

SPIS TREŚCI

NUMERU 8(55)

Zasada Fermata, czyli o przyczynowości i celowości w fizyce <i>Dr hab. Andrzej Szymacha</i>	str. 2
O metodzie najmniejszych kwadratów <i>Doc. dr Ryszard Zieliński</i>	str. 4
Zasady ekstremalne w termodynamice <i>Dr Bogdan Cichocki</i>	str. 6
Środki ciężkości, średnie, mediany <i>Dr Tadeusz B. Iwiński</i>	str. 8
Kącik filatelistyczny	str. 9
Nierówności elementarne i ekstrema <i>Dr Arkadiusz Płoski</i>	str. 10
Zadania	str. 12
Mała Delta	str. 13
Mechanika, komputer, człowiek <i>Prof. dr Dominik Rogula</i>	str. 16

W następnym numerze:

**Zegar chemiczny
i
Liczby p -adyczne**

Konkurs prac maturalnych z matematyki — rozstrzygnięty.

Final Konkursu organizowanego przez Polskie Towarzystwo Matematyczne oraz Redakcję „Delta” odbył się w Poznaniu 5 czerwca b.r. Autorzy prac zakwalifikowanych do finału na specjalnym posiedzeniu XVI Sesji Naukowej PTM wygłosili krótkie referaty, w których omówili najistotniejsze, ich zdaniem, informacje o pracach: genezę, sposób wykorzystania źródeł, własny wkład w opracowanie tematu oraz to, co uważali za najciekawsze lub najtrudniejsze. Po każdym referacie następowała dyskusja (na ogół dość ożywiona), w której brali udział uczestnicy Sesji Naukowej i Walnego Zgromadzenia PTM oraz goście. Szczególną uwagę w dyskusji zwracano na oryginalność wyników lub ujęcia.

Pośród sześciu prac finałowych za najlepszą uznana została praca „Liczby Fibonacciego”, napisana przez Pawła Domańskiego z Poznania (pod opieką prof. D. Wiśniewskiej) — otrzymał on złoty medal i I nagrodę MOiW. Srebrny medal przyznano Urszuli Łach z Wodzisławia Śl. za pracę „Równania różniczkowe i niektóre ich zastosowania w fizyce” (opiekun — prof. A. Hajok), a medal brązowy — Bogusławie Grzywacz z Wrocławia za pracę „Przekształcenia afiniczne płaszczyzny (opiekun — prof. C. Terlikowska). Obie medalistki otrzymały ponadto nagrody II stopnia. Irenie Kranc z Kopienicy przyznano wyróżnienie za pracę „Rachunek zdań i niektóre jego zastosowania w dowodzeniu”.

Medale, nagrody i dyplomy wręczone były — obok tzw. Wielkich Nagród PTM i Nagród dla młodych matematyków — w trakcie Walnego Zgromadzenia PTM w dniu 8 czerwca. Walne Zgromadzenie oceniło Konkurs jako udany i zaleciło organizowanie go w latach następnych. Skróit zwycięskiej pracy opublikujemy w pierwszym (stycziowym) numerze „Delta” w 1979 roku.

„Delta”
 matematyczno-fizyczny miesięcznik
 popularny
 Polskiego Towarzystwa
 Matematycznego i Polskiego
 Towarzystwa Fizycznego
 wydawany przy poparciu
 Ministerstwa Oświaty i Wychowania

Komitet Redakcyjny
 doc. dr J. Bartke
 doc. dr A. Bączyński
 doc. dr B. Gleichgewicht
 prof. dr K. Goebel
 doc. dr B. Iwazkiewicz
 doc. dr T. Iwiński
 doc. dr A. Januszajtis
 prof. dr Leon Jeśmanowicz —
 wiceprzewodniczący
 mgr H. Kaczorek
 prof. dr B. Karczewski
 prof. dr M. Kuczma
 mgr A. Mąkowski
 prof. dr Z. Pawlak
 prof. dr A. Piekara
 prof. dr Z. Semadeni
 prof. dr J. Stankowski

Wydano z pomocą finansową Polskiej Akademii Nauk

WARUNKI PRENUMERATY Cena prenumeraty rocznej zł 60, — cena prenumeraty półrocznej zł 30, —

Prenumeratę na kraj przyjmują Oddziały RSW „Prasa—Książka—Ruch” oraz urzędy pocztowe i doręczyciele — w terminach:
 — do 25 listopada na styczeń, I kwartał, I półrocze roku następnego i cały rok następnny
 — do dnia 10 miesiąca, poprzedzającego okres prenumeraty na pozostałe okresy roku bieżącego.
 Jednostki gospodarki społecznoniej instytucje i organizacje społeczno-polityczne składają zamówienie w miejscowych Oddziałach RSW „Prasa—Książka—Ruch”.
 Zakłady pracy i instytucje w miejscowościach, w których nie ma Oddziałów RSW, oraz prenumeratorzy indywidualni zamawiają prenumeratę w urzędach pocztowych lub u doręczycieli.
 Prenumeratę ze zleceniem wysyłki za granicę, która jest o 50% droższa od prenumeraty krajowej, przyjmuje RSW „Prasa—Książka—Ruch”. Centrala Kolportażu Prasy i Wydawnictw, ul. Towarowa 28, 00-958 Warszawa, konto PKO nr 1531-71 w terminach podanych dla prenumeraty krajowej

Sprzedaż numerów bieżących i uprzednich

Instytucje państwowe i społeczne, zakłady pracy, szkoły i czytelnicy indywidualni mogą nabywać „DELTE”:

w Księgarni Ośrodka Rozpowszechniania Wydawnictw Naukowych PAN.
 Sprzedaż gotówkowa i wysyłkowa, numerów bieżących i archiwalnych; płatność gotówką, przelewem lub za zaliczeniem pocztowym.
 Adres: ORPAN 00-901 Warszawa, Pałac Kultury i Nauki, Konto PKO I OM W-wa 1531-912

w Księgarni Ossolineum, Rynek 8, 50-106 Wrocław
 w Głównej Księgarni Naukowej, Krakowskie Przedmieście 7, 00-068 Warszawa
 w Księgarni Naukowej, ul. Podwale 6, 31-118 Kraków

Orders for this periodical from abroad can be placed with „Ars Polona” Krakowskie Przedmieście 00-068 Warszawa, Poland or with
 — Kubon & Sagner, Inhaber Otto Sagner, D8 Munchen 34, Postfach 68, Bundesrepublik Deutschland.
 — Earls Court Publications Ltd., 130 Shephard Bush Centre, London W 12, Great Britain,
 — Licosa Commissionaria Sansoni, Via Lamarmora 45, 50 121 Firenze, Italia

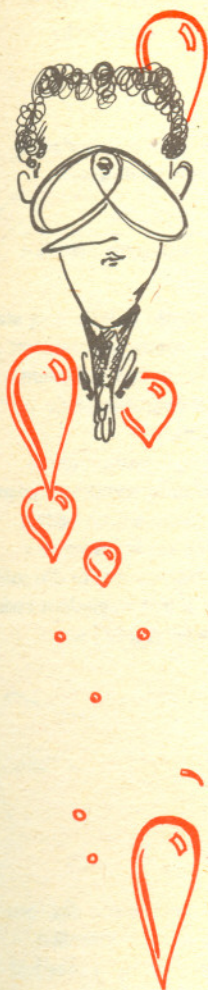
Cena 1 egzemplarza zł 5, — nr indeksu 35723/35550

prof. dr M. Subotowicz
 doc. dr S. Turnau
 doc. dr J. Wdowczyk
 prof. dr Janusz Zakrzewski —
 przewodniczący

Redaguje Kolegium w składzie:
 doc. dr T. Hofmokl — z-ca red. nac.
 dr T. B. Iwiński
 B. Jaworska-Kordos — ilustracje
 dr M. Kordos — red. nac.
 mgr K. Prażmowski — red. techn. graf.
 mgr K. Szypcio — sekr. red.
 doc. dr M. Święcki
 Adres Redakcji
 ul. Hoża 69 pok. 151,
 00-681 Warszawa

Zakład Narodowy im.
 Ossolińskich — Wydawnictwo
 Wrocław, Oddział w Warszawie
 Nakład 20 000 egz. Objętość 2 ark.
 wyd.; 2,50 ark. druk.;
 papier offsetowy III kl. 80 g. 61 × 86
 Wydrukowano w Drukarni im.
 Rewolucji Październikowej
 Warszawa ul. Mińska 65.
 Nr zam. 696/78 S-83

Co można zobaczyć w butelce z szamponem?



Odpowiedź zależy od tego, jaki to szampon i co chce się ujrzeć. Historia zaczęła się od kupienia szamponu cytrynowego firmy Palmolive za 60 zł. Płyn był żółty, lekko opalizujący i o dużym współczynniku lepkości (patrz dalej). Nie schowaliśmy go do łazienki, został na stole i wszyscy zaczęli się kolejno bezmyślnie nim bawić. W butelce było trochę powietrza. Tyle, że starczało na jeden porządny bąbel o średnicy około półtora centymetra. Przewracając butelkę można obserwować, jak bąbel dziarsko wędruje do góry. Najpierw zafascynowała nas strona wizualna zjawiska i stąd powstanie zdjęć, które reprodukuje na okładce. Zdjęcia robiliśmy w okresie, gdy opanowaliśmy trudną sztukę rozdzielania bąbla na mniejsze pęcherzyki. Obserwując zachowanie się całej rodziny pęcherzyków spostrzegliśmy, że

1. Kształt pęcherzyka zależy od jego rozmiarów. Małe przybierają kształt kulisty, większe są bardziej kropłowate z ostrym ogonem.

2. Prędkość wznoszenia się zależy od rozmiarów. Im większy pęcherzyk, tym prędzej się porusza.

3. Pęcherzyki poruszające się dostatecznie blisko siebie przyciągają się. Mały pęcherzyk, który w pojedynkę porusza się wolniej od dużego, potrafi dogonić duży, jeżeli znajdzie się przypadkiem w dostatecznie małej odległości. Po dogonieniu następuje połączenie pęcherzyków, ale zdarza się również, że pomimo połączenia widać wyraźnie granicę obu bąbli. Spróbujcie przeprowadzić opisane obserwacje. Jeżeli nie macie do dyspozycji odpowiedniego szamponu, możecie sporządzić roztwór wody z denaturatem, w którym będzie pływać kropla oliwy.

Zaobserwujecie podobne zjawiska. Jeżeli uwierzycie naszym obserwacjom lub przeprowadziliście własne, spróbujemy je wyjaśnić. Pęcherzyk jest lżejszy od cieczy, więc wypływa na powierzchnię. Jeżeli ma kształt kuli o promieniu r , to wypiera $\frac{4}{3}\pi r^3$ cieczy. Oznaczając gęstość pęcherzyka (w naszym przypadku powietrza) przez d_1 a gęstość cieczy (szamponu) przez d_2 , otrzymamy siłę wyporu F_w

$$F_w = \frac{4}{3}\pi \cdot r^3 \cdot (d_2 - d_1) \cdot g \quad (g \text{ — przyspieszenie ziemskie}).$$

Gdyby była to jedyna siła działająca na pęcherzyk, jego ruch byłby jednostajnie przyspieszony. Tak jednak nie jest, a więc coś przeciwstawia się jego ruchowi. To coś to opór lepkości cieczy. Siła oporu lepkości F_s zależy od prędkości ruchu v , od promienia pęcherzyka r oraz od liczby charakteryzującej ciecz, zwanej współczynnikiem lepkości k .

$$F_s = 6\pi k r v.$$

Jest to tak zwane prawo Stokesa, stosujące się do ruchów na tyle powolnych, aby nie powstawały w cieczy wiry. Pęcherzyk nabierze takiej prędkości, przy której $F_w = F_s$, siła wyporu i siła oporu zrównoważą się. Stąd łatwo obliczyć prędkość ruchu jednostajnego pęcherzyka

$$v = \frac{2r^2}{9k} (d_2 - d_1)g.$$

Prędkość ta zależy od promienia — im większy promień, tym szybciej porusza się pęcherzyk.

Możemy teraz przejść do wyjaśnienia różnicy w kształcie. Im większa prędkość pęcherzyka, tym większy napotyka on opór. Dlatego duże pęcherzyki unoszące się ku górze względnie szybko przybierają kształt kropli zapewniający im najmniejszy opór. Małe pęcherzyki prawie nieruchome odbiegają od kształtu kulistego prawie niezauważalnie. Stąd widzimy zawieszony w cieczy malutki pęcherzyk o kształcie kulistym.

Pozostaje jeszcze do wyjaśnienia zjawisko przyciągania się pęcherzyków. Moglibyśmy je wyjaśnić, rozważając rozkład ciśnień w cieczy za i obok poruszającego się ciała. Jakościowo możemy powiedzieć, że ciśnienie w pobliżu poruszającego się ciała zależy od rodzaju prądów cieczy i w pewnych obszarach, na przykład za ciałem jest mniejsze niż przed nim. Ta różnica ciśnień powoduje łączenie się przepływających w pobliżu pęcherzyków, a na morzu nawet może doprowadzić do kolizji mijających się statków.

Namawiamy gorąco, pobawcie się szamponem przed umyciem włosów — można w nim zobaczyć ciekawe rzeczy.

Zasada Fermata, czyli o przyczynowości i celowości w fizyce

Dr hab. Andrzej SZYMACHA

Zapewne większość Czytelników zna, przynajmniej powierzchownie, zasadę Fermata. Mówi ona, że światło spośród wszystkich możliwych torów łączących dwa ustalone punkty wybiera ten, którego przebieg wymaga najkrótszego czasu. Formalne znaczenie tej zasady, bez wnikaną w głębsze subtelności, polega na tym, że ujmuje ona bardzo elegancko wszystkie prawa optyki geometrycznej. Wykazanie tego jest dość proste. Przypomnijmy tylko, że dla wszystkich zastosowań optyki geometrycznej (np. konstrukcja obrazów w różnych przyrządach) wystarczy wiedzieć, że

- a) w ośrodku jednorodnym światło rozchodzi się po liniach prostych,
- b) w przypadku odbicia światła kąt padania jest równy kątowi odbicia,
- c) przekraczając powierzchnię oddzielającą dwa ośrodki, w których światło rozchodzi się z różnymi prędkościami promień światła załamuje się zgodnie z prawem Snelliusa:

$$\frac{\sin \alpha}{\sin \beta} = \frac{n_2}{n_1}, \quad \text{gdzie } n_i = \frac{c}{v_i} = \frac{\text{prędkość światła w próżni}}{\text{prędkość światła w danym ośrodku}}$$

Wykażemy teraz, że z zasady Fermata wynikają prawa a, b i c.

a) W ośrodku jednorodnym współczynnik załamania n , a zatem i prędkość światła v są stałe i czas przebiegu jest wprost proporcjonalny do odległości. Zasada najkrótszego czasu jest zatem równoważna w tym przypadku zasadzie najmniejszej odległości. Jest oczywistym faktem geometrycznym, że właśnie odcinek prostej łączący dane punkty A i B charakteryzuje się najmniejszą długością w porównaniu z innymi liniami łączącymi te punkty.

b) Odbicie. Rozpatrzmy długość toru promienia od A do B wzdłuż drogi, która choć w jednym punkcie dotyka powierzchni zwierciadła. Zgodnie z poprzednimi wynikami wiemy, że od punktu A do zwierciadła i od zwierciadła do punktu B światło porusza się będzie po prostej. Wyznaczenie toru promienia sprowadza się w tym wypadku do wyznaczenia punktu C' , dla którego łamana $AC'B$ będzie najkrótsza. W tym celu zauważmy, że długość łamanej $AC'B$ jest identyczna z długością łamanej $A'C'B$, jeśli punkt A' jest względem zwierciadła położony symetrycznie do punktu A . Łamana $A'C'B$ jest najkrótsza wtedy, gdy jest odcinkiem $A'B$. Jeżeli $A'CB$ jest odcinkiem, to

$$(1) \quad \sphericalangle A'CS' = \sphericalangle SCB.$$

Ponieważ $\triangle APC$ jest przystający do $\triangle A'PC$, to $\sphericalangle PCA = \sphericalangle PCA'$. Kąty dopełniające tych kątów są też równe.

$$(2) \quad \sphericalangle A'CS' = \sphericalangle ACS.$$

Porównując równości (1) i (2) widzimy, że $\sphericalangle SCB = \sphericalangle ACS$.

A to właśnie chcieliśmy wykazać.

c) Załamanie. I w tym wypadku najkrótszy czas przebiegu realizuje się na pewno dla toru będącego łamaną (po każdej stronie płaszczyzny załamującej ośrodek jest z założenia jednorodny). W celu wyznaczenia warunku na położenie punktu C , przy którym czas jest najkrótszy, obliczmy ten czas dla położenia C scharakteryzowanego odległością x , a następnie zminimalizujmy to wyrażenie ze względu na x .

$$T = \frac{AC}{v_1} + \frac{CB}{v_2} = \frac{\sqrt{h_A^2 + x^2}}{c/n_1} + \frac{\sqrt{h_B^2 + (l-x)^2}}{c/n_2}$$

W celu wyznaczenia minimum przyrównajmy do zera pochodną $\frac{dT}{dx}$.

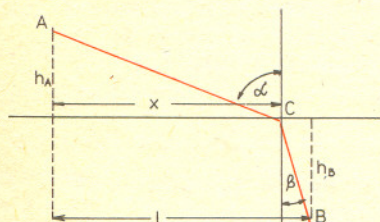
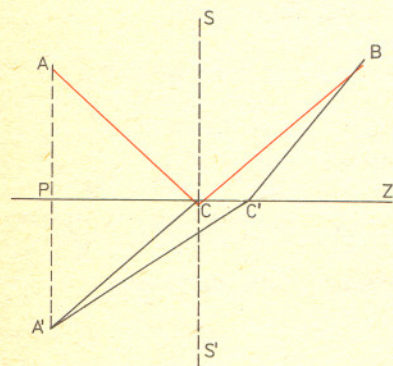
$$0 = \frac{dT}{dx} = \frac{1}{c} \left[n_1 \frac{d}{dx} \sqrt{h_A^2 + x^2} + n_2 \frac{d}{dx} \sqrt{h_B^2 + (l-x)^2} \right] = \frac{1}{c} \left[\frac{n_1 x}{\sqrt{h_A^2 + x^2}} - \frac{n_2 (l-x)}{\sqrt{(l-x)^2 + h_B^2}} \right] = \frac{1}{c} \left[n_1 \frac{x}{AC} - n_2 \frac{l-x}{CB} \right].$$

Ponadto

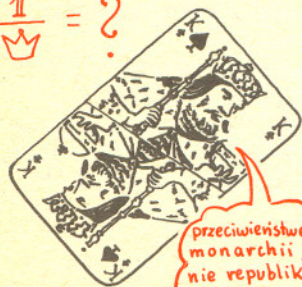
$$\frac{x}{AC} = \sin \alpha, \quad \frac{l-x}{CB} = \sin \beta. \quad \text{Zatem } n_1 \sin \alpha = n_2 \sin \beta \quad \text{i} \quad \frac{\sin \alpha}{\sin \beta} = \frac{n_2}{n_1} = n_{12}.$$

Widzimy więc, że istotnie zasada najkrótszego czasu prowadzi do prawa Snelliusa.

Odkrycie zasady Fermata wywołało w swoim czasie poważny spór filozoficzny: Czy Przyroda działa przyczynowo, czy celowo? Gdybyśmy znali tylko zasadę Fermata, bez wniosków z niej wynikających, to skłanialibyśmy się do drugiego poglądu. Przyroda wybiera dla promienia taki przebieg, by wypełnić pewien cel globalny, zużyć na przejście światła jak najmniej czasu. Gdyby od dziecka wychowywano nas w tym przekonaniu, to łatwo uznałibyśmy je w końcu za oczywiste i nie wymagające dalszych komentarzy. Z drugiej strony prawa optyki a, b, c mają charakter lokalny, przyczynowy. Prawo a mówi, że światło wysyłane z punktu A



$\frac{1}{\omega} = ?$



w określonym kierunku porusza się nieco do przodu i nie napotyka żadnej przeszkody kontynuując swój prostoliniowy bieg na podobieństwo swobodnego punktu materialnego. Powierzchnie odbijające bądź załamujące zmieniają kierunek tego biegu na bardzo małym obszarze (lokalnie) według określonych praw, prowadząc krok po kroku do takiego, a nie innego ostatecznego biegu promienia.

Dylemat powyższy nazwałem filozoficznym raczej niż fizycznym, gdyż, jak wskazują zanalizowane przykłady, oba sformułowania są matematycznie równoważne. Jednakże, gdy przejdziemy do sytuacji bardziej skomplikowanych, jedno podejście może okazać się bardziej wartościowe od drugiego. We współczesnej fizyce uważa się, że rozumowanie przyczynowe, lokalne, wnika istotnie w naturę rzeczy, podczas gdy efekt globalny jest tylko uboczną konsekwencją, którą zresztą warto lepiej zrozumieć.

Argumentów za tym, że podejście przyczynowe jest pierwotne, a efekt globalny (w postaci najkrótszego czasu) wtórny, można dostarczyć pozostając w ramach optyki geometrycznej, ale stają się one szczególnie dobitne jeśli uwzględnić, że optyka geometryczna jest tylko pewnym przypadkiem granicznym (idealizacją) optyki falowej.

W artykule tym ograniczymy się do argumentów pierwszego rodzaju. Polegają one na zauważeniu, że czas przebiegu światła — wbrew podanemu powyżej sformułowaniu — nie zawsze jest najkrótszy. Rozpatrzmy następujący przykład. Mamy dwie odbijające koncentryczne powierzchnie sferyczne i rozpatrujemy ruch światła między nimi. Przechodząc z odległością między sferami do zera uzyskamy graniczny przypadek promienia światła zmuszonego poruszać się (ze stałą prędkością) po powierzchni kuli. Nie ulega wątpliwości, że w tym granicznym przypadku kształty promieni będą pokrywały się z łukami kół wielkich na powierzchni kuli (Czytelnik powinien postarać się wykazać to formalnie). Jest to zgodne z zasadą Fermata, ale tylko dla łuków niezbyt dużych. Dla łuków większych od połowy obwodu koła wielkiego łuk AB wcale nie jest najkrótszą odległością między punktami A i B , ale na pewno jest to możliwy tor promienia. W otoczeniu tego łuku istnieją zarówno łuki dłuższe (1) jak też krótsze (2). Widać, że z celowością jesteśmy w tym przypadku na bakier. Prawdą natomiast jest, że po podzieleniu tego wielkiego łuku na mniejsze fragmenty uzyskujemy dla każdego z tych fragmentów prawdziwe minimum. Dowodzi to, że prawa lokalne są zawsze spełnione — natomiast minimum czasu przebiegu jest ich matematyczną konsekwencją jedynie dla torów dostatecznie krótkich.

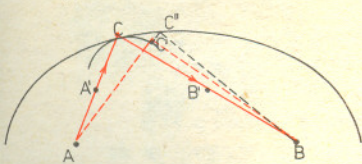
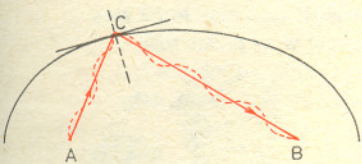
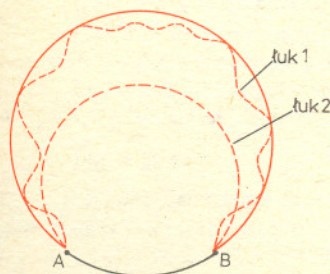
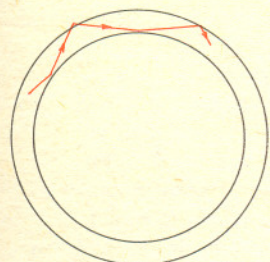
A oto inny, bardziej realistyczny przykład.

Rozpatrzmy elipsoidę obrotową. Matematyczną własnością elipsy jest, że $r_1 + r_2 = \text{const}$ (jeśli A i B są ogniskami elipsy). Prawdą jest również, że styczna do elipsy jest prostopadła do dwusiecznej kąta ACB . Zatem łamana ACB jest jednym z możliwych torów promienia światła. Zgodnie z pierwszą własnością w otoczeniu tego toru istnieją tory o identycznej długości (a zatem o identycznym czasie przebiegu). Istnieją również tory o dłuższym czasie przebiegu (linie krzywe, np. ta zaznaczona linią przerywaną). Znowu zasada minimum nie działa. Jest ona naruszona jeszcze wyraźniej, jeśli wprowadzić zwierciadło kuliste Z styczne do elipsoidy w punkcie C . Linia ACB jest znowu możliwym torem światła (spełnia prawo odbicia). Ale w otoczeniu łamanej ACB istnieją tory krótsze od niej. Takim torem jest łamana $AC'B$. Z rysunku widać, że

$$AC'B < AC'B = ACB.$$

I znowu, podobnie jak w przypadku z powierzchnią kuli, można dowieść, że dla punktów A' i B' dostatecznie bliskich łamana $A'CB'$ jest najkrótsza. A więc dla małych fragmentów toru zasada Fermata jest słuszna, a globalnie nie. Podważa to pogląd o jakiegokolwiek „celowości” i świadczy na rzecz przyczynowości.

Pozostaje do wyjaśnienia, dlaczego w rozpatrywanych na początku przykładach (a, b, c) zasada Fermata okazała się równoważna prawom optyki geometrycznej. Otóż we wszystkich tych trzech przykładach nie było w geometrii problemu niczego, co definiowałoby jakąś skalę odległości. Tym samym odcinki o dowolnych długościach są równoprawne, każdy może być zatem uważany za „dostatecznie krótki”. Ograniczając się do płaskich powierzchni odbijających i załamujących udowodniliśmy więc w istocie słuszność zasady Fermata dla lokalnych fragmentów toru. Kiedy przeszliśmy do kul i elipsoid, owa równoważność się załamała. Na zakończenie dodam, że choć zasada Fermata w najprostszej oryginalnej wersji okazała się za mało ogólna (i w sensie globalnym nieprawdziwa), to istnieje jej uogólnienie wolne od owych ograniczeń. W tym ogólniejszym sformułowaniu zasady Fermata należy słowo „najmniejszy” zastąpić słowem „stacjonarny”. Mówimy, że czas przebiegu dla jakiegoś toru jest stacjonarny, jeśli dla bliskich mu torów czas przebiegu różni się mało. Tor minimalny jest torem stacjonarnym, ale nie na odwrót. Owa stacjonarność odgrywa decydującą rolę przy analizowaniu zasady Fermata z punktu widzenia optyki falowej. Tor stacjonarny ma tę własność, że w jego otoczeniu istnieją tory, wzdłuż których propagujące się (zgodnie z zasadą Huyghensa) fale interferują konstruktywnie (mała zmiana czasu przelotu odpowiada małej zmianie fazy fali). Otoczenie torów stacjonarnych optyki geometrycznej odgrywa więc szczególną i wyróżnioną rolę w obrazie falowym światła. Zasada Fermata (uogólniona) jest tego naturalną konsekwencją. Ale podejście falowe jest lokalne, przyczynowe, a nie celowe.



$$\frac{1}{\text{crown}} = ?$$



Przez trzy kolejne dni mierzyłem odległość od mojego domu do najbliższego kiosku z gazetami i otrzymałem następujące wyniki (w krokach): 348, 352 i 342. Pytanie: jaka jest odległość od mojego domu do tego kiosku?

A oto mniej błahy problem o dokładnie takiej samej strukturze. Dla ustalenia ładunku elektronu dokonano pięciu niezależnych pomiarów i otrzymano wyniki (w absolutnych jednostkach elektrostacyjnych): $4,781 \cdot 10^{-10}$; $4,795 \cdot 10^{-10}$; $4,769 \cdot 10^{-10}$; $4,792 \cdot 10^{-10}$ i $4,779 \cdot 10^{-10}$. Pytanie: ile wynosi ładunek elektronu? Co w ogóle można na ten temat powiedzieć mając takie wyniki pomiarów?

Wróćmy do pomiarów, które każdy z nas może wykonać samodzielnie i rozpatrzmy trochę prostszą wersję naszego pierwszego zadania. Przypuśćmy, że zmierzylismy odległość od domu do najbliższego kina i otrzymaliśmy wynik 812 kroków. Co na tej podstawie możemy powiedzieć o „prawdziwej” odległości do kina, skoro wiemy, że drugi pomiar dałby już inny wynik? Oto pewne, może nieco zbyt egocentryczne, rozumowanie prowadzące do pewnej odpowiedzi. W wyniku pomiaru interesującej nas wielkości, powiedzmy a , otrzymaliśmy wynik, powiedzmy x . Popelniliśmy błąd $\varepsilon = x - a$, ale oczywiście nie znamy wielkości tego błędu. Chcieliśmy wykonać nasz pomiar jak najstaranniej, to znaczy chcieliśmy wykonać go tak, żeby błąd był jak najbardziej bliski zeru. Przypuśćmy, że udało się nam to osiągnąć. W konsekwencji tego założenia za a powinniśmy przyjąć taką liczbę, która przy ustalonym wyniku pomiaru x minimalizuje $|x - a|$. Taką liczbą jest oczywiście x , a więc w naszym przykładzie stwierdzamy, że odległość od domu do najbliższego kina wynosi 812 kroków. Zasadę, na której oparliśmy oszacowanie interesującej nas wielkości, mogliśmy z oczywistych powodów nazwać zasadą najmniejszego błędu. Sprawa nieco komplikuje się, gdy — jak w pierwszym naszym zadaniu — wykonujemy więcej niż jeden pomiar. Jeżeli przez a oznaczymy odległość od domu do kiosku oraz przez $\varepsilon_1, \varepsilon_2, \varepsilon_3$ — błędy jakie popelniliśmy w kolejnych pomiarach, to mamy

$$a + \varepsilon_1 = 348$$

$$a + \varepsilon_2 = 352$$

$$a + \varepsilon_3 = 342$$

Jak teraz oszacować wielkość a ? Postąpimy znowu zgodnie z zasadą najmniejszego błędu. Zauważmy, że teraz — w trzech pomiarach — popelniliśmy trzy błędy: $\varepsilon_1, \varepsilon_2, \varepsilon_3$, wygodniej jednak będzie mówić o błędzie $\varepsilon = (\varepsilon_1, \varepsilon_2, \varepsilon_3)$ i o jego „składowych” $\varepsilon_1, \varepsilon_2, \varepsilon_3$. Ta terminologia sugeruje, żeby w tym przypadku błąd reprezentować za pomocą punktu w przestrzeni trójwymiarowej i żeby za wielkość błędu przyjąć odległość tego punktu od zera (od początku układu współrzędnych). Jeżeli umówimy się, że odległość w tej przestrzeni trójwymiarowej będzie zwykłą odległością euklidesową, to zasada najmniejszego błędu nakazuje nam ustalić a tak, żeby $\varepsilon_1^2 + \varepsilon_2^2 + \varepsilon_3^2$ było możliwie najmniejsze, to znaczy tak, żeby zminimalizować wielkość:

$$(a - 348)^2 + (a - 352)^2 + (a - 342)^2.$$

Taka metoda szacowania nieznannej wielkości a nazywa się — znowu z oczywistych powodów — metodą najmniejszych kwadratów. Czytelnik zechce sprawdzić, że metoda ta daje wynik $a = (348 + 352 + 342)/3$.

Uogólnienie na n pomiarów jest proste: jeżeli x_1, x_2, \dots, x_n są wynikami pomiaru pewnej wielkości a , to za oszacowanie tej wielkości metodą najmniejszych kwadratów przyjmuje się po prostu średnią arytmetyczną $(x_1 + x_2 + \dots + x_n)/n$. Jest to zgodne z oczekiwaniem. A oto mniej banalne przykłady zastosowania metody najmniejszych kwadratów.

Przykład 1. Ze szkolnego kursu fizyki wszyscy znamy prawo Boyle’a-Mariotte’a: iloczyn objętości i ciśnienia gazu ma w tej samej temperaturze wartość stałą. Zapisuje się to za pomocą znanego wzoru: $p \cdot v = c$, gdzie p jest ciśnieniem, v — objętością oraz c — pewną stałą. Stała c zależy od różnych czynników, na przykład od temperatury. Przypuśćmy, że zadanie polega na oszacowaniu tej stałej dla pewnego konkretnego naczynia i pewnego konkretnego gazu. Przypuśćmy dalej, że możemy zmieniać objętość naszego naczynia (na przykład za pomocą manipulowania tłokiem) i że możemy mierzyć ciśnienie, przy czym pomiar ciśnienia obarczony jest pewnym błędem przypadkowym. Jeżeli przy objętości v_1 zarejestrujemy

ciśnienie p_1 , to błąd wynosi $p_1 - \frac{c}{v_1}$. Jeżeli powtórzymy nasz pomiar n razy i przy

objętościach v_1, v_2, \dots, v_n otrzymamy wyniki p_1, p_2, \dots, p_n , to zgodnie z metodą najmniejszych kwadratów za oszacowanie stałej c należy przyjąć taką wartość, która minimalizuje wielkość

$$\left(p_1 - \frac{c}{v_1}\right)^2 + \left(p_2 - \frac{c}{v_2}\right)^2 + \dots + \left(p_n - \frac{c}{v_n}\right)^2$$

Czytelnik mógłby zauważyć, że można wykonać tylko jeden pomiar i za c przyjąć po prostu iloczyn $p_1 v_1$. Albo że można wykonać kilka razy pomiar ciśnienia przy tej samej objętości v i że wtedy należałoby minimalizować prostszą wielkość

$$\left(p_1 - \frac{c}{v}\right)^2 + \left(p_2 - \frac{c}{v}\right)^2 + \dots + \left(p_n - \frac{c}{v}\right)^2,$$

$$\frac{1}{\text{?}} = ?$$



?



Rozwiązanie zadania F 55.

W stanie równowagi energia swobodna musi mieć ekstremum. W rozważanym problemie energia swobodna jest energią powierzchniową kryształu i równa się

$$E = 2a^2\sigma_2 + 4ac\sigma_1.$$

Zgodnie z treścią zadania składnik energii swobodnej równy energii potencjalnej został pominięty.

Aby wyznaczyć stosunek a/c należy znaleźć ekstremum E przy założeniu dodatkowym, że objętość kryształu $V = a^2c$ jest stała. Innymi słowy, należy znaleźć ekstremum funkcji

$$E(a) = 2a^2\sigma_2 + 4V\sigma_1/a,$$

gdzie $V = a^2c = \text{const}$.

Obliczając pochodną i przyrównując ją do zera otrzymujemy

$$4a\sigma_2 - 4V\sigma_1/a^2 = 0$$

$$a\sigma_2 - \sigma_1 = 0,$$

a zatem

$$\frac{a}{c} = \frac{\sigma_1}{\sigma_2}.$$

a to zadanie już rozwiązaliśmy i wiemy, że jego rozwiązaniem jest średnia arytmetyczna: $c = v(p_1 + p_2 + \dots + p_n)/n$. Czytelnik ma rację, ale okazuje się, że dokładność oszacowania c jest tym większa, im więcej wykonuje się pomiarów i im zreczniej dobiera się wartości v_1, v_2, \dots, v_n . Wyjaśnienie, dlaczego tak jest, wykracza jednak daleko poza ramy naszej krótkiej rozprawki.

Przykład 2. Na pewno każdy z Czytelników usłyszał kiedyś od lekarza, że „jak na swój wiek” jest zbyt wysoki, zbyt niski lub że ma wzrost „w normie”. Na pewno jednak nie każdy starał się dociec, co to znaczy. Spróbuję to wyjaśnić. Będzie to zarazem ilustracją pewnego dość powszechnego, ale odmiennego od poprzednich, sposobu zastosowania metody najmniejszych kwadratów. Postępuje się mianowicie w następujący sposób. „Bierze się” dużą grupę N ludzi w różnym wieku i mierzy się wzrost każdej osoby w tej grupie. Otrzymuje się wyniki: $(t_1, w_1), (t_2, w_2), \dots, (t_N, w_N)$, gdzie t_i jest wiekiem oraz w_i — wzrostem i -tej osoby. Po naniesieniu na wykres wyglądają one mniej więcej tak, jak na rys. 1 (każda kropka na tym rysunku reprezentuje jedną osobę.)

Następnie rysuje się krzywą, która ma przebiegać „mniej więcej” tak, jak układają się nasze punkty (linia ciągła na rysunku). Z kolei „wymyśla się” wzór, który byłby równaniem tej krzywej. Dla krzywych wzrostu często przyjmuje się wzór

$$(*) \quad w = \frac{a}{1 + b \cdot c^t}$$

gdzie a, b, c są pewnymi stałymi. Jest to tzw. *krzywa logistyczna*. Zgódźmy się na to równanie; od tej pory staje się ono dla nas „empirycznym prawem wzrostu”. Różnym wartościom stałych a, b, c (podobnie jak różnym wartościom stałej c w prawie Boyle’a-Mariotte’a) odpowiadają różne krzywe (rys. 2) i zadanie polega na znalezieniu takich a, b, c , które najlepiej odpowiadają wynikom naszych pomiarów. Zgodnie z naszym prawem wzrostu, wzrost osoby w wieku t_i powinien wynosić $a/(1 + b \cdot c^t)$, natomiast w wyniku pomiaru otrzymaliśmy dla tej osoby wynik w_i . Wielkość $w_i - \frac{a}{1 + b \cdot c^t}$ jest odpowiednikiem tego, co w poprzednich przykładach nazywaliśmy błędem ε_i . Zatem — zgodnie z metodą najmniejszych kwadratów — stałe a, b, c należy dobrać tak, żeby wielkość

$$\sum_{i=1}^N \left(w_i - \frac{a}{1 + b \cdot c^t} \right)^2$$

osiągała minimum. To zadanie jest dość skomplikowane rachunkowo i praktycznie rozwiązuje się je za pomocą komputerów (nawet dla małych wartości N). Wzór (*) z ustalonymi wartościami a, b, c opisuje właśnie to, co nazywa się „normalnym wzrostem w danym wieku”. Takie krzywe wzrostu można oczywiście konstruować oddzielnie dla mężczyzn i dla kobiet, czasami oddzielnie dla różnych środowisk (np. miasto — wieś), a na przykład powiedzenie, że „Szwedzi są średnio wyżsi od Polaków” oznacza, że krzywa wzrostu Szwedów leży powyżej polskiej krzywej wzrostu.

Metoda najmniejszych kwadratów jest codziennym narzędziem fizyków, inżynierów, przyrodników, ekonomistów, a jej teoria jest przedmiotem zainteresowania statystyki matematycznej.

Przykład (konkretny). Zakładamy, że wielkości X i Y łączy zależność liniowa $Y = aX + b$. W czterech pomiarach tych wielkości uzyskano wyniki

X	0	10	20	30
Y	14	15	37	41

Wyznaczyć metodą najmniejszych kwadratów przybliżenia a_0 i b_0 współczynników a i b .

Rozwiązanie. Załóżmy, ogólniej, że dysponujemy wynikami n pomiarów $(x_1, y_1), \dots, (x_n, y_n)$. Należy wyznaczyć liczby a_0 i b_0 tak, by wyrażenie

$$W(a, b) = \sum_{i=1}^n (y_i - ax_i - b)^2$$

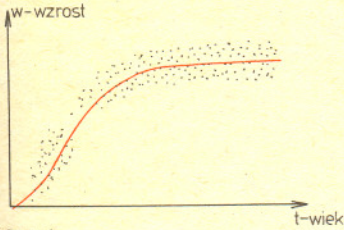
osiągnęło wartość najmniejszą dla $a = a_0$ i $b = b_0$. Ponieważ funkcja ta ma tylko jedno minimum, to można je wyznaczyć elementarnie w sposób następujący. Funkcja $W(a_0, b)$ jest trójmianem kwadratowym względem b :

$$W(a_0, b) = nb^2 - \left[2 \sum_i (y_i - a_0 x_i) \right] b + \sum_i (y_i - a_0 x_i)^2$$

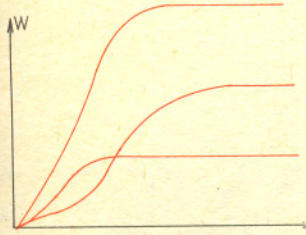
i minimum osiąga w punkcie $b = \frac{1}{2} \frac{2 \sum_i (y_i - a_0 x_i)}{n}$, który musi być identyczny z b_0 , skąd

$$(*) \quad b_0 = \bar{y} - a_0 \bar{x}$$

gdzie \bar{x}, \bar{y} są odpowiednio średnimi arytmetycznymi liczb x_1, \dots, x_n i y_1, \dots, y_n .



Rys. 1



Rys. 2



Podobnie, minimum trójmianu kwadratowego względem a

$$W(a, b_0) = \left[\sum_i x_i^2 \right] a^2 - \left[2 \sum_i (y_i - b_0) x_i \right] a + \sum_i (y_i - b_0)^2$$

osiągane jest w punkcie
$$a_0 = \frac{1}{2} \frac{2 \sum_i (y_i - b_0) x_i}{\sum_i x_i^2},$$
 skąd

$$(**) \quad a_0 \sum_i x_i^2 = \sum_i x_i y_i - b_0 \sum_i x_i.$$

Podstawiając (*) do (**) otrzymamy

$$a_0 \left[\sum_i x_i^2 - \bar{x} \sum_i x_i \right] = \sum_i x_i y_i - \bar{y} \sum_i x_i,$$

skąd po przekształceniach

$$(***) \quad a_0 = \frac{\sum_i (x_i - \bar{x})(y_i - \bar{y})}{\sum_i (x_i - \bar{x})^2}.$$

Wzory (***) i (*) dają szukane rozwiązanie. W naszym konkretnym przykładzie otrzymamy

$$a_0 = 1,03, \quad b_0 = 11,3.$$

Dla znających rachunek prawdopodobieństwa: wystarczy skorzystać ze wzorów $E[(X-EX)^2] = E(X^2) - [E(X)]^2$ i $E[(X-EX)(Y-EY)] = E(XY) - EY \cdot EX$.

Zasady ekstremalne w termodynamice

Dr Bogdan CICHOCKI

Zanim przejdziemy do omówienia zagadnienia wspomnianego w tytule niniejszego artykułu, przypomnimy podstawowe pojęcia termodynamiki. Wiemy na podstawie doświadczenia, że każde ciało makroskopowe, niezależnie od stanu początkowego, osiąga po pewnym czasie stan równowagi. Stan taki można scharakteryzować przez podanie niewielkiej liczby parametrów. Przykładowo, stan równowagi określonej liczby moli gazu jednoskładnikowego opisują w pełni dwie wielkości (np. ciśnienie p i objętość V). Na skutek działania czynników zewnętrznych układ fizyczny może przechodzić z jednego stanu równowagi do innego. Przejściem takim zwanym *procesem termodynamicznym* rządzą pewne prawa. Zgodnie z treścią pierwszej zasady termodynamiki w procesie termodynamicznym zmiana energii wewnętrznej U układu (tzn. różnica między wartościami U w stanach końcowym i początkowym) jest równa sumie ciepła Q dostarczonego układowi i pracy L wykonanej nad układem, czyli

$$(1) \quad U_k - U_p = Q + L.$$

Jeżeli ciało fizyczne składa się z dwóch części, powiedzmy 1 i 2, to jego energia wewnętrzna U jest równa

$$(2) \quad U = U_1 + U_2 + U_{12},$$

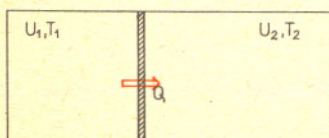
gdzie U_1, U_2 są energiami wewnętrznymi poszczególnych części, zaś U_{12} energią oddziaływania tych części. Ponieważ dla ciał makroskopowych $U_1 \gg U_{12}$ i $U_2 \gg U_{12}$, więc w przybliżeniu zachodzi związek

$$(3) \quad U = U_1 + U_2.$$

Mówimy, że wielkość U ma własność addytywności.

Pierwsza zasada termodynamiki nie charakteryzuje w pełni procesów występujących w przyrodzie. Aby się o tym przekonać, rozważmy następujący przykład. Wyobraźmy sobie układ fizyczny składający się z dwóch podukładów 1 i 2, o energiach wewnętrznych i temperaturach odpowiednio U_1, T_1 i U_2, T_2 . Niech między tymi podukładami znajduje się ścianka przewodząca ciepło. Z doświadczenia wiemy, że układ taki znajduje się w stanie równowagi, o ile $T_1 = T_2$. Jeżeli tak nie jest, to z układu o wyższej temperaturze przepływa ciepło do układu o niższej temperaturze, aż do momentu wyrównania się temperatur. Zgodnie z pierwszą zasadą każdy proces, dla którego zachodzi równość (1), jest możliwy. Również taki, w którym ciepło przepływa w kierunku przeciwnym niż ten, o którym wspomnieliśmy powyżej. Wynika z tego, że pierwsza zasada wymaga uzupełnienia. Zadanie to spełnia druga zasada termodynamiki.

Istnieje wiele równoważnych sformułowań drugiej zasady termodynamiki. Właściwie to już podaliśmy jedno z nich. Mianowicie, jest to twierdzenie dotyczące kierunku przepływu ciepła. Sformułowanie to podane po raz pierwszy przez Clausiusa jest bardzo poglądowe, jednakże droga prowadząca od niego do wniosków o charakterze ogólnym jest dość zawiła. Omówimy teraz inne sformułowanie drugiej zasady termodynamiki, które przybiera postać zasady



$$T_1 > T_2$$

$$\frac{1}{\text{crown}} = ?$$

ekstremalnej. Brzmi ono następująco: „Istnieje addytywna funkcja stanu S zwana entropią określona przez związek

$$\left(\frac{\partial S}{\partial U}\right)_v = \frac{1}{T} \quad \text{i} \quad \left(\frac{\partial S}{\partial V}\right)_u = \frac{p}{T}$$

która osiąga wartość maksymalną w stanie równowagi”.

Symbol $\left(\frac{\partial S}{\partial U}\right)_v$ oznacza pochodną funkcji S (dwa zmiennych) względem U .

Wyrażenie „osiąga wartość maksymalną” wymaga wyjaśnienia.

Rozważmy izolowany układ fizyczny składający się z wielu podukładów, z których każdy znajduje się w stanie równowagi. W ogólności stan całego układu nie jest stanem równowagi, o czym mogliśmy się przekonać przy rozpatrywaniu przykładu dotyczącego przepływu ciepła. Mimo to dla takiego stanu możemy określić każdą z rozpatrywanych wielkości, jeśli ma ona własność addytywności. Nie jesteśmy w stanie określić temperatury, gdyż może być ona różna w różnych podukładach, ale możemy określić np. energię wewnętrzną. Dotyczy to także entropii zgodnie z jej postulowaną addytywnością. Możemy zatem porównywać wartość entropii w stanie, w którym tylko podukłady są w stanach równowagi, z jej wartością, gdy cały układ jest w równowadze. Wyrażenie „osiąga wartość maksymalną” odnosi się do tego typu porównania.

Wykażemy teraz, że sformułowana przez nas zasada jest zgodna z doświadczeniem.

Rozpatrzymy ponownie układ przedstawiony na rysunku. Przyjmijmy, że przegroda jest nieruchoma. Zatem objętości podukładów 1 i 2 są ustalone. Ponieważ cały układ jest izolowany,

$$U_1 + U_2 = U = \text{const.}$$

Możemy rozpatrywać różne podziały energii U pomiędzy układami, ale tylko jeden odpowiada stanowi równowagi. Podziały te można opisać w pełni podając np. tylko wartość U_1 .

Zgodnie z własnością addytywności entropia S całego układu dla danej wartości U_1 jest równa

$$S(U_1) = S_1(U_1) + S_2(U - U_1),$$

gdzie $S_i(x)$ oznacza entropię i -tego podukładu o energii wewnętrznej x . Jeżeli w stanie równowagi osiąga ona wartość maksymalną, to (traktujemy S jako funkcję jednej zmiennej U):

$$0 = \frac{dS}{dU_1} = \frac{dS_1}{dU_1} + \frac{dS_2}{dU_1} = \frac{dS_1}{dU_1} - \frac{dS_2}{dU_2}, \quad \text{czyli}$$

$$(4) \quad \frac{dS_1}{dU_1} = \frac{dS_2}{dU_2}.$$

Przypomnijmy, że równowaga zostanie osiągnięta, gdy $T_1 = T_2$. Widzimy zatem, że pochodna entropii S względem energii wewnętrznej U (przy ustalonej objętości) ma własność analogiczną do temperatury. Przyjmijmy

$$(5) \quad \frac{\partial S}{\partial U} = \frac{1}{T}$$

(dlaczego $1/T$, a nie T wyjaśni się za chwilę). Warunek (4) jest warunkiem na ekstremum entropii. Sprawdźmy, że przy przyjęciu związku (5) jest to maksimum. W tym celu wyliczmy zmianę entropii ΔS przy nieznacznym odchyleniu od stanu równowagi. Możemy napisać

$$\Delta S = \Delta S_1 + \Delta S_2 = \frac{dS_1}{dU_1} \Delta U_1 + \frac{dS_2}{dU_2} \Delta U_2 = \frac{1}{T_1} \Delta U_1 + \frac{1}{T_2} \Delta U_2 = \left(\frac{1}{T_1} - \frac{1}{T_2}\right) \Delta U_1.$$

gdź $\Delta U_1 = -\Delta U_2$.

Jeżeli $\Delta U_1 > 0$, to zgodnie z doświadczeniem $T_1 > T_2$ i $\Delta S < 0$, jeżeli zaś $\Delta U_1 < 0$, to $T_1 < T_2$ i również $\Delta S < 0$. Widzimy zatem, że rzeczywiście entropia osiąga maksimum w stanie równowagi (gdybyśmy przyjęli, że $\frac{\partial S}{\partial U} = T$ otrzymalibyśmy minimum). Zwróćmy

jeszcze uwagę, że zamiast T mogliśmy w powyższych wzorach podstawić dowolną rosnącą funkcję T . Zmieniłoby się wtedy jedynie funkcyjną zależność entropii od temperatury. Sama entropia nie jest bowiem określona jednoznacznie przez własność addytywności oraz

maksymalności. Bardziej istotny jest natomiast fakt, że np. $\frac{\partial S}{\partial U}$ zależy jedynie od temperatury.

Przeprowadzenie analogicznego rozumowania w przypadku układu składającego się z dwóch podukładów przedzielonych ruchomym tłokiem pozostawiamy Czytelnikowi. Okazuje się wtedy,

że pochodna $\frac{\partial S}{\partial V}$ ma własności analogiczne do ciśnienia i przyjmujemy, że jest równa $\frac{p}{T}$.

Przytoczone sformułowanie II zasady termodynamiki jest bardzo wygodnym punktem startowym do wyprowadzenia szeregu wniosków natury ogólnej. Warto również wspomnieć, że odgrywa ono ważną rolę przy wytłumaczeniu praw termodynamiki w ramach fizyki statystycznej.

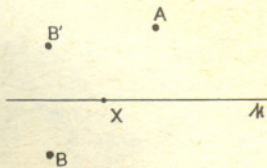


elekcja?

Rozwiązanie zadania M 167.

Niech B' będzie obrazem punktu B w symetrii względem prostej k , X zaś dowolnym jej punktem. Mamy $|AX - BX| = |AX - B'X| \leq AB'$ i równość zachodzi tu tylko w przypadku, gdy punkty X , B' i A są współliniowe. Tak więc szukanym punktem jest punkt X_0 przecięcia prostych AB' i k .

Jeżeli prosta AB' jest równoległa do k , to zadanie nie ma rozwiązania, jeżeli zaś punkty A i B są symetryczne względem k (tzn. $B' = A$), to k jest symetralną odcinka AB i wobec tego dla każdego jej punktu X zachodzi równość $AX - BX = 0$.



Pojęcie środka masy (tradycyjnie zwanego środkiem ciężkości) jest tak intuicyjne i jednoznaczne, że dość trudno wyobrazić sobie celowość wprowadzenia tu jakichkolwiek komplikacji.

Postaramy się właśnie to zrobić.

Rozważmy na początek

Dwa zadania

1. Na osi liczbowej dane są punkty x_1, \dots, x_n . W każdym z nich umieszczono masę jednostkową $m = 1$. Znaleźć środek ciężkości x_0 .

2. Na osi liczbowej dany jest (skończony) zbiór punktów $X = \{x_1, \dots, x_n\}$. Wyznaczyć na osi punkt x'_0 tak, by suma kwadratów odległości x'_0 od punktów zbioru X była możliwie najmniejsza.

Zadania te są równoważne: $x_0 = x'_0$. Rzeczywiście, rozwiązanie zadania 1 polega na wyznaczeniu punktu x_0 takiego, by suma momentów względem x_0 mas umieszczonych w punktach x_i była równa zero, tj. by zachodziła równość

$$\sum_{i=1}^n m(x_i - x_0) = 0, \quad \text{czyli}$$

$$(*) \quad \sum_{i=1}^n (x_i - x_0) = 0.$$

Zadanie drugie sprowadza się do znalezienia minimum trójmianu kwadratowego

$$s(x) = \sum_{i=1}^n (x_i - x)^2 \quad (x_i - \text{liczby, } x - \text{zmienna!}).$$

Trójmian ten ma jedno minimum — jest to punkt, w którym jego pochodna względem x

$$s'(x) = 2 \sum_{i=1}^n (x_i - x)$$

przyjmuje wartość zero, a więc punkt x'_0 spełniający warunek

$$(**) \quad \sum_{i=1}^n (x_i - x'_0) = 0.$$

Warunki (*) i (**) są identyczne, co dowodzi równoważności zadań 1 i 2. Przy tym

$$\text{oczywiście} \quad x_0 = x'_0 = \frac{1}{n} \sum_{i=1}^n x_i,$$

a więc rozwiązaniem każdego z tych zadań jest średnia arytmetyczna liczb x_1, \dots, x_n . Okazało się więc, że w rozważanym przypadku pojęcie środka ciężkości pokrywa się z pojęciem średniej (arytmetycznej), przy czym każde z tych pojęć można zdefiniować jako minimum pewnej funkcji.

W przestrzeni kartezjańskiej wielowymiarowej

sytuacja jest analogiczna. Można tu sformułować odpowiedniki zadań 1 i 2. Nietrudno udowodnić, że zadania te są równoważne (warto to zrobić). Wspólne rozwiązanie tych zadań można — przez analogię z przypadkiem 1-wymiarowym — nazwać średnią punktów x_1, \dots, x_n . Średnia jest więc z definicji tym samym, co środek ciężkości i jest rozwiązaniem następującego problemu:

Dany jest zbiór $X = \{x_1, \dots, x_n\}$. Wyznaczyć punkt x_0 taki, by suma kwadratów odległości x_0 od punktów zbioru X była możliwie najmniejsza.

W statystyce występuje inne jeszcze ważne pojęcie, będące — intuicyjnie — swego rodzaju „środkiem ciężkości”. Jest to

Mediana.

Nie będziemy tu podawać ogólnej definicji tego pojęcia. W sytuacji z zadań 1 i 2 medianą będzie każdy taki spośród punktów x_i , że ilość punktów leżących na prawo od niego możliwie mało różni się od ilości punktów leżących na lewo od niego. Wyznaczenie median jest w tym przypadku proste: jeśli n jest nieparzyste, to medianą jest jeden („środkowy”) punkt x_i , jeśli zaś n jest parzyste, to medianami są dwa „środkowe” punkty.

Zastanówmy się teraz nad możliwością uogólnienia pojęcia mediany na przypadek wielowymiarowy. Bezpośrednie przeniesienie definicji nie jest tu oczywiście możliwe, bowiem już na płaszczyźnie pojęcia „prawo” i „lewo” tracą sens. Można jednak spróbować drogi analogicznej, jak w przypadku uogólnienia pojęcia średniej: sformułować problem znalezienia mediany jako problem wyznaczenia minimum pewnej funkcji w taki sposób, by to drugie sformułowanie przenosiło się już na przypadek wielowymiarowy.

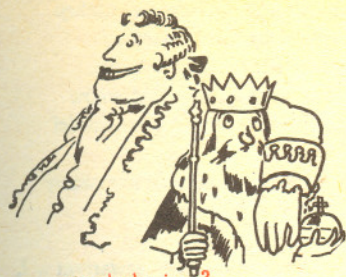
Porównaj artykuł R. Zielińskiego o metodzie najmniejszych kwadratów.

Zadanie można oczywiście rozwiązać elementarnie: współczynnik przy x^2 wynosi n , współczynnik przy x wynosi $2 \sum_{i=1}^n x_i$, a więc minimum osiągnięte jest

$$\text{w punkcie } x'_0 = (2 \sum_{i=1}^n x_i) / 2n = \frac{1}{n} \sum_{i=1}^n x_i.$$

W przestrzeni l -wymiarowej przyjmujemy $x_i = (x_i^1, \dots, x_i^l)$. W zadaniu 1 otrzymujemy warunek formalnie identyczny z (*), będący układem l równań określających niezależne l współrzędnych punktu x_0 , a w zadaniu 2 otrzymujemy sumę l trójmianów kwadratowych postaci $\sum (x_i^j - x^j)^2$, które minimalizujemy niezależnie.

$$\frac{1}{\text{crown}} = ?$$



protekcja?

Por. artykuły M. Moszyńskiej w Delcie: 1/1975, 5/1975 i 8/1975.

Okazuje się, że jest to możliwe. Rozważmy bowiem na osi liczbowej ciąg różnych punktów x_1, x_2, \dots, x_n oraz dowolny punkt x i utwórzmy wyrażenie

$$d(x) = \sum_{i=1}^n |x_i - x|.$$

Jak łatwo zauważyć, funkcja $d(x)$ jest ciągła oraz jest malejąca, gdy x jest mniejszy od większości x_i (bo wtedy współczynnik przy x jest ujemny), i rosnąca, gdy x jest większy od większości x_i . A to właśnie oznacza, że minimum funkcji $d(x)$ osiągnięte jest w tych punktach x_i , które są medianami. Jeśli jeszcze weźmiemy pod uwagę fakt, że $|x_i - x|$ jest odległością punktów x oraz x_i , to stwierdzimy, że mediana jest rozwiązaniem następującego problemu:

Dany jest zbiór $X = \{x_1, \dots, x_n\}$. Wyznaczyć punkt x_i taki, by suma odległości punktu x_i od punktów zbioru X była możliwie najmniejsza.

A to sformułowanie przenosi się już na dowolną przestrzeń wielowymiarową, może więc służyć jako definicja mediany w takiej przestrzeni.

Ale na tym możliwości uogólnień wcale się nie kończą. Zauważmy, że zarówno w wielowymiarowym problemie średniej, jak i wielowymiarowym problemie mediany, operuje się pojęciami zbioru, punktu i odległości. Łatwo więc widać, że oba problemy można sformułować w dowolnej

Przestrzeni metrycznej

i — o ile problemy te mają rozwiązania — otrzymywać bardzo już odbiegające od intuicji „środki ciężkości”: średnie i mediany.

Można jednak posunąć się jeszcze dalej. Zauważmy mianowicie, że zarówno występująca w problemie średniej suma kwadratów odległości, jak i występująca w problemie mediany suma odległości są pewnego rodzaju miarami rozproszenia zbioru X wokół punktu x . Są to funkcje bardzo specjalnej postaci. Można by pokusić się o określenie innych jeszcze miar rozproszenia i poszukiwanie ich minimów. Minima takie byłyby też swojego rodzaju „środkami ciężkości”. Jak więc widać, można drogą kolejnych uogólnień sformułować następujący

Problem środka ciężkości:

W przestrzeni metrycznej (A, ρ) dany jest skończony zbiór X oraz miara rozproszenia $r_X(x)$, $x \in A$. Wyznaczyć minimum funkcji r_X .

Rozwiązanie takiego problemu — o ile istnieje — nazwiemy środkiem ciężkości zbioru X w przestrzeni (A, ρ) względem miary rozproszenia r_X .

Tak rozumiane pojęcie środka ciężkości jest nie tylko bezpośrednim uogólnieniem pojęć, od których startowaliśmy. Powstaje ono również w naturalny sposób w zastosowaniach matematyki w naukach społecznych. Ale o tym kiedy indziej.

Problem: czy można podać na prostej taką metrykę i taką miarę rozproszenia, by rozwiązaniem problemu środka ciężkości była (a) średnia geometryczna, (b) średnia harmoniczna? (Autor nie zna rozwiązania.)

Zadanie 1. Wyznaczyć średnie i mediany zbioru X złożonego z punktów $(0, 0)$, $(1, 0)$, $(0, 1)$ i $(2, 3)$ na płaszczyźnie z metrykami: kartezjańską, miejską, kolejową (por. Delta 1/1975).

Zadanie 2. Z badać problem środka ciężkości na prostej ze zwykłą metryką i miarą rozproszenia postaci

$$r_X(x) = \sum_{i=1}^n [\rho(x, x_i)]^\alpha, \quad \alpha > 0.$$

Kącik filatelistyczny (5)

Johannes Kepler (1571—1630), niemiecki astronom i matematyk, sformułował prawa ruchu planet znane do dziś jako trzy prawa Keplera. Prawa te miały zasadnicze znaczenie dla zrozumienia ruchu ciał przyciągających się siłami grawitacyjnymi i zostały uzasadnione przez Newtona.

Skonstruował także lunetę astronomiczną, odmienną od lunety Galileusza.

Reprodukuje znaczek wydany przez Republikę Dahomeju w roku 1971, na którym obok portretu uczonego przedstawione są symboliczne orbity planet. Znaczki poświęcone Keplerowi wydano także w Austrii (w roku 1953), Meksyku (1971), NRD (1971), RFN (1971), i Rumunii (1971).

Jerzy BARTKE



Niektóre zadania prowadzące do wyznaczania ekstremów (tzn. maksimów i minimów) funkcji wielu zmiennych można rozwiązać stosując pewne nierówności, a więc środkami bardziej elementarnymi niż te, które są zwykle wykładane w podręcznikach analizy matematycznej. Nie mając tu pretensji do opisanie wszystkich elementarnych metod wyznaczania ekstremów opartych na nierównościach, ograniczymy się do podania przykładów zastosowań nierówności mających swe źródło w wypukłości pewnych funkcji.

Funkcję rzeczywistą f zdefiniowaną na pewnym podzbiore ciała liczb rzeczywistych nazywamy wypukłą w dół na przedziale I (zawartym w dziedzinie tej funkcji), gdy spełnia ona następujące warunki:

(W₁) funkcja f jest ograniczona na każdym przedziale domkniętym i ograniczonym zawartym w I ,

(W₂) dla dowolnych różnych liczb $x_1, x_2 \in I$ zachodzi nierówność

$$f\left(\frac{1}{2}(x_1+x_2)\right) < \frac{1}{2}(f(x_1)+f(x_2)).$$

Definicję funkcji wypukłej w górę otrzymujemy zatrzymując warunek (W₁) i zastępując w warunku (W₂) znak $<$ znakiem $>$. Zatem funkcja g jest wypukła w górę, gdy funkcja $-g$ jest wypukła w dół. Czytelnik może łatwo zinterpretować geometrycznie wypukłość funkcji. W podanych niżej przykładach sprawdzenie warunku (W₁) nie przedstawia trudności, dlatego zajmujemy się sprawdzeniem tylko drugiego warunku.

Przykłady.

a) Funkcja sinus jest wypukła w górę na przedziale $\langle 0, \pi \rangle$. Istotnie jeśli weźmiemy pod uwagę $x_1, x_2 \in \langle 0, \pi \rangle$ takie, że $x_1 \neq x_2$, to $0 < \cos\left(\frac{1}{2}(x_1-x_2)\right) < 1$, a więc

$$\frac{1}{2}(\sin x_1 + \sin x_2) = \sin\left(\frac{1}{2}(x_1+x_2)\right) \cos\left(\frac{1}{2}(x_1-x_2)\right) < \sin\left(\frac{1}{2}(x_1+x_2)\right).$$

b) Funkcje $x \rightarrow x^m$, gdzie $m > 1$ jest liczbą całkowitą, są wypukłe w dół dla $x \geq 0$. Sprawdzenie, że funkcja $x \rightarrow x^2$ jest wypukła w dół pozostawiamy Czytelnikowi. Załóżmy, że funkcja $x \rightarrow x^m$ jest wypukła w dół na rozważanym przedziale. Wykażemy wypukłość funkcji $x \rightarrow x^{m+1}$. Niech x_1, x_2 będą różnymi liczbami nieujemnymi. Ponieważ $(x_1^m - x_2^m)(x_1 - x_2) > 0$ zatem $x_1 x_2^m + x_2 x_1^m < x_1^{m+1} + x_2^{m+1}$.

Korzystając z tej nierówności i z wypukłości funkcji $x \rightarrow x^m$ otrzymujemy

$$\left(\frac{1}{2}(x_1+x_2)\right)^{m+1} < \frac{1}{4}(x_1^m+x_2^m)((x_1+x_2) < \frac{1}{2}(x_1^{m+1}+x_2^{m+1}).$$

c) Funkcje wykładnicze $x \rightarrow a^x$ ($0 < a \neq 1$) są wypukłe w dół w zbiorze wszystkich liczb rzeczywistych: gdy $x_1 \neq x_2$ to $(a^{\frac{x_1}{2}} - a^{\frac{x_2}{2}})^2 > 0$ a więc $a^{\frac{1}{2}(x_1+x_2)} < \frac{1}{2}(a^{x_1}+a^{x_2})$.

Zadanie. Pokazać, że funkcje $x \rightarrow \sqrt{x^2+h^2}$, dla $h \neq 0$ są wypukłe w dół w zbiorze liczb rzeczywistych.

W rachunku różniczkowym funkcji zmiennej rzeczywistej podaje się warunki dostateczne na wypukłość funkcji; powyższe przykłady pokazują, że w prostych przypadkach można wypukłość funkcji sprawdzić bezpośrednio. Podstawową nierównością, z której będziemy korzystali, jest podana niżej nierówność Jensena.

Twierdzenie Jensena. Niech f będzie funkcją wypukłą w dół na przedziale I i niech p_1, \dots, p_n będzie ciągiem liczb rzeczywistych spełniającym warunki:

$$p_i > 0 \quad (1 \leq i \leq n) \quad \text{oraz} \quad p_1 + \dots + p_n = 1.$$

Wtedy

$$f(p_1 x_1 + \dots + p_n x_n) \leq p_1 f(x_1) + \dots + p_n f(x_n)$$

dla dowolnych liczb $x_1, \dots, x_n \in I$. Równość ma miejsce tylko wtedy, gdy $x_1 = \dots = x_n$.

Dowód powyższego twierdzenia wymaga bliższego zbadania pojęcia wypukłości i wykorzystuje pojęcie ciągłości. Tutaj udowodnimy nierówność Jensena tylko dla przypadku, gdy liczby p_1, \dots, p_n są wymierne. Dowód przeprowadzimy w trzech krokach.

I. Załóżmy, że $p_1 = \dots = p_n = 1/2^m$, gdzie $m \geq 1$ jest liczbą całkowitą. W tym przypadku

nierówność Jensena ma postać:
$$f\left(\frac{1}{2^m}(x_1 + \dots + x_{2^m})\right) \leq \frac{1}{2^m}(f(x_1) + \dots + f(x_{2^m})).$$

Powyższą nierówność udowodnimy indukcyjnie względem m . Gdy $m = 1$ nierówność redukuje się do warunku (W₂). Załóżmy więc, że nierówność jest prawdziwa dla $m \geq 1$ i niech $x_1, \dots, x_{2^{m+1}} \in I$.

Mamy
$$f\left(\frac{1}{2^{m+1}}(x_1 + \dots + x_{2^{m+1}})\right) = f\left(\frac{1}{2}\left(\frac{x_1 + \dots + x_{2^m}}{2^m} + \frac{x_{2^m+1} + \dots + x_{2^{m+1}}}{2^m}\right)\right) \leq \frac{1}{2}\left(f\left(\frac{x_1 + \dots + x_{2^m}}{2^m}\right) + f\left(\frac{x_{2^m+1} + \dots + x_{2^{m+1}}}{2^m}\right)\right) \leq \frac{1}{2^{m+1}}(f(x_1) + \dots + f(x_{2^{m+1}})).$$

$\frac{1}{\omega} = ?$



lekcja?

Rozwiązanie zadania M 166.

Zauważmy, że $2 \cdot 10 + 5 \cdot 10 = 70$, a więc jeżeli liczby naturalne x i y spełniają równanie $2x + 5y = 70$, to $2(x-10) = 5(10-y)$. Ponieważ liczba $5(10-y)$ jest podzielna przez 2 oraz 2 i 5 są względnie pierwsze, więc $2|10-y$ czyli $10-y = 2t$, gdzie t jest pewną liczbą całkowitą. Jest więc $2(x-10) = 5 \cdot 2 \cdot t$, $x-10 = 5t$, skąd $x = 10+5t$, a poprzednio otrzymaliśmy równość $y = 10-2t$. Zadanie sprowadza się więc do wyznaczenia największej wartości sumy $(10+5t) + (10-2t) = 20+3t$, gdzie t jest liczbą całkowitą spełniającą nierówność $10+5t > 0$ i $10-2t > 0$ czyli $-2 < t < 5$. Funkcja $t \rightarrow 20+3t$ jest rosnąca, a więc dla całkowitych wartości t z przedziału $(-2, 5)$ przyjmuje największą wartość dla $t = 4$. Wartością tą jest 32.



Rozwiązanie zadania M 168.

Zadanie to można rozwiązać stosując rachunek różniczkowy. Pokażemy tu jednak, jak można uniknąć stosowania go, wykorzystując twierdzenie o średniej geometrycznej i arytmetycznej:

Dla dowolnych liczb dodatnich a_1, a_2, \dots, a_n zachodzi (por. art. A. Płockiego) nierówność

$$(a_1 a_2 \dots a_n)^{1/n} \leq \frac{a_1 + a_2 + \dots + a_n}{n}, \text{ przy}$$

czym równość ta staje się równością tylko w przypadku, gdy wszystkie liczby a_i są równe.

Zastosowanie tego twierdzenia do liczb $x, x, 1-x, 1-x$ nie dażądanego efektu:

$$x^2(1-x)^3 \leq \left[\frac{1}{5}(3-x) \right]^5 \leq \left(\frac{3}{5} \right)^5,$$

aby zaś zachodziła równość $x^2(1-x)^3 = \left(\frac{3}{5} \right)^5$ musiałoby być $x = 1-x$, czyli

$x = \frac{1}{2}$, wtedy zaś badana funkcja

przyjmuje wartość $\left(\frac{1}{2} \right)^5$.

Zastosujemy więc pewien trik: zachodzi równość

$$x^2(1-x)^3 = ax \cdot ax \cdot \frac{1-x}{a^{2/3}} \cdot \frac{1-x}{a^{2/3}} \cdot \frac{1-x}{a^{2/3}}.$$

Dobierzemy teraz odpowiednią wartość a . Zastosujemy twierdzenie o średniej geometrycznej i arytmetycznej

$$(ax)^2 \left(\frac{1-x}{a^{2/3}} \right)^3 \leq \left(\frac{2ax + 3 - 3x}{5a^{2/3}} \right)^5,$$

dla takiej wartości a , dla której po prawej stronie nierówności współczynnik przy x

staje się zerem, czyli dla $a_0 = \left(\frac{3}{2} \right)^{2/5}$.

Dla tej wartości a ostatnia nierówność staje się równością tylko dla takiej

wartości x , dla której $a_0 x = \frac{1-x}{a_0^{2/3}}$,

czyli $(a_0^{5/3} + 1)x = 1$, $\left(\frac{3}{2} + 1 \right) x = 1$,

$x = \frac{2}{5}$. Dla innych wartości $x \in (0, 1)$

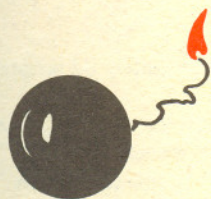
mamy $(a_0 x)^2 \left(\frac{1-x}{a_0^{2/3}} \right)^3 < \left(\frac{3}{5a_0^{2/5}} \right)^5$.

Tak więc funkcja f przyjmuje w przedziale

$(0, 1)$ maksymalną wartość dla $x = \frac{2}{5}$.

Jest nią $f\left(\frac{2}{5}\right) = \frac{2^2 \cdot 3^3}{5^5}$.

$$\frac{1}{\text{crown}} = ?$$



anarchia?

Posłużyliśmy się tu warunkiem (W₂) a następnie założeniem indukcyjnym. Nierówność \leq można zastąpić przez równość tylko wtedy, gdy $x_1 + \dots + x_{2^m} = x_{2^{m+1}} + \dots + x_{2^{m+1}}$, $x_1 = \dots = x_{2^m}$ oraz $x_{2^{m+1}} = \dots = x_{2^{m+1}}$, a więc wtedy, gdy $x_1 = \dots = x_{2^{m+1}}$. Ta uwaga kończy dowód twierdzenia Jensena w rozważanym przypadku.

II. Pokażemy, że $f\left(\frac{x_1 + \dots + x_n}{n}\right) < \frac{1}{n}(f(x_1) + \dots + f(x_n))$ dla dowolnego układu

nierównych liczb $x_1, \dots, x_n \in I$. W tym celu obierzmy liczbę całkowitą m taką, że $2^m > n$ i zastosujmy udowodnioną w I nierówność do układu 2^m liczb:

$$x_1, \dots, x_n, \frac{1}{n}(x_1 + \dots + x_n), \dots, \frac{1}{n}(x_1 + \dots + x_n),$$

w którym ostatnia liczba występuje $2^m - n$ razy. Mamy:

$$f\left(\frac{1}{2^m}(x_1 + \dots + x_n + \frac{2^m - n}{n}(x_1 + \dots + x_n))\right) < \frac{1}{2^m}(f(x_1) + \dots + f(x_n) + (2^m - n)f\left(\frac{x_1 + \dots + x_n}{n}\right)).$$

Stąd po prostych przekształceniach otrzymujemy żadaną nierówność.

III. Niech teraz p_1, \dots, p_n będzie ciągiem liczb wymiernych dodatnich takim, że $p_1 + \dots + p_n = 1$ i niech $x_1, \dots, x_n \in I$ będzie ciągiem liczb rzeczywistych nierównych. Możemy napisać:

$$p_i = \frac{k_i}{l_i} \quad (1 \leq i \leq n) \text{ gdzie } 0 < k_i < l_i \text{ są liczbami naturalnymi. Oznaczmy } l = l_1 \cdot \dots \cdot l_n$$

i zastosujmy nierówność udowodnioną w części II dowodu do ciągu $x_1, \dots, x_1, \dots, x_n, \dots, x_n$,

w którym liczba x_i występuje $\frac{k_i l}{l_i} = k_i l_1 \cdot \dots \cdot l_{i-1} l_{i+1} \cdot \dots \cdot l_n$ razy.

$$\text{Mamy} \quad f\left(\sum_{i=1}^n p_i x_i\right) = f\left(\frac{1}{l} \sum_{i=1}^n \frac{k_i l}{l_i} x_i\right) < \frac{1}{l} \sum_{i=1}^n \frac{k_i l}{l_i} f(x_i) = \sum_{i=1}^n p_i f(x_i),$$

co kończy dowód nierówności Jensena dla przypadku, gdy p_1, \dots, p_n są liczbami wymiernymi. W dowodzie powyższym korzystaliśmy tylko z warunku (W₂) definicji funkcji wypukłej w dół, warunek (W₁) jest niezbędny do dowodu twierdzenia w całej ogólności.

Rozważmy teraz kilka zadań o maksimach i minimach.

Przykłady.

a) W zbiorze wszystkich n -kątów wpisanych w dane koło o promieniu $r > 0$ wyznaczyć wielokąt o polu maksymalnym.

Rozważmy dowolny n -kąt wpisany w koło o promieniu r i niech x_1, \dots, x_n będą miarami kątów środkowych opartych na kolejnych bokach wielokąta. Zatem $x_i \in \langle 0, \pi \rangle$ oraz

$$x_1 + \dots + x_n = 2\pi. \text{ Pole rozważanego wielokąta dane jest wzorem } \frac{1}{2} r^2 (\sin x_1 + \dots + \sin x_n).$$

Stosując nierówność Jensena do funkcji \sin wypukłej w górę na przedziale $\langle 0, \pi \rangle$ i liczb

$$p_1 = \dots = p_n = \frac{1}{n} \text{ otrzymujemy: } \frac{1}{n} (\sin x_1 + \dots + \sin x_n) \leq \sin \frac{1}{n} (x_1 + \dots + x_n) = \sin \frac{2\pi}{n},$$

zatem pole n -kąta wpisanego w koło o promieniu r jest nie większe od $\frac{1}{2} r^2 n \sin \frac{2\pi}{n}$, przy czym

maksimum jest osiągnięte gdy $x_1 = \dots = x_n$, a więc rozwiązaniem zadania jest n -kąt foremny.

b) Na płaszczyźnie dany jest trójkąt o bokach a, b, c . Niech $h > 0$ będzie liczbą rzeczywistą i rozważmy zbiór wszystkich ostrosłupów o wysokości h , których podstawą jest dany trójkąt. Należy wyznaczyć ostrosłup należący do rozważanego zbioru o minimalnym polu powierzchni bocznej.

Ostrosłup z rozważanego zbioru jest jednoznacznie wyznaczony przez podanie rzutu M jego wierzchołka na płaszczyznę trójkąta. Położenie punktu M jest określone przez podanie odległości x, y, z z punktu M od odpowiednich boków trójkąta, przy czym każdą z tych odległości bierzemy ze znakiem plus, gdy punkt M leży z tej samej strony odpowiedniego boku, co cały trójkąt, ze znakiem minus — w przeciwnym wypadku. Mamy: $ax + by + cz = 2P$, gdzie P oznacza pole trójkąta. Natomiast pole powierzchni bocznej ostrosłupa dane jest wzorem:

$$\frac{1}{2} (a\sqrt{x^2 + h^2} + b\sqrt{y^2 + h^2} + c\sqrt{z^2 + h^2}).$$

Funkcja $\sqrt{x^2 + h^2}$ jest wypukła w dół, zatem na mocy nierówności Jensena

$$\sqrt{(px + qy + rz)^2 + h^2} \leq p\sqrt{x^2 + h^2} + q\sqrt{y^2 + h^2} + r\sqrt{z^2 + h^2}$$

dla liczb dodatnich p, q, r spełniających warunek $p + q + r = 1$. Podstawiając $p = \frac{a}{a+b+c}$,

$q = \frac{b}{a+b+c}$ oraz $r = \frac{c}{a+b+c}$ i korzystając z nierówności $ax + by + cz \leq 2P$ otrzymujemy:

$$(a+b+c) \sqrt{\frac{4P^2}{(a+b+c)^2} + h^2} \leq a\sqrt{x^2 + h^2} + b\sqrt{y^2 + h^2} + c\sqrt{z^2 + h^2},$$

przy czym równość zachodzi tylko w przypadku gdy $x = y = z$, a więc minimalne pole powierzchni bocznej ma ostrosłup, którego rzut wierzchołka M jest środkiem koła wpisanego w dany trójkąt.

Zadanie. Wyznaczyć minimum sumy m -tych potęg: $x_1^m + \dots + x_n^m$ liczb nieujemnych x_i zakładając, że ich suma a jest stała i dodatnia.

$$\frac{1}{\text{crown}} = ?$$

Niech teraz p_1, \dots, p_n będzie ciągiem liczb spełniających założenia twierdzenia Jensena.

Wykażemy następującą nierówność:

Nierówność dla średniej geometrycznej. Dla dowolnych $x_1 > 0, \dots, x_n > 0$: $x_1^{p_1} \dots x_n^{p_n} \leq p_1 x_1 + \dots + p_n x_n$ przy czym równość ma miejsce tylko wtedy, gdy $x_1 = \dots = x_n$.

Dla dowodu wystarczy obrać liczby t_i ($1 \leq i \leq n$) tak, że $x_i = 10^{t_i}$ i zastosować nierówność Jensena do funkcji wykładniczej $t \rightarrow 10^t$ wypukłej w dół w całym zbiorze liczb rzeczywistych.

Zauważmy, że dla $p_1 = \dots = p_n = \frac{1}{n}$ otrzymujemy nierówność $\sqrt[n]{x_1 \dots x_n} \leq \frac{1}{n}(x_1 + \dots + x_n)$, która

orzeka, że średnia geometryczna skończonego układu liczb nie przekracza średniej arytmetycznej.

Możemy teraz łatwo udowodnić następujące:

Twierdzenie (maksimum iloczynu przy stałej sumie) *Jeśli suma n liczb dodatnich*

x_1, \dots, x_n *jest stała, to iloczyn $x_1^{q_1} \dots x_n^{q_n}$ gdzie q_1, \dots, q_n są dowolnymi liczbami dodatnimi*

ma największą wartość wtedy i tylko wtedy, gdy $\frac{x_1}{q_1} = \dots = \frac{x_n}{q_n}$.

Dowód: Połóżmy $c = x_1 + \dots + x_n$, $q = q_1 + \dots + q_n$; stosując nierówność dla średniej

geometrycznej otrzymujemy: $\left(\frac{x_1}{q_1}\right)^{q_1} \dots \left(\frac{x_n}{q_n}\right)^{q_n} \leq \frac{c}{q}$, stąd $\left(\frac{x_1}{q_1}\right)^{q_1} \dots \left(\frac{x_n}{q_n}\right)^{q_n} \leq \left(\frac{c}{q}\right)^q$;

iloczyn $x_1^{q_1} \dots x_n^{q_n}$ jest największy, gdy największy jest iloczyn po lewej stronie ostatniej

nierówności, a więc gdy $\frac{x_1}{q_1} = \dots = \frac{x_n}{q_n}$.

Dowód twierdzenia podanego niżej przebiega podobnie.

Twierdzenie (minimum sumy przy stałym iloczynie). *Jeżeli iloczyn $x_1^{q_1} \dots x_n^{q_n}$ jest stały, to*

suma $x_1 + \dots + x_n$ ma najmniejszą wartość, gdy $\frac{x_1}{q_1} = \dots = \frac{x_n}{q_n}$.

Przykłady.

a) *Spośród wszystkich trójkątów o danym obwodzie wyznaczyć trójkąt o największym polu*
Niech $2p$ będzie obwodem trójkąta o bokach x, y, z ; kwadrat jego pola jest dany wzorem Herona: $p(p-x)(p-y)(p-z)$. Ponieważ $(p-x) + (p-y) + (p-z) = p$ i $p-x > 0, p-y > 0, p-z > 0$, zatem iloczyn $(p-x)(p-y)(p-z)$ jest największy, gdy $p-x = p-y = p-z$, a więc rozwiązaniem zadania jest trójkąt równoboczny.

b) *Spośród wszystkich prostopadłościanów o danej objętości wyznaczyć prostopadłościan o najmniejszym polu powierzchni.* Niech v będzie objętością prostopadłościanu o krawędziach x, y, z . Pole powierzchni prostopadłościanu jest zatem równe $2(xy + xz + yz)$. Ponieważ $(xy)(xz)(yz) = v^2$ oraz $xy > 0, xz > 0, yz > 0$ zatem suma $yx + yz + xz$ jest najmniejsza, gdy $xy = xz = yz$, a więc gdy $x = y = z$. Rozwiązaniem zadania jest sześcian.

Proponujemy Czytelnikowi rozwiązanie podanych niżej zadań.

Zadania.

- 1) Znaleźć największą wartość funkcji $\sin x + \sin y - \sin(x+y)$ w trójkącie ograniczonym osiami x, y i prostą $x+y = 2\pi$.
- 2) Spośród wszystkich prostopadłościanów o danym polu powierzchni wyznaczyć prostopadłościan o największej objętości.
- 2) Niech p, q będą liczbami dodatnimi. Znaleźć maksimum funkcji $(\cos x)^p (\sin x)^q$ w przedziale $(0, 2\pi)$.

Redaguje mgr Andrzej MAKOWSKI

M 166. Wyznaczyć największą wartość sumy liczb naturalnych x, y spełniających równanie $2x + 5y = 70$.

Rozwiązanie na str. 10

M 167. Na płaszczyźnie dana jest prosta k oraz punkty A i B leżące po różnych jej stronach. Wyznaczyć na prostej k punkt X , dla którego $|AX - BX|$ przyjmuje wartość największą.

Rozwiązanie na str. 7.

M 168. Funkcja f jest określona na przedziale $(0, 1)$ wzorem $f(x) = x^2(1-x)^3$. Obliczyć największą wartość tej funkcji.

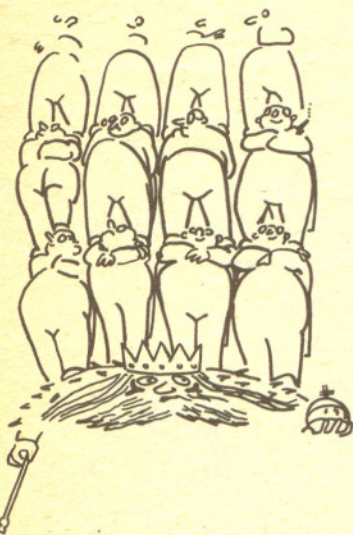
Rozwiązanie na str. 11

Redaguje dr Waldemar GORZKOWSKI

F 55. Napięcie powierzchniowe na płaskiej granicy kryształu i roztworu zależy od orientacji tej granicy względem kierunków krystalograficznych.

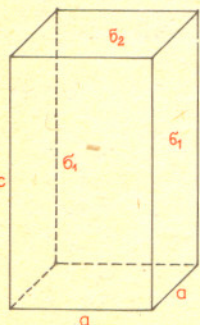
Napięcia powierzchniowe kryształu pokazanego na rysunku, znajdującego się w roztworze nasyconym tej samej substancji, z której zbudowany jest kryształ, na wzajemnie prostopadłych ścianach wynoszą odpowiednio σ_1, σ_1 i σ_2 . Kryształ znajduje się w stanie równowagi termodynamicznej z roztworem. Należy wyznaczyć stosunek krawędzi a/c . Wpływ siły ciężkości na kryształ i roztwór zaniedbujemy.

Rozwiązanie na str. 4



hierarchia?

Zadania



Mała delta



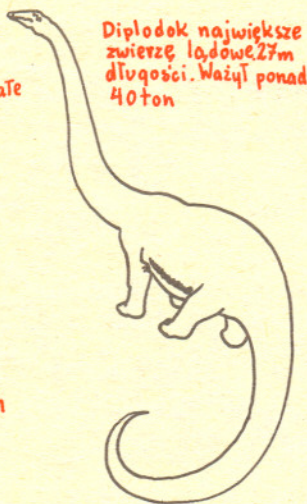
Lepiej być małym czy dużym?



Wieloręby są największymi zwierzętami, jakie kiedykolwiek żyły na Ziemi



Stawonogi są raczej małe za to najliczniejsze



Diplodok największe zwierzę lądowe 27m długości. Ważył ponad 40 ton



Zębiek karliczek jeden z najmniejszych ssaków

Jaka kulka	Masa	Po- wierzchnia	Ile gramów masy przypada na 1 cm ² powierzchni
kulka o średnicy 1 cm	0,5 g	3 cm ²	0,17 g/cm ²
kulka o średnicy 2 cm	4 g	12 cm ²	0,33 g/cm ²

Porównajmy sylwetkę słonia z sylwetką myszy. Dlaczego małe zwierzęta są zgrabniejsze od dużych? Okazuje się, że wszystko tłumaczy proste rachunki. Oto dwie kulki z plasteliny — jedna o średnicy 1 cm, druga o średnicy 2 cm. Przyjmując, że gęstość plasteliny jest równa w przybliżeniu 1 g/cm³, obliczmy, ile ważą te kulki i jaka jest ich powierzchnia. Oto odpowiednie wzory.

Objętość kuli o promieniu $R = \frac{4}{3} \pi R^3$.

Pole powierzchni kuli o promieniu $R = 4\pi R^2$.

Dla pierwszej z kulek $R = \frac{1}{2}$ cm, stąd jej objętość

jest równa $\frac{4}{3} \pi \cdot \frac{1}{8} \text{ cm}^3 = \frac{1}{6} \pi \text{ cm}^3$. A więc kulka ta ma

masę $\frac{\pi}{6} \text{ g} \approx 0,5 \text{ g}$. Pole jej powierzchni jest równe

$4\pi \frac{1}{4} \text{ cm}^2 = \pi \text{ cm}^2 \approx 3 \text{ cm}^2$. Dla drugiej z kulek

$R = 1$ cm, stąd jej objętość jest równa $\frac{4}{3} \pi \text{ cm}^3$. A więc

kulka ta ma masę $\frac{4}{3} \pi \text{ g} \approx 4 \text{ g}$ (8 razy więcej niż

poprzednia). Pole jej powierzchni jest równe $4\pi \text{ cm}^2 \approx 12 \text{ cm}^2$ (tylko czterokrotnie więcej niż poprzednio).

Wyniki obliczeń zestawmy w tabelce, której ostatnia rubryka informować będzie, ile gramów masy przypada na 1 cm² powierzchni danej kulki.

Wniosek: im mniejsza kulka, tym mniej gramów jej masy przypada na centymetr kwadratowy powierzchni.

Tyranozaur-największy drapieżnik lądowy Ważył 8 ton



Najmniejszy ptak świata - Koliber- waży zaledwie 1,6 grama



Diatryma-największy ptak drapieżny. Wzrost około 2 m.



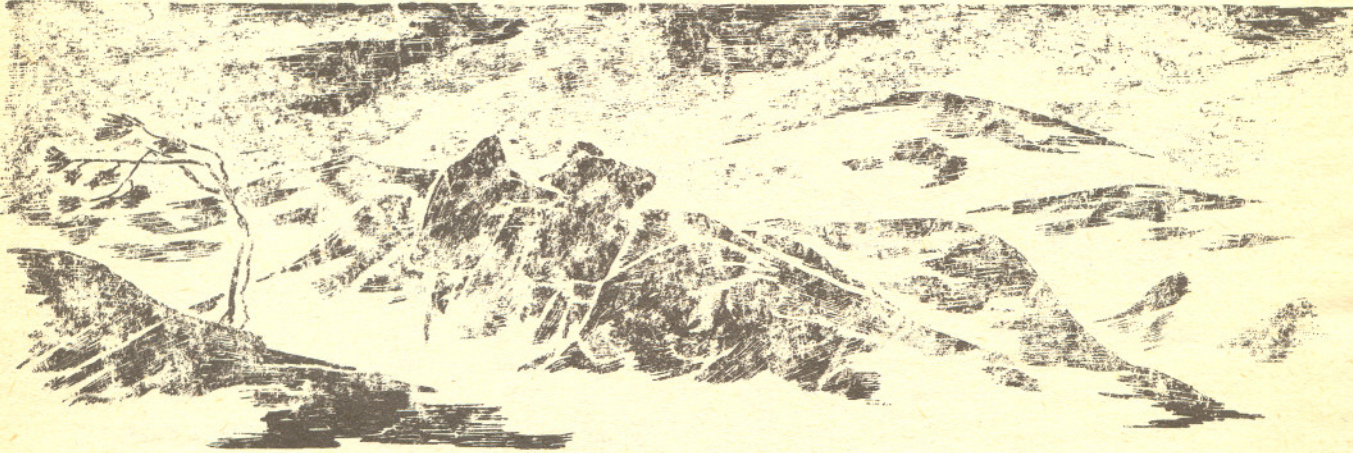


Gigantopitek -
-największa
znana małpa.
Ważył około 300kg.
Wzrost jego docho-
dził prawdopodobnie
do 2,7m.

Spróbujmy teraz wyjaśnić, co by się stało, gdyby małą mysz powiększyć proporcjonalnie do wielkości słonia? Otóż na 1 cm² powierzchni przekroju poprzecznego jej kości przypadałoby po tym zabiegu prawie stokrotnie więcej masy ciała!

A przecież wytrzymałość kości na złamanie jest proporcjonalna do jej przekroju poprzecznego! Mała mysz, dzięki swym małym wymiarom, może być zbudowana lżej i zgrabniej. Małe wymiary mają jednak i swoje minusy. Gdy zależy nam, by na mrozie stracić jak najmniej ciepła, lepiej gdy na jednostkę masy przypada jak najmniej powierzchni. Tłumaczy to, dlaczego zwierzęta żyjące w chłodniejszym klimacie są większe niż osobniki tego samego gatunku z okolic cieplejszych.

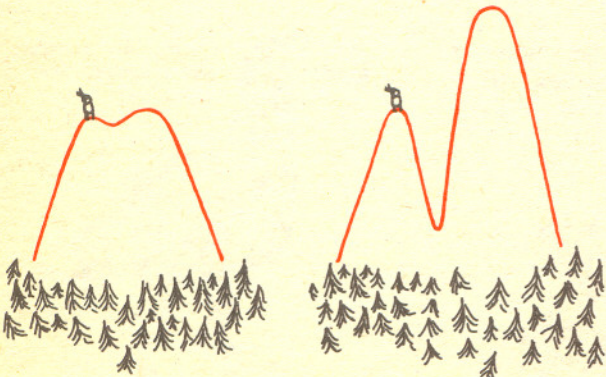
A swoją drogą, jakie to dziwne, że świat wokół nas nie jest wcale jednorodny — to, co małe, rządzi się innymi prawami niż to, co duże. A matematyka przekonuje nas, że gigantyczna małpa King Kong, wielkości sporego domu, to zupełnie nierealna fantazja.



Wycieczki po maksimach, minimach i punktach przegięcia

„Burza szalała bez przerwy. Dwaj alpinisci byli już u kresu sił, ale krok po kroku pięli się w górę. Śnieg zasypywał im oczy, a wiatr chciał zepchnąć ich w dolinę. Szli już trzy godziny od obozu VII. Nagle droga pod górę skończyła się. Mimo zadymki widzieli, że przed nimi tylko droga w dół. Byli na szczycie!”. Alpinisci rozumowali tak: skoro wszystkie drogi prowadzą już w dół, jesteśmy na szczycie. Czy to jest poprawne rozumowanie? Zależy to, niestety, od tego, co nazwiemy szczytem (rys. 1). Na pytanie postawione na rysunku 1 większość ludzi odpowie „nie”, przeciwnie niż na rysunku 2.

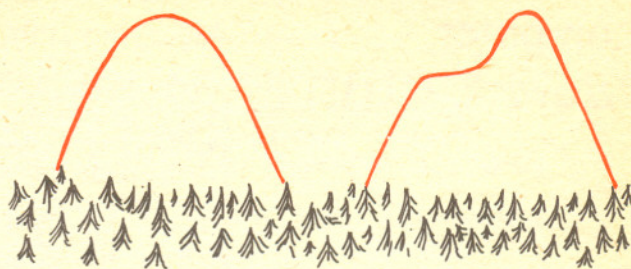
W matematyce odróżniamy pojęcie „maksimum” funkcji od jej „wartości największej”. Funkcja ma maksimum w punkcie x , jeżeli blisko tego punktu nie ma wartości większych, a „największą wartość” — jeżeli w ogóle nie przyjmuje większych wartości. Każdy wierzchołek górski tworzy maksimum (funkcji „wysokość”). Z wierzchołka nie ma drogi pod górę.



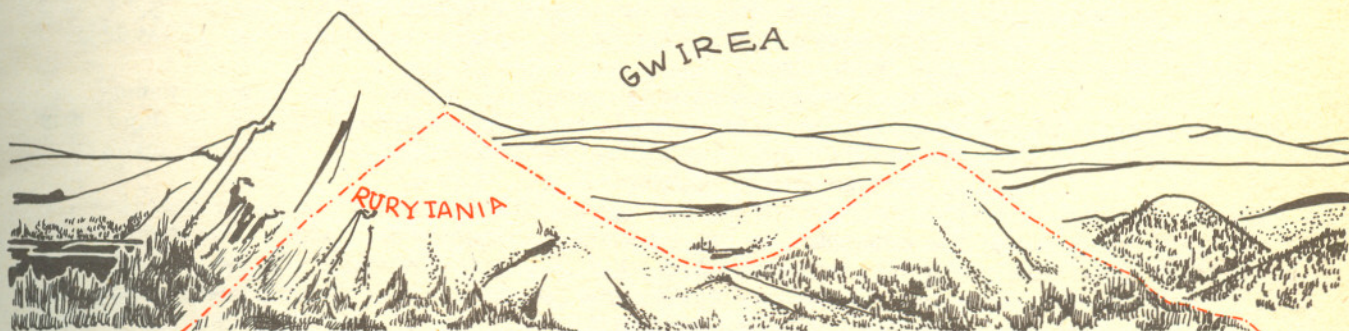
Rys. 1. Czy alpinista jest na szczycie?

Rys. 2. A teraz?

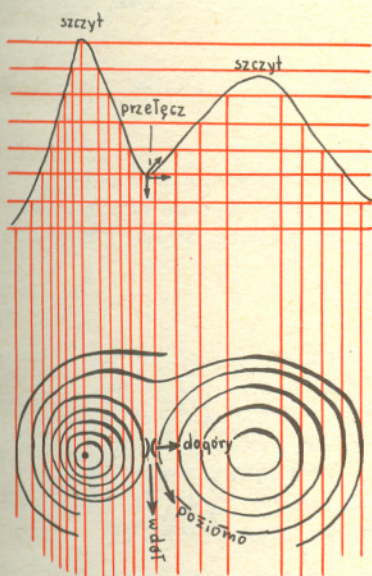
Wśród szczytów tatrzańskich tylko jeden jest największy, ale są funkcje, które przyjmują swą największą wartość w wielu punktach (rys. 3). Jeżeli zawężymy obszar, na którym rozpatrujemy funkcję, możemy dostać inną wartość największą. To zrozumiałe, że funkcja „wysokość” kulminuje gdzie indziej w Tatrach Polskich, całych Tatrach, a jeszcze gdzie indziej na obszarze całej Ziemi. Przy ograniczaniu obszaru może się nawet zdarzyć, że największa wartość funkcji w jednym obszarze nie będzie nawet maksimum w większym. Widzimy to na rysunku 4. Najwyższe wzniesienie Rurytania nie jest w ogóle szczytem dla Gwirczyków. Ale przy dowolnych zmianach granic każdy szczyt zostanie szczytem.



Rys. 3. Te dwie góry mają te same wysokości.



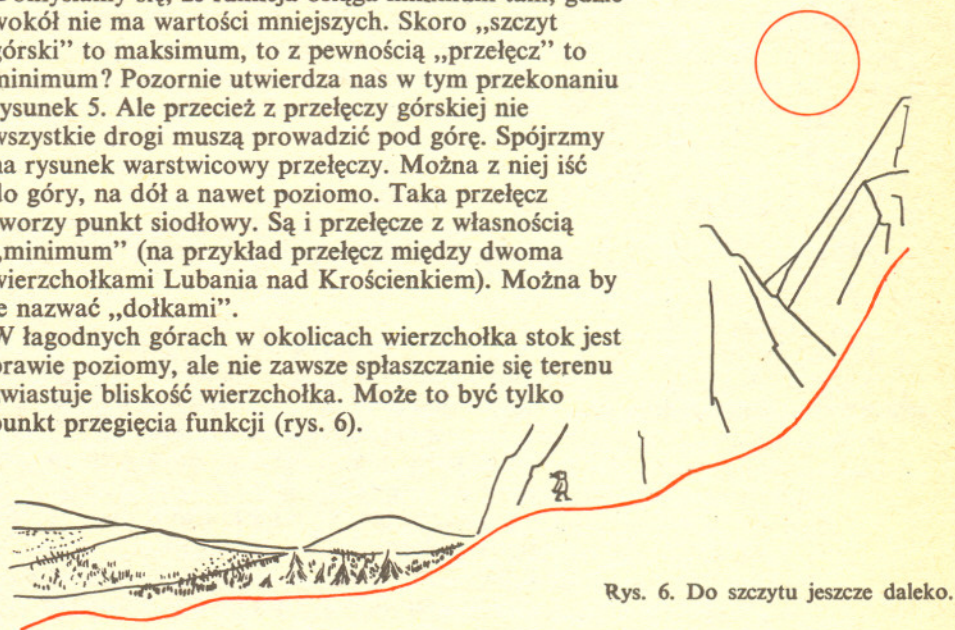
Rys. 4.



Rys. 5. Przełęcz-siodło.

Domyślamy się, że funkcja osiąga minimum tam, gdzie wokół nie ma wartości mniejszych. Skoro „szczyt górski” to maksimum, to z pewnością „przełęcz” to minimum? Pozornie utwierdza nas w tym przekonaniu rysunek 5. Ale przecież z przełęczy górskiej nie wszystkie drogi muszą prowadzić pod górę. Spójrzmy na rysunek warstwicowy przełęczy. Można z niej iść do góry, na dół a nawet poziomo. Taka przełęcz tworzy punkt siodłowy. Są i przełęcze z własnością „minimum” (na przykład przełęcz między dwoma wierzchołkami Lubania nad Krościenkiem). Można by je nazwać „dołkami”.

W łagodnych górach w okolicach wierzchołka stok jest prawie poziomy, ale nie zawsze spłaszczanie się terenu zwiastuje bliskość wierzchołka. Może to być tylko punkt przegięcia funkcji (rys. 6).



Rys. 6. Do szczytu jeszcze daleko.

Maksima i minima noszą wspólną nazwę ekstremów funkcji. To bardzo ważne pojęcie. Każdego z nas interesuje osiągnięcie maksymalnych efektów przy minimalnym (a nawet najmniejszym) nakładzie pracy i nie ma w tym nic złego. Minimum funkcji szuka i Natura. Ciała poruszają się po najkrótszych torach, choć w warunkach zakłóceń niekoniecznie są to linie proste, bańki mydlane starają się zawrzeć maksymalną objętość w minimalnej powierzchni, promienie świetlne wybierają najkrótsze drogi. Niemal w każdym dziale fizyki znajduje zastosowanie zasada minimum (sformułowana w połowie zeszłego wieku przez matematyka irlandzkiego Williama Rowana Hamiltona). Ta zasada mówi, że Natura działa w taki sposób, by pewna (ściśle określona, rzecz jasna) całka osiągała minimum. Jest to chyba najpotężniejsza pojedyncza zasada nauk ścisłych.

Algorytmy, obliczalność i rozstrzygalność

Przykładów algorytmów dostarcza w obfitości arytmetyka — wykonywanie działań podstawowych, wyszukiwanie liczb pierwszych itp. Jako inny przykład przytoczymy pewien algorytm rysowania na kartce papieru.

Niech operacjami podstawowymi będą:

up — podnieś rysik (tak by nie dotykał kartki),

down — opuść rysik (tak by dotykał kartki),

move (*a*, *b*) — przesuń (prostoliniowo) rysik do punktu o współrzędnych (*a*, *b*),

end — zakończ.

Wówczas ciąg instrukcji na marginesie przedstawia algorytm rysowania znaku \times o środku w punkcie (0, 0) i długości ramienia *a*.

Intuicyjne pojęcie algorytmu jest dość jasne: algorytm to zupełnie jednoznacznie określona procedura, czyli sposób postępowania. Algorytm nie może — w żadnej sytuacji, której dotyczy — pozostawiać wątpliwości, czy dalej postąpić tak, czy inaczej. Musi też jednoznacznie określać swe zakończenie.

Poszczególne „kroki” algorytmu muszą być efektywnie wykonalne, dając zawsze ten sam wynik, przynajmniej w zakresie istotnym dla zadania, którego algorytm dotyczy.

Mimo to, pojęcie algorytmu jest na tyle pierwotne w ludzkim pojmowaniu spraw tego świata, że nie oczekujemy, iż jakaś uznana gałąź nauki dostarczy nam jego ścisłej definicji. Zamiast tego mamy szereg formalnych odpowiedników intuicyjnego pojęcia algorytmu.

Mówiąc o algorytmie, abstrahujemy od materialnej strony jego realizacji (liczydło, palce, napisy). Dlatego zamiast mówić o obiektach, których dotyczy algorytm, i o sytuacjach, w jakich obiekty te uczestniczą w kolejnych stadiach algorytmu, możemy mówić o znakach, które w jakikolwiek umowny sposób oznaczają te obiekty i sytuacje. Algorytm widziany od tej formalno-lingwistycznej strony, to dokładnie określony przepis na dokonywanie zmian w ciągu znaków. Przyjmuje się przy tym, że zbiór znaków dopuszczalnych w tych napisach (czyli tzw. *alfabet*) jest skończony, same napisy są skończone, no i oczywiście opis reguł określających algorytm też jest skończony. Rodzaj znaków alfabetu, ani nawet ich liczba, nie są przy tym istotne, a to wskutek możliwości jednoznacznych i prostych przekodowań napisów.

Możliwa jest także „arytmetyzacja” algorytmów. Mając dowolny alfabet

$$A = \{a_1, a_2, \dots, a_n\}$$

możemy rozpatrywać (nieskończony) zbiór *słów* nad tym alfabetem, to znaczy zbiór wszystkich skończonych ciągów utworzonych ze znaków a_j należących do *A*. Zbiór słów nad alfabetem *A* oznaczmy przez A^* . Wprowadzając do alfabetu *A* dowolne uporządkowanie, np. takie że

$$a_i < a_j \equiv i \leq j,$$

możemy już jednoznacznie uporządkować zbiór A^* . Słowa o tej samej ilości znaków porządkujemy w sposób leksykograficzny i umawiamy się, że słowo krótsze jest zawsze wcześniejsze od dłuższego. Każde słowo nad *A* ma wówczas dokładnie określony numer $k > 0$, a każdy taki numer wyznacza dokładnie jedno słowo nad *A*. Nadając słowu pustemu numer 0, otrzymujemy wzajemnie jednoznaczny relację między ciągami znaków a liczbami naturalnymi. Działanie dowolnego algorytmu można więc również jednoznacznie przedstawić jako przekształcenie liczb naturalnych.

Z pojęciem algorytmu wiąże się cały szereg pojęć pochodnych. Należy do nich przede wszystkim pojęcie funkcji obliczalnej. Niech *A*, *B* będą alfabetami i niech

$$f: A^* \rightarrow B^*$$

będzie funkcją częściową (tzn. określoną na ogół tylko dla niektórych słów nad *A*, nie koniecznie na całym zbiorze A^*). Funkcję tę nazywamy *obliczalną*, jeżeli istnieje taki algorytm, że dla każdego $s \in A^*$, dla którego funkcja *f* jest określona, wynikiem działania tego algorytmu jest $f(s)$.

Podobnie określamy pojęcie obliczalności dla funkcji arytmetycznych $f: N \rightarrow N$, tzn. określonych w zbiorze liczb naturalnych i przybierających wartości naturalne. Z uwagi na możliwość wzajemnego „przekładu” klasa formalno-lingwistycznych funkcji obliczalnych wyznacza klasę arytmetycznych funkcji obliczalnych

up
move (*a*, *a*)
down
move ($-a$, $-a$)
up
move ($-a$, *a*)
down
move (*a*, $-a$)
up
end

$$\frac{1}{\text{crown}} = ?$$

1
B
E
O
M
E
S
O
E
R
C
H
E
-
B
N
I



$$\frac{1}{\cup} = ?$$

i vice versa. Możemy więc uznać, że mamy w gruncie rzeczy jedno ogólne pojęcie obliczalności funkcji. Dotyczy to także funkcji wielu zmiennych, a to dzięki istnieniu efektywnych wzajemnie jednoznacznych odwzorowań postaci

$$N \times N \leftrightarrow N, \quad N \times N \times N \leftrightarrow N$$

etc., pozwalających wzajemnie jednoznacznie odwzorować zbiory par, trójek, etc. liczb naturalnych na zbiór liczb naturalnych. Innym ważnym pojęciem jest pojęcie rozstrzygalności predykatów. Predykat to nazwa własności; zdanie orzekające, że: „przedmiot x ma własność P ”, zapisujemy w skrócie jako $P(x)$.

Predykat P o dziedzinie D nazywamy *rozstrzygalnym*, jeżeli istnieje algorytm, który dla każdego $x \in D$ pozwala rozstrzygnąć czy $P(x)$, czy nie $P(x)$.

O rozstrzygalnym predykatie możemy mówić, że wyraża on własność lub relację rozstrzygalną.

Mając zdefiniowane pojęcie rozstrzygalności predykatu możemy natychmiast zdefiniować pojęcie zbioru rozstrzygalnego. Mianowicie mówimy, że zbiór $X \subset D$ jest rozstrzygalny (względem D), jeżeli predykat P_X zdefiniowany przez

$$P_X(x) \stackrel{\text{df}}{=} x \in X$$

jest rozstrzygalny. Równoważną definicję otrzymujemy korzystając wprost z pojęcia obliczalności: zbiór jest rozstrzygalny wtedy i tylko wtedy, gdy jego funkcja charakterystyczna (równa 1 dla elementów należących do zbioru i 0 dla pozostałych) jest obliczalna.

Pojęcie rozstrzygalności jest więc szczególnym przypadkiem pojęcia obliczalności, sprowadza się ono bowiem do obliczalności funkcji o wartościach zero-jedynkowych (logicznych, bądź arytmetycznych).

Wszystkie powyższe pojęcia mają charakter nieformalny. Jednakże wszystkie one mogą być uściślone. Oczywiście uściślone pojęcia zawsze różnią się od ich pierwowzorów intuicyjnych. Za triumf teorii algorytmów można jednakże uznać fakt, że dokonane na wiele różnych sposobów uściślenia okazują się formalnie wzajemnie równoważne. Takimi są np. pojęcia obliczalności w sensie Turinga, λ -definiowalności Churcha, μ -rekursywności Kleene'go, ogólnej rekursywności i wiele innych.

Algorytm jest zawsze algorytmem czegoś, przepisem na wykonywanie pewnego zadania, jak rozwiązywanie równania kwadratowego, tłumaczenie programu z języka FORTRAN na kody wewnętrzne maszyny cyfrowej, wykonywanie pewnych rysunków. W trakcie realizacji algorytmu następuje przekształcenie pewnej sytuacji wejściowej w wyjściową, czyli „danych” w „wyniki”. W sytuacji, kiedy algorytm jest narzędziem działania, a nie obiektem studiów, sama struktura algorytmu jest mniej ważna niż wykonywane zadanie, czyli funkcja

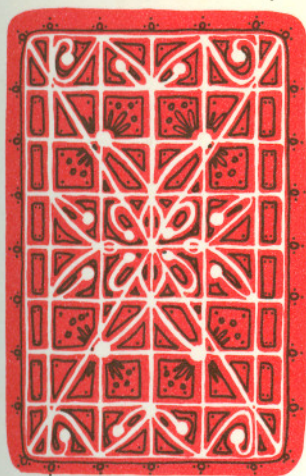
dane \rightarrow wyniki.

Każda taka funkcja określona przez wyniki realizacji pewnego algorytmu jest, zgodnie z podanymi wyżej definicjami, funkcją obliczalną.

Takie okoliczności powodują, że oprócz tego rygorystycznego pojęcia algorytmu, które omówiliśmy powyżej, celowe może być wprowadzanie pojęć opartych na złagodzonych wymaganiach co do zdefiniowania postępowania w każdej sytuacji, która może się zdarzyć w trakcie wykonywania algorytmu. W pewnych sytuacjach możliwe są bowiem różne warianty postępowania, które jednak różnią się w sposób nieistotny w tym sensie, że prowadzą do tego samego wyniku. Zdarza się to już w najprostszych działaniach arytmetycznych: np. instrukcja „oblicz sumę liczb a , b , c ” jest całkiem jednoznaczna, jeśli chodzi o wynik, lecz z punktu widzenia działań elementarnych algorytmem nie jest, bo nie określa kolejności poszczególnych operacji dodawania.

Dopuszczenie tego rodzaju swobody wyboru postępowania w szczegółach nie mających wpływu na wynik prowadzi do rozszerzenia pojęcia algorytmu. Posługując się tak rozszerzonym pojęciem algorytmu należy jednak pamiętać, że jest ono różne od pojęcia rygorystycznego, i że ta różnica czasem może okazać się ważna. Jeżeli przez maszynę rozumiemy układ, który nie jest w stanie dokonać wyboru — ani w sposób przypadkowy, ani metodyczny — to maszyna taka może realizować algorytmy tylko w sensie rygorystycznym.

Natomiast pojęcia takie jak „zadanie algorytmizowalne” (tzn. mające algorytm wykonywania), „funkcja obliczalna” itp. nie ulegają zmianie przy takim rozszerzeniu pojęcia algorytmu. Każdy bowiem algorytm rygorystyczny jest algorytmem w sensie rozszerzonym, a każdy algorytm w sensie rozszerzonym można „przeoibić” na algorytm rygorystyczny, realizujący tę samą funkcję.



?

Przepisy obliczeniowe, zawierające działania na dowolnych liczbach rzeczywistych, nie są algorytmami w sensie określonym wyżej. Ścisłe rozumiane działania na dowolnych liczbach rzeczywistych nie spełniają bowiem warunku efektywności, gdyż nie są wykonalne w skończonej liczbie kroków. Rachunki przybliżone są oczywiście algorytmizowalne, ale wymaga to sprecyzowania sposobu traktowania przybliżeń. Po takim zabiegu przepis obliczeniowy staje się algorytmem i wówczas daje się reprezentować przez pewne operacje na liczbach naturalnych.