Praktycznie o problemie niezdolności do studiowania

Marek KORDOS, Warszawa

Wielu spośród ludzi, których opinie szczególnie cenię, wyraża pogląd, że współczesni studenci nie są zdolni do studiowania, czego dowodem jest fakt, że nijak się ich do tego skłonić nie da. Można ich skłonić aby się uczyli – uczą się tego, co jest im przekazywane na wykładach, ćwiczeniach, seminariach itd. Ale do studiowania nakłonić się nie dają. Przynajmniej w masie. Jako przyczyna podawany bywa najczęściej brak wprawy, a więc strukturalny analfabetyzm, czyli brak nawyku (a w konsekwencji i umiejętności) do posiedzenia nad książką. Uważam tę diagnozę za prawdziwą.

Skoro jednak tak, to powstaje pytanie, czy przypadkiem nie ma jakiejś praprzyczyny. I wydaje się że jest. Dostępna matematyczna literatura dla nastolatka to przeważnie teksty refleksyjne, których autorzy zapoznają nas z myślami, uczuciami i doznaniami, jakie towarzyszą ich kontaktom z matematyką, opisują dusze, bądź też poczucie humoru matematyków. A samej matematyki jakby nie ma. To są teksty, których poznanie pozwoli niematematykowi podczas spotkań towarzyskich na subtelne sugestie, iż i matematyczna materia nie jest mu obca. Z całą pewnością nie ma to jednak nic wspólnego z propedeutyką studiowania. Godnym odnotowania wyjątkiem są *Okruchy matematyki* Górnickiego – tekst jest jednak dostępny dla tych, którzy studiowania się nie boją, a nawet mają w jego kierunku inklinacje.

Zapewne będzie miała miejsce reforma edukacji, która – w najbardziej minimalistycznej wersji – spowoduje reorganizację szkoły. Podstawówka będzie o rok dłuższa i podzielona – nie jak dotąd na dwie części (nauczanie początkowe i reszta) – lecz na trzy (j.w. i gimnazjum). Da to w konsekwencji nieco starszą młodzież w (krótszych) liceach. Być może będzie też ona intelektualnie sprawniejsza. Można by, jak sądzę, wykorzystać ten moment do tego, aby w liceach uczono nie z podręczników, lecz z książek o innej organizacji tekstu.

Można by tę propozycję teoretycznie rozwijać, ale lepiej chyba zaproponować przykład. To, co niżej, jest próbką napisania takiego tekstu. Nie pisałem go na żadne zamówienie, więc doskonale nadaje się do krytycznej oceny. Stanowi fragment większej o rząd wielkości całości, której poszczególne części są na zupełnie różne tematy. Będę bardzo wdzięczny za wszelkie uwagi, w zamian służę wyjaśnieniami.

Niezależnie od przykładu chciałbym zachęcić wszystkich do pisania tekstów mogących proponować propedeutykę studiowania, tym bardziej, że znane mi są przykłady autorskiego nauczania poprzez takie teksty (przeważnie pisane przez nauczyciela). Pisanie takich tekstów nie jest kopalnią złota (jak np. podręcznik do podstawówki) i nie wiem, czy kiedykolwiek będzie. Ale nie samym złotem człowiek żyje.

6. Do czego może służyć dzielenie z resztą

Każdy z nas, gdy po raz pierwszy w swojej edukacji stykał się z dzieleniem, miał do czynienia z dzieleniem z resztą. I większość z nas po ukończeniu szkoły podstawowej już więcej z takim dzieleniem do czynienia nie miała. Tymczasem, takie dzielenie jest źródłem kilku ważnych konstrukcji matematycznych. Wyliczę tu trzy z nich.

Odmienne liczby

Dzielenie z resztą liczb naturalnych a przez b polega na znalezieniu takiej liczby naturalnej n, że $nb \leq a < (n+1)b$ – i to jest wynik dzielenia, a także takiej liczby naturalnej r, że $a = nb + r$ – i to jest reszta z dzielenia. Jak łatwo zauważyć, r jest jedną z liczb $0,\ 1, \ldots,\ (b-1)$. Tak to wygląda w podstawówce.

Nietrudno zauważyć, że można dzielić również liczby całkowite ujemne. Dla wszystkich liczb całkowitych używa się pojęcia *kongruencji o module m*. Mówi się, że liczby a i b *są w kongruencji modulo m*, lub że *przystają modulo m* – symbolicznie $a \equiv b \pmod m$ – gdy $a - b$ dzieli się przez m z resztą 0, lub (co na jedno wychodzi) gdy a i b dzielą się przez m z tą samą (obojętnie już jaką) resztą.

Relacja kongruencji jest bardzo podobna do zwyczajnej równości, np. kongruencje można stronami mnożyć, dzielić, potęgować, dodawać, odejmować, uzyskując za każdym razem znowu liczby będące w kongruencji. Nie jest to trudno sprawdzić. Dla porządku sprawdźmy, powiedzmy, dodawanie i mnożenie.

Jeśli $a \equiv b \pmod m$ i $c \equiv d \pmod m$, to istnieją takie liczby n_a, n_b, n_c, n_d oraz r_1 i r_2, że $a = n_a m + r_1$, $b = n_b m + r_1$, $c = n_c m + r_2$ i $d = n_d m + r_2$. Zatem

$$a + c = n_a m + r_1 + n_c m + r_2 = m(n_a + n_c) + (r_1 + r_2)$$
$$\text{i} \quad b + d = n_b m + r_1 + n_d m + r_2 = m(n_b + n_d) + (r_1 + r_2),$$

z czego wynika, że zarówno $a + c$, jak i $b + d$ przy dzieleniu przez m dają taką samą resztę jak $r_1 + r_2$, czyli równą, a zatem $a + c \equiv b + d \pmod{m}$.

