Jeszcze inne uzasadnienie niewymierności $\sqrt{2}$ można znaleźć w artykule Adama Barańskiego *Uogólnienie algorytmu Euklidesa*, Δ^5_{21}.

O metodzie Herona można przeczytać w *Delcie*, na przykład w artykule Piotra Krzyżanowskiego i Grzegorza Łukaszewicza *Przez wieki z metodą Newtona*, Δ^9_{21}.

Rozumowanie 2. Załóżmy, że $\sqrt{2}$ jest liczbą wymierną. Istnieje wtedy najmniejsza liczba naturalna k taka, że $k\sqrt{2}$ jest liczbą naturalną. Wówczas $k\sqrt{2} - k$ jest *mniejszą* liczbą o tej własności. Sprzeczność! $\square$

Pozostał problem wyznaczenia przybliżonej wartości $\sqrt{2}$. Heron z Aleksandrii około 60 roku n.e. opisał rekurencyjne postępowanie: $x_1 = 2$ i $x_{n+1} = \frac{1}{2}(x_n + \frac{2}{x_n})$ dla $n = 1, 2, \ldots$, w którym każda kolejna iteracja coraz dokładniej przybliża wartość $\sqrt{2}$, podwajając liczbę cyfr znaczących po przecinku:

$$x_2 = \tfrac{3}{2} = 1{,}5,$$

$$x_3 = \tfrac{17}{12} = \mathbf{1{,}41}66\ldots,$$

$$x_4 = \tfrac{577}{408} = \mathbf{1{,}41421}56\ldots,$$

$$x_5 = \tfrac{665857}{470832} = \mathbf{1{,}41421356}23745\ldots,$$

$$x_6 = \tfrac{886731088897}{627013566048} = \mathbf{1{,}4142135623730950}488016896\ldots \text{ itd.}$$

Podstawą tej idei jest pomysł, że jeśli x_n jest przeszacowaniem, a $\frac{2}{x_n}$ jest niedoszacowaniem, lub odwrotnie, to średnia arytmetyczna tych liczb może być lepszym przybliżeniem $\sqrt{2}$ niż każda z nich z osobna. Pokażemy, że procedura Herona jest zbieżna.

(1) Ciąg $\{x_n\}$ jest ograniczony z dołu przez $\sqrt{2}$, bo

$$x_{n+1}^2 - 2 = \left(\frac{x_n^2+2}{2x_n}\right)^2 - 2\left(\frac{2x_n}{2x_n}\right)^2 = \left(\frac{x_n^2-2}{2x_n}\right)^2 \geqslant 0.$$

(2) $x_{n+2} \leqslant x_{n+1}$ dla $n = 1, 2, \ldots$, bo

$$x_{n+2} - x_{n+1} = \frac{1}{2}\left(x_{n+1} + \frac{2}{x_{n+1}}\right) - x_{n+1} = \frac{2 - x_{n+1}^2}{2x_{n+1}} \leqslant 0,$$

gdyż $x_{n+1}^2 \geqslant 2$.

Skoro ciąg $\{x_n\}$ jest ograniczony z dołu i nierosnący, to jest zbieżny. Jego granicą jest pierwiastek równania $g = \frac{1}{2}\big(g + \frac{2}{g}\big)$ spełniający nierówność $g \geqslant \sqrt{2}$, czyli w tym przypadku $g = \sqrt{2}$.

Zbieżność ciągu $\{x_n\}$ do granicy $\sqrt{2}$ jest *kwadratowa*, tzn. różnica między kolejnym przybliżeniem a granicą $\sqrt{2}$ maleje jak kwadrat poprzedniej różnicy:

(3) $|x_{n+1} - \sqrt{2}| = \left|\frac{1}{2}\big(x_n + \frac{2}{x_n}\big) - \sqrt{2}\right| = \frac{|x_n - \sqrt{2}|^2}{2|x_n|} \leqslant \frac{1}{2\sqrt{2}}|x_n - \sqrt{2}|^2.$

To właśnie kwadratowa zbieżność jest odpowiedzialna za to, że liczba cyfr znaczących podwaja się przy każdej iteracji.

Algorytm Herona jest skuteczny w przypadku obliczania pierwiastka kwadratowego $\sqrt{A}$ dla dowolnego $A > 0$, gdzie jako pierwsze przybliżenie można wybrać dowolną liczbę dodatnią.

Prostota, pomysłowość, piękno i użyteczność w jednym, uzyskana w wyniku pracy wielu pokoleń! To wszystko dzięki temu, że matematyka przemieszcza się z pokolenia na pokolenie, z miejsca na miejsce i rozwija, jest „żywym organizmem".

Rozwiązanie zadania M 1745.
Dzieląc obustronnie równość z treści zadania przez $a - b \neq 0$, otrzymamy

$$(a-b)^3 = a^2 + ab + b^2.$$

Niech $k = a - b$, wówczas

$$k^3 = k^2 + 3ab = k^2 + 3bk + 3b^2.$$

Równanie kwadratowe

$$3X^2 + 3kX + k^2 - k^3 = 0$$

ma współczynniki całkowite oraz pierwiastek całkowity b, zatem wyróżnik tego równania $\Delta = k^2(12k - 3)$ jest kwadratem liczby całkowitej. Stąd $12k - 3 = \ell^2$ dla pewnej liczby całkowitej ℓ. Rozpatrując uzyskane równanie modulo 6, łatwo zauważymy, że $\ell = 6m - 3$ dla pewnej liczby całkowitej m. Wobec tego

$$k = \frac{1}{12}(l^2 + 3) = 3m^2 - 3m + 1$$

i kolejno

$$b = \frac{-3k + \sqrt{\Delta}}{6} = \frac{-3k + kl}{6} = (m-1)(3m^2 - 3m + 1),$$

$$a = b + k = m(3m^2 - 3m + 1)$$

oraz $9a - 1 = (3m - 1)^3$.

Pitagoras wywodzi swoją matematykę z Jonii, gdzie jeszcze 100 lat przed nim Tales odkrywał pierwsze twierdzenia i dostrzegał wielką wartość dowodzenia ich prawdziwości. Obaj, zanim zaczną zajmować się matematyką, odbywają podróże do Egiptu i Mezopotamii. Od ostatniego Pitagorejczyka Archytasa z Tarentu uczy się jej Eudoksos oraz Platon, i w ten sposób matematyka trafia do Aten. Stamtąd wędruje do Aleksandrii. Tam pisze swoje *Elementy* Euklides, tam studiują Archimedes i Apoloniusz, a w późniejszych wiekach działają Heron, Ptolemeusz, Diofantos, Pappus. *Pax Romana*, powstanie chrześcijaństwa i islamu zmieniają świat. Justynian edyktem z 529 roku likwiduje Akademię Platona, a matematyków umieszcza wśród złoczyńców i innych podobnych. Matematyka przenosi się do Indii (Arjabhata, Brahmagupta, Bhaskara), a później z powrotem do Mezopotamii i Azji Środkowej (Muhammad al-Chwarizmi, Omar Chajjam), i koło się zamyka. Spuściznę arabską przejmuje nauka europejska, matematyka ponownie rozkwita w Europie. Wspaniały owoc pojawia się w XVII wieku, to rachunek różniczkowy i całkowy.

Wartość $\sqrt{2}$ ma praktyczne zastosowanie! Rozmiary papieru formatu A0, A1, A2,... (normy ISO 216) zostały tak zaprojektowane, żeby po podzieleniu na dwie równe części wzdłuż dłuższego boku uzyskać dwa arkusze o tych samych proporcjach długości do szerokości. Jest to możliwe tylko, jeśli ten stosunek wynosi $\sqrt{2}$. Oczywiście rzeczywiste wymiary są zaokrąglone do pełnych milimetrów.

Nie czytaj, jeśli nie wiesz, czym jest pochodna

*Piotr KRZYŻANOWSKI**

*Wydział Matematyki, Informatyki i Mechaniki, Uniwersytet Warszawski

No dobrze. Mimo że nadal czytasz, na wszelki wypadek przypomnienie: pochodną funkcji $f : \mathbb{R} \to \mathbb{R}$ w punkcie x nazywamy wartość granicy

$$\lim_{h\to 0} \frac{f(x+h)-f(x)}{h}$$

i oznaczamy $f'(x)$. Możemy wspomóc się komputerem, by ją obliczyć w sposób przybliżony: przecież dla bardzo małego h wartość ilorazu

$$f'_h(x) := \frac{f(x+h)-f(x)}{h} \tag{1}$$

powinna być bardzo bliska wartości granicznej. Wykorzystanie przybliżenia (1) może być w praktycznych zastosowaniach jedyną rozsądną opcją numerycznego wyznaczenia wartości $f'(x)$: często bowiem się zdarza, że nasza funkcja f nie jest zadana jawnym wzorem, tylko przez „czarną skrzynkę" – tzn. program komputerowy o *nieznanej* treści, który dla zadanego x po prostu zwraca $f(x)$ (i nic więcej).

Intuicja podpowiada, że im mniejsze będzie h, tym lepsza powinna być ta aproksymacja – i rzeczywiście, korzystając ze wzoru Taylora, można łatwo pokazać, że

$$|f'_h(x) - f'(x)| \approx C \cdot |h|, \tag{2}$$

gdzie $C = |f''(x)|/2$.

Zgoda. Tytuł artykułu powinien brzmieć: *Nie czytaj, jeśli nie wiesz, co to jest wzór Taylora.*

Zróbmy więc szybki eksperyment i, korzystając z (1), obliczmy na komputerze przybliżenia wartości pochodnej, powiedzmy w $x = 1$, następujących funkcji:

$$F(x) = x^2, \qquad G(x) = x^3, \qquad S(x) = \sin x.$$

Oto wyniki uzyskane dla kilku całkiem małych wartości h:

h	$F'_h(1)$	$G'_h(1)$	$S'_h(1)$
10^{-20}	0,0000000000	0,0000000000	0,0000000000
10^{-30}	0,0000000000	0,0000000000	0,0000000000
10^{-40}	0,0000000000	0,0000000000	0,0000000000

Rachunki przeprowadziliśmy w GNU Octave, ale to samo dostaniesz, prowadząc obliczenia w podwójnej precyzji w dowolnym języku (C/C++, Pythonie, MATLAB-ie, Julii itd.).

Jak się okazuje, otrzymamy identyczny wynik, nawet jeśli *jeszcze bardziej* zmniejszymy h... Wszystko wskazuje więc na to, że dla każdej z trzech funkcji „graniczną" wartością ilorazu (1) jest zero – tyle że to oczywista nieprawda! Powinniśmy byli przecież otrzymać:

funkcja	jej pochodna	wartość pochodnej w $x = 1$
$F(x) = x^2$	$F'(x) = 2x$	2
$G(x) = x^3$	$G'(x) = 3x^2$	3
$S(x) = \sin x$	$S'(x) = \cos x$	$\approx 0{,}5403023058681397$

Czytelnik Dotychczas Nieznający Pochodnej (a mimo to śledzący nasz wywód) łatwo sprawdzi, że dla $h \to 0$

$$\frac{F(x+h)-F(x)}{h} = 2x + h \to 2x,$$

zatem $F'(x) = 2x$ i podobnie $G'(x) = 3x^2$. Ponadto ze wzoru

$$\sin(x+h) = \sin x \cos h + \cos x \sin h$$

wywnioskuje, że $S'(x) = \cos x$, gdyż $(\sin h)/h \to 1$ i $(\cos h - 1)/h \to 0$ dla $h \to 0$.

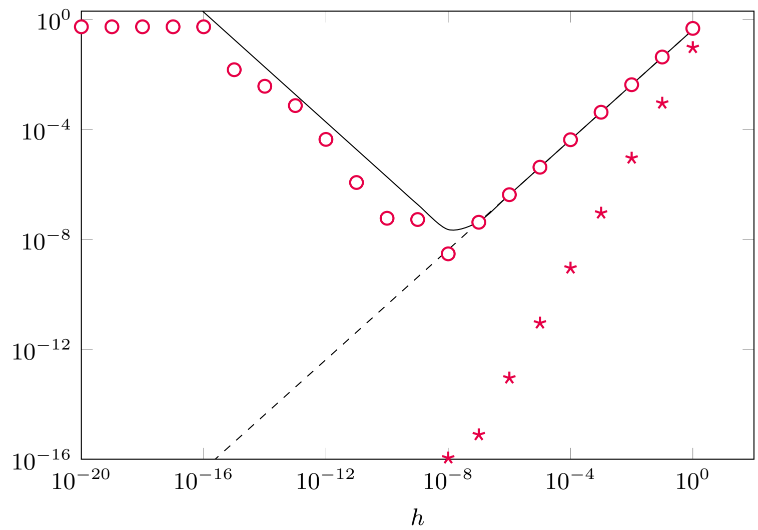
Co się stało? Przyczyna tkwi głęboko we wnętrznościach komputera. Namiastka liczb rzeczywistych, z której korzysta procesor – tzw. liczby maszynowe – to zbiór *skończony*. W przedziale $[1, 2)$ jest ich dokładnie 2^{52}, a najmniejszą liczbą maszynową większą od 1 jest $1 + 2^{-52} \approx 1 + 2{,}2 \cdot 10^{-16}$.

Mówimy tu o tzw. maszynowych liczbach zmiennopozycyjnych podwójnej precyzji, domyślnie wykorzystywanych przez różne pakiety obliczeniowe.

Wobec tego dla dowolnego $0 < h < 2^{-53} \approx 10^{-16}$ obliczona w komputerze wartość $1 + h < 1 + 2^{-53}$ zostanie zaokrąglona do najbliżej leżącej liczby maszynowej, czyli... do 1. W efekcie, wyznaczając dla dostatecznie małego h różnicę $f(1 + h) - f(1)$, komputer w rzeczywistości obliczał $f(1) - f(1)$, co oczywiście *musiało* dać w wyniku *zero*.

Czy więc przybliżenie $f'(x)$ przez $f'_h(x)$ jest bezwartościowe? Nie przesadzajmy – metoda (1) nie jest wcale taka zła – chociaż, jak przekonaliśmy się na początku, ma *pewne ograniczenia*. Jak widać z tabelki na następnej stronie, dla każdej z funkcji F, G, S, gdy h maleje (ale nie za bardzo), obliczane ilorazy stabilizują się do pewnego momentu na poziomie prawdziwej wartości pochodnej.

h	$F'_h(1)$	$G'_h(1)$	$S'_h(1)$
10^{-5}	2,0000100000	3,0000300001	0,5402980985
10^{-6}	2,0000009999	3,0000029998	0,5403018851
10^{-7}	2,0000001011	3,0000003015	0,5403022640
10^{-8}	1,9999999878	3,0000000040	0,5403023029
10^{-9}	2,0000001655	3,0000002482	0,5403023584
10^{-10}	2,0000001655	3,0000002482	0,5403022474
10^{-11}	2,0000001655	3,0000002482	0,5403011372
10^{-12}	2,0001778012	3,0002667017	0,5403455461



Dla F i G zależność ta wyglądałaby bardzo podobnie.

Omawiany eksperyment często pokazujemy studentom MIM UW jako przykład zjawiska tzw. *redukcji cyfr przy odejmowaniu.*

*Właściwie słowo *wina* powinniśmy wziąć w cudzysłów, bo procesor specjalnie zaprojektowano, żeby *właśnie tak* działał: nieco niedokładnie, ale za to piekielnie szybko.

Jeszcze lepiej zobaczymy to na wykresie pokazującym, jak od h zależy błąd $|S'(1) - S'_h(1)|$ (na rysunku powyżej zaznaczyliśmy go kolorowymi kółeczkami). Dodatkowo czarna linia przerywana obrazuje przybliżony poziom błędu przewidywany przez (2). Wygląda więc na to, że przez jakiś czas – gdy h jest w miarę duże – aproksymacja pochodnej przez (1) rzeczywiście *słucha się* matematyki. Jednak potem, gdzieś tak dla $h \approx 10^{-8}$, zaczyna się dziać coś dziwnego, powodującego, że wbrew (2) błąd aproksymacji przestaje maleć – a nawet zaczyna wyraziście rosnąć! Jak można się domyślić, wina* znów leży po stronie komputera – a dokładniej tego, jak wykonywane są w nim działania matematyczne.

Jak więc wybierać h? Dla ustalenia uwagi przyjmijmy, że $h > 0$, i przyjrzyjmy się najprostszemu przypadkowi, czyli $F(x) = x^2$. W związku z tym, że liczb maszynowych jest skończenie wiele, wynik każdego działania arytmetycznego podlega potencjalnemu zaokrągleniu (do najbliższej liczby maszynowej). Dla uproszczenia przyjmiemy, że tak się zdarzy tylko raz: wyłącznie w przypadku podnoszenia do kwadratu wartości $(x + h)$. Z tego, jak została zaprojektowana komputerowa arytmetyka, wynika, że zamiast dokładnej wartości $z = (x + h)^2$ komputer obliczy $\tilde{z} = z(1 + \eta)$, przy czym $|\eta| \leqslant 2^{-53} =: \nu$. Wobec tego na koniec dostaniemy z komputera wcale nie $F'_h(x)$, tylko raczej

$$\widetilde{F'_h(x)} = \frac{(x + h)^2(1 + \eta) - x^2}{h} = F'_h(x) + F(x + h)\frac{\eta}{h},$$

skąd błąd między faktycznie obliczoną na komputerze wartością $\widetilde{F'_h(x)}$ a prawdziwą $F'(x)$ da się z grubsza oszacować, na mocy (2), przez

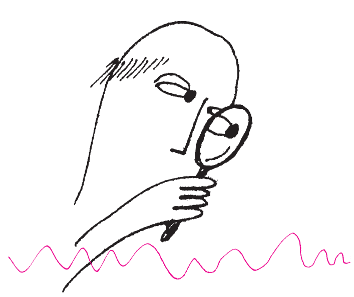
Tak, daje się to zrobić ściśle, ale nam tu zależy tylko na odpowiedzi o charakterze jakościowym – wszak już na początku poczyniliśmy pewne uproszczenia.

$$|F'(x) - \widetilde{F'_h(x)}| \lesssim \frac{|F''(x)|}{2}h + \frac{\nu|F(x)|}{h}. \tag{3}$$

Dla ustalonego x wyrażenie po prawej stronie nierówności (3) jest funkcją h, której wykres zaznaczyliśmy linią ciągłą na rysunku powyżej. Jak widać, całkiem trafnie przewiduje realistyczne zachowanie się błędu, a jej minimum wypada w punkcie

$$h = \sqrt{\frac{2|F(x)|}{|F''(x))|} \cdot \nu}$$

– co potwierdza, że optymalna wartość h powinna być rzędu $\sqrt{\nu} \approx 10^{-8}$, zgodnie z tym, co widzieliśmy na wykresie powyżej. Podstawiając do (3), błąd aproksymacji $F'(x)$ przez $F'_{\sqrt{\nu}}(x)$ powinien wówczas być również rzędu $\sqrt{\nu}$.



Czy można lepiej numerycznie przybliżać pochodną? Owszem, można – i to na kilka sposobów – ale tutaj wspomnimy tylko o jednym, za to zaskakującym (przynajmniej na pierwszy rzut oka). Gdyby bowiem szczęśliwie zdarzyło się, że funkcja f daje się rozszerzyć do funkcji analitycznej w dziedzinie *zespolonej* – a przypadkiem nasze funkcje F, G, S właśnie takie są – to można skorzystać z przybliżenia

W. Squire i G. Trapp, *Using Complex Variables to Estimate Derivatives of Real Functions*, SIAM Review 40 (1), 1998.

$$(\star) \qquad f'(x) \approx \partial_h f(x) := \frac{1}{h}\Im\left(f(x+ih)\right),$$

gdzie $\Im(z)$ oznacza część urojoną liczby zespolonej z, natomiast $i = \sqrt{-1}$ to jednostka urojona. To przybliżenie *nie psuje się* nawet przy bardzo małych h, a dodatkowo błąd maleje w *znacznie szybszym* tempie, co pokazuje poniższa tabela.

Błędy przybliżenia $S'(1)$ przez $\partial_h S(1)$, zaznaczone kolorowymi gwiazdkami, można też zobaczyć na wykresie na poprzedniej stronie.

h	$\partial_h F(1)$	$\partial_h G(1)$	$\partial_h S(1)$
10^{-1}	2,0000000000000000	2,9900000000000000	0,5412032600703925
10^{-2}	2,0000000000000000	2,9999000000000000	0,5403113109515962
10^{-4}	2,0000000000000000	2,9999999900000000	0,5403023067686435
10^{-8}	2,0000000000000000	3,0000000000000000	0,5403023058681399
10^{-30}	2,0000000000000000	3,0000000000000000	0,5403023058681398
10^{-100}	2,0000000000000000	3,0000000000000000	0,5403023058681398
10^{-300}	2,0000000000000000	3,0000000000000000	0,5403023058681398
10^{-310}	2,0000000000000000	3,0000000000000000	0,5403023058681550

Ostatni wiersz tabeli powinien zwrócić Twoją uwagę. Dla *patologicznie* małych h, poniżej około 10^{-308}, coś jednak znowu... zaczyna się psuć! Czy wiesz, dlaczego?

