ZWIĘZŁA HISTORIA MATEMATYKI DLA FILOZOFÓW

Przedmowa

Od czasów starożytnych matematyka i jej filozofia toczyły walkę. Gdy tylko wydawało się, że istnieje ustalony pogląd na naturę matematyki, pojawiało się nowe odkrycie matematyczne, które go burzyło. Tak więc pitagorejski pogląd, że „wszystko jest liczbą", został zakłócony przez odkrycie niewymiernych długości, a filozofia matematyki musiała się rozszerzyć, aby objąć odrębną dziedzinę geometrii. Ale to podniosło pytanie: Czy pogląd geometryczny można pogodzić z poglądem numerycznym? Jeśli tak, to w jaki sposób? I tak to trwało przez tysiąclecia. W wielu przypadkach postęp w matematyce zmieniał idee dotyczące matematyki, wymuszając akceptację koncepcji wcześniej uważanych za niemożliwe lub paradoksalne. Tak więc matematyka zakłóciła filozofię. W przeciwnym kierunku filozofia utrzymywała matematykę uczciwą, wskazując sprzeczności i sugerując, w jaki sposób można wyjaśnić koncepcje, aby je rozwiązać. Czasami filozof i matematyk byli jedną i tą samą osobą – jak Kartezjusz, Leibniz czy Bolzano – więc można by niemal powiedzieć, że matematyka jest szczególnie bogatą i stabilną gałęzią filozofii. W każdym razie, jeśli ktoś chce zrozumieć przeszłą i obecną sytuację filozofii matematyki, musi najpierw zrozumieć matematykę i jej historię. Celem niniejszego Elementu jest krótkie wprowadzenie do matematyki i jej historii, ze szczególnym naciskiem na wydarzenia, które wstrząsnęły jej filozofią. Jeśli chcesz, jest to książka o „matematyce dla filozofów". Staram się nie zajmować żadnego konkretnego stanowiska filozoficznego, poza stwierdzeniem, że uważam, że matematyka kieruje filozofią bardziej niż odwrotnie. Jako wniosek, uważam, że matematycy wnieśli ważny wkład do filozofii, nawet jeśli nie było to ich intencją. Każda sekcja zaczyna się od zapowiedzi tematów do omówienia i kończy się sekcją podkreślającą kwestie filozoficzne podnoszone przez matematykę. Te same tematy powtarzają się w kolejnych sekcjach – intuicja i logika, znaczenie i istnienie oraz dyskretność i ciągłość – ale ewoluują pod wpływem nowych odkryć matematycznych. Eksperci mogą być zaskoczeni, że niewiele lub wcale nie wspomina się o filozofiach matematyki, które były widoczne w XX wieku – na przykład platonizm, logicyzm, formalizm, nominalizm i intuicjonizm. Dzieje się tak częściowo dlatego, że nie uważam żadnej z nich za odpowiednią, ale głównie dlatego, że mam nadzieję spojrzeć na filozofię matematyki bez wpływu etykiet. Chcę przedstawić jak najwięcej filozoficznie pouczającej matematyki i pozostawić czytelnikom decyzję, jak powinna być sortowana i etykietowana w kategoriach filozoficznych. Mam nadzieję, że ten Element wyposaży czytelników w „soczewkę matematyczną", przez którą będą mogli postrzegać wiele kwestii filozoficznych.

Liczby niewymierne i geometria

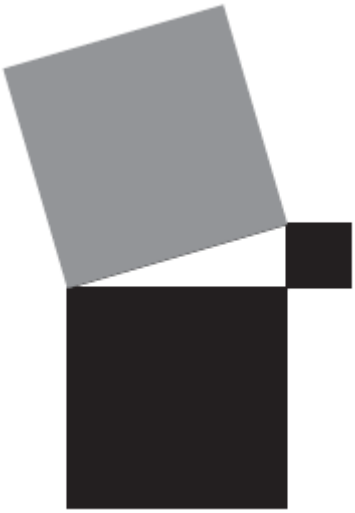
PODGLĄD

Źródłem wielu zagadnień w filozofii matematyki – natury dowodu i prawdy; znaczenia i istnienia liczb; roli nieskończoności; i relacji między geometrią, algebrą i arytmetyką – są Elementy Euklidesa z około 300 r. p.n.e. Elementy są najbardziej znane z geometrii aksjomatycznej – geometrii euklidesowej – która obejmuje dowody charakterystycznych wyników, takich jak twierdzenie Pitagorasa i istnienie dokładnie pięciu regularnych wielościanów. Jednak Elementy obejmują również podstawowe idee teorii liczb, takie jak istnienie nieskończenie wielu liczb pierwszych, algorytm euklidesowy dla największego wspólnego dzielnika i (odpowiednik) jednoznacznego rozkładu na czynniki pierwsze. W czasach Euklidesa, podobnie jak teraz, istniała przepaść koncepcyjna między geometrią a teorią liczb – między mierzeniem a liczeniem lub między ciągłością a dyskretnością. Głównym powodem tej przepaści było istnienie liczb niewymiernych, odkrytych przed czasami Euklidesa przez pitagorejczyków, a w czasach Elementów, będących przedmiotem wyrafinowanej „teorii proporcji". Teoria ta, w Księdze V Elementów, stworzyła kruchy pomost między ciągłością a dyskretnością. Pomost ten był stopniowo wzmacniany przez wieki przez prace późniejszych matematyków, ale nie bez

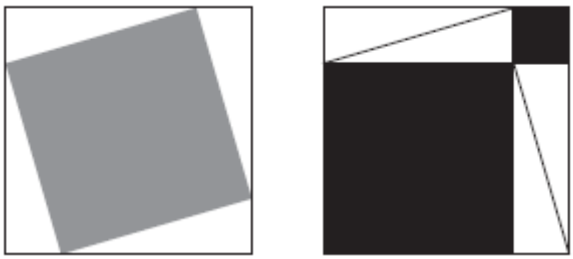
konfliktów filozoficznych i matematycznych niespodzianek. Te kwestie są tematem tej sekcji i następnej.

Twierdzenie Pitagorasa

Twierdzenie Pitagorasa zostało odkryte niezależnie kilka razy w historii ludzkości i w kilku różnych kulturach. Jeśli więc jakieś twierdzenie jest typowe dla matematyki – i jej uniwersalności – to jest to właśnie ono. Rysunek 1 ilustruje twierdzenie: (szary) kwadrat na przeciwprostokątnej (białego) trójkąta prostokątnego jest równy sumie (czarnych) kwadratów na pozostałych dwóch bokach.



Rysunek 2 przedstawia prawdopodobny „dowód za pomocą obrazu" twierdzenia: szary kwadrat jest równy dużemu kwadratowi minus cztery kopie trójkąta, który z kolei jest równy sumie dwóch czarnych kwadratów.



Większość niezależnych odkryć twierdzenia była prawdopodobnie taka, a ludzki układ wzrokowy ma na swoim koncie wiele odkryć matematycznych. Niemniej jednak to radykalnie odmienna aksjomatyczna ścieżka do twierdzeń, wyznaczyła kierunek matematyki na następne 2000 lat. Ale zanim ustalono ścieżkę aksjomatyczną, twierdzenie Pitagorasa wywołało inny ważny rozwój koncepcyjny: rozróżnienie długości i liczby. Legenda głosi, że filozofia pitagorejczyków brzmiała „wszystko jest liczbą", co zostało zainspirowane odkryciem, że stosunki liczb całkowitych rządzą harmonią muzyczną. Ta filozofia została obalona, gdy w geometrii odkryto stosunki niewymierne – z powodu twierdzenia Pitagorasa.

Nieracjonalność

Twierdzenie Pitagorasa mówi o sumie kwadratów – operacji, o której powiemy więcej poniżej – ale pośrednio mówi nam też coś o długościach. W szczególności mówi, że jeśli trójkąt ma prostopadłe boki o długości 1, to jego przeciwprostokątna ma długość l, której kwadrat wynosi 2. Używając współczesnej

notacji l^2 do oznaczenia kwadratu boku l, mamy $l^2 = 1^2 + 1^2 = 2$. Teraz (ponownie używając współczesnej notacji) załóżmy, że l jest wymierne, w takim przypadku możemy założyć, że l = m/n, gdzie m i n są liczbami całkowitymi. Możemy również założyć, że m i n nie mają wspólnego dzielnika poza 1, ponieważ każdy inny wspólny dzielnik można podzielić z m i n z góry, bez zmiany l. W tych warunkach możemy wyprowadzić sprzeczność z następującej serii implikacji (prawdopodobnie sięgają one Pitagorejczyków, ale pierwsza znana wskazówka takiego dowodu znajduje się w Analityce Arystotelesa 1.23):

l = m/n) -> $2 = m^2/n^2$ (podniesienie do kwadratu obu stron)

-> $2n^2 = m^2$ (mnożenie obu stron przez n^2)

-> m^2 jest parzyste

-> m jest parzyste; powiedzmy; = 2p (ponieważ kwadrat liczby nieparzystej jest nieparzysty)

-> $2n^2 = (2p)^2 = 4p^2$ (podstawienie m = 2p w $2n^2 = m^2$)

-> $n^2 = 2p^2$

-> n^2 jest parzyste

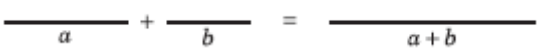
-> n jest parzyste

-> 2 dzieli zarówno m, jak i n; wbrew założeniu braku wspólnego dzielnika:

Ponieważ założenie, że l jest stosunkiem liczb całkowitych, jest sprzeczne, l jest długością niewymierną. Grecy często wyrażali to, mówiąc, że bok i przeciwprostokątna trójkąta prostokątnego o równych bokach są niewspółmierne – nie są wielokrotnościami liczb całkowitych żadnej wspólnej jednostki miary. W opinii pitagorejczyków brak wspólnej jednostki długości oznaczał, że długości nie są liczbami, ponieważ „liczby" były dla nich liczbami całkowitymi i ich stosunkami. W szczególności sumy i iloczyny długości niekoniecznie są takie same jak sumy i iloczyny liczb, więc koncepcja „sumy kwadratów" wymaga wyjaśnienia. W następnej sekcji zobaczymy, jak Euklides radził sobie z sumami i iloczynami długości.

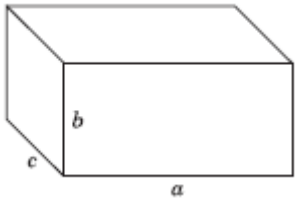
Działania na długościach i liczbach

Zaprzeczając, że niewymierne długości mogą być liczbami, a jednocześnie dopuszczając, że można je podnieść do kwadratu i dodać, greccy matematycy po Pitagorasie musieli zdefiniować sumę i iloczyn w czysto geometrycznych terminach. Suma dwóch długości jest zdefiniowana w oczywisty sposób sugerowany przez Rysunek 3.



Długości są reprezentowane przez dwa odcinki a i b, a a + b jest uzyskiwane przez połączenie tych odcinków koniec do końca. Łatwo z tego wynika, że a + b = b + a i a +(b+c) = (a+b) + c (prawa przemienności i łączności). Jest również jasne, ponieważ suma długości jest długością, że można dodać dowolną liczbę długości. Tak więc długości zachowują się dokładnie jak liczby, jeśli chodzi o dodawanie. Zachowanie iloczynów nie jest takie proste. Iloczyn długości a i b nie jest długością, ale prostokątem o

prostopadłych bokach a i b. A iloczyn długości a, b i c jest prostokątnym pudełkiem o prostopadłych bokach a, b i c (Rysunek 4).



Z tych definicji jasno wynika, że ab = ba i a(bc) = (ab)c, a także można zauważyć, że a(b+c) = ab + ac (to ostatnie jest w rzeczywistości szczególnym przypadkiem propozycji 1 Euklidesa z Księgi II Elementów). Tak więc, w zakresie, w jakim zdefiniowano sumę i iloczyn, długości spełniają te same prawa, co liczby dodatnie. Problem polega na tym, że są one zdefiniowane tylko w ograniczonym zakresie, więc algebra długości jest sparaliżowana. Iloczyny o więcej niż trzech długościach nie są dopuszczalne, ponieważ nie mają odpowiednika geometrycznego. Podobnie, iloczyny można dodawać tylko wtedy, gdy każdy ma ten sam „wymiar", to znaczy, jest iloczynem tej samej liczby długości. Wreszcie istnieje skomplikowana, choć geometrycznie naturalna, koncepcja równości. Mówi ona na przykład, że dwa prostokąty R i S są równe, jeśli R można rozciąć na skończoną liczbę trójkątów, które można złożyć ponownie, aby utworzyć S. Więcej o teorii równości Euklidesa dla prostokątów i innych wielokątów powiemy w następnej sekcji. Co ciekawe, teoria ta jest całkowicie odpowiednia dla wielokątów, ponieważ dowolne dwa wielokąty o równej powierzchni (w nowoczesnym sensie) są w rzeczywistości równe w sensie Euklidesa. Jednak teoria ta nie jest odpowiednia dla wielościanów, jak wykazał Dehn (1900). Dehn wykazał, że sześcian i czworościan foremny o równej objętości nie są równe w sensie Euklidesa.

Aksjomatyka

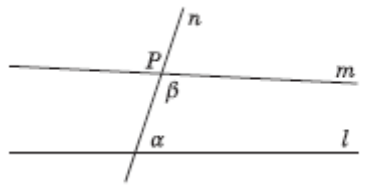
Moc metody aksjomatycznej została czarująco opisana przez Johna Aubreya w jego Krótkich życiach, mówiąc o Thomasie Hobbesie:

Miał 40 lat, kiedy zaczął interesować się geometrią; co zdarzyło się przypadkowo. Będąc w bibliotece pewnego dżentelmena… Elementy Euklidesa leżały otwarte, a to była Księga I 47 Elementów. Przeczytał to twierdzenie. „Na Boga –" powiedział, „to niemożliwe!" Więc przeczytał dowód tego, który odesłał go z powrotem do takiego twierdzenia; które to twierdzenie przeczytał. To odesłało go z powrotem do innego, które również przeczytał… że w końcu był demonstracyjnie przekonany o tej prawdzie. To sprawiło, że zakochał się w geometrii.

Propozycja 47 Księgi I, nawiasem mówiąc, jest twierdzeniem Pitagorasa. Elementy to pierwsze systematyczne sprawozdanie z twierdzeń i dowodów, jakie do nas dotarło, i stało się standardowym sposobem przedstawiania matematyki w świecie zachodnim (a później w świecie islamskim) przez następne 2000 lat. Euklides zaczyna od niewielkiej liczby podstawowych założeń (aksjomatów) i wyprowadza z nich wszystkie twierdzenia za pomocą logiki. Jego aksjomaty obejmują proste stwierdzenia dotyczące punktów, linii, długości i kąta. Istnieją również stwierdzenia dotyczące równości, dodawania i odejmowania, takie jak „rzeczy, które są równe tej samej rzeczy, są równe sobie nawzajem" i „jeśli równe zostaną dodane do równych, to całości będą równe". Zasady logiki nie są wyraźnie określone. Najważniejszym aksjomatem, potrzebnym do twierdzenia Pitagorasa i wielu innych, jest aksjomat równoległości. Euklides stwierdza to następująco w tłumaczeniu Heatha (1956):

Że jeśli prosta padająca na dwie proste sprawia, że kąty wewnętrzne po tej samej stronie są mniejsze od dwóch kątów prostych, to dwie proste, jeśli są prowadzone w nieskończoność, spotykają się po tej stronie, po której kąty są mniejsze od dwóch kątów prostych.

To dość długie stwierdzenie zilustrowano na Rysunku 5.



Linia n pada na linie l i m, tworząc kąty α i β po prawej stronie, przy czym α + β < π. Wniosek jest taki, że l i m spotykają się gdzieś po prawej stronie. Tak więc aksjomat równoległości faktycznie daje warunek, aby linie nie były równoległe. Wynika z tego (nie do końca oczywiste), że istnieje dokładnie jedna linia równoległa do danej linii l przechodząca przez dany punkt P poza l, mianowicie linia m, dla której α + β = π. Skomplikowany charakter aksjomatu równoległości wywołał wiele prób jego wyeliminowania poprzez wykazanie, że wynika on z innych aksjomatów Euklidesa. Jednak wszystkie takie próby zakończyły się niepowodzeniem. Doprowadziło to w XIX wieku do gruntownego zbadania metody aksjomatycznej i późniejszej analizy jej zakresu i ograniczeń. Kontynuujemy tę historię później.

Zagadnienia filozoficzne

Według legendy pitagorejczycy jako pierwsi zaproponowali filozofię matematyki, w rzeczywistości bardzo prostą „teorię wszystkiego": wszystko jest liczbą. Mówi się, że zaobserwowali rolę liczb całkowitych w harmonii muzycznej i doszli do wniosku, że cały wszechświat jest rządzony przez liczby całkowite i ich stosunki. Echa tej filozofii są nadal słyszalne w zwrotach takich jak „harmonia sfer". Bez względu na jej szczegóły, filozofia pitagorejska została zakłócona przez odkrycie wielkości niewymiernych, takich jak $\sqrt{2}$. Liczby niewymierne były niedopuszczalne jako liczby, ale nieuniknione w geometrii, ponieważ nikt nie mógł zaprzeczyć, że jeśli istnieje kwadrat, to istnieje również jego przekątna. Doprowadziło to do rozdzielenia teorii liczb i geometrii, co widać w Elementach Euklidesa, ale także do teorii zawartej w Księdze V Elementów. „Teoria proporcji" zawarta w Księdze V ustanawia punkt styczny między liczbami (racjonalnymi) a wielkościami geometrycznymi, choć nie godząc ich w pełni. Znaczna część późniejszej historii matematyki i jej filozofii wyrasta z walki o pogodzenie pojęć liczby i ilości, albo dyskretnego i ciągłego, albo racjonalnego (logicznego) i wizualnego (intuicyjnego). Rozwój filozofii matematycznej towarzyszy tej walce, jak zobaczymy w kolejnych sekcjach. Na końcu każdej sekcji przedstawię historyczną aktualizację, jak to się mówi, rozwoju filozoficznego, pod nagłówkami logiki i intuicji, znaczenia i istnienia oraz dyskretnego i ciągłego. Jako matematyk wolę myśleć w tych kategoriach, ale mam nadzieję, że filozofowie będą w stanie przełożyć filozoficzną treść moich uwag na swoje własne preferowane terminy.

Intuicja i logika. Prawdopodobnie w próbie precyzyjnej pracy z wielkościami geometrycznymi, Elementy Euklidesa są pierwszym znanym przykładem aksjomatycznego podejścia do prawdy, w którym twierdzenia są wyprowadzane z aksjomatów za pomocą logiki. Jednak jego aksjomaty są niekompletne i często odwołują się do intuicji, nawet w jego pierwszym twierdzeniu. Tak więc Elementy nieumyślnie ilustrują, jak trudno jest uniknąć nieświadomych założeń w rozumowaniu matematycznym.

Znaczenie i istnienie. Euklides podważa również to, co obecnie uważamy za metodę aksjomatyczną, próbując zdefiniować prymitywne pojęcia, takie jak „punkt" i „linia". Ogranicza również pojęcie „liczby" zasadniczo do liczb naturalnych i ich stosunków. Uważa się, że nieracjonalność $\sqrt{2}$ dyskwalifikuje ją z bycia liczbą, ale Euklides nie przepisał, jakie powinny być właściwości liczb.

Dyskretne i ciągłe. Ze względu na wrażliwość na wielkości niewymierne Elementy generalnie wyraźnie rozdzielają pojęcia liczby i ilości lub dyskretności i ciągłości. Jednak Księga V rozpoczyna możliwe połączenie tych dwóch, jak zobaczymy później. Ilustruje to czasami przeciwstawne tendencje matematyki i filozofii. Matematycy generalnie podzielają pogląd wyrażony przez Poincarégo w 1908 r.:

Myślę, że gdzieś już powiedziałem, że matematyka jest sztuką nadawania tej samej nazwy różnym rzeczom. Wystarczy, że te rzeczy, choć różnią się materią, są podobne pod względem formy, aby ich istnienie, że tak powiem, mogło biec w tej samej formie. Gdy język został dobrze dobrany, człowiek jest zdziwiony, gdy odkrywa, że wszystkie demonstracje dokonane dla znanego obiektu odnoszą się natychmiast do wielu nowych obiektów: nic nie wymaga zmiany, nawet terminy, ponieważ nazwy stały się takie same.

Innymi słowy, matematycy uważają, że rzeczy są takie same, jeśli zachowują się tak samo. Filozofowie jednak lubią dokonywać rozróżnień: szukają powodów, dla których rzeczy nie powinny być uważane za takie same. Czasami rozróżnienie wydaje się uzasadnione, tak jak greckie rozróżnienie między liczbami a wielkościami geometrycznymi, takimi jak długość. Jednak matematyka stara się wymazać rozróżnienia, gdzie to możliwe. Długą ewolucję koncepcji liczby rzeczywistej można postrzegać jako projekt mający na celu wymazanie rozróżnienia między liczbą a ilością, a wraz z nią rozróżnienia między teorią liczb rzeczywistych a geometrią.

Nieskończoność w matematyce greckiej

PODGLĄD

Chociaż liczba i długość są w większości rozdzielone w Elementach, istnieje jeden proces, który Euklides zastosował do obu – algorytm Euklidesa, który działa na parze poprzez „wielokrotne odejmowanie mniejszej od większej”. Gdy zastosuje się go do pary dodatnich liczb całkowitych (lub ogólniej, do pary dodatnich wielokrotności całkowitych długości jednostkowej), algorytm kończy się, ponieważ liczby całkowite dodatnie nie mogą się zmniejszać w nieskończoność. Jednak gdy zastosuje się go do pary długości w stosunku niewymiernym, algorytm się nie kończy. Rzeczywiście, Euklides użył nieskończoności swojego algorytmu jako kryterium niewymierności, wprowadzając w ten sposób nieskończoność do dyskusji na temat niewymierności. Bezliczbowa teoria pola, używana przez Euklidesa do udowodnienia twierdzenia Pitagorasa, działa całkiem płynnie w przypadku pól wielokątów. Jednak podobne podejście do objętości zawodzi nawet w przypadku prostych wielościanów, takich jak czworościan. Euklides był w stanie znaleźć objętość czworościanu, rozkładając go na nieskończenie wiele graniastosłupów, wprowadzając w ten sposób nieskończoność do teorii objętości. Grecka teoria pola miała również trudności z obszarami zakrzywionymi, których oczywiście nie można rozłożyć na skończenie wiele wielokątów. Jednak Archimedes był w stanie znaleźć pole odcinka parabolicznego, rozkładając go na nieskończenie wiele trójkątów. Niemniej jednak Grecy starali się „uniknąć nieskończoności”, rozważając dowolne sumy skończone zamiast sum nieskończonych.

Nieracjonalność i nieskończenie

Algorytm Euklidesa został wprowadzony w Księdze VII Elementów jako metoda znajdowania największego wspólnego dzielnika dwóch dodatnich liczb całkowitych. Jak mówi Euklides, „ciągle odejmuje się mniejszą liczbę od większej”; dokładniej, jeśli a > b, to zastępuje się parę a;b przez a - b;b. Ponieważ liczby całkowite dodatnie nie mogą się zmniejszać w nieskończoność, algorytm zawsze się kończy. Na przykład, z parą 13, 8, otrzymuje się 13; 8 → 5; 8 → 5; 3 → 2; 3 → 2; 1 → 1; 1: Gdy otrzymuje

się parę identycznych liczb, liczbą tą jest gcd(a,b), ponieważ wszystkie wspólne dzielniki pary są zachowywane przez odejmowanie. Tak więc nasz przykład pokazuje, że gcd(13,8) = 1. W Księdze X Elementów Euklides uogólnia algorytm do długości a i b, w którym to przypadku algorytm może się nie zakończyć. Na przykład, jeśli (używając współczesnej notacji) długości wynoszą a = √2 i b = 1, to pierwsze dwa kroki to

$$\sqrt{2}, 1 \rightarrow 1, \sqrt{2} - 1 \rightarrow 2 - \sqrt{2}, \sqrt{2} - 1$$

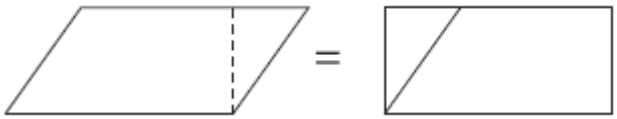
W tym momencie można zauważyć, że 2 - √2 i √2 – 1 są w tym samym stosunku co √2 i 1. Nie jest jasne, czy Grecy to zauważyli (choć prawdopodobnie byli świadomi czegoś podobnego), ale jest to jasne na podstawie podstawowej algebry, ponieważ

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

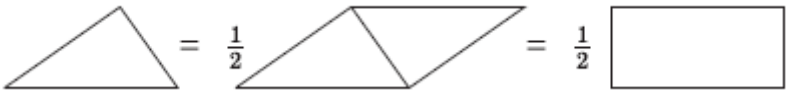
Ponieważ 2 - √2 i √2 – 1 są w tym samym stosunku co √2 i 1, zastosowanie do nich algorytmu Euklidesa wytworzy w dwóch krokach kolejną parę w tym stosunku – i tak dalej, w nieskończoność. Niezależnie od tego, czy Euklides znał ten konkretny przykład, zdał sobie sprawę, że algorytm nie kończy się na parze długości w stosunku niewymiernym (Księga X, Propozycja 2). W ten sposób algorytm Euklidesa elegancko oddziela to, co wymierne, od tego, co niewymierne, poprzez oddzielenie tego, co skończone, od tego, co nieskończone.

Pola i objętości

W Księdze I Elementów Euklides wykazuje równość różnych obszarów poprzez dodawanie lub odejmowanie równych trójkątów. Na przykład, Rysunek 6 pokazuje, że równoległobok jest równy prostokątowi o tej samej podstawie i wysokości.



A Rysunek 7 pokazuje, że trójkąt jest równy połowie prostokąta o tej samej podstawie i wysokości.



Istnieją również rozkłady pokazujące, że każdy prostokąt jest równy prostokątowi o danej podstawie. Korzystając z tego faktu, można znaleźć prostokąt równy dowolnemu wielokątowi, dzieląc wielokąt na skończenie wiele trójkątów. Teraz, jeśli (jak w przypadku trójkąta) jeden obszar R jest równy wymiernej wielokrotności r pewnego standardowego obszaru, który możemy przyjąć jako jednostkę, to jest zgodne z Euklidesem, aby pozwolić liczbie r „mierzyć" obszar R w taki sam sposób, w jaki mierzymy pole R. W tych warunkach będziemy od teraz mówić o liczbowych „powierzchniach" i „objętościach". Teraz zakrzywionego obszaru oczywiście nie można podzielić na skończenie wiele trójkątów. Najlepsze, na co możemy liczyć, to rozkład na nieskończenie wiele trójkątów, który, jeśli będziemy mieli szczęście, może być zrozumiały. Archimedes odniósł dzięki tej metodzie błyskotliwy sukces, znajdując pole odcinka parabolicznego. Odcinek paraboliczny jest wypełniony trójkątami w sposób pokazany na rysunku 8: najpierw czarny trójkąt, następnie dwa ciemnoszare trójkąty pod nim, następnie cztery jaśniejsze szare trójkąty pod nimi i tak dalej.

Każdy trójkąt jest o połowę szerszy od trójkąta nad nim, a obliczenia pokazują, że każda grupa (jeden, dwa, cztery,… trójkąty) ma całkowitą powierzchnię równą jednej czwartej powierzchni grupy nad nim. Tak więc, jeśli czarny trójkąt ma powierzchnię 1, to całkowita powierzchnia trójkątów jest nieskończoną sumą

$$1 + \frac{1}{4} + \left(\frac{1}{4}\right)^2 + \left(\frac{1}{4}\right)^3 + \left(\frac{1}{4}\right)^4 + \ldots.$$

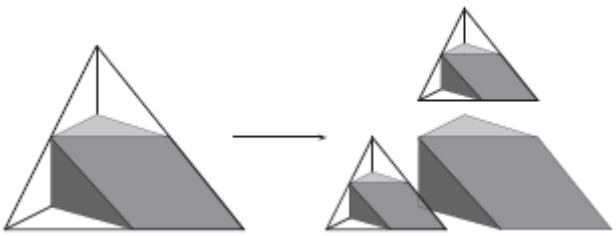
Jeśli ustawimy

$$S = 1 + \frac{1}{4} + \left(\frac{1}{4}\right)^2 + \left(\frac{1}{4}\right)^3 + \left(\frac{1}{4}\right)^4 + \ldots,$$

wtedy wyraźnie

$$4S = 4 + 1 + \frac{1}{4} + \left(\frac{1}{4}\right)^2 + \left(\frac{1}{4}\right)^3 + \left(\frac{1}{4}\right)^4 + \ldots,$$

więc 4S - S = 4 i dlatego S = 4/3. (To ryzykowne obliczenie z nieskończonymi sumami daje szybki sposób na odgadnięcie odpowiedzi. Zobaczymy, jak Archimedes zrobił to dokładniej w następnej sekcji.) Czworościan jest wypełniony graniastosłupami, jak pokazano na rysunku 9.



Po usunięciu dwóch graniastosłupów (jednego jasnoszarego, drugiego ciemnoszarego) pozostają dwa mniejsze czworościany, z których ponownie usuwamy graniastosłupy i kontynuujemy. W ten sposób Euklides odkrył, że objętość czworościanu jest podobną nieskończoną sumą,

$$\frac{1}{4} + \left(\frac{1}{4}\right)^2 + \left(\frac{1}{4}\right)^3 + \left(\frac{1}{4}\right)^4 + \ldots,$$

co równa się 1/3.

Metoda wyczerpania

Ostatnie dwie sekcje pokazały przykłady procesów nieskończonych w matematyce greckiej. Te w poprzedniej sekcji są znane ze znajdowania racjonalnego obszaru lub objętości, których najwyraźniej nie można uzyskać za pomocą procesów skończonych. Niemniej jednak Grecy byli podejrzliwi wobec nieskończoności i starali się jej unikać tak dalece, jak to możliwe. W wielu przypadkach byli w stanie to zrobić za pomocą tego, co nazywa się metodą wyczerpania. Metoda ta jest tak nazywana, ponieważ potwierdza się wynik x, uzyskany za pomocą procesu nieskończonego, poprzez wyczerpanie wszystkich innych możliwości (mniejszych niż x lub większych niż x). Zazwyczaj inne możliwości są wyczerpujące przez założenie, że proces przebiega przez dowolną, ale skończoną liczbę kroków. Możemy zilustrować metodę wyczerpania w przypadku oceny pola odcinka parabolicznego przez Archimedesa, zastępując procesy nieskończone dowolnymi procesami skończonymi. Procesy nieskończone zaangażowane w tym przykładzie są następujące:

1 Wypełnienie odcinka parabolicznego nieskończenie wieloma trójkątami. Możemy argumentować, że trójkąty wypełniają segment (lub „wyczerpują" go), pokazując, że dowolny punkt wewnątrz paraboli mieści się wewnątrz pewnego trójkąta. Jest to czysto skończony, choć żmudny, argument.

2 Pokazanie, że nieskończony szereg

$$1 + \frac{1}{4} + \left(\frac{1}{4}\right)^2 + \left(\frac{1}{4}\right)^3 + \left(\frac{1}{4}\right)^4 + \ldots$$

ma sumę 4/3

Można to zrobić, znajdując skończoną sumę

$$S_n = 1 + \frac{1}{4} + \left(\frac{1}{4}\right)^2 + \ldots + \left(\frac{1}{4}\right)^n = \frac{4}{3} - \frac{1}{3}\left(\frac{1}{4}\right)^n.$$

Widzimy, że S_n „wyczerpuje" wszystkie liczby < 4/3, przyjmując dowolnie duże wartości n, ponieważ $(1/4)^n$ staje się wówczas dowolnie małe. I oczywiście S_n nie może być > 4/3, więc wartość, która pozostaje, 4/3, jest koniecznie sumą nieskończonego szeregu. Zatem pole odcinka parabolicznego (i podobnie objętość czworościanu) można znaleźć, używając dowolnych sum skończonych zamiast sum nieskończonych. Jest to typowe dla sposobu, w jaki Grecy unikali faktycznego użycia nieskończoności.

Teoria proporcji

Powszechnym problemem dla czytelników Elementów była Księga V, dotycząca tak zwanej teorii proporcji. Uważa się, że teoria ta pochodzi od Eudoksosa, a dotyczy pojęcia długości i jego związku z pojęciem liczby. Jej główna trudność to w zasadzie trudność, z jaką spotykamy się dzisiaj, próbując zrozumieć oś liczb rzeczywistych. Mniejsza trudność wynika z greckiego zwyczaju pracy ze stosunkami (liczbami całkowitymi lub długościami), który powstrzymywał ich przed pracą z liczbami wymiernymi i jednostką długości. Pominę to jednak, aby skupić się na głównej trudności – którą jest oczywiście istnienie liczb niewymiernych – poprzez dopuszczenie liczb wymiernych i jednostki długości. Gdy dopuszczone zostaną liczby wymierne, staje się jasne, że są one kluczowym składnikiem teorii proporcji. Biorąc pod uwagę długość jednostki 1, mamy również długość wymierną l = m/n dla dowolnych dodatnich liczb całkowitych m, n. Jest to długość taka, że n kopii l jest równe m kopii 1. Te wymierne długości określają, dla dowolnych długości a i b, czy a < b; a = b, czy a > b. Mianowicie

a < b ⇔ istnieje wymierne m/n z a < m/n < b;

$a > b \Leftrightarrow$ istnieje wymierne m/n z $a > m/n > b$;

$a = b \Leftrightarrow$ nie jest ani $a < b$, ani $a > b$

Tak więc jeśli $a < b$ lub $a > b$, istnieje para m, n, która „świadczy" o tym fakcie: albo dlatego, że $a < m/n < b$ lub $a > m/n > b$. Ale jeśli $a = b$, nie ma pojedynczej liczby wymiernej m/n, która świadczyłaby o tym fakcie (chyba że a i b są wymierne): a równa się b tylko wtedy, gdy wszystkie liczby wymierne mniejsze od a są mniejsze od b i odwrotnie. Tak więc równość jest trudniejsza do uchwycenia niż nierówność, a długości niewymierne są trudniejsze do uchwycenia niż długości wymierne. Z powodu podejrzeń co do nieskończoności Grecy nie podjęli kroku, aby powiedzieć, że długość niewymierna jest określona przez długości wymierne po obu jej stronach, ponieważ to określenie obejmuje nieskończenie wiele długości wymiernych. Wkrótce zobaczymy, co się stało, gdy podjęto ten krok w XIX wieku. Nawet test nierówności podnosi interesujący punkt filozoficzny. Załóżmy, że $0 < b$, a zatem $0 < m/n < b$ dla pewnych dodatnich liczb całkowitych m i n. Wynika z tego, pomnożenie przez n, że $0 < m < nb$. Tak więc teoria proporcji zakłada to, co później nazwano aksjomatem Archimedesa lub nieistnieniem nieskończenie małych: jeśli $b > 0$, to pewna liczba całkowita wielokrotna $nb > 1$. Stało się to gorącym tematem w XVII wieku, kiedy matematycy uznali za wygodne zakładanie istnienia nieskończenie małych; to znaczy, założyli, że istnieje $b > 0$ takie, że żadna liczba całkowita wielokrotna $nb > 1$.

Archimedes i rzeczywista nieskończoność

Podobnie jak metoda wyczerpania, teoria proporcji unika nieskończoności, zajmując się zamiast tego „dowolną skończonością". W szczególności wykorzystuje dowolne skończone liczby całkowite m, n, aby wykazać, że $a < b$, ponieważ $a < m/n < b$. Rzeczywiście, gdy $a = b$ jest udowadniane, odbywa się to przez wyczerpanie; mianowicie poprzez wykazanie, że $a < m/n < b$ i $a > m/n > b$ nie spełniają wymagań dla dowolnej możliwej pary m,n. Nieskończoność, którą można „wyczerpać" w ten sposób za pomocą skończonych części, Grecy nazywali „potencjalną nieskończonością", a w argumentacji matematycznej dopuszczano tylko potencjalne nieskończoności. Celem metody wyczerpania było uniknięcie rzeczywistej nieskończoności; to znaczy postrzeganie nieskończonego zbioru lub procesu jako ukończonej całości. Przez ponad 2000 lat po Euklidesie oficjalną praktyką matematyków było akceptowanie tylko potencjalnych nieskończoności i unikanie rzeczywistych nieskończoności poprzez odwoływanie się do metody wyczerpania. Taka była teoria, ale w praktyce rzeczywiste nieskończoności były często używane jako skrót do wyników, które (miejmy nadzieję) można było później rygorystycznie udowodnić metodą wyczerpania. O ile nam wiadomo, pierwszym, który użył rzeczywistych nieskończoności, był Archimedes w dziele zatytułowanym Metoda. Dzieło to zostało utracone i nieznane przez stulecia, zanim zostało ponownie odkryte w 1906 r., więc nie wpłynęło na rozwój matematyki ani jej filozofii. Pokazuje jednak, że greckie myślenie o nieskończoności wiązało się z czymś więcej, niż można by wywnioskować z Euklidesa. Archimedes poszedł znacznie dalej niż myślenie, że sekwencja dyskretnych kroków może zostać ukończona; był nawet skłonny postrzegać kontinuum punktów na linii jako skończony zbiór (choć prawdopodobnie nie zdając sobie sprawy, że kontinuum jest nowym rodzajem nieskończoności).

Zagadnienia filozoficzne

Euklides dotarł do sedna rozróżnienia między tym, co racjonalne, a tym, co nieracjonalne: jest to różnica między tym, co skończone, a tym, co nieskończone. Tak więc geometria, która akceptuje wielkości nieracjonalne, musi również akceptować nieskończoność w pewnej formie. Tak więc logika geometrii musiała znaleźć sposób na rozumowanie o nieskończoności.

Intuicja i logika. Grecy starali się, o ile to możliwe, unikać rozumowania o nieskończoności. Pochodzenie ich strachu przed nieskończonością nie jest do końca jasne, chociaż paradoksy Zenona pokazują, że strach ten istniał przed czasami Euklidesa i że Arystoteles próbował go obalić. Metoda wyczerpania była matematyczną odpowiedzią na te filozoficzne debaty: unikaj rozumowania o nieskończoności, rozumując o skończonych (ale dowolnych) etapach nieskończonego procesu i argumentuj, że „wyczerpują wszystkie możliwości" z wyjątkiem jednej – którą to możliwość można zatem uznać za wynik nieskończonego procesu.

Znaczenie i istnienie. Tak więc metoda wyczerpania pośrednio nadaje sens wynikowi nieskończonego procesu, bez zobowiązania do istnienia rzeczywistej nieskończoności.

Ciągłe i dyskretne. Procesy z nieskończoną liczbą dyskretnych kroków można by uznać za „potencjalną nieskończoność". Jednak odrzucając rzeczywistą nieskończoność, Grecy generalnie trzymali się z daleka od idei „ciągłej nieskończoności". Być może Archimedes był wyjątkiem, ponieważ wydawał się skłonny zaakceptować kontinuum punktów na linii jako całość skończoną. Jak wspomniano, tendencją filozofii jest dokonywanie rozróżnień, podczas gdy tendencją matematyki jest ich wymazywanie, gdzie to możliwe. Jednak matematyka czasami odkrywa rozróżnienia, których filozofia (i wcześniejsza matematyka) nie przewidziała. Nieskończoność jest tego przykładem. Przez tysiąclecia matematyka akceptowała rozróżnienie między potencjalną a rzeczywistą nieskończonością i, podążając za filozofią, uważała rzeczywistą nieskończoność za niedopuszczalną. Ale od końca XIX wieku rzeczywista nieskończoność nie tylko stała się akceptowalna (dla większości matematyków), ale także podlega nieprzewidzianym i skomplikowanym rozróżnieniom. Doprowadziło to do nowych kontrowersji dotyczących tego, gdzie wyznaczyć granicę między akceptowalną a nieakceptowalną nieskończonością

Liczby urojone

PODGLĄD

Opór przed traktowaniem wielkości takich jak √2 jako liczb stopniowo zanikał na przestrzeni wieków, prawdopodobnie z powodu rozwoju algebry w Indiach i świecie islamskim. W Indiach Brahmagupta około 600 r. n.e. podał zasadniczo współczesne rozwiązanie $x = \frac{-b \pm \sqrt{b^2 - 4ac}}{2a}$ równania kwadratowego $ax^2 + bx + c = 0$, a zaczęto akceptować, że pierwiastek kwadratowy liczby dodatniej jest sam w sobie liczbą. Jednocześnie istniała niechęć do używania liczb ujemnych (choć Brahmagupta je akceptował), a pierwiastek kwadratowy liczby ujemnej wydawał się bezsensowny. Tak więc, gdy $b^2 - 4ac < 0$, wydawało się naturalne stwierdzenie, że $ax^2 + bx + c = 0$ nie ma rozwiązania. Sytuacja zmieniła się w XVI wieku, gdy włoscy matematycy odkryli rozwiązanie równania sześciennego $x^3 = px + q$ w postaci

$$x = \sqrt[3]{\frac{q}{2} + \sqrt{\left(\frac{q}{2}\right)^2 - \left(\frac{p}{3}\right)^3}} + \sqrt[3]{\frac{q}{2} - \sqrt{\left(\frac{q}{2}\right)^2 - \left(\frac{p}{3}\right)^3}}$$

Wzór ten wymagał nie tylko akceptacji pierwiastków kwadratowych i sześciennych, ale także pierwiastków kwadratowych liczb ujemnych – ponieważ istnieją równania, dla których istnieje oczywiste rozwiązanie rzeczywiste, ale $(q/2)^2 - (p/3)^3 < 0$. Przez długi czas liczby takie jak √-1 były nazywane niemożliwymi i nadal są nazywane „urojonymi". Mimo to były akceptowane w matematyce, przynajmniej do udowodnienia wyników dotyczących liczb rzeczywistych, ponieważ były użyteczne i (zwykle) nie prowadziły do sprzeczności. Ostatecznie układ liczb rzeczywistych i urojonych zaczęto postrzegać jako naturalny, zarówno algebraicznie, jak i geometrycznie.

Równania kwadratowe i sześcienne

Jak widzieliśmy , algebra była sparaliżowana w matematyce greckiej przez geometryczną interpretację iloczynu. Zgodnie z interpretacją geometryczną iloczyny więcej niż trzech wyrazów nie mają znaczenia, a iloczyny różnych wymiarów nie mogą być dodawane. Ograniczenia te nie dotyczą iloczynu liczb, więc w rywalizacji między algebrą liczb a algebrą długości, algebra liczb wyraźnie wygrywa. To mniej więcej to, co wydarzyło się w rozwoju algebry, która początkowo była symboliką rozwiązywania problemów (zwłaszcza równań) dotyczących liczb. Do około 1600 roku algebra była dyscypliną rozwiązującą równania w liczbach, ale powróciła do geometrii, aby uzasadnić swoje posunięcia – ponieważ Elementy Euklidesa były nadal modelem dowodu matematycznego. Przejście od geometrii do algebry liczb rozpoczęło się od Diofantosa, około 200 r. n.e., w ostatniej fazie klasycznej matematyki greckiej. Diofantos użył symboliki, która dopuszczała iloczyny czterech lub więcej elementów. Ponieważ jednak interesowało go znalezienie racjonalnych rozwiązań równań, używał wyłącznie operacji racjonalnych $+, -, \times, \div,$ a nie operacji pierwiastka kwadratowego √. Operacja √ występuje w rozwiązaniu ogólnym

$$x = \frac{-b \pm \sqrt{b^2 - 4ac}}{2a} \qquad (*)$$

równania kwadratowego $ax^2 + bx + c = 0$. Zasadniczo rozwiązanie to podał Brahmagupta w Indiach około 600 r. n.e., choć w słowach, a nie symbolach. Badanie równań rozprzestrzeniło się z Indii do świata islamskiego, gdzie około 800 r. n.e. al-Khwarizmi nadał mu nazwę „algebra". Stamtąd trafiło do Włoch, gdzie nastąpił kolejny duży postęp: rozwiązanie równań sześciennych. Rozwiązanie $x^3 = px + q$ zostało odkryte przez Scipione del Ferro około 1500 r., ale było przechowywane jako „tajna broń" w popularnych wówczas konkursach matematycznych. Zostało ponownie odkryte przez Tartaglię w latach trzydziestych XVI wieku i po raz pierwszy opublikowane w Ars Magna of Cardano (1545). Tak zwany wzór Cardano na rozwiązanie to

$$x = \sqrt[3]{\frac{q}{2} + \sqrt{\left(\frac{q}{2}\right)^2 - \left(\frac{p}{3}\right)^3}} + \sqrt[3]{\frac{q}{2} - \sqrt{\left(\frac{q}{2}\right)^2 - \left(\frac{p}{3}\right)^3}}. \qquad (**)$$

Matematycy w tym czasie byli skłonni zaakceptować pierwiastki kwadratowe i sześcienne liczb dodatnich, ale wzbraniali się przed pierwiastkami kwadratowymi liczb ujemnych. Oczywiście pierwiastki kwadratowe liczb ujemnych występują już we wzorze kwadratowym (*), gdy $b^2 < 4ac$. Ale w tym przypadku można śmiało powiedzieć, że równanie $ax^2 + bx + c = 0$ nie ma rozwiązania. Inaczej było z rozwiązaniem (**) równania sześciennego.

Algebra liczb urojonych Bombellego

Równanie $x^3 = px + q$ może mieć oczywiste rozwiązanie, nawet jeśli wzór Cardano zawiera pierwiastek kwadratowy z liczby ujemnej $(q/2)^2 - (p/3)^3$. Tak jest na przykład w przypadku $x^3 = 15x + 4$, które ma oczywiste rozwiązanie $x = 4$. Dla tego równania wzór Cardano daje (po pewnym uproszczeniu)

$$x = \sqrt[3]{2 + 11\sqrt{-1}} + \sqrt[3]{2 - 11\sqrt{-1}}.$$

Aby pogodzić to wyrażenie z wartością $x = 4$, Bombelli (1572) założył, że $\sqrt{-1}$ podlega tym samym regułom algebraicznym, co zwykłe liczby. Utrzymywał swoje obliczenia w tajemnicy, ale łatwo je

+zrekonstruować. Używając współczesnej notacji i dla √-1 , czyli $i^2 = -1$, można sprawdzić, że $(2 + i)^3 = 2 + 11i$ a $(2-i)^3 = 2 -11i$.

i dlatego

$$\begin{aligned}\sqrt[3]{2+11\sqrt{-1}}+\sqrt[3]{2-11\sqrt{-1}} &= \sqrt[3]{2+11i}+\sqrt[3]{2-11i}\\ &= \sqrt[3]{(2+i)^3}+\sqrt[3]{(2-i)^3}\\ &= (2+i)+(2-i)\\ &= 4.\end{aligned}$$

Jak zauważyło wiele osób, algebra wydaje się być mądrzejsza od nas!

W każdym razie przykład Bombelliego i jemu podobne ostatecznie przekonały matematyków, że używanie liczb „urojonych" lub „niemożliwych" jest bezpieczne. Cokolwiek oznacza √-1, jeśli w ogóle cokolwiek oznacza, to wydaje się zachowywać jak zwykła liczba i dawać poprawne wyniki dotyczące zwykłych liczb.

Wygoda liczb urojonych

W XVIII i XIX wieku matematycy odkryli wiele sytuacji, w których znane właściwości liczb zwykłych są łatwiejsze do wyrażenia lub wyjaśnienia za pomocą liczb urojonych. Oto kilka przykładów.

1 Wzory trygonometryczne

$$\cos(\theta+\varphi) = \cos\theta\cos\varphi - \sin\theta\sin\varphi$$

$$\sin(\theta+\varphi) = \sin\theta\cos\varphi + \cos\theta\sin\varphi$$

są wyrażone bardziej zwięźle (i zapadają w pamięć) za pomocą jednego wzoru

$$\cos(\theta+\varphi) + i\,\sin(\theta+\varphi) = (\cos\theta + i\,\sin\theta)(\cos\varphi + i\,\sin\varphi).$$

2 Tę ostatnią formułę można wyrazić jeszcze zwięźlej w następujący sposób:

$$e^{i(\theta+\varphi)} = e^{i\theta}e^{i\varphi},$$

jeśli przyjmiemy $e^{ix} = \cos x + i \sin x$. Nadaje to nowe znaczenie funkcjom sinus i cosinus – jako częściom urojonej funkcji wykładniczej.

3 Jeśli a, b, c, d są liczbami całkowitymi dodatnimi, to iloczyn $a^2 + b^2$ i $c^2 + d^2$ jest sam w sobie sumą dwóch kwadratów liczb całkowitych. Aby znaleźć te ostatnie kwadraty, używamy $i^2 = -1$, aby utworzyć rozkłady urojone

$$\begin{aligned}a^2+b^2 &= a^2-(ib)^2 = (a+ib)(a-ib)\\ c^2+d^2 &= c^2-(id)^2 = (c+id)(c-id),\end{aligned}$$

używając tożsamości $x^2 - y^2 = (x + y)(x - y)$. Łącząc czynniki i ponownie używając tożsamości, otrzymujemy

$$
\begin{aligned}
(a^2+b^2)(c^2+d^2) &= [(a+ib)(c+id)][(a-ib)(c-id)] \\
&= [ac-bd+i(ad+bc)][ac-bd-i(ad+bc)] \\
&= (ac-bd)^2+(ad+bc)^2.
\end{aligned}
$$

Te przykłady sugerują, że liczby urojone powinny być akceptowane, choćby dla wygody. Można jednak zrobić coś lepszego. Możemy podać przekonującą interpretację liczb urojonych, która pokazuje, że są one tak samo „rzeczywiste" jak liczby zwykłe (i przy okazji wyjaśnia ich rolę w geometrii i trygonometrii). Bardziej konserwatywnie, można pokazać, jak wyeliminować liczby urojone z każdego argumentu, który ich używa, na rzecz liczb zwykłych.

Uświadomienie sobie urojenia

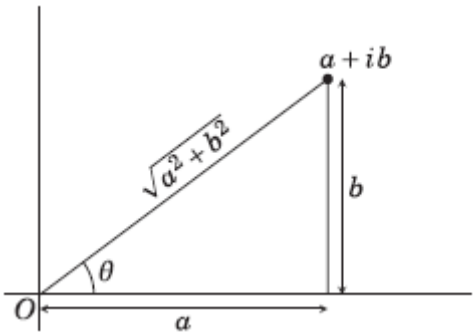
Prosty sposób na wyeliminowanie liczb urojonych z matematyki wprowadził Hamilton (1835): zastąp a + ib uporządkowaną parą (a,b) liczb rzeczywistych a,b i zdefiniuj sumę i iloczyn par za pomocą

$$
\begin{aligned}
(a_1,b_1)+(a_2,b_2) &= (a_1+a_2,b_1+b_2) \\
(a_1,b_1)\cdot(a_2,b_2) &= (a_1a_2-b_1b_2,a_1b_2+a_2b_1).
\end{aligned}
$$

Wówczas para dla sumy $(a_1 + ib_1) +(a_2 + ib_2)$ jest sumą par dla $a_1 + ib_11$ i $a_2 + ib_2$, a para dla iloczynu jest również iloczynem par. Wynika z tego, że każde stwierdzenie dotyczące sumy i iloczynu liczb w postaci a + ib jest równoważne stwierdzeniu dotyczącemu liczb rzeczywistych, na przykład

$$(a_1+ib_1)(a_2+ib_2)=c+id \Leftrightarrow a_1a_2-b_1b_2=c \text{ and } a_1b_2+a_2b_1=d.$$

Stąd każdy argument dotyczący i może być zastąpiony argumentem dotyczącym samych liczb rzeczywistych. Z tego powodu mówimy, że teoria liczb zespolonych (jak nazywane są liczby a + ib) jest konserwatywnym rozszerzeniem teorii liczb rzeczywistych. Jest „konserwatywna" w tym sensie, że każdy wynik dotyczący liczb rzeczywistych udowodniony przy użyciu i można udowodnić bez niego. Konstrukcja Hamiltona pokazuje, że założenie, że liczby urojone istnieją, jest nieszkodliwe, ale jednocześnie pokazuje, że nie ma potrzeby zakładania, że istnieją. Wszystko, co możemy z nimi zrobić, możemy zrobić bez nich, choć być może nie tak łatwo. Dla większości matematyków tym, co zmusza do wiary w liczby zespolone, jest to, że dają one więcej, niż żądaliśmy. To tak, jakby zawsze były częścią struktury matematyki, ale na początku zauważyliśmy tylko jedną małą nić w rozwiązaniu równań sześciennych. W rzeczywistości i daje rozwiązania nie tylko równań sześciennych, ale wszystkich równań wielomianowych. To jest podstawowe twierdzenie algebry, o którym powiemy więcej w rozdziale 5. Co więcej, chociaż i nie może leżeć na linii liczb rzeczywistych, to ma sens, aby leżało na prostopadłej linii liczb urojonych. Wystarczy wyobrazić sobie płaszczyznę liczb zespolonych, z a + ib reprezentowanym przez punkt (a,b) w odległości poziomej a od początku O i w odległości pionowej b. Zgodnie z tą interpretacją mnożenie przez a + ib ma znaczenie geometryczne, które jest szerokim uogólnieniem $i^2 = -1$. Mianowicie mnożenie przez a + ib powiększa płaszczyznę o $\sqrt{a^2+b^2}$ (odległość a + ib od O) i obraca ją wokół O o kąt $\theta = \tan^{-1}b/a$. Rysunek 10 przedstawia a + ib w jego kontekście geometrycznym.



W szczególności mnożenie przez i obraca płaszczyznę wokół O o kąt prosty, a mnożenie przez cos θ + i sin θ obraca ją o kąt θ. Tak więc mnożenie liczb zespolonych obejmuje całą podstawową geometrię i trygonometrię. A to dopiero początek. Jak wspomniano w poprzedniej sekcji,

$$e^{i\theta} = \cos\theta + i\sin\theta,$$

więc liczby zespolone łączą trygonometrię z funkcją wykładniczą. Krótko mówiąc, liczby zespolone są prawdopodobnie najpotężniejszą siłą unifikującą i upraszczającą w wyższej matematyce. Dlatego matematycy w nie wierzą.

Zagadnienia filozoficzne

Pojawienie się algebry w Indiach i świecie islamskim początkowo nie wpłynęło na filozofię matematyki. W rzeczywistości autorytet Euklidesa rządził algebrą aż do XVI wieku. Algebraicy islamscy uzasadniali swoje dowody odwołując się do diagramów geometrycznych, a takie dowody pojawiają się dopiero u Cardano (1545). Dopiero pod koniec XVI wieku, kiedy rozwinął się nowoczesny symbolizm algebraiczny, obliczenia algebraiczne stały się nową siłą w matematyce. Algebra symboliczna utorowała drogę innym gałęziom matematyki symbolicznej, takim jak rachunek różniczkowy i całkowy, które filozofia matematyki musiała ostatecznie uwzględnić. Niemniej jednak przez długi czas żywiono nadzieję, że Euklides pozostanie fundamentem matematyki.

Intuicja i logika. Algebraicy włoscy początkowo uzasadniali zasady algebry (podobnie jak ich islamscy poprzednicy) odwołując się do logiki geometrycznej. Jednak obliczenia Bombelliego z √-1 sugerowały, że algebra ma własną niezależną logikę.

Znaczenie i istnienie. Liczby urojone początkowo miały jedynie „symboliczne istnienie", co pozwalało na ich wykorzystanie w obliczeniach dających rzeczywiste wyniki. Stając się systemem reguł manipulowania symbolami, algebra zaczęła uwalniać się od swoich geometrycznych fundamentów. Jednak, co zaskakujące, liczby urojone dostarczają nowych spostrzeżeń na temat geometrii.

Dyskretne i ciągłe. Greckie przekonanie, że tylko liczby wymierne są rzeczywiście liczbami, stopniowo zanikało pod wpływem algebry, która nalegała na akceptację pierwiastków kwadratowych i sześciennych, oraz trygonometrii, która nalegała na akceptację funkcji sinus i cosinus. Jednak nie istniała jeszcze spójna teoria liczb rzeczywistych – istniało jedynie przekonanie, że można je modelować za pomocą punktów linii.

Rachunek różniczkowy i całkowy oraz nieskończenie małe

Podgląd

Wpływ algebry wzrósł w XVII wieku, najpierw w geometrii algebraicznej Fermata i Kartezjusza, a następnie w rachunku nieskończenie małych Newtona i Leibniza. W geometrii algebra szybko

rozwiązała starożytne problemy dotyczące linii i przekrojów stożkowych i dała łatwy dostęp do ogromnej klasy krzywych, których Grecy ledwo dotknęli. Algebraiczne podejście do geometrii stało się możliwe dzięki arytmetyzacji linii i płaszczyzny: identyfikowaniu punktów linii z liczbami rzeczywistymi, a punktów płaszczyzny z parami liczb rzeczywistych. Zgodnie z tą identyfikacją wiele krzywych można było opisać równaniami wielomianowymi, p(x,y) = 0, co pozwalało na wyodrębnianie własności geometrycznych poprzez manipulację algebraiczną. Rachunek różniczkowy i całkowy rozszerzył tę ideę, umożliwiając operacje algebraiczne na nieskończenie małych – wielkościach, które zachowywały się jak liczby niezerowe w obliczeniach, ale były w innym przypadku pomijalne. Na przykład nachylenie krzywej p(x,y) = 0 można obliczyć jako iloraz dy/dx nieskończenie małych dx i dy, gdzie dx przyjęto jako nieskończenie mały wzrost x, a dy jako odpowiadający mu wzrost y. Właściwości przypisywane nieskończenie małym były bliskie, jeśli nie w rzeczywistości, niespójne. Jednak, podobnie jak liczby urojone, nieskończenie małe wydawały się łatwe i bezpieczne w użyciu. Matematycy wierzyli, że jeśli zostaną wyzwani, będą mogli odtworzyć wyniki rachunku nieskończenie małych za pomocą bardziej rygorystycznej metody wyczerpania. Magia rachunku nieskończenie małych polegała na jego zdolności do zastępowania skomplikowanych argumentów wyczerpania rutynowymi obliczeniami, więc po raz kolejny wygoda pokonała wszelkie wątpliwości co do istnienia używanych obiektów matematycznych.

Szeregi nieskończone

Widzieliśmy już, że Euklides i Archimedes używali sum szeregów nieskończonych do znajdowania pewnych pól i objętości. Szeregi, o których mowa, były przykładami nieskończonych szeregów geometrycznych $a + ar + ar^2 + ar^3 + \ldots$; które mają sumę a/1-r, gdy $|r| < 1$. Wartość tę można rygorystycznie potwierdzić metodą wyczerpania. Odkrywamy, że szeregi skończone

$$a + ar + ar^2 + ar^3 + \ldots + ar^n = \frac{a - ar^{n+1}}{1 - r}$$

a ta skończona suma (dla $a > 0$ i $|r| < 1$) jest wyraźnie mniejsza niż a/1-r, ale może przekroczyć dowolną liczbę mniejszą niż a/1-r, ponieważ r^{n+1} można uczynić dowolnie małym, wybierając n wystarczająco duże. Dlatego też sumy skończone „wyczerpują" wszystkie liczby mniejsze niż a/1-r, a zatem suma nieskończona musi być równa a/1-r. Kiedy wynaleziono rachunek różniczkowy i całkowy, około 1665 r., szereg geometryczny był punktem wyjścia dla wielu innych wyników dotyczących szeregów nieskończonych. Jednak zanim wynaleziono rachunek różniczkowy i całkowy, w piętnastowiecznych Indiach odkryto niezwykłe wyniki dotyczące szeregów nieskończonych w trygonometrii. Głównym współautorem tych odkryć był Madhava (ok. 1340–ok. 1425 r.), a jego metody były w dużej mierze algebraiczne. Punktem wyjścia był ponownie szereg geometryczny, ale nowe szeregi były również pomysłowo wykorzystywane, zwłaszcza szereg

$$1^k + 2^k + 3^k + \ldots + n^k \text{ for } k = 1, 2, 3, \ldots .$$

Ta ostatnia seria odegrała rolę, którą później przejął rachunek różniczkowy i całkowy, dowodząc, że

$$\tan^{-1} x = x - \frac{x^3}{3} + \frac{x^5}{5} - \frac{x^7}{7} + \ldots \text{ for } -1 < x \leq 1.$$

Madhava odkrył również szereg dla funkcji sinus i cosinus:

$$\sin x = x - \frac{x^3}{3!} + \frac{x^5}{5!} - \frac{x^7}{7!} + \ldots$$

$$\cos x = 1 - \frac{x^2}{2!} + \frac{x^4}{4!} - \frac{x^6}{6!} + \ldots.$$

Ta ostatnia seria, a także pokrewne serie, zostały ponownie odkryte w Europie w XVII wieku i odegrały ważną rolę w rozwoju rachunku różniczkowego i całkowego. Niezależne odkrycie tych wyników w Indiach i Europie było prawdopodobnie najbardziej niezwykłym przykładem kulturowej uniwersalności matematyki od czasów twierdzenia Pitagorasa.

Geometria algebraiczna

Nieskończone procesy na liczbach były jednym z warunków wstępnych rachunku różniczkowego i całkowego. Innym było zastosowanie algebry do geometrii, w skrócie geometrii algebraicznej. To ostatnie stało się możliwe po ugruntowaniu się symboliki algebraicznej w XVI wieku, co pozwoliło na wykonywanie obliczeń z wielomianami z taką samą łatwością, jak z liczbami. W latach trzydziestych XVII wieku Fermat i Kartezjusz byli w stanie podać algebraiczne rozwiązanie problemu, który obecnie jest zwykle rozwiązywany przez rachunek różniczkowy i całkowy: znalezienie stycznych do krzywej algebraicznej. Konfiguracja tego problemu jest obecnie znana uczniom szkół średnich. Każdy punkt P na płaszczyźnie jest dany przez uