. Zbiór A , z kolei, jest induktywny, jeśli zbiór pusty \emptyset (który uznamy za liczbę naturalną „zero”) doń należy ($0 \in A$) oraz spełnia on następujący warunek:

$$a \in A \Rightarrow a \cup \{a\} \in A.$$

W dalszym ciągu zamiast $a \cup \{a\}$ pisać będziemy $S(a)$ i nazywać będziemy to **następnikiem** a , czymkolwiek a jest. Wówczas, wedle pomysłu J. von Neumanna, $1 = S(0) = \{0\}$, $2 = S(1) = \{0, 1\}$, $3 = S(2) = \{0, 1, 2\}$, $4 = S(3) = \{0, 1, 2, 3\}$, etc.

Definicja zbioru liczb naturalnych jako najmniejszego zbioru induktywnego implikuje znaną ze szkoły zasadę **indukcji matematycznej**: jeśli jakaś własność Φ przysługuje liczbie zero oraz z tego, że przysługuje dowolnej liczbie n , wynika, że przysługuje liczbie następnej $S(n)$, to własność ta przysługuje wszystkim liczbom naturalnym. Niech bowiem A będzie zbiorem tych i tylko tych liczb naturalnych, które mają własność Φ . Wówczas, oczywiście, $0 \in A$ oraz jeśli $a \in A$, to także i $S(a) \in A$, czyli A jest zbiorem induktywnym. Skoro ω zawarty jest w każdym zbiorze induktywnym, oznacza to, że w szczególności zawiera się w A . I tym sposobem dla dowolnego $a \in \omega$ zachodzi $\Phi(a)$.

Rozważmy teraz rodzinę \mathfrak{R}_\oplus wszystkich relacji $\rho \subseteq (\omega \times \omega) \times \omega$ (czyli relacji wiążących pary liczb naturalnych z pojedynczymi liczbami naturalnymi), które spełniają następujące dwa warunki: $\langle a, 0 \rangle \rho a$, dla dowolnego $a \in \omega$, oraz jeśli już $\langle a, b \rangle \rho c$, to także $\langle a, S(b) \rangle \rho S(c)$, gdzie $a, b, c \in \omega$. Oczywiście, rodzina ta jest niepusta, bo relacja pełna $(\omega \times \omega) \times \omega$ do niej należy. Niech dalej $\rho_\oplus = \bigcap \mathfrak{R}_\oplus$. Nietrudno wykazać za pomocą zasady indukcji, że dla dowolnej pary liczb $a, b \in \omega$ istnieje $c \in \omega$, dla którego $\langle a, b \rangle \rho_\oplus c$. Nieco więcej zachodu wymaga zauważenie, że owo c jest jedyne. Gdyby bowiem dla jakiejś pary a, b tych „ c ” było dwa, powiedzmy c_1 i c_2 , to wówczas relacja $\rho' = \rho_\oplus \setminus \{\langle \langle a, b \rangle, c_1 \rangle\}$ należałaby nadal do rodziny \mathfrak{R}_\oplus , skąd wynikałoby, że $\rho_\oplus \subseteq \rho'$, co nie jest prawdą. Skoro tak, to znaczy, że relacja ρ_\oplus jest funkcją, a to jedyne c , dla którego zachodzi $\langle a, b \rangle \rho_\oplus c$, nazywać będziemy **sumą** liczb a i b i oznaczać je symbolem $a + b$. Policzmy sobie coś. Skoro dla dowolnej relacji $\rho \in \mathfrak{R}_\oplus$ zachodzi $\langle a, 0 \rangle \rho a$, oznacza to, że $a + 0 = a$ (co zresztą doskonale „wiemy”). Skoro $\langle a, 0 \rangle \rho a$, to także $\langle a, S(0) \rangle \rho S(a)$, dla $\rho \in \mathfrak{R}_\oplus$. Czyli, innymi słowy, $a + 1 = S(a)$. Tj. dodanie jedynki, to po prostu wzięcie następnika danej liczby. Posiłkując się zasadą indukcji matematycznej (czyli właściwie definicją liczb naturalnych), można wykazać doskonale nam znane własności operacji dodawania liczb naturalnych – przemienność i łączność:

$$a + b = b + a \quad \text{i} \quad a + (b + c) = (a + b) + c, \quad a, b, c \in \omega.$$

Warunki ujęte w definicji rodziny \mathfrak{R}_\oplus przekładają się na znane skądinąd formuły:

$$\begin{cases} a + 0 = a \\ a + (b + 1) = (a + b) + 1 \end{cases}$$

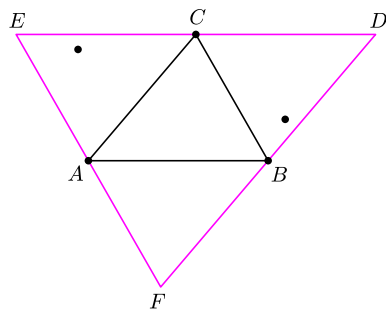
dla $a, b \in \omega$.

John von Neumann (ur. 28 grudnia 1903 r. w Budapeszcie – zm. 8 lutego 1957 r. w Waszyngtonie), inżynier chemik, fizyk, matematyk i informatyk. Wniósł znaczący wkład do wielu dziedzin matematyki, szczególnie teorii gier i uporządkował formalizm matematyczny mechaniki kwantowej. Uczestniczył w projekcie Manhattan. Przyczynił się do rozwoju numerycznych prognoz pogody.



Rozwiązanie zadania M 1162.

Spośród wszystkich $\binom{5}{3} = 10$ trójkątów o wierzchołkach w danych punktach wybierzmy taki, który ma największe pole. Oznaczmy jego wierzchołki przez A, B, C . Przez każdy wierzchołek trójkąta ABC poprowadźmy prostą równoległą do przeciwległego boku, uzyskując trójkąt DEF , o polu niewiększym niż 4.



Ponieważ każdy trójkąt o wierzchołkach w danych punktach ma pole niewiększe od pola trójkąta ABC , więc pozostałe dwa punkty leżą w trójkącie DEF . Ponadto w jednym z trójkątów ABF, BCD, CEA nie leży żaden spośród danych punktów. Odcinając trójkąt ten od trójkąta DEF , uzyskujemy trapez o polu niewiększym od 3 i zawierający dany 5 punktów.

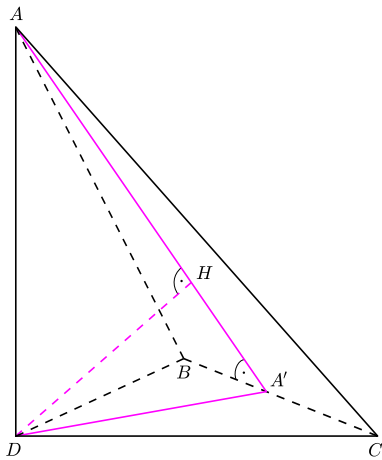
*Wydział Matematyki i Informatyki, Uniwersytet Jagielloński



Rozwiązanie zadania M 1163.

W rozwiązaniu skorzystamy z następującego twierdzenia: *prosta p jest prostopadła do płaszczyzny π wtedy i tylko wtedy, gdy jest prostopadła do pewnych dwóch nierównoległych prostych zawartych w płaszczyźnie π .*

Niech A' będzie rzutem prostokątnym punktu A na prostą BC .



Prosta AD jest prostopadła do płaszczyzny BCD , skąd $AD \perp BC$. Ponadto $AA' \perp BC$, a więc z zacytowanego twierdzenia wynika, że prosta BC jest prostopadła do płaszczyzny $AA'D$. Zatem każda prosta zawarta w płaszczyźnie $AA'D$ jest prostopadła do prostej BC .

Niech H będzie spodkiem wysokości w trójkącie $AA'D$ opuszczonej z wierzchołka D . Wtedy $DH \perp AA'$ oraz $DH \perp BC$. Zatem prosta DH jest prostopadła do płaszczyzny ABC . Wobec tego spodek wysokości czworostianu $ABCD$, opuszczonej z wierzchołka D leży na wysokości AA' trójkąta ABC .

Analogicznie dowodzimy, że spodek ten leży na pozostałych wysokościach trójkąta ABC , co kończy dowód.



Rozwiązanie zadania F 687.

Temperatura pary jest proporcjonalna do średniego kwadratu prędkości molekuł:

$$T = \rho \frac{v_{sr}^2}{3R}.$$

Zatem dla oszacowania temperatury trzeba znaleźć średni kwadrat prędkości. Ruch cząstek w pionie odbywa się pod wpływem siły ciężkości. Z jednej strony czas ruchu cząstki to $\tau = \frac{l}{v_{sr}}$, z drugiej jest to czas spadku swobodnego równy $\sqrt{\frac{2h}{g}}$. Stąd $v_{sr} = l\sqrt{\frac{g}{2h}}$, co ostatecznie daje:

$$T = \frac{\rho l^2 g}{6Rh} \approx 600 \text{ K}.$$

Zostawmy jednak dodawanie, bo interesuje nas przecież mnożenie liczb.