Zadania

W zeszłym miesiącu umieściliśmy w tym miejscu matematyczne „zadania z myszką", które ukazały się w *Delcie* również rok wcześniej. Chcielibyśmy uznać to za nieśmieszny primaaprilisowy żart, jednak niestety była to po prostu nasza pomyłka. Opiekuna działu, dra Dominika Burka, oraz wszystkich naszych Czytelników najmocniej za nią przepraszamy.

Redakcja

Przygotował Dominik BUREK

M 1744. Rozstrzygnąć, czy istnieje funkcja $f:\mathbb{R}\to\mathbb{R}$ taka, że dla dowolnych liczb rzeczywistych x, y ($x \geqslant y$) spełniona jest nierówność:

$$f(x) - f(y) \geqslant \sqrt{x-y}.$$

Rozwiązanie na str. 2

M 1745. Dane są liczby całkowite dodatnie $a \neq b$ takie, że

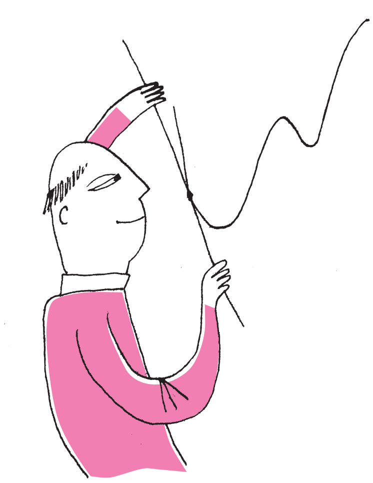
$$(a-b)^4 = a^3 - b^3.$$

Udowodnić, że liczba $9a-1$ jest sześcianem liczby całkowitej.
Rozwiązanie na str. 3

M 1746. Dane są trzy szkoły, w każdej z nich uczy się 200 uczniów. Każdy uczeń ma przynajmniej jednego znajomego w każdej ze szkół (znajomość jest wzajemna oraz nikt nie jest znajomym samego siebie). Przypuśćmy, że istnieje zbiór A, zawierający 300 uczniów, o następującej własności: dla dowolnej szkoły S oraz dwóch uczniów x, y ze zbioru A, którzy nie uczą się w S, x i y mają różną liczbę znajomych uczących się w szkole S. Udowodnić, że istnieje trzech uczniów z trzech różnych szkół, którzy wzajemnie się znają.
Rozwiązanie na str. 17

Przygotował Andrzej MAJHOFER

F 1071. Podczas jazdy z prędkością v na samochód działa siła oporu powietrza opisana wzorem: $F_{op} = \frac{1}{2}C\rho S v^2$, w którym ρ jest gęstością powietrza, S powierzchnią przekroju samochodu prostopadłą do prędkości, a C jest współczynnikiem związanym z kształtem samochodu. Na jakim odcinku drogi l energia W potrzebna do pokonania oporu powietrza podczas jazdy ze stałą prędkością równa jest energii kinetycznej E_k samochodu? W obliczeniach przyjmij $\rho = 1{,}3$ kg/m^3 oraz wartości typowe dla samochodu osobowego: $m = 1{,}3\cdot 10^3$ kg, $S = 2$ m^2, $C = 0{,}3$.
Rozwiązanie na str. 14



F 1072. Ciężar Q wskazywany przez wagę na powierzchni Ziemi jest pomniejszony o siłę wyporu powietrza. Gdyby nie było atmosfery, to na jaką wysokość h nad powierzchnią Ziemi musielibyśmy się wznieść, żeby nasz ciężar (siła przyciągania przez Ziemię) był równy Q? W obliczeniach przyjmij: gęstość powietrza na powierzchni Ziemi $\rho = 1{,}3$ kg/m^3, promień Ziemi $R = 6400$ km, gęstość wody $\rho_0 = 10^3$ kg/m^3.
Rozwiązanie na str. 16