Analogicznie dla mnożenia mamy

$$ac = n_a n_b m^2 + n_a m r_2 + n_c m r_1 + r_1 r_2 =$$
$$= m(n_a n_c m + n_a r_2 + n_c r_1) + r_1 r_2$$
$$\text{i podobnie} \quad bd = m(n_b n_d m + n_b r_2 + n_d r_1) + r_1 r_2,$$

a więc ac i bd dają z dzielenia przez m taką samą resztę, jak $r_1 r_2$, czyli mamy $ac \equiv bd \pmod{m}$. Podobnie sprawdza się pozostałe wymienione własności kongruencji.

Dla zwolenników konkretnych przykładów mam propozycję prześledzenia, jak z pierwszej kongruencji wynikają wszystkie wypisane dalej:
$16 \equiv 9 \pmod{7}$, bo $16 - 9 = 7$;
$13 \equiv 6 \pmod{7}$, bo odjęliśmy $3 \equiv 3 \pmod{7}$;
$-39 \equiv 24 \pmod{7}$, bo pomnożyliśmy przez $-3 \equiv 4 \pmod{7}$;
$1521 \equiv 576 \pmod{7}$, bo podnieśliśmy do kwadratu;
$946 \equiv 1 \pmod{7}$, bo odjęliśmy $575 \equiv 575 \pmod{7}$.
Oczywiście, każdą z kongruencji można by sprawdzić bezpośrednio.

Kolejną – bardzo oczywistą, ale ogromnie ważną własność kongruencji poprzedzi

Dygresja o klasach abstrakcji

Równość ma trzy interesujące własności

$$a = a, \quad \text{gdy } a = b, \text{ to } b = a, \quad \text{gdy } a = b \text{ i } b = c, \text{ to } a = c,$$

zwane odpowiednio: *zwrotność, symetria, przechodniość*. Każdą relację o takich własnościach nazywa się *relacją równoważności* lub po prostu *równoważnością*. Własności te pozwalają na podzielenie zbioru, w którym jest określona relacja, na części, zwane dumnie *klasami równoważności*, lub *klasami abstrakcji*.

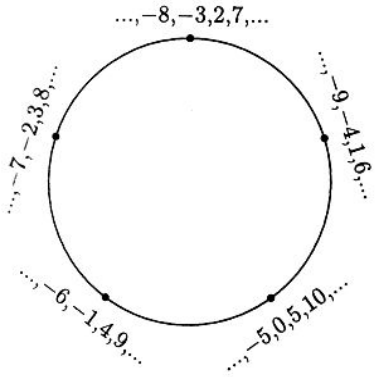
Ta ostatnia nazwa pokazuje istotną cechę myślenia matematycznego – każde pojęcie, nawet tak zdawałoby się nieokreślone, jak abstrakcja, matematyk musi dokładnie opisać, jednoznacznie zdefiniować. Tak więc abstrakcja to tworzenie klas jakichś obiektów za pomocą mającej wymienione własności relacji.

Klasa abstrakcji wyznaczona przez element a jakiegoś zbioru, w którym określona jest pewna zwrotna, symetryczna i przechodnia relacja R, to zbiór tych wszystkich elementów, które są z a w relacji R. Większość pojęć matematyki (co nie zawsze widać) to klasy abstrakcji. Choćby najpospolitsze nazwy figur: koło, prostokąt, kwadrat to klasy abstrakcji relacji podobieństwa w zbiorze figur. Faktycznie, gdy mówimy „koło" czy „kwadrat", nie mamy na myśli konkretnej figury, tylko dowolną z pewnego zbioru, który zawiera istotnie tylko obiekty wzajemnie podobne. Jeśli się zastanowić, to właściwie prawie wszystkie nazwy nie są nazwami konkretnych obiektów, lecz dowolnego obiektu z pewnej zbiorowości: „okno", „pies", „człowiek", „miłość" to abstrakty, nazwy klas. Tyle że jedynie w matematyce zrobiono z tej obserwacji porządną teorię. Na przykład posługując się nawet taką gadaną definicją klasy abstrakcji, jaka rozpoczyna ten akapit, można udowodnić, że klasy abstrakcji są rozłączne – gdy o jakimś punkcie stwierdzimy, że należy do dwóch klas abstrakcji tej samej relacji, oznaczać to będzie, że te klasy są równe, że to jest jedna klasa. Proszę udowodnić.

Często używa się także nazw klas abstrakcji, które się różnią od nazw obiektów składających się na taką klasę. Na przykład *kierunek*: jest to nazwa klasy abstrakcji prostych (odcinków, wektorów) ze względu na relację równoległości. A ta relacja, oczywiście, ma potrzebne własności:

$$k \parallel k; \quad \text{gdy } k \parallel l, \text{ to } l \parallel k; \quad \text{gdy } k \parallel l \text{ i } l \parallel m, \text{ to } k \parallel m;$$

jest więc zwrotna, symetryczna i przechodnia. Mówiąc, że prosta ma jakiś kierunek, mówimy, że należy do odpowiedniej klasy relacji równoległości.



Rys. 6.1

Jeszcze częściej jest stosowane pojęcie wektora swobodnego: jest to klasa abstrakcji relacji posiadania tego samego kierunku, zwrotu i długości. O ile zwykły wektor (czasem nazywany związanym) to para punktów, o tyle wektor swobodny to nic innego, jak przesunięcie – prawda ?

Czasami dobrze jest wyobrazić sobie jakiś mechanizm powstawania klasy abstrakcji. Dla klas wyznaczonych wśród liczb całkowitych przez relację kongruencji modulo m dogodnie jest wyobrazić sobie, że jest to wynik nawijania osi liczbowej na okrąg o obwodzie m – wszystkie liczby całkowite znajdą się tylko w jego m różnych punktach – wszystkie, które trafiają w ten sam punkt, należą do tej samej klasy abstrakcji (rys. 6.1 ilustruje przypadek kongruencji modulo 5). Oczywiście, najpierw należałoby się przekonać, że kongruencja jest relacją równoważności – ale to chyba nikomu trudności nie sprawi.

Dostrzeżone przez nas własności działań dadzą się w tym języku opisać tak, że jeśli którąkolwiek z liczb danej klasy dodamy do którejkolwiek z liczb innej danej klasy, to otrzymamy wynik należący stale do tej samej klasy. Innymi słowy, nawijanie odbywa się w tak regularny sposób, że wyniki zależą jedynie od klas, z których zostały wzięte argumenty. Działania są zgodne z nawijaniem. Albo inaczej – można określić w ten sposób działania na klasach: wykonujemy zwykłe działania na zwykłych liczbach, na zwykłej osi liczbowej, a potem oś nawijamy na odpowiedni okrąg i wynik mamy gotowy.

Napisałem „na zwykłych liczbach", bo w ten sposób otrzymaliśmy zupełnie nowe liczby – właśnie te klasy. Często klasy abstrakcji relacji kongruencji modulo m nazywa się tą liczbą spośród $\{0, 1, 2, \ldots, (m-1)\}$, która należy do tej właśnie klasy. Trzeba jednak pamiętać, że są to zupełnie inne liczby niż tak samo się nazywające liczby naturalne – przecież zupełnie inaczej się na nich rachuje. Zbiór nowych liczb, czyli klas abstrakcji kongruencji modulo m oznaczany jest na ogół przez $\mathbf{Z}_m$. Zobaczymy, że własności liczb $\mathbf{Z}_m$ zależą od własności m.

Rachunki w $\mathbf{Z}_m$ można by, oczywiście, opisać nie używając nawijania. Ot, po prostu, umawiamy się, że liczby Z_m są to liczby $\{0, 1, 2, \ldots, (m-1)\}$, tylko, że liczy się na nich w ten sposób, iż zamiast wyniku działania pisze się jego resztę z dzielenia zwykłej liczby przez m. Każdy może patrzeć na to tak, aby jak najwięcej zobaczyć.

Aby się przekonać, że są $\mathbf{Z}_m$ lepsze i gorsze, spróbujmy w $\mathbf{Z}_6$ pomnożyć 2 przez 3 – wyjdzie 0. Taka sytuacja jest zdecydowanie niedobra – otrzymaliśmy *dzielniki zera*, różne od zera liczby, których iloczyn jest zerem. To wyklucza możliwość, aby dało się w $\mathbf{Z}_6$ określić przyzwoite dzielenie – ani 2, ani 3 nie mają tu swoich odwrotności. Stąd nie ma co marzyć np. o normalnym rozwiązywaniu równań. Zauważmy jednak, że gdy m jest liczbą pierwszą, żadnych kłopotów nie będzie. Faktycznie, dla dowolnej liczby różnej od zera jest liczba odwrotna, co zapewnia możliwość dzielenia (bo to przecież mnożenie przez odwrotność). Oto uzasadnienie. Mnożąc liczbę a, różną od 0, kolejno przez wszystkie liczby $\mathbf{Z}_m$, gdy m jest liczbą pierwszą, otrzymujemy m różnych liczb, a więc wśród nich i 1. Gdybyśmy bowiem uzyskali mniej wyników, to dla pewnych dwóch liczb b i c byłoby $ab = ac$, czyli $a(b-c) = 0$, a więc byłyby dzielniki zera, czyli – wracając do zwykłych liczb – byłoby $a(b-c) = k \cdot m$, podczas gdy zarówno a, jak i $b-c$ i $c-b$ są od m mniejsze: sprzeczność z założeniem, że m jest liczbą pierwszą.

$\mathbf{Z}_p$, gdzie p jest liczbą pierwszą (zwyczajowo używa się takiej litery) to liczby mające podstawowe własności algebraiczne takie same, jak np. liczby rzeczywiste czy wymierne. Dla przykładu, wzory na rozwiązywanie równań kwadratowych są w $\mathbf{Z}_p$ takie same, jak dla liczb rzeczywistych – działa wzór

$$x = \frac{-b \pm \sqrt{\Delta}}{2a}, \quad \text{gdzie} \quad \Delta = b^2 - 4ac,$$

opisujący pierwiastki równania $ax^2 + bx + c = 0$, choć wyniki liczbowe są

zupełnie inne. Zobaczmy to na przykładzie równania

$$x^2 + x + 1 = 0,$$

które, jak wiadomo, nie ma pierwiastków rzeczywistych, a to dlatego, że

$$\Delta = 1 - 4 = -3$$

nie jest wśród liczb rzeczywistych kwadratem żadnej liczby. W $\mathbf{Z}_3$ jednak ta sama Δ jest równa 0 i, rzeczywiście, równanie ma jeden pierwiastek

$$\frac{-1 \pm 0}{2 \cdot 1} = \frac{2}{2} = 1.$$

W $\mathbf{Z}_5$ znowu pierwiastków nie ma – też się zgadza, bo jedyne kwadraty wśród tych liczb to 0, 1 i 4, a $\Delta = 2$.

W $\mathbf{Z}_7$ z kolei są dwa pierwiastki, $\Delta = 4 = 2^2 = 5^2$ i pierwiastkami są

$$\frac{-1+2}{2 \cdot 1} = \frac{1}{2} = 4 \quad \text{oraz} \quad \frac{-1+5}{2 \cdot 1} = \frac{4}{2} = 2$$

(czterech pierwiastków nie ma, bo przecież $-2 = 5$).

Kongruencje są wykorzystywane nie tylko do konstrukcji nowych liczb, lecz przede wszystkim do rozstrzygania najrozmaitszych problemów związanych z podzielnością. Tu nie będzie o tym mowy, ale dla zachęty polecam wypróbowanie swych sił przy wykazaniu twierdzenia, które nosi nazwę *Małe Twierdzenie Fermata*. Można je sformułować tak:
Jeśli p jest liczbą pierwszą, to dla każdej liczby całkowitej a liczba $a^p - a$ jest podzielna przez p
(jak kto woli: a^p przystaje do a modulo p).

Dowód tego twierdzenia najłatwiej przeprowadzić indukcyjnie, a konkretnie – ponieważ $1^p - 1 = 0$ jest podzielne przez p – wykazać, że jeśli $a^p - a$ jest podzielne przez p, to również $(a+1)^p - (a+1)$ dzieli się przez p. To, co stanowi główny argument dowodu, to fakt, że wszystkie współczynniki wielomianu $(x+1)^p$ – poza pierwszym i ostatnim, które są jedynkami – dzielą się przez p (gdy p jest liczbą pierwszą, oczywiście).

Na koniec wypada jeszcze napisać, że najczęściej jako Małe Twierdzenie Fermata podaje się wniosek z przytoczonego twierdzenia, a mianowicie
Jeśli p jest liczbą pierwszą i liczba a nie jest podzielna przez p, to liczba $a^{p-1} - 1$ jest podzielna przez p.

Teraz przejdziemy do zapowiedzianych jeszcze dwóch kierunków wykorzystania dzielenia z resztą. W każdym z nich użyty jest algorytm Euklidesa. Teraz więc będzie

Dygresja o algorytmie Euklidesa

Algorytm Euklidesa to następujące postępowanie, które stosuje się do dowolonych dwóch liczb, nazwijmy je a i b. Dzielimy z resztą a przez b i otrzymujemy wynik w_1 i resztę r_1. Następnie dzielmy b przez r_1, otrzymując wynik w_2 i resztę r_2. Z kolei r_1 dzielimy przez r_2, otrzymując wynik w_3 i resztę r_3. I tak dalej, co oznacza, że tak długo, dopóki nie otrzymamy reszty 0, powtarzamy tę operację, otrzymując z dzielenia r_{n-1} przez r_n wynik w_{n+1} i resztę r_{n+1}. Gdy nie natrafiamy na resztę 0, traktujemy całą operację jako ciągnącą się w nieskończoność. Tak więc algorytm Euklidesa z pary liczb a i b produkuje skończoną lub nieskończoną liczbę par w_i i r_i. Na marginesie przykłady.

$r = 1517,\ s = 1073$

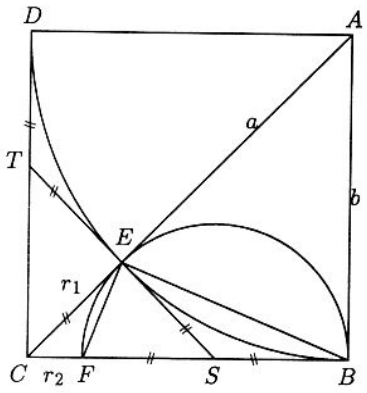
$1517 = 1 \cdot 1073 + 444,$
$1073 = 2 \cdot 444 + 185,$
$444 = 2 \cdot 185 + 74,$
$185 = 2 \cdot 74 + 37,$
$74 = 2 \cdot 37 + 0.$

$c = 771,\ d = 146$

$771 = 5 \cdot 146 + 41,$
$146 = 3 \cdot 41 + 23,$
$41 = 1 \cdot 23 + 18,$
$23 = 1 \cdot 18 + 5,$
$18 = 3 \cdot 5 + 3,$
$5 = 1 \cdot 3 + 2,$
$3 = 1 \cdot 2 + 1,$
$2 = 2 \cdot 1 + 0.$

Pierwsza refleksja, jaka się nasuwa, to spostrzeżenie, że ten algorytm zawsze się musi zakończyć – przecież za każdym razem otrzymujemy, jako resztę, coraz mniejszą liczbę naturalną – od jak wielkiej byśmy nie zaczęli, po skończonej liczbie kroków musimy uzyskać zero.

Spostrzeżenie byłoby prawdziwe, gdybyśmy algorytm Euklidesa stosowali tylko do liczb naturalnych. Możemy go jednak stosować do dowolnych liczb, bo przecież określenie dzielenia z resztą da się do nich również zastosować.



Rys. 6.2. $\frac{BC}{CE} = \frac{EC}{CF}$ – dlaczego ?

Co więcej – to wcale nie muszą być liczby: historycznie algorytm Euklidesa najpierw zastosowano do odcinków.

Spróbujmy więc podać taką definicję dzielenia z resztą, aby stosowała się i do dowolnych liczb, i do odcinków (i może do jeszcze paru wielkości): dzielenie z resztą wielkości a przez tego samego rodzaju wielkość b polega na znalezieniu takiej liczby naturalnej n, że wielkość a zawiera się pomiędzy wielkością nb i wielkością $(n+1)b$ – i to jest wynik dzielenia, a także takiej wielkości r, że $a = nb + r$ – i to jest reszta z dzielenia.

Algorytm Euklidesa będzie więc nam dawał w ogólności nie parę liczb w każdym kroku, lecz parę złożoną z liczby naturalnej w_i i wielkości tego samego rodzaju, co a i b, będącej resztą r_i. Aby obejrzeć choć jeden przykład tego rodzaju, proszę wziąć jako a przekątną kwadratu, a jako b jego bok – odkładając cyrklem odcinki, przekonamy się, że – dopóki to narzędzie nie okaże się zbyt grube – otrzymywać będziemy takie same wyniki w_1, w_2, w_3,..., jak w pierwszym z przykładów liczbowych (rys. 6.2). Można by nawet postarać się o (geometryczne) uzasadnienie, że poczynając od w_2 wszystkie w_i będą równe 2 i że algorytm nigdy się nie zatrzyma.

Sprawą algorytmu Euklidesa traktowanego tak ogólnie zajmiemy się w trzeciej części tego rozdziału (tam też ściągawka dla tych, którzy nie zrobili przykładu geometrycznego).

$$
\begin{aligned}
444 &= 1 \cdot r - 1 \cdot s,\\
185 &= 1 \cdot s - 2 \cdot 444 =\\
&= -2 \cdot r + (1+2) \cdot s =\\
&= -2 \cdot r + 3 \cdot s,\\
74 &= 1 \cdot 444 - 2 \cdot 185 =\\
&= (1+4) \cdot r + (-1-6) \cdot s =\\
&= 5 \cdot r - 7 \cdot s,\\
37 &= 1 \cdot 185 - 2 \cdot 74 =\\
&= (-2-10) \cdot r + (3+14) \cdot s =\\
&= -12 \cdot r + 17 \cdot s.
\end{aligned}
$$

$$
\begin{aligned}
41 &= 1 \cdot c - 5 \cdot d,\\
23 &= 1 \cdot d - 3 \cdot 41 =\\
&= -3 \cdot c + (1+15) \cdot d =\\
&= -3 \cdot c + 16 \cdot d,\\
18 &= 1 \cdot 41 - 1 \cdot 23 =\\
&= (1+3) \cdot c + (-5-16) \cdot d =\\
&= 4 \cdot c - 21 \cdot d,\\
5 &= 1 \cdot 23 - 1 \cdot 18 =\\
&= (-3-4) \cdot c + (16+21) \cdot d =\\
&= -7 \cdot c + 37 \cdot d,\\
3 &= 1 \cdot 18 - 3 \cdot 5 =\\
&= (4+21) \cdot c + (-21-111) \cdot d =\\
&= 25 \cdot c - 132 \cdot d,\\
2 &= 1 \cdot 5 - 1 \cdot 3 =\\
&= (-7-25) \cdot c + (37+132) \cdot d =\\
&= -32 \cdot c + 169 \cdot d,\\
1 &= 1 \cdot 3 - 1 \cdot 2 =\\
&= (25+32) \cdot c + (-132-169) \cdot d =\\
&= 57 \cdot c - 301 \cdot d.
\end{aligned}
$$

Największy wspólny dzielnik

Tymczasem powróćmy do algorytmu Euklidesa wśród liczb naturalnych. Mamy więc pewność, że algorytm się zatrzymuje i istnieje pierwsza reszta równa zeru. Zauważmy, że

jeśli w algorytmie Euklidesa, zastosowanym do liczb naturalnych a i b, pierwszą resztą równą zeru jest r_n, to r_{n-1} jest największym wspólnym dzielnikiem a i b.

Uzasadnienie jest bardzo proste: jeśli r_n jest równa 0, to r_{n-1} dzieli bez reszty r_{n-2}, a ponieważ $r_{n-3} = w_{n-1}r_{n-2} + r_{n-1}$, więc r_{n-1} dzieli również r_{n-3}. Powtarzając to rozumowanie, stwierdzamy, że r_{n-1} dzieli wszystkie r_i o mniejszych numerach oraz a i b. Jest więc ich wspólnym dzielnikiem.

Pozostaje jeszcze tylko przekonać się, że największym. Ale tu uzasadnienie jest bardzo podobne do poprzedniego. Jeśli bowiem jakaś liczba d jest dzielnikiem i a, i b, to wobec $a = w_1 \cdot b + r_1$ jest dzielnikiem również r_1, wobec $b = w_2 \cdot r_1 + r_2$ jest również dzielnikiem r_2 itd., aż w końcu stwierdzamy, że jest dzielnikiem r_{n-1}. Skoro więc każdy wspólny dzielnik a i b jest dzielnikiem r_{n-1}, więc ten dzielnik jest największy.

Metoda poszukiwania największego wspólnego dzielnika za pomocą algorytmu Euklidesa jest na ogół (czyli jeśli przykład nie jest specjalnie dobrany) szybsza od poszukiwania go metodą rozkładu na czynniki – proszę tradycyjnie poszukać największego wspólnego dzielnika dla r i s czy c i d z przykładów podanych wyżej. Ma jednak jeszcze jedną zaletę: jej uzasadnienie pozwala stwierdzić, że *dla dowolnych liczb naturalnych a i b istnieją takie liczby całkowite k i l, że $k \cdot a + l \cdot b$ jest równe największemu wspólnemu dzielnikowi tych liczb.*

Zamiast dowodu (którego przemyślenie polecam) przedstawiam tylko na marginesie sposób obliczenia tych liczb dla podanych wyżej r i s oraz c i d – porównanie go z przebiegiem algorytmu Euklidesa wskazuje drogę dowodu w ogólnym przypadku. Ostatecznie więc

$37 = (-12) \cdot 1517 + 17 \cdot 1073$ $(= -18204 + 18241)$ i $1 = 57 \cdot 771 + (-301) \cdot 146$ $(= 43947 - 43946)$.

W ten sposób okazało się, że każdą całkowitą wielokrotność największego wspólnego dzielnika dwóch liczb naturalnych można uzyskać jako sumę ich iloczynów przez jakieś liczby całkowite.

Szczególnie ciekawie przedstawia się to dla liczb o największym wspólnym dzielniku równym 1, czyli liczb *względnie pierwszych*. W tym przypadku można,

mnożąc każdą z nich przez odpowiednią liczbę całkowitą i dodając, uzyskać dowolną liczbę całkowitą. Na przykład c i d są liczbami względnie pierwszymi – wykorzystując przeprowadzone rachunki, możemy napisać, że dla dowolnej liczby całkowitej k jest

$$k = 57 \cdot k \cdot c - 301 \cdot k \cdot d \ (= 57 \cdot k \cdot 771 - 301 \cdot k \cdot 146).$$

Trudno powiedzieć, aby to było widać „na oko".

Polecam dalsze figle z tak wykorzystywanym algorytmem Euklidesa. Sam powrócę do ostatniego spostrzeżenia w rozdziale 10. Teraz pora na trzeci sposób wykorzystania dzielenia z resztą, a dokładniej – na kolejne wykorzystanie algorytmu Euklidesa.

Arytmetyczne ułamki łańcuchowe

W poprzedniej części interesowały nas bardziej reszty uzyskiwane za pomocą algorytmu Euklidesa niż wyniki poszczególnych dzieleń. Tu będzie przeciwnie.

Zauważmy, że iloraz r i s, wziętych z przeliczanego przykładu, można zapisać jako

$$\frac{1517}{1073} = 1 + \cfrac{1}{2 + \cfrac{1}{2 + \cfrac{1}{2 + \cfrac{1}{2}}}}.$$

Tak zapisany ułamek nazywamy *arytmetycznym ułamkiem łańcuchowym* lub *arytmetycznym ułamkiem ciągłym*. Ogólniejsze pojęcie ułamka łańcuchowego dotyczy wyrażenia różniącego się od arytmetycznego ułamka łańcuchowego tym, że występujace w nim liczby (również jedynki!) nie muszą być liczbami naturalnymi. W przypadku arytmetycznych ułamków łańcuchowych dopuszcza się jeszcze, aby liczba stojąca na pierwszym miejscu była ujemną liczbą całkowitą – robi się tak dlatego, aby również liczby ujemne można było przedstawiać jako arytmetyczne ułamki łańcuchowe.

Ogólnie: jeśli z zastosowania algorytmu Euklidesa do liczb a i b uzyskamy wyniki w_i, to prawdą jest, że

$$\frac{a}{b} = w_1 + \cfrac{1}{w_2 + \cfrac{1}{w_3 + \cfrac{1}{w_4 + \cfrac{1}{w_5 + \ldots}}}}.$$

W przypadku, gdy a i b są liczbami całkowitymi, rzecz jest nietrudno sprawdzić – ułamek taki ma skończoną liczbę wyników w_i i kresek ułamkowych.
W przykładzie, od którego zaczęliśmy, mamy

$$1 + \cfrac{1}{2 + \cfrac{1}{2 + \cfrac{1}{2 + \frac{1}{2}}}} = 1 + \cfrac{1}{2 + \cfrac{1}{2 + \frac{1}{\frac{5}{2}}}} = 1 + \cfrac{1}{2 + \cfrac{1}{2 + \frac{2}{5}}} = 1 + \cfrac{1}{2 + \frac{1}{\frac{12}{5}}} =$$

$$= 1 + \cfrac{1}{2 + \frac{5}{12}} = 1 + \frac{1}{\frac{29}{12}} = 1 + \frac{12}{29} = \frac{41}{29} = \frac{1517}{1073};$$

ostatnia równość bierze się z pomnożenia licznika i mianownika przez 37, znaleziony poprzednio największy wspólny dzielnik 1517 i 1073. Jak widać, korzystając tylko z wyników dzieleń w algorytmie Euklidesa, można znaleźć wartość skróconego ułamka inaczej, niż dzieląc licznik i mianownik przez ich największy wspólny dzielnik. Można się też pokusić o przeprowadzenie ogólnego dowodu zauważonych tu prawidłowości.

Odmiennie ma się sprawa, gdy stosujemy algorytm Euklidesa w ogólnym przypadku. Tu można otrzymać ułamek łańcuchowy nieskończony. Wartość ułamka łańcuchowego (arytmetycznego, ale o innych nie będzie tu mowy, więc

będę to określenie pomijał) wzięta tylko do pewnego miejsca, rozważenie tylko wartości w_i dla i mniejszych od pewnego n, nazywa się n-tym reduktem ułamka. Można wykazać, że te n-te redukty są przybliżeniami wartości ułamka, a więc, że można nimi wartość ułamka przybliżyć z dowolną dokładnością.

Ciekawe, że dowolnie wypisany ciąg liczb naturalnych może być wzięty za ciąg w_i. Oznacza to tyle, że zawsze znajdą się takie dwie liczby, dla których będzie to odpowiadający im ułamek łańcuchowy. Pomińmy dowód tego faktu, ale korzystajmy z tego, że jest prawdziwy.

Aby nie pisać takich ogromnych ułamków, będziemy je zapisywali jako $(w_1; w_2, w_3, \ldots)$. Tak więc ułamek $\frac{1517}{1073}$ (czy, oczywiście, $\frac{41}{29}$) możemy zapisać jako $(1; 2, 2, 2, 2)$, a ułamek $\frac{771}{146}$ jako $(5; 3, 1, 1, 3, 1, 1, 2)$.

Jeżeli ułamek łańcuchowy jest skończony, to przedstawia liczbę wymierną – możemy przecież go zwinąć do postaci ułamka zwykłego. Mając z kolei liczbę wymierną, dodatnią, czyli ułamek o naturalnym liczniku i mianowniku, możemy zastosować do nich algorytm Euklidesa – zatrzyma się on po skończonej liczbie kroków, a więc otrzymamy skończony ułamek łańcuchowy. Mając liczbę wymierną ujemną, możemy potraktować ją jako sumę jej *cechy* (czyli części całkowitej, a więc największej liczby całkowitej, która nie przekracza danej liczby) i nieujemnej *mantysy* (czyli różnicy między liczbą a jej cechą); ułamek łańcuchowy równy mantysie zaczyna się od zera – gdy zamiast niego wpiszemy cechę, dostaniemy rozwinięcie w skończony ułamek łańcuchowy również ujemnej liczby wymiernej.

Wynika z tego, że nieskończone ułamki łańcuchowe odpowiadają liczbom niewymiernym (jest tu więc większy porządek niż np. wśród rozwinięć dziesiętnych). Dawniej zajmowano się tą sprawą znacznie bardziej i powstała nawet terminologia, która dotyczyła nie tylko liczb, ale także odcinków czy innych wielkości (ciężarów, pól itp.). Wielkości, dla których algorytm Euklidesa się nie zatrzymywał, nazywano *niewspółmiernymi*, a te dla których się zatrzymywał *współmiernymi*. Aby wytłumaczyć te nazwy w przypadku odcinków zauważmy, że dla odcinków współmiernych istnieje mniejszy od nich odcinek, który mieści się całkowitą liczbę razy tak w jednym z nich, jak w drugim (taki wspólny dzielnik) – gdyby przyjąć go za jednostkę długości, to oba odcinki miałyby w tych jednostkach długość całkowitą. Liczby wymierne np. to liczby współmierne z 1.

Wśród nieskończonych rozwinięć dziesiętnych te, które od pewnego miejsca są okresowe, odpowiadają liczbom wymiernym (podobnie, jak te, które mają rozwinięcia skończone). Czym charakteryzują się (od pewnego miejsca) okresowe ułamki łańcuchowe, jakim liczbom niewymiernym odpowiadają ? Okazuje się, że są to rozwinięcia *niewymierności kwadratowych*, to znaczy liczb będących pierwiastkami równań drugiego stopnia o współczynnikach całkowitych. Zamiast dowodu będą przykłady, z których można odtworzyć ten dowód. Najpierw jednak otrzymajmy w praktyce choćby jeden ułamek łańcuchowy nieskończony – będzie to

Dygresja o przekątnej kwadratu

Biorąc pod uwagę fakt, że nie wszyscy czytający te słowa skorzystali z propozycji zawartej w zakończeniu dygresji o algorytmie Euklidesa, przeprowadzimy sprawę od początku. Interesuje nas stosunek

$$\frac{AC}{AB},$$

czyli stosunek przekątnej kwadratu do jego boku (oznaczenia z rysunku 6.2). Rysunek powstaje w następujący sposób. W kwadracie $ABCD$ kreślimy okrąg o środku A i promieniu AB. Przecina on przekątną w punkcie E, przez który kreślimy styczną, otrzymując na bokach kwadratu punkty S i T. Kreślimy teraz okrąg o środku S i promieniu SB przecina on BC w punkcie F. To, co trzeba

zauważyć, to $SB = SE = ET = TD = SF = EC$ (rysunek) oraz podobieństwo trójkątów BCE i ECF (bo $\angle CBE = \angle CEF$, jako kąt wpisany i kąt dopisany, kąt C wspólny). Gdy to już wiemy, mamy:

$$\frac{AC}{AB} = 1 + \frac{CE}{CB} = 1 + \frac{1}{\frac{CB}{CA}} = 1 + \frac{1}{2 + \frac{CF}{CE}} =$$

$$= 1 + \frac{1}{2 + \frac{CE}{CB}} = \ldots = 1 + \cfrac{1}{2 + \cfrac{1}{2 + \cfrac{1}{2 + \cfrac{1}{2 + \ldots}}}},$$

czyli tak, jak poprzednio obiecaliśmy, ułamek $u = (1; 2, 2, 2, 2, \ldots)$, co zapisuje się dobitniej $(1; \overline{2})$.

A teraz to samo dla tych, dla którzy nie lubią geometrii, a za to wiedzą, że przekątna kwadratu o boku 1 ma długość $\sqrt{2}$.

$$\sqrt{2} = 1 + (\sqrt{2} - 1) = 1 + \frac{1}{\sqrt{2} + 1} = 1 + \frac{1}{2 + (\sqrt{2} - 1)} =$$

$$= 1 + \frac{1}{2 + \frac{1}{\sqrt{2} + 1}} = \ldots = 1 + \cfrac{1}{2 + \cfrac{1}{2 + \cfrac{1}{2 + \cfrac{1}{2 + \ldots}}}}.$$

I tyle dygresji.

Metodę wykazania, że okresowe ułamki łańcuchowe opisują niewymierności kwadratowe obejrzymy najpierw na przykładzie tegoż ułamka $(1; \overline{2})$. Dla wszystkiego zapiszmy go w zwykłej postaci

$$u = 1 + \cfrac{1}{2 + \cfrac{1}{2 + \cfrac{1}{2 + \cfrac{1}{2 + \ldots}}}}$$

i jego część okresową oznaczmy przez x. Zatem

$$u = 1 + \frac{1}{x} \quad \text{i} \quad x = 2 + \frac{1}{x},$$

w tej drugiej równości skorzystaliśmy z nieskończoności ułamka – prawda ? Obliczmy teraz z niej x i u.

$$x^2 - 2x - 1 = 0, \quad \text{czyli} \quad x = \frac{-(-2) + \sqrt{8}}{2} = 1 + \sqrt{2},$$

gdzie wybraliśmy dodatni pierwiastek równania kwadratowego, bo przecież z dodawania liczb dodatnich nie można otrzymać liczby ujemnej. Dalej mamy

$$u = 1 + \frac{1}{1 + \sqrt{2}} = 1 + \frac{\sqrt{2} - 1}{1} = \sqrt{2},$$

a więc wszystko się zgodziło. Warto tu zwrócić uwagę, że skorzystaliśmy w tym dowodzie z faktu, że ułamek łańcuchowy nieskończony odpowiada jakiejś liczbie – bez tego wprowadzony wyżej symbol x byłby bez sensu. Tu wiedzieliśmy o tym z poprzedzającej dygresji. Jeszcze raz podkreślę, że każdy ciąg liczb naturalnych można potraktować jako ciąg w_i otrzymując zawsze ułamek łańcuchowy o określonej wartości, co jest wielką (choć tu podaną bez dowodu) zaletą ułamków łańcuchowych.

Można iść dalej poprzednim tropem. Ułamek łańcuchowy $(1; \overline{2})$ dał nam pierwiastek z liczby wymiernej, a $(2; \overline{2})$, mimo że jeszcze bardziej regularny, wyrażenie „mieszane". Może istnieje prosty sposób, by zgadnąć, kiedy otrzymamy sam pierwiastek z liczby wymiernej, a kiedy nie ? Oto kolejny przykład ułamka łańcuchowego równego takiemu pierwiastkowi:

$w = (2; \overline{1, 3, 1, 4})$. Mamy

$$w = 2 + \frac{1}{y} \quad \text{gdzie} \quad y = 1 + \cfrac{1}{3 + \cfrac{1}{1 + \cfrac{1}{4 + \cfrac{1}{y}}}}.$$

Obliczamy

$$1 + \cfrac{1}{3 + \cfrac{1}{1 + \cfrac{1}{4 + \cfrac{1}{y}}}} = 1 + \cfrac{1}{3 + \cfrac{1}{1 + \cfrac{y}{4y+1}}} = 1 + \cfrac{1}{3 + \cfrac{4y+1}{5y+1}} = 1 + \frac{5y+1}{19y+4} =$$

$$= \frac{24y+5}{19y+4} = y, \quad \text{czyli} \quad 19y^2 - 20y - 5 = 0.$$

Stąd mamy

$$y = \frac{20 + 2\sqrt{195}}{2 \cdot 19} = \frac{10 + \sqrt{195}}{19},$$

a zatem

$$w = 2 + \frac{19}{10 + \sqrt{195}} = 2 + \frac{19(\sqrt{195} - 10)}{95} = \sqrt{195}.$$

Wyszło ! Jaka jest jednak reguła ? Jak się wydaje, niemożliwa do odgadnięcia. Jest ona taka:
ułamek łańcuchowy $(a; \overline{b_1, b_2, \ldots, b_2, b_1, 2a})$, *którego okres po odrzuceniu ostatniej liczby jest symetryczny (przy czym nie ma znaczenia, czy jego długość jest parzysta, czy nie), przedstawia niewymierny pierwiastek z liczby wymiernej większej od 1.*
Co więcej, jest prawdziwe i twierdzenie odwrotne.

Czy udowodnienie takiego faktu jest trudne, jest podobno kwestią wprawy. Przysłowiowe jest zdanie umieszczone w *Teorii liczb* Wacława Sierpińskiego, że przekonać się, iż

$$\sqrt{991} = (31; \overline{2, 12, 10, 2, 2, 2, 1, 1, 2, 6, 1, 1, 1, 1, 3, 1, 8, 4, 1, 2, 1, 2, 3, 1, 4, 1, 20, 6, 4, 31}$$
$$\overline{4, 6, 20, 1, 4, 1, 3, 2, 1, 2, 1, 4, 8, 1, 3, 1, 1, 1, 1, 6, 2, 1, 1, 2, 2, 2, 10, 12, 2, 62}),$$

można za pomocą niezbyt długich rachunków.

Rozwinięcia stosunku dwóch wielkości (czyli liczby rzeczywistej) na ułamek łańcuchowy są dużo starsze od np. rozwinięć dziesiętnych. Jest to historycznie pierwszy sposób radzenia sobie z liczbami rzeczywistymi. Wymyślony został w IV wieku p.n.e. przez Greka Teaitetosa. Jednak matematyka poszła inną drogą i ułamki łańcuchowe znalazły się z dala od jej głównego nurtu. Wielkie znaczenie zapewne miał tu fakt, że podczas gdy dla układu dziesiętnego zostały wymyślone bardzo sprawne algorytmy działań – te, które nazywamy rachunkami pisemnymi – to dla ułamków łańcuchowych przepisów takich nie ma.

Starożytni mieli do ułamków łańcuchowych bardzo pozytywny stosunek – np. proporcję odcinków, daną przez „najprostszy" ułamek nieskończony $(1; \overline{1})$ uznano za dominujący kanon piękna – proszę sprawdzić, że jest to *złota proporcja*, czyli stosunek odcinka a do takiej jego części b, że

$$\frac{a}{b} = \frac{b}{a-b}.$$

Jak dalece dominujący był to kanon, można przekonać się na rzeźbach greckich, gdzie proporcje ciała i jego części właściwie wszystkie są złote.

Współczesny pozytywny stosunek do ułamków łańcuchowych zasadza się na spostrzeżeniu, że dają one najlepsze przybliżenia wymierne liczb niewymiernych. Znaczy to tyle, że jeśli początkowy fragment ułamka łańcuchowego (czyli jego redukt), dla jakiejś liczby niewymiernej a, jest zwykłym ułamkiem nieskracalnym, to lepsze przybliżenia wymierne liczby a można otrzymać tylko używając ułamków o większym mianowniku. Tak więc np.

(patrz początek rozdziału) ułamek $\frac{41}{29}$ jest najlepszym przybliżeniem $\sqrt{2}$ spośród ułamków o mianownikach mniejszych od 30.

Warto przy okazji zwrócić uwagę na istotną różnicę między rozwijaniem w jakimkolwiek systemie pozycyjnym a rozwijaniem w ułamek łańcuchowy. Tutaj na każdym miejscu może się pojawić duża liczba. I to dowolnie duża.

Łatwo spostrzec, że w rozwinięciu liczb wymiernych, jak też w rozwinięciach niewymierności kwadratowych pojawia się skończenie wiele liczb. Jest też oczywiste, że można napisać liczbę, w której rozwinięciu pojawiać się będzie nieskończenie wiele liczb naturalnych, a nawet taką, że pojawią się wszystkie – choćby

$$(1; 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, \ldots).$$

Gdy jednak chce się zaatakować ten problem z przeciwnej strony, to sprawa staje się bardzo trudna. Do tej pory np. nie wiadomo, czy w rozwinięciu $\sqrt[3]{2}$ występuje tylko skończenie wiele różnych liczb. Bo przecież nieokresowy ciąg można zbudować nawet z dwóch różnych liczb.

I na tym zakończę przykłady wykorzystania zwykłego dzielenia z resztą, choć oczywiście przykłady można mnożyć – powiedzmy, dzielenie wielomianów.