Niech \mathfrak{R}_\odot będzie rodziną wszystkich relacji $\rho \subseteq (\omega \times \omega) \times \omega$, które spełniają następujące warunki: $\langle a, 0 \rangle \rho 0$ dla $a \in \omega$ oraz jeśli $\langle a, b \rangle \rho c$, to także $\langle a, S(b) \rangle \rho (c + a)$. Oczywiście, znowu relacja pełna jest w \mathfrak{R}_\odot i znowu możemy mówić o relacji $\rho_\odot = \bigcap \mathfrak{R}_\odot$, o której bez trudu, podobnie jak było w przypadku relacji ρ_\oplus , pokażemy, że jest funkcją, czyli że dla dowolnych $a, b \in \omega$ istnieje **jedyn**e c , dla którego $\langle a, b \rangle \rho_\odot c$. To jedyne c nazywać będziemy **iloczynem** liczb a i b i oznaczać będziemy symbolem $a \cdot b$. Teraz, skoro $\langle a, 0 \rangle \rho 0$ dla $a \in \omega$ i $\rho \in \mathfrak{R}_\odot$, możemy to zapisać jako $a \cdot 0 = 0$. Ponadto dostajemy stąd, że $\langle a, 1 \rangle \rho (0 + a) = a$ dla $\rho \in \mathfrak{R}_\odot$, czyli że $a \cdot 1 = a$ dla $a \in \omega$. Bezpośrednio z definicji rodziny \mathfrak{R}_\odot dostajemy także, że $a \cdot (b + 1) = a \cdot b + a$ dla $a, b \in \omega$, co stanowi uszczuploną wersję prawa rozdzielności mnożenia względem dodawania. Pełną wersję tego prawa, jak i też inne własności mnożenia, takie jak przemienność i łączność, dostaniemy, pracowicie korzystając z zasady indukcji matematycznej.

Ile zatem jest $2 \cdot 2$? Mamy:

$$2 \cdot 2 = 2 \cdot (1 + 1) = 2 \cdot 1 + 2 = 2 + 2 = 2 + (1 + 1) = (2 + 1) + 1 = 3 + 1 = 4.$$

Uważny Czytelnik zauważył pewien schemat w naszym dotychczasowym wywodzie. Rozważaliśmy rodziny relacji ρ spełniające pewne dwa warunki. Pierwszy warunek stwierdzał, że pary postaci $\langle a, 0 \rangle$ są w relacji z jakąś liczbą $g(a)$. Drugi warunek był implikacją: jeśli $\langle a, b \rangle$ był w relacji z c , to wtedy także $\langle a, S(b) \rangle$ jest w relacji z „czymś tam” zależnym od a, b i c . Uściślenie tej obserwacji prowadzi do następującego twierdzenia.

Twierdzenie (o rekursji). *Niech A i P będą dowolnymi zbiorami i niech $g : P \rightarrow A$ oraz $h : P \times \omega \times A \rightarrow A$. Wówczas istnieje dokładnie jedna funkcja $f : P \times \omega \rightarrow A$, spełniająca równania:*

$$\begin{cases} f(p, 0) = g(p) \\ f(p, n + 1) = h(p, n, f(p, n)) \end{cases}$$

dla $p \in P, n \in \omega$.

Dwa przykłady zastosowania tego twierdzenia widzieliśmy wyżej ($A = P = \omega$, jakie g i h ?).

Przyjrzyjmy się dokładniej innym.

Przykład

- **(potęgowanie)** *Niech $A = P = \omega$ i niech $g(p) = 1$ oraz $h(p, n, w) = p \cdot w$ dla $p, n, w \in \omega$. Wówczas funkcja f z twierdzenia o rekursji to $f(p, n) = p^n$, gdzie $p, n \in \omega$*

- **(silnia)** *Niech $P = \{0\}$, $A = \omega$ oraz $g(0) = 1$ i $h(p, n, w) = w \cdot (n + 1)$, dla $p \in P, n, w \in \omega$. Wówczas $f(0, n)$, gdzie f jest funkcją, o której mowa w twierdzeniu o rekursji, to $n!$.*

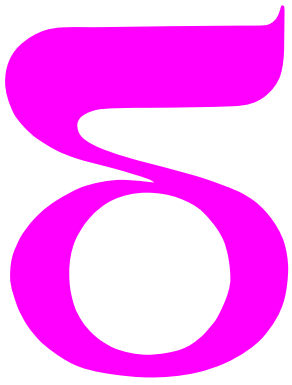
- **(poprzednik)** *Niech $P = \{0\}$, $A = \omega$ oraz $g(0) = 0$ i $h(p, n, w) = n$, dla $p \in P, n, w \in \omega$. Wówczas $f(0, n)$, gdzie f jest funkcją, o której mowa w twierdzeniu o rekursji, to 0 dla $n = 0$, a dla $n > 0$ to liczba, której następnikiem jest n .*

Funkcję tę nazywać będziemy **funkcją poprzednika** i oznaczymy ją symbolem \mathbf{p} .

- **(odejmowanie)** *Niech $P = A = \omega$ oraz $g(a) = a$ i $h(p, n, w) = \mathbf{p}(w)$, dla $p \in P, n, w \in \omega$. Wówczas $f(p, n)$, gdzie f jest funkcją, o której mowa w twierdzeniu o rekursji, to 0 dla $p \leq n$, a dla $p > n$ to jedyna liczba k , dla której $n + k = p$ (czyli różnica p i n , oznaczamy $p \dot{-} n$)*

Twierdzenie o rekursji można na wiele sposobów uogólniać, ale to temat na całkiem inny artykuł.

Pozostaje tylko zapytać, czy wszystkie „sensowne” działania na liczbach naturalnych można tak zdefiniować. Np. dzielenie z resztą? Okazuje się, że nie wszystkie, chociaż dzielenie z resztą akurat tak.



mała delta

Wyświetlacze ciekłokrystaliczne

W ciągu ostatnich kilku lat kolorowe wyświetlacze ciekłokrystaliczne (LCD) zdobyły szturmem rynek elektronicznych urządzeń użytkowych. Są obecne w zegarkach, telefonach komórkowych, aparatach cyfrowych po monitory i telewizory, o najrozmaitszych kształtach i wielkościach. Są wciąż znacznie tańsze i prostsze w produkcji niż wyświetlacze plazmowe i znacznie mniejsze niż kineskopowe.

Kolorowy wyświetlacz to nic innego jak matryca złożona z ogromnej liczby małych, różnokolorowych lampek. Niby proste, lecz projektanci każdego wyświetlacza muszą rozwiązać kilka problemów: pojedynczy świecący element (piksel) musi być bardzo mały, ale jednocześnie musi mieć możliwość wysyłania światła w trzech podstawowych barwach (np. czerwonej, zielonej i niebieskiej), o bardzo różnej intensywności (aby uzyskać dobry kontrast), i wreszcie świecące elementy muszą szybko reagować na żądanie zmiany jasności (aby na obrazie nie powstawały smugi podczas wyświetlania czegoś w ruchu). Pożądane jest też, by urządzenie było jak najmniejsze i zużywało możliwie mało energii.

Widać, że sprawa nie jest prosta. Zanim wyjaśnimy, jak poradzili sobie twórcy LCD, proponuję mały eksperyment. Proszę spojrzeć na włączony wyświetlacz LCD przez polaryzator. Jeśli ktoś nie ma go w domu, można w zastępstwie użyć np. kawałka szkła, folii lub plastiku i obserwować odbicie obrazu przy niewielkim kącie padania. Gdy kręcimy polaryzatorem lub plastikiem, łatwo dostrzec, że widoczna jasność wyświetlacza zmienia się, co wskazuje, że światło z wyświetlacza jest całkowicie spolaryzowane liniowo (zdjęcia).

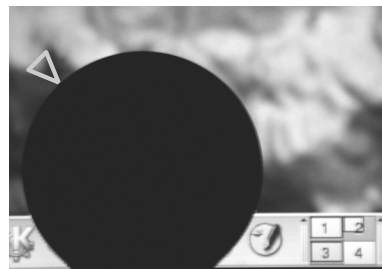
Zjawisko polaryzacji światła służy w matrycy LCD do kontrolowania jasności piksela. Zasada działania jest prosta: światło wysyłane np. przez diodę lub dowolne źródło dające białe światło o dużej, stałej jasności przechodzi przez cienką folię polaryzującą je liniowo. Dalej znajduje się „magiczna substancja” skręcająca płaszczyznę polaryzacji przechodzącego światła o kąt

regulowany przyłożonym do niej napięciem. Wreszcie drugi stały polaryzator liniowy, skrzyżowany z pierwszym, wygasza światło w mniejszym lub większym stopniu w zależności od kąta skręcenia. Na koniec kolorowy filtr zabarwia światło na któryś z podstawowych kolorów.

Wspomnianą „magiczną substancją” jest tzw. ciekły kryształ. Jest to substancja, której cząsteczki mają postać długich, raczej sztywnych łańcuchów. W odpowiedniej temperaturze i ciśnieniu cząsteczki mają tendencję do ustawiania się mniej więcej w jednym kierunku. W kryształcie użytym w LCD kierunek uporządkowania skręca o 90 stopni wzdłuż drogi światła. Dzięki temu kryształ skręca płaszczyznę polaryzacji przechodzącego przezeń światła. Jeśli jednak włączymy pole elektryczne, siły elektrostatyczne będą starały się ustawić wszystkie cząsteczki wzdłuż linii sił pola i skręcenie zmniejszy się. Przy odpowiednio dużym napięciu skręcenie zniknie i światło będzie całkiem wygaszane przez drugi polaryzator.

Ciekłe kryształy reagują szybko na zmiany przyłożonego napięcia, a użycie polaryzatorów pozwala regulować jasność piksela w bardzo szerokim zakresie: układ wygasza niemal całe światło, gdy światło ma płaszczyznę polaryzacji prostopadłą do drugiego polaryzatora, a przepuszcza większość, gdy równoległą.

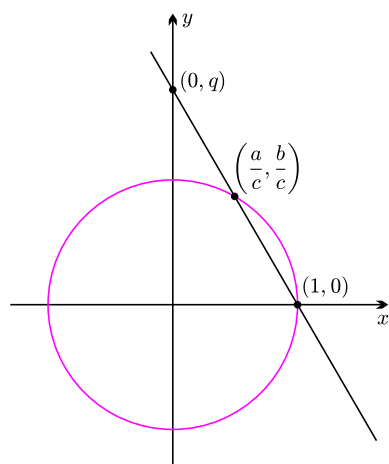
Stare, czarno-białe wyświetlacze ciekłokrystaliczne znane z kalkulatorów, działają według podobnych zasad, ale zamiast własnego podświetlenia od tyłu korzystają z odbitego od wewnętrznej części wyświetlacza światła zewnętrznego.



Trójki pitagorejskie

Stwierdzenie, czy dla ustalonej trójki liczb całkowitych dodatnich (a, b, c) istnieje trójkąt, którego boki mają taką właśnie długość, jest proste. Wystarczy, aby suma dowolnych dwóch liczb z tej trójki była większa od trzeciej. Jeśli ten warunek jest spełniony, to od razu możemy powiedzieć, czy trójkąt będzie prostokątny. Kwadrat największej liczby musi być równy sumie kwadratów pozostałych. Takie trójki nazywamy pitagorejskimi. Każda trójka pitagorejska daje nam nowe trójki postaci (ka, kb, kc) , dlatego wystarczy rozpatrywać tylko trójki, które nie są wielokrotnością jakiejś innej. Będziemy je nazywać pierwotnymi, a do ich wyznaczenia przyda nam się okrąg.

Każdej pierwotnej trójce pitagorejskiej (a, b, c) przyporządkujemy punkt $(a/c, b/c)$ na okręgu zadanym równaniem $x^2 + y^2 = 1$. Obie współrzędne tego punktu są wymierne. Ponieważ tę procedurę możemy odwrócić (ułanki sprowadzamy do wspólnego mianownika), wszystkie pierwotne trójki pitagorejskie są zakodowane przez takie właśnie punkty na okręgu.



Idźmy jednak dalej. Prosta przechodząca przez punkty $(1,0)$ i $(a/c, b/c)$ przetnie oś OY w punkcie $(0, q)$, gdzie q jest liczbą wymierną większą od 1 (rys.). Odwrotnie, weźmy względnie pierwsze liczby naturalne m i n , tak by $m > n$ i poprowadźmy prostą przez $(1,0)$ i $(0, m/n)$. Przecina ona okrąg w dwóch punktach. Jeden z nich to oczywiście $(1,0)$. Współrzędne drugiego wyznaczamy, rozwiązując układ równań:

$$\begin{cases} y = -\frac{m}{n}x + \frac{m}{n} \\ x^2 + y^2 = 1 \end{cases}$$

Pamiętając, że punkt $(1,0)$ jest jednym z rozwiązań i korzystając ze wzoru na sumę pierwiastków równania kwadratowego, znajdujemy drugie:

$$\left(\frac{m^2 - n^2}{m^2 + n^2}, \frac{2mn}{m^2 + n^2} \right).$$

Gdy m i n są różnej parzystości, ułamki te są nieskracalne, stąd ogólna postać pierwotnej trójki pitagorejskiej to $(m^2 - n^2, 2mn, m^2 + n^2)$. Uzyskany wzór umożliwia podanie warunku koniecznego na to, by trójka (a, b, c) była trójką pitagorejską bez obliczania kwadratów. Mianowicie, w każdej trójce pitagorejskiej musi wystąpić liczba podzielna przez 3, liczba podzielna przez 4 oraz liczba podzielna przez 5. Mogą to być różne liczby, tak jak w $(3, 4, 5)$, a może to być jedna wspólna liczba dla wszystkich dzielników, tak jak w $(11, 60, 61)$. Przykładowy dowód: jeśli 3 dzieli m lub n , to jest to oczywiste. W przeciwnym przypadku, m^2 i n^2 przystają do 1 modulo 3 i $m^2 - n^2$ jest szukaną liczbą. Podobnie dowodzimy dla 5. Dla 4 wynika to bezpośrednio z faktu, że jedna z liczb m i n jest parzysta.

Powyższa metoda może być użyta do poszukiwania rozwiązań w liczbach wymiernych także innych równań wielomianowych stopnia dwa z całkowitymi współczynnikami (wykorzystała ją, na przykład, M. Brambilla w *Delcie* 10/2006). Najważniejsze, żeby najpierw znaleźć przynajmniej jeden punkt o współrzędnych wymiernych na zadanej przez równanie krzywej. Ale czy taki punkt zawsze istnieje?

Małą Deltę przygotowali Mikołaj KORZYŃSKI i Marcin HAUZER

Ciąg, który lubi nierówności

Jan SZEJKO

Praca nagrodzona złotym medalem
w Konkursie Uczniowskich Prac z Matematyki, Gdańsk 2006.

Prace na tegoroczny konkurs można nadsyłać do 1 maja 2007 roku.
Regulamin KUPzM zamieściliśmy w numerze 1(392)/2007.

Rozważmy ciąg (a_n) zdefiniowany rekurencyjnie dla $n \in \mathbb{Z}_+$ w następujący sposób:

$$a_1 = 1, \quad a_n = \sum_{0 < k < n, (n,k)=1} a_k$$

dla $n > 1$ (gdzie (n, k) oznacza największy wspólny dzielnik n i k). Zatem a_n to suma tych poprzednich wyrazów, które noszą numery względnie pierwsze z n ; np. $a_6 = a_1 + a_5$, $a_8 = a_1 + a_3 + a_5 + a_7$.

Wartości tego ciągu dla początkowych liczb naturalnych przedstawiam poniżej.

n	1	2	3	4	5	6	7	8	9	10
a_n	1	1	2	3	7	8	22	32	66	91

n	11	12	13	14	15	16	17	18
a_n	233	263	729	1038	2059	3119	7674	8666

Będą nas interesować nierówności spełniane przez wyrazy ciągu a_n . Zaczniemy od najprostszych, które natychmiast wynikają z definicji:

$$(1) \quad a_{n+1} \geq a_n,$$

$$(2) \quad a_{2n+1} \geq a_{2n} + a_{2n-1}.$$

Pierwsza z nich wynika stąd, że $(n+1, n) = 1$, druga zaś stąd, że $(2n+1, 2n-1) = 1$ dla każdego n . Równie oczywiste są nierówności w drugą stronę:

$$(3) \quad a_{2n+1} \leq a_{2n} + a_{2n-1} + a_{2n-2} + \dots + a_1,$$

$$(4) \quad a_{2n} \leq a_{2n-1} + a_{2n-3} + a_{2n-5} + \dots + a_1$$

(druga wynika stąd, że wszystkie liczby względnie pierwsze z liczbą parzystą są nieparzyste). Zauważmy, że w dwóch ostatnich nierównościach równość zachodzi w nieskończenie wielu przypadkach dla n dodatnich: w pierwszej, zawsze gdy $2n+1$ jest liczbą pierwszą, w drugiej zaś dla każdego n będącego potęgą dwójki.

Mniej trywialna jest nierówność

$$(5) \quad a_{2n+1} > a_{2n} + a_{2n-2} + \dots + a_2.$$

Jej dowód przebiega przez indukcję względem n . Dla $n = 1$ teza oczywiście zachodzi. Przypuśćmy, że zachodzi również dla $n = k$, gdzie $k \in \mathbb{Z}_+$. Wówczas, korzystając z (2) i z założenia indukcyjnego, otrzymujemy

$$a_{2k+3} \geq a_{2k+2} + a_{2k+1} > a_{2k+2} + a_{2k} + a_{2k-2} + \dots + a_2,$$

co kończy dowód kroku indukcyjnego.

Z (5) oraz (1) natychmiast wynika kolejna nierówność:

$$(6) \quad a_{2n+1} > a_{2n-1} + a_{2n-3} + \dots + a_1.$$

Zauważmy jeszcze, że korzystając z (4) oraz (6), dostajemy

$$a_{2n} \leq a_{2n-1} + a_{2n-3} + \dots + a_1 < a_{2n-1} + a_{2n-1},$$

czyli

$$(7) \quad a_{2n} < 2a_{2n-1}.$$

Kiedy zacząłem zajmować się tytułowym ciągiem, zainteresowało mnie, jak zachowują się ilorazy kolejnych wyrazów: czy rosną do nieskończoności, czy może zbiegają do jakiejś liczby, czy też zachowują się w jeszcze inny sposób. Szybko zauważyłem, że choć ilorazy te bywają bardzo bliskie liczby 3, nie wyglądają na to, by ją kiedykolwiek przekraczały. Z początku próbowałem udowodnić tę zależność indukcyjnie, korzystając z powyższych nierówności, jednak bezskutecznie. Na trop właściwej metody wpadłem dopiero wtedy, gdy usiłowałem udowodnić nierówność

$$(8) \quad a_{2n} + a_{2n-2} > a_{2n-1} + a_{2n-3} + \dots + a_1.$$

Tej nierówności również nie sposób dowieść przez prostą indukcję. Jednak gdy oznaczymy przez p najmniejszą liczbę pierwszą nieparzystą, która nie dzieli n , to liczba $2n-2$ nie dzieli się przez żadną liczbę pierwszą nieparzystą mniejszą niż p , a więc jest względnie pierwsza z każdą liczbą nieparzystą mniejszą od p . Zatem $2n$ jest względnie pierwsza z liczbami $2n-1$ i $2n-p$, a $2n-2$ z liczbami $2n-3, 2n-5, \dots, 2n-p$. Wobec tego

$$a_{2n} \geq a_{2n-1} + a_{2n-p},$$

$$a_{2n-2} \geq a_{2n-3} + a_{2n-5} + \dots + a_{2n-p}.$$

Sumując powyższe nierówności i korzystając z (6), dostajemy

$$\begin{aligned} a_{2n} + a_{2n-2} &\geq a_{2n-1} + a_{2n-3} + \dots + a_{2n-p} + a_{2n-p} > \\ &> a_{2n-1} + a_{2n-3} + \dots + a_{2n-p} + \\ &\quad + a_{2n-p-2} + a_{2n-p-4} + \dots + a_1, \end{aligned}$$

czyli dowodzoną nierówność.

Podobnej, choć nieco subtelniejszej, sztuczki użyłem w dowodzie (który tutaj pominię) nierówności

$$(9) \quad a_{2n} > a_{2n-2} + a_{2n-3} + \dots + a_1,$$

z której pożądana przez nas nierówność

$$(10) \quad a_{n+1} < 3a_n$$

jest prostym wnioskiem. Dla n nieparzystych mamy bowiem mocniejszą nierówność (7), a dla $n = 2k$, korzystając z (3) oraz (9), dostajemy

$$\begin{aligned} a_{2k+1} &\leq a_{2k} + a_{2k-1} + a_{2k-2} + \dots + a_1 < \\ &< a_{2k} + a_{2k-1} + a_{2k} \leq 3a_{2k}. \end{aligned}$$

Warto przy okazji zauważyć, że nierówność (9) można łatwo uogólnić, ponieważ z (2) dostajemy

$$a_{2n+1} \geq a_{2n-1} + a_{2n} > a_{2n-1} + a_{2n-2} + \dots + a_1.$$

Zatem dla każdego $n > 2$

$$(11) \quad a_n > a_{n-2} + a_{n-3} + \dots + a_1.$$

Pojawia się pytanie, czy nierówność (10) można poprawić, tzn. czy istnieje stała mniejsza niż 3, przy której nierówność byłaby nadal prawdziwa. Okazuje się jednak, że nie, bo w dowodzie nierówności (10) wszystkie istotne oszacowania odbywały się na poziomie wyrazów rzędu a_{2n-p} , które przy dużym p są małe w porównaniu z a_{2n} . Ścisły dowód pomijam.

Z (10) natychmiast wynika, że $a_{n+2} < 9a_n$. Tę nierówność można jednak znacząco poprawić. Pokażemy teraz, że dla $n \in \mathbb{Z}_+$ zachodzi

$$(12) \quad a_{n+2} < 5a_n.$$

Skorzystamy kolejno z (3), (9), (7), (4), (10) oraz (8):

$$a_{2k+1} \leq a_{2k} + a_{2k-1} + a_{2k-2} + \dots + a_1 < \\ < 2a_{2k} + a_{2k-1} < 5a_{2k-1},$$

$$a_{2k+2} \leq a_{2k+1} + a_{2k-1} + a_{2k-3} + \dots + a_1 < \\ < 3a_{2k} + a_{2k} + a_{2k-2} < 5a_{2k}.$$

Okazuje się, że współczynnik 5 w tej nierówności wciąż nie jest optymalny, jednak dowód tego faktu tutaj pominę. Z kolei ze stałą 4 nierówność nie będzie prawdziwa, ponieważ mamy $a_{32}/a_{30} \approx 4,3$. Stąd widać, że optymalna stała dla tej nierówności będzie niecałkowita. Można przypuszczać, że będzie to liczba niewymierna, a nawet przestępna, jednak ciężko wyobrazić sobie, jak można by to udowodnić.

Interesujące jest, czy dla jakiegoś $k > 2$ optymalna stała c w nierówności $a_{n+k} < ca_n$ będzie liczbą całkowitą. Obliczenia na komputerze pokazują, że największy iloraz a_{n+6}/a_n , dla n mniejszych niż 10 000, to około 35,996. Dość długo usiłowałem udowodnić nierówność $a_{n+6} < 36a_n$, jednak bezskutecznie. W końcu zacząłem sprawdzać ją dla jeszcze większych n i okazało się, że

$$\frac{a_{30036}}{a_{30030}} \approx 36,02.$$

Widać więc, że nie zawsze warto polegać na hipotezach, które działają dla dziesięciu tysięcy przypadków.

Teraz przejdę do hipotezy, którą obliczenia potwierdzają z dużym zapasem, i co do której nie wyobrażam sobie, żeby mogła się kiedyś okazać fałszywa. Stwierdza ona, że dla wszelkich $n, k \in \mathbb{Z}_+$ zachodzą nierówności

$$(13) \quad a_k a_n \leq a_{n+k} < 5a_k a_n.$$

W rzeczywistości sądzę, że po lewej stronie mogłoby być $a_{k+1}a_n$, ale nie mam za taką nierównością żadnych argumentów, poza danymi doświadczalnymi i dowodami szczególnych przypadków dla małych k . Najważniejsze jednak w mojej hipotezie jest to, że skrajne strony różnią się o stały czynnik.

Okazuje się, że obie powyższe nierówności wynikają z mocniejszej, choć – przy pewnym obyciu z naszym ciągiem – dosyć intuicyjnej hipotezy. Niech $P(n)$ oznacza najmniejszą liczbę pierwszą, która nie dzieli n . Wspomniana hipoteza stwierdza, że dla każdego $k \in \mathbb{Z}$ istnieje takie $l > 0$, że dla wszelkich $n, m \in \mathbb{Z}_+$ takich, że $P(n) + l \leq P(m)$ i $n + k > 1$ zachodzi nierówność

$$(14) \quad \frac{a_{n+k}}{a_n} \leq \frac{a_{m+k}}{a_m}.$$

Pokażę teraz szkic dowodu (bez korzystania z tej hipotezy) lewej nierówności z (13) w przypadku, gdy $P(n+k) \geq k$.

Dla wygody zmieńmy trochę oznaczenia. Zapiszmy nierówność w postaci

$$(15) \quad a_n \geq a_k a_{n-k} \quad \text{dla } n > k > 0, P(n) \geq k.$$

W dowodzie posłużyłem się pomocniczym ciągiem dwuwymiarowym $(b_{i,j})$, zdefiniowanym dla $i \geq j \geq 1$ w następujący sposób:

$$b_{1,1} = 1, \quad b_{i,j} = \sum_{0 < k < j} b_{k,k}, \quad \text{gdy } i \cdot j \neq 1.$$

Można pokazać przez indukcję względem j , że zachodzi nierówność

$$a_n \geq \sum_{i=j}^k b_{i,j} a_{n-i}, \quad \text{gdy } P(n) \geq k, 1 \leq j \leq k.$$

Z niej w szczególności dla $j = k$ wynika, że przy $P(n) \geq k$

$$a_n \geq b_{k,k} a_{n-k}.$$

Jednak wprost z definicji ciągu $(b_{i,j})$ wynika, że dla każdego $i \in \mathbb{Z}_+$ mamy $b_{i,i} = a_i$. Stosując to do ostatniej nierówności, dostajemy nierówność (15). W przypadku, gdy $P(n) < k$, pożądana nierówność wynika łatwo z udowodnionego przypadku oraz z hipotezy (14).

W przypadku prawej nierówności z (13) nie udało mi się wiele udowodnić bez korzystania z hipotezy (14), przy założeniu zaś jej prawdziwości dowód przebiega dość łatwo, indukcyjnie względem k , przy czym trzeba w kroku indukcyjnym rozważyć osobno przypadki, gdy $P(n) > k$ oraz gdy $P(n) \leq k$. Stąd widać, że (13) wynika z hipotezy (14).

Zastanówmy się teraz nad konsekwencjami hipotezy (13). Zauważmy najpierw, że jeśli mamy nierówność

$$a_{n+k} \leq ca_n$$

dla każdego $n \in \mathbb{Z}_+$ i ustalonych $k \in \mathbb{Z}_+$ i $c \in \mathbb{R}_+$, to możemy stąd wnioskować, że ciąg (a_n) można oszacować z góry przez ciąg geometryczny o ilorazie $\sqrt[k]{c}$. Jest tak, ponieważ każdy z ciągów $(a_i, a_{k+i}, a_{2k+i}, \dots)$, dla $i = 1, 2, \dots, k$, można – co bardzo łatwo wykazać indukcyjnie – słabo oszacować z góry przez ciąg $(a_i, ca_i, c^2 a_i, \dots)$.

Zachodzi więc nierówność

$$a_{mk+i} \leq a_i (\sqrt[k]{c})^{mk} = \frac{a_i}{(\sqrt[k]{c})^i} (\sqrt[k]{c})^{mk+i}$$

dla każdego $m \in \mathbb{Z}_+$. Oznaczmy $a_i / (\sqrt[k]{c})^i$ przez c_i . Widać więc, że jeśli oznaczymy przez c' największą z liczb c_i dla $i = 1, 2, \dots, k$, to zachodzić będzie nierówność $a_n \leq c' (\sqrt[k]{c})^n$. Analogiczne rozumowanie można przeprowadzić dla nierówności w przeciwną stronę.

Jeśli więc hipoteza (13) jest prawdziwa, to dla dowolnego $k \in \mathbb{Z}_+$ ciąg (a_n) można ograniczyć z dołu i z góry przez dwa ciągi geometryczne o wyrazach dodatnich i o ilorazach odpowiednio $\sqrt[k]{a_k}$ oraz $\sqrt[k]{5a_k}$.

Jest oczywiste, że jeśli dany ciąg można ograniczyć z dołu przez ciąg geometryczny o ilorazie q , to nie można go ograniczyć z góry przez ciąg o ilorazie $q' < q$ (i *vice versa*). W przeciwnym bowiem przypadku mielibyśmy dla każdego $n \in \mathbb{Z}_+$ i dla pewnych c, c' dodatnich

$$cq^n < a_n < c'q'^n,$$

czyli

$$c \left(\frac{q}{q'} \right)^n < c'.$$

Jednak po lewej stronie mamy rosnącą funkcję wykładniczą, zatem otrzymaliśmy sprzeczność.

Łatwo zauważyć, że wraz ze wzrostem k różnica liczb $\sqrt[k]{a_k}$ i $\sqrt[k]{5a_k}$ dąży do zera. Wynika to szybko stąd, że $(\sqrt[k]{a_n})$ jest ograniczony z góry. Z poprzedniego spostrzeżenia wynika, że jeśli moja hipoteza jest prawdziwa, to kres górny $\sqrt[k]{a_k}$ nie przekracza kresu dolnego $\sqrt[k]{5a_k}$, zatem musi istnieć taka liczba a , że dla każdego $k \in \mathbb{Z}_+$ będzie

$$\sqrt[k]{a_k} \leq a \leq \sqrt[k]{5a_k}.$$

Skoro jednak skrajne strony mają wspólną granicę przy $k \rightarrow \infty$, to z twierdzenia o trzech ciągach tą granicą musi być a . Widzimy więc, że istnieje taka liczba, że ciąg (a_n) możemy ograniczyć z dołu przez ciąg geometryczny o dowolnym ilorazie dodatnim mniejszym od tej liczby, z góry zaś przez ciąg o dowolnym ilorazie od niej większym. Pokazuje to, że choć iloraz a_{n+1}/a_n nie dąży do żadnej liczby przy $n \rightarrow \infty$, to (przy

założeniu prawdziwości hipotezy (13)) zachowanie ciągu (a_n) przypomina trochę zachowanie ciągu geometrycznego o ilorazie a .

Na nierównościach spełnianych przez wyrazy ciągu nie kończą się jego ciekawe własności. Szukając informacji o nim w Internecie, wpisałem jego początkowe wyrazy na stronie *The On-Line Encyclopedia of Integer Sequences* (<http://www.research.att.com/~njas/sequences/>). Okazało się, że istnieje on w tamtejszej bazie danych jako ciąg A045545. Znalazłem tam też informację, że Benoit Cloitre udowodnił, iż równość

$$a_{n+2} = 2a_n + a_{n+1}$$

zachodzi wtedy i tylko wtedy, gdy n i $n+2$ są liczbami pierwszymi bliźniaczymi. Widać stąd, jak szeroką gamę problemów otwiera omawiany ciąg. Zmagania z nim dały mi wiele radości i satysfakcji.

Czytelników zainteresowanych tematem przedstawionym w niniejszym artykule zapraszam do przeczytania całej pracy, której aktualna wersja znajduje się pod adresem <http://students.mimuw.edu.pl/~js248325/praca.pdf>.



Zadania

Redaguje Waldemar POMPE

M 1162. Na płaszczyźnie danych jest 5 punktów o tej własności, że pole każdego trójkąta o wierzchołkach w tych punktach nie przekracza 1. Wykazać, że punkty te leżą w pewnym trapezie o polu nieprzekraczającym 3. Rozwiązanie na str. 6

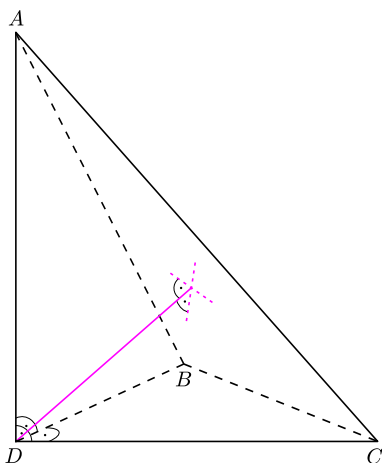
M 1163. Dany jest czworościan $ABCD$, w którym kąty płaskie przy wierzchołku D są proste (rys. 1). Wykazać, że spodek wysokości tego czworościanu, opuszczonej z wierzchołka D , pokrywa się z punktem przecięcia wysokości trójkąta ABC . Rozwiązanie na str. 7

M 1164. Niech $n \geq 2$ będzie liczbą naturalną. Wykazać, że każdy wyraz ciągu $n! + 1, n! + 2, \dots, n! + n$ ma dzielnik pierwszy, który nie jest dzielnikiem żadnego innego wyrazu tego ciągu. Rozwiązanie na str. 16

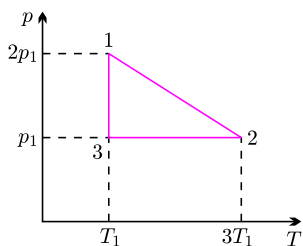
Redaguje Ewa CZUCHRY

F 687. Z butli, w której znajdują się silnie zagęszczone pary potasu, ucieka przez wąską poziomą rurkę wiązka atomów. Oszacować temperaturę par, wiedząc, że średnie obniżenie wiązki w odległości $l = 50$ cm od butli wynosi $h = 3 \mu\text{m}$. Rozwiązanie na str. 7

F 688. Nad gazem idealnym wykonano cykl przemian pokazany na rysunku 2. Znaleźć stosunek maksymalnej objętości gazu do minimalnej, osiągniętych w tym cyklu przemian, objętości. Rozwiązanie na str. 16



Rys. 1



Rys. 2

Śpiewające piaski

Urodziliśmy się odkrywcami, a potem posłano nas do szkoły. . .

Mądrzy nauczyciele i uczniowie z wytrwałością sięgają ferment wątpliwości, ale mogą być co najwyżej ziarnami piasku w trybach młynów powoli przerabiających młodzież we wszystkich wiekach, mentalnych starców. Ten proces otepiania nie wynika z niczyjej złej woli. Jest, po prostu, jednym z dowodów działania doboru naturalnego.

Sposoby ogłupiania również ewoluują. Ostatnie lata to era nieomyślności internetu.

Proszę tylko nie myśleć, że uważam powyższy proces za coś złego. Przecież nie może być złe coś, co jest naturalne. To dzięki niemu trwamy jako gatunek. Ewolucja nie jest jednak możliwa bez mutacji. Mogą one być korzystne lub nie. Naszym (np. *Delty*) zadaniem jest wspomaganie genu chronicznej ciekawości. Jak przekonać kogoś, że, wbrew temu, czego uczą (bądź uczyli) go w szkole, nie wszystko jeszcze wiemy, a jeszcze mniej rozumiemy (w dodatku prawdopodobnie błędnie)? Może posłużyć się przykładem odkrycia, które było w zasięgu ręki od kilkuset lat?

Marco Polo w rozdziale LVII *Opisanie Świata* wspomina o donośnych odgłosach, o stałym tonie, dochodzących z pustyni. Choć ten, jeden z najsłynniejszych podróżników, być może nigdy nie dotarł dalej niż nad Morze Czarne, to fenomen śpiewających piasków jest jak najbardziej rzeczywisty i przez wieki pozostawał, najpierw w ogóle nie, a później nie w pełni wyjaśnionym spektakularnym zjawiskiem naturalnym.

Tak długo czekaliśmy na zadowalające rozwiązanie częściowo ze względu na rzadkość zjawiska. Tylko niektóre wydmy potrafią śpiewać. Historycznie pierwszym wytłumaczeniem, przytoczonym już przez Marco Polo, było przypisanie tych dziwnych dźwięków duchom pustyni. Obecnie wyjaśnienie to uważa się za nienaukowe, tzn. doświadczalnie niesprawdzalne. Nawet jeżeli są to duchy, to jak to robią?

Praktycznie wszystkie instrumenty muzyczne grają, wykorzystując rezonans. Nic dziwnego, że jakiegoś rezonansu poszukiwano również na pustyni, ale nie udało się go znaleźć. Wysokość dźwięku nie zależy od wielkości wydmy. Również przypisanie dźwięku samemu wiatrowi okazało się błędne. Ostatecznie zrozumiano, że dźwięk wywołują przesypujące się lawiny piasku, ale nadal nie było wiadomo, jaki jest mechanizm jego powstawania.

Pierwszą ilościową zależność udało się przewidzieć Bagnoldowi w 1966 roku. Zaproponował on proporcjonalność częstości do pierwiastka z ilorazu przyspieszenia ziemskiego i średnicy ziaren d : $f \sim \sqrt{g/d}$. Grające pola piasku składają się z obłych ziaren praktycznie tej samej wielkości i właśnie od tej wielkości zależy wysokość dźwięku. Nadal jednak nie wiadomo było, w jaki sposób dźwięk powstaje.

Wydaje się, że wyjaśnienie opublikował wreszcie w zeszłym roku [1] zespół naukowców, który zajmuje się tym problemem od kilku lat. Pomogło zbadanie śpiewających piasków „na czterech krańcach świata”: w Maroku, Omanie, Chinach i Chile, oraz przeprowadzenie dwóch eksperymentów, jednego w regionie Tarfayah w Maroku,

a drugiego w laboratorium (ale z użyciem piasku zebranego w rejonie pierwszego eksperymentu).

Eksperyment laboratoryjny bezapelacyjnie pokazał, że do śpiewu piasek nie potrzebuje pustyni. Doświadczenie przeprowadzono za pomocą urządzenia przypominającego przemysłową dziezę do mieszania ciasta. W toroidalnym, od góry otwartym naczyniu umieszczono piasek, który przesypywała obracająca się ze stałą, ale regulowaną prędkością łopatką o dodatkowo regulowanym zanurzeniu. Okazało się, że w ten sposób można wydobyć dźwięk w zakresie od 100 Hz do 350 Hz oraz że jego częstość jest proporcjonalna do stosunku prędkości i wysokości hałdki przesypującego się piasku. Dodatkowo stwierdzono, że brzmienie nie powstaje, jeżeli prędkość łopatki jest mniejsza niż około 47 cm/s. Podobny efekt progowy zaobserwowano w eksperymencie przeprowadzonym na pustyni, polegającym na kontrolowanym wywoływaniu lawin piaskowych w dużym, drewnianym korycie. Tym razem jednak brzmienie nie występowało, o ile prędkość osypywania się piasku nie przekraczała około 23 cm/s, czyli zaobserwowano próg dla dwa razy mniejszej prędkości.

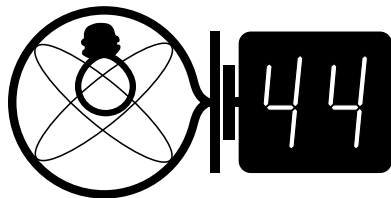
Spójny obraz zjawiska można uzyskać, przyjmując hipotezę zamieszczoną w podręczniku z 1922 roku napisanym przez J.H. Poyntinga i J.J. Thomsona. Według niej powstający dźwięk jest związany ze względnym ruchem ziaren zsypanych się warstwy piasku. Opublikowane w zeszłym roku wyniki pozwalają lepiej zrozumieć to zjawisko. Dźwięk wywołuje oscylująca powierzchnia osypującego się piasku. Żeby efekt występował, ziarna muszą poruszać się na tyle szybko, aby charakterystyczny czas ruchu pojedynczego ziarna był krótszy niż czas rozchodzenia się zaburzenia w zaangażowanej w przesypywanie części piasku. W przeciwnym przypadku wywołująca dźwięk fala „płynącego” piasku nie powstaje. W każdym razie taka interpretacja pozwala wyjaśnić czynnik dwa w prędkości krytycznej między wymuszonym osypywaniem (laboratorium) i spontanicznym (wydmy), jeżeli przyjmiemy, że w pierwszym przypadku fala obejmuje dwa razy dłuższą, w jednostkach długości fali, część piasku. Różnica ta jest, według autorów, spowodowana różnymi warunkami brzegowymi. Wyjaśnienie jest dość mętne, ale jedyną możliwością jest przyjęcie, że w pierwszym przypadku grająca warstwa piasku ma dwa węzły, a w drugim może mieć strzałkę na jednym z końców.

Zależność częstości od grubości ziaren w przypadku naturalnego śpiewu piasku jest związana z tym, że grubość zsuwającej się lawiny piasku ma zawsze mniej więcej tyle samo warstw, a więc jest proporcjonalna do średnicy ziaren. Śpiewające piaski muszą być bardzo sypkie, a więc jak najbardziej suche oraz pokryte warstwą zwietrzenia, w każdym razie zbyt intensywnie szuflowany w laboratorium piasek tracił muzykalność.

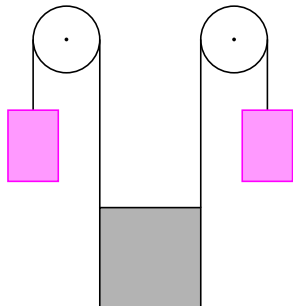
Autorzy wydali również płytę z muzyką piasków. Bycie naukowcem naprawdę może być ciekawym zajęciem.

Piotr ZALEWSKI

[1] S. Douady, A. Manning, P. Hersen, E. Elbelrhiti, S. Protière, A. Daerr i B. Kabbachi, *Song of the Dunes as a Self-Synchronized Instrument*, Phys. Rev Lett. **97**, 018002 (2006)



Termin nadsyłania rozwiązań:
31 V 2007



Rys. 1

Dobrocią obwodu drgającego nazywa się wielkość $2\pi E/\Delta E$, gdzie E – energia drgań, ΔE – strata energii w ciągu jednego okresu. Dla drgań słabo tłumionych jest ona równa ilorazowi częstości ω przez podwojony wykładnik tłumienia (parametr α w czynniku $e^{-\alpha t}$).



Rys. 2

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 przesyłać 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 <http://www.mimuw.edu.pl/delta/regulamin.html>.

Zadania z fizyki nr 434, 435

Redaguje Jerzy B. BROJAN

434. Jednorodna kwadratowa płytko o boku a wisi na dwóch długich niciach przełożonych przez bloki, a na drugim końcu każdej nici wisi ciężarek o masie równej połowie masy płytki (rys. 1). Obliczyć częstotliwość małych obrotów płytki w płaszczyźnie rysunku.

Wskazówka: moment bezwładności płytki względem osi przechodzącej przez jej środek i prostopadłej do jej płaszczyzny jest równy $\frac{1}{6}ma^2$ (m – masa płytki).

435. Z kondensatora o pojemności C i zwojnicy zestawiono obwód. Drgania elektryczne w tym obwodzie okazały się tłumione, z dwóch powodów:

- kondensator charakteryzuje się pewną niewielką upływnością, co można przedstawić jako równoległe dołączone do niego opornik o dużym oporze R_1 ,
- uzwojenie zwojnicy ma pewną niewielką oporność R_2 .

Aby poprawić dobroć obwodu, możemy wsuwać do zwojnicy lub wysuwać rdzeń ferromagnetyczny, zmieniając w ten sposób indukcyjność L . Straty energii w rdzeniu (wynikające z histerezy i prądów wirowych) są pomijalnie małe. Jak należy wybrać wartość L , aby dobroć osiągnęła wartość maksymalną?

Rozwiązania zadań z fizyki z numeru 11/2006

Przypominamy treść zadań:

426. Cienki jednorodny pręt położono na poziomym stole kończącym się pionową krawędzią (rys. 2) w pozycji prostopadłej do tej krawędzi. Następnie powoli przesuwno pręt w stronę krawędzi, aż zaczął się przechylać i ześlizgiwać. Tarcie między prętem a stołem nie występuje, tzn. siła reakcji stołu jest stale prostopadła do pręta. Który punkt pręta jako ostatni utraci kontakt z podłożem? Ile w tym momencie będzie wynosił kąt przechyłu pręta? Dopuszczalna jest odpowiedź oparta na obliczeniach numerycznych.

Jak zmieniłyby się odpowiedzi na powyższe pytania, gdyby wartość przyspieszenia ziemskiego g uległa podwojeniu?

427. Przypuśćmy, że masa neutronu byłaby o 0,1% mniejsza od rzeczywistej, przy niezmienionej masie protonu i elektronu. Jak wpłynęłoby to na właściwości atomu wodoru? Niezbędne dane wzięć z tablic.

426. Wprowadźmy oznaczenia: x i y – współrzędne wektora położenia środka pręta względem krawędzi stołu (przy czym oś y ma zwrot w dół), s – długość tego wektora, α – kąt przechyłu pręta, m – jego masa, I – moment bezwładności, R – siła reakcji krawędzi stołu. Równania ruchu postępowego i obrotowego pręta mają postać: $m = \ddot{x}R \sin \alpha$, $m\ddot{y} = mg - R \cos \alpha$, $I\ddot{\alpha} = Rs$, gdzie kropką oznaczono pochodną względem czasu. Wykorzystując związki $x = s \cos \alpha$, $y = s \sin \alpha$, można wyeliminować zmienne x , y i R , dochodząc do układu dwóch równań:

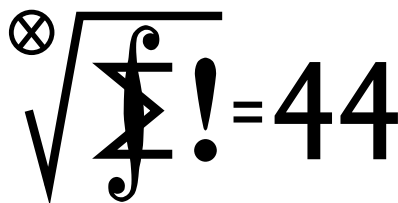
$$\ddot{s} = g \sin \alpha + s\dot{\alpha}^2, \quad \ddot{\alpha} = \frac{ms(g \cos \alpha - 2\dot{s}\dot{\alpha})}{I + ms^2}.$$

Autorowi udało się scałkować te równania jedynie numerycznie. (Można, oczywiście, przywołać stałą wartość całkowitej energii, ale nie wystarcza to do pełnego rozwiązania układu). Należy podstawić $I = \frac{1}{12}ml^2$

(l – długość pręta), jako warunki początkowe przyjmując trzy spośród wielkości s , \dot{s} , α , $\dot{\alpha}$ równe zero, a czwartą bardzo małą, i zakończyć obliczenia w chwili osiągnięcia punktu, w którym $\ddot{\alpha} = 0$ (wtedy $R = 0$, tzn. następuje oderwanie). Wynik jest następujący: oderwanie pręta od krawędzi stołu nastąpi przy wartości $s = 0,25l$, $\alpha = 0,686 \text{ rad} = 39,3^\circ$.

Nietrudno sprawdzić, że podwojenie g odpowiada przeskalowaniu czasu o czynnik $\sqrt{2}$ i nie ma wpływu na powyższe wyniki.

427. Gdyby masa neutronu była o 0,1% mniejsza od rzeczywistej, byłaby mniejsza od sumy mas protonu i elektronu. Zatem atom wodoru nie byłby trwały – przekształciłby się w neutron wskutek wychwytu elektronu orbitalnego, zgodnie ze schematem $p + e^- \rightarrow n + \nu$. Materia w Kosmosie składałaby się głównie z neutronów.



Termin nadsyłania rozwiązań:
31 V 2007

Czołówka ligi zadaniowej
Klub 44 F
po uwzględnieniu ocen rozwiązań zadań
422 (WT = 3,16) i **423** (WT = 3,40)
z numeru 9/2006

Mateusz Łacki	– Kraków	42,38
Tomasz Tkocz	– Rybnik	38,12
Marian		
Lupieżowicz	– Zebrzydowice	35,00
Konrad		
Kapcia	– Częstochowa	33,49
Tomasz Wietecha	– Tarnów	27,66
Jerzy Witkowski	– Radlin	27,36
Krzysztof Magiera	– Łosów	21,94
Andrzej		
Nowogrodzki	– Chocianów	17,01
Jacek Konieczny	– Poznań	16,87

Zadania z matematyki nr 537, 538

Redaguje Marcin E. KUCZMA

537. Liczby a, b, c są długościami boków trójkąta. Udowodnić, że

$$\frac{\sqrt{b+c-a}}{\sqrt{b}+\sqrt{c}-\sqrt{a}} + \frac{\sqrt{c+a-b}}{\sqrt{c}+\sqrt{a}-\sqrt{b}} + \frac{\sqrt{a+b-c}}{\sqrt{a}+\sqrt{b}-\sqrt{c}} \leq 3.$$

538. Pięciokąt $ABCDE$ jest wpisany w okrąg. Punkty A', C', D' są odpowiednio ortocentrami trójkątów BAE, BCE, BDE . Dowieść, że trójkąty ACD i $A'C'D'$ są przystające.

Zadanie 538 zaproponował pan Michał Kieza z Warszawy.

Rozwiązania zadań z matematyki z numeru 11/2006

Przypominamy treść zadań:

529. Dane są liczby całkowite x_1, x_2, x_3, x_4, x_5 oraz liczba całkowita dodatnia m , będąca dzielnikiem zarówno sumy liczb x_i , jak i sumy ich kwadratów. Wyjaśnić, czy z tych założeń wynika, że m jest także dzielnikiem liczby

$$x_1^5 + x_2^5 + x_3^5 + x_4^5 + x_5^5 - 5x_1x_2x_3x_4x_5.$$

530. Wysokości trójkąta ostrokątnego ABC przecinają się w punkcie H . Proste poprowadzone z punktu A są styczne do okręgu o średnicy BC w punktach P i Q . Udowodnić, że punkty P, Q i H są współliniowe.

529. Liczby x_i są pierwiastkami wielomianu

$$P(x) = \prod_{i=1}^5 (x - x_i) = x^5 + Ax^4 + Bx^3 + Cx^2 + Dx + E$$

o współczynnikach

$$A = -\sum_{i=1}^5 x_i,$$

$$B = \sum_{i < j} x_i x_j = \frac{1}{2} \left(A^2 - \sum_{i=1}^5 x_i^2 \right),$$

...

$$E = -\prod_{i=1}^5 x_i.$$

W myśl założenia liczby A oraz $2B$ są podzielne przez m . Należy zaś wyjaśnić, czy liczba $F = \sum_{i=1}^5 x_i^5 + 5E$ dzieli się przez m . Odpowiedź łatwo wynika z zależności

$$0 = \sum_{i=1}^5 P(x_i) = \sum_{i=1}^5 x_i^5 + A \sum_{i=1}^5 x_i^4 + B \sum_{i=1}^5 x_i^3 + C \sum_{i=1}^5 x_i^2 + D \sum_{i=1}^5 x_i + 5E \equiv F + B \sum_{i=1}^5 x_i^3 \pmod{m}.$$

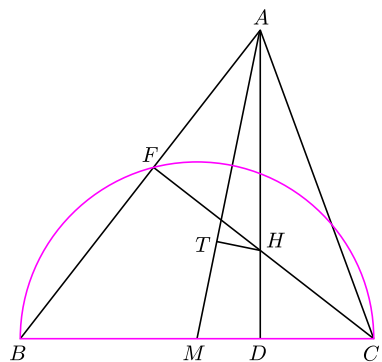
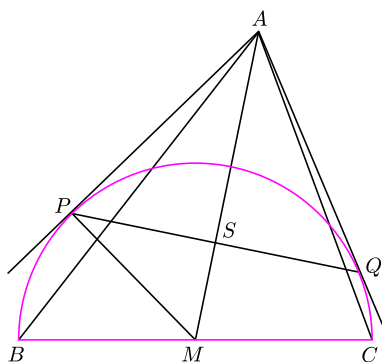
Jeśli $B \equiv 0$, to oczywiście $F \equiv 0 \pmod{m}$. Niech teraz $B \not\equiv 0 \pmod{m}$. Wiemy, że $2B \equiv 0$, więc liczba m musi być parzysta. Zatem suma $\sum x_i$ (podzielna przez m) jest liczbą parzystą, więc i suma $\sum x_i^3$ jest liczbą parzystą. Wobec tego

$$F \equiv -B \sum_{i=1}^5 x_i^3 = -2B \cdot \left(\frac{1}{2} \sum_{i=1}^5 x_i^3 \right) \equiv 0 \pmod{m}.$$

Wniosek: niezależnie od przypadku, badana liczba F dzieli się przez m .

530. Oznaczmy przez M środek boku BC , przez S – punkt przecięcia odcinków AM i PQ , przez D – rzut prostokątny punktu A na odcinek BC , przez F – rzut prostokątny punktu C na odcinek AB , a przez T – rzut prostokątny punktu H na odcinek AM . Ponieważ $PQ \perp AS$, zadanie sprowadza się do wykazania, że punkty S i T pokrywają się – czyli że $|AS| = |AT|$.

Przyjmijmy, że punkty P, Q leżą odpowiednio po tych stronach prostej AM , co punkty B, C . Z podobieństw trójkątów prostokątnych $\triangle ATH \sim \triangle ADM$, $\triangle AFH \sim \triangle ADB$ oraz $\triangle ASP \sim \triangle APM$ wynikają równości $|AT| \cdot |AM| = |AH| \cdot |AD| = |AF| \cdot |AB|$ oraz $|AS| \cdot |AM| = |AP|^2$. Ich prawe strony są równe, bowiem AP jest odcinkiem stycznej do okręgu (BCF) . Zatem $|AT| = |AS|$.





Rozwiązanie zadania F 688.

Na diagramie pT izochory to proste przechodzące przez początek układu współrzędnych, przy czym tangens kąta ich nachylenia jest proporcjonalny do objętości gazu. Rysując izochory na wykresie, widzimy, że maksymalna objętość gazu będzie w punkcie 2, a minimalna – w 1. Z równania Clapeyrona mamy:

$$\frac{2p_1 V_{\min}}{T_1} = \frac{p_1 V_{\max}}{3T_1},$$

zatem $V_{\max}/V_{\min} = 6$.



Rozwiązanie zadania M 1164.

Oznaczmy $a_j = n! + j$ (dla $j = 1, 2, \dots, n$) oraz przypuścimy, że istnieje taka liczba k , że każdy dzielnik pierwszy p liczby a_k jest dzielnikiem pewnej liczby a_l ($l \neq k$). Wtedy również $p \mid (a_k - a_l) = k - l$. Ponieważ $1 \leq |k - l| < n$, więc $p < n$ i $p \mid n!$, a skoro p jest dzielnikiem liczb a_k i a_l , to p musi dzielić obie liczby k i l .

Zatem każdy dzielnik pierwszy p liczby $a_k = n! + k$ jest dzielnikiem pewnej takiej liczby $l \neq k$, że $1 < l \leq n$.

Zapiszmy

$$a_k = k \cdot (1 \cdot \dots \cdot (k-1)(k+1) \cdot \dots \cdot n + 1)$$

oraz przyjmijmy, że p jest dzielnikiem pierwszym liczby

$$1 \cdot \dots \cdot (k-1)(k+1) \cdot \dots \cdot n + 1.$$

Ponieważ w iloczynie

$$1 \cdot \dots \cdot (k-1)(k+1) \cdot \dots \cdot n$$

występuje liczba l , więc iloczyn ten dzieli się przez p . Stąd p musi być dzielnikiem liczby 1, co jest niemożliwe.

Patrz w niebo

Jedną z najważniejszych spraw w astronomii jest umiejętność wyznaczania odległości kosmicznych obiektów. Znajomość bowiem odległości przekłada się na znajomość fizycznych cech tych obiektów – a o to przecież chodzi. Można powiedzieć, że jeżeli wyznaczanie odległości jest sprawą pierwszej wagi, to sprawą drugiej wagi jest **dokładne** wyznaczanie odległości. W przypadku gwiazd powszechną metodą jest porównanie jasności obserwowanej z tzw. absolutną. Formalnie jasność absolutna to taka, jaką miałyby gwiazda, gdyby można było ją zobaczyć z umownie przyjętej odległości 10 pc. Określa się ją na podstawie wyglądu jej widma, który zależy głównie od temperatury gwiazdy. Niestety, gwiazdy o tej samej temperaturze mogą silnie różnić się jasnością absolutną, gdyż mogą mieć różne rozmiary. Co prawda, widmo zawiera też informacje o rozmiarach gwiazdy, informacje te jednak są dużo trudniejsze do uzyskania, przez co nie zawsze można je wydobyć i nie tak dokładnie, jak by się chciało.

Tymczasem grupie amerykańskich astronomów (w której znalazł się wtedy polski astronom Janusz Kałużny) udało się kilka lat temu wyznaczyć odległość pewnej gwiazdy w najjaśniejszej gromadzie kulistej omega Centauri z wyjątkowo wysoką dokładnością. Była to gwiazda podwójna zaćmieniowa, a więc taka, której płaszczyzna obiegania się gwiazd prawie przechodzi przez Ziemię – to ważne. Pomierzono u niej prędkości radialne, dla każdego składnika z osobna (w km/s). Znając okres obiegu gwiazd (bo to przecież okres zmian jasności), można wyznaczyć wtedy wzajemną odległość gwiazd (w km). Z krzywej jasności można ocenić stosunek rozmiarów gwiazd do rozmiarów orbity, czyli również poznać rozmiary gwiazd (w km). Z widm znamy temperatury obu gwiazd, również więc moc emitowaną z jednostki powierzchni, a ponieważ rozmiary już znamy, to znamy pełne moce gwiazd, czyli ich jasności absolutne – i to z wysoką dokładnością. W ten sposób dzięki szczęśliwemu zbiegowi okoliczności odległość gromady omega Centauri oceniono na 5460 pc z błędem podobno zaledwie kilku procent. Dane uzyskane przy okazji dla gwiazd (masy: 0,76 i 0,81 mas Słońca, promienie: 0,9 i 1,9 promienia Słońca, jasności: 1 i 3,5 mocy słonecznej) umożliwiły również wyjątkowo dokładne określenie ich wieku (przez porównanie z modelami) na 11,8 mld lat (z błędem 0,6 mld lat). Jest to zarazem dolna granica wieku samej gromady omega Centauri.

Tomasz KWAST

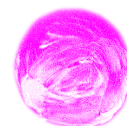
Marzec

W marcowe wieczory niebo jest jeszcze wygwieżdżone zimowymi okazałymi gwiazdozbiorami, spróbujmy jednak zajrzeć jak najdalej na południe. Rufa jest rozległym południowym gwiazdozbiorem, którego jedynie północną część widać z terenu Polski. Przed laty gwiazdozbiór ten stanowił część jeszcze większego gwiazdozbioru o nazwie Okręt Argo. Właśnie w tej północnej części Rufy znajdują się trzy gromady otwarte (M46, M47 – najjaśniejsza z nich, i NGC 2423) o jasnościach na granicy zasięgu nieuzbrojonego oka. Ponieważ znajdują się one najwyżej 25° nad horyzontem, to samym okiem raczej ich nie zobaczymy z powodu nieprzejrzystości powietrza, a zobaczenie ich (lub nie) przez lornetkę może być sprawdzianem stanu zanieczyszczenia atmosfery.

Wenus jest w Rybach i widać ją krótko po zachodzie Słońca. Mars jest w Wodniku, a więc zbyt blisko Słońca, by go zobaczyć. Jowisz jest w Wężowniku i wschodzi dopiero koło północy. Saturn jest na granicy Raka i Lwa, więc widać go praktycznie przez całą noc.

Merkury znajdzie się 22 III najdalej od Słońca i można go szukać wieczorem na zachodnim niebie. Pełnia Księżyca wypada 3 III i nastąpi wtedy jego całkowite zaćmienie widoczne również w Polsce w nocy 3/4 III. Nów Księżyca wypada 19 III i nastąpi wtedy częściowe zaćmienie Słońca, widoczne od Alaski do centralnej Azji. Oprócz Słońca Księżyc zakryje jeszcze Saturna (1/2 III, zakrycie widoczne w Polsce), Regulusa (2 III, widoczne w Azji i w Arktyce), Antaresa (11 III, widoczne na Antarktydzie), Merkurego (17 III, widoczne na oceanie na południe od Nowej Zelandii), ponownie Saturna (29 III, widoczne w Skandynawii i na Grenlandii) i ponownie Regulusa (30 III, widoczne w Zachodniej Europie, Skandynawii i w Arktyce). Razem będzie w marcu sześć zakryć i dwa zaćmienia – wyjątkowa obfitość. Przewidywalnych rojów meteorów nie będzie, za to na pewno przewiduje się nadejście wiosny, mianowicie około północy 20/21 III – odtąd dni będą już dłuższe od nocy.

T. K.



Markow rządzi (nawet buldogami pod dywanem)

Rafał SZTENCEL

ORRR RORR OROO RRRO ORRR
OROR RORO ROOR RROR RROR
OORR OOOO RROR ROOR ROOR
OOOR RORO OORR RROR ROOR

OORR RORO RRRO ROOR ORRR
ROOR RORR ORRO RROR RROR
OORR ROOR ROOR RROR OORR
OORR RRRO ORRR OORR ROOR

Czy można rozpoznać, który z widniejących na marginesie ciągów orłów i reszek powstał w wyniku rzutów monetą, a który został napisany przez człowieka?

Nie jest to trudne, bowiem ludzie mają na ogół mylne wyobrażenie o losowości.

Nie każdy wie, że w ciągu 100 rzutów monetą bardzo często pojawia się seria pięciu jednakowych wyników. Jeszcze skuteczniejszej metody dostarczają łańcuchy Markowa. Jeśli potraktować pojedyncze symbole jako stany procesu, to w przypadku ciągu losowego wszystkie prawdopodobieństwa przejścia p_{OO} , p_{OR} , p_{RO} i p_{RR} powinny być równe $\frac{1}{2}$. Tymczasem zbadanie 250 ciągów o długości 50, wygenerowanych przez respondentów autora [2] daje

$$p_{OR} \approx p_{RO} \approx 0,57, \quad p_{OO} \approx p_{RR} \approx 0,43.$$

Już tu widać wrzuszająco naiwne próby przywracania równowagi; a gdy za stany procesu przyjmiemy dwa sąsiednie symbole (np. dla ciągu OOROR kolejnymi stanami będą OO, OR, RO, OR), to szansa przejścia z OO do OO wyniesie tylko 36%, za to z RR do RO – 59%, a z OO do OR aż 64%.

Krótki spacer po Internecie i lektura znalezionej tam pracy [3] prowadzą do wniosku, że niewiele jest dziedzin działalności ludzkiej, gdzie nie stosuje się łańcuchów Markowa. Za pomocą tak zwanych ukrytych modeli Markowa można rozpoznawać znaki (IBM, 1967), twarze, zagrania w tenisie, a nawet przewidywać wydarzenia polityczne. My ograniczymy się do krótkiego przedstawienia modelu, który pozwala prognozować czas pomiędzy kolejnymi erupcjami gejzeru Old Faithful (drugiej, po misiu Yogi, atrakcji parku Yellowstone).

Azzalini i Bowman [1] stwierdzili, że zarówno czasy trwania erupcji, jak i odstępy pomiędzy nimi, dzielą się na „długie” i „krótkie” (odpowiednie histogramy mają dwa ostre i wyraźnie rozdzielone maksima). Co więcej, długość erupcji i czas oczekiwania na następną są silnie skorelowane.

Pierwszą składową modelu jest łańcuch Markowa o macierzy przejścia

$$\begin{bmatrix} 0,00 & 1,00 \\ 0,83 & 0,17 \end{bmatrix}$$

który opisuje wewnętrzny stan układu, bezpośrednio nieobserwowalny.

Zauważmy, że po stanie „krótkim” musi nastąpić „długi”. Z formalnego punktu widzenia łańcuch Markowa jest ciągiem zmiennych losowych X_0, X_1, X_2, \dots o wartościach w zbiorze $\{d, k\}$. Jeśli znamy wartość zmiennej losowej X_n , to zmienną Y_n (także o wartościach w zbiorze $\{d, k\}$) otrzymujemy za pomocą losowania, tak by

$$P(Y_n = d | X_n = k) = 0,23, \quad P(Y_n = d | X_n = d) = 1.$$

Drugą składową modelu stanowi zatem wektor warunkowych prawdopodobieństw wylosowania stanu d : $[0,23; 1]$. Zmienne losowe Y_n są obserwowane bezpośrednio i służą do estymacji parametrów modelu, podczas gdy zmienne X_n opisują tytułową walnę buldogów pod dywanem.

Prognoza zachowania się gejzeru jest niezwykle ważna z punktu widzenia turystów. Pod www.nps.gov/archive/yell/oldfaithfulcam.htm można obejrzeć bieżący stan gejzeru. Od 20 grudnia 2006 miała być dostępna prognoza czasu najbliższej erupcji.

Łańcuch Markowa można krótko opisać jako proces, w którym przyszłość zależy tylko od teraźniejszości, a nie od przeszłości. W języku zmiennych losowych o wartościach w zbiorze stanów $\{s_0, s_1, s_2, \dots\}$ wyraża się to tak: dla każdego $n = 1, 2, \dots$

$$P(X_n = s_n | X_{n-1} = s_{n-1}) = P(X_n = s_n | X_{n-1} = s_{n-1}, X_{n-2} = s_{n-2}, \dots, X_0 = s_0),$$

jeśli tylko prawdopodobieństwo warunkowe po prawej stronie jest dobrze określone. Trudno nie zadać pytania: czy ciąg (Y_n) jest łańcuchem Markowa?

Errata. Prof. Wojciech Guzicki zwrócił mi uwagę na przykry i poważny błąd w artykule *Kod Huffmana*, (*Delta* 1/2007). Zdanie „Pierwszy sekretarz KC PZPR, Edward Gierek,...” powinno brzmieć „Pierwszy Sekretarz KC PZPR, towarzysz Edward Gierek,...”.

Przepraszam. RS

Literatura

[1] A. Azzalini, A. W. Bowman, *A look at some data on the Old Faithful geyser*, Appl. Statist. 39 (1990), 357–365.

[2] P. Pacewicz, *Jak odróżnić ciąg losowy od nielosowego? Przykład zastosowania sieci neuronowej*, Praca magisterska, WMiM UW, Warszawa 2002.

[3] P. W. Zwiniak, *Ukryte modele Markowa w analizie danych dotyczących koniunktury gospodarczej*, Praca magisterska, SGH, Warszawa 2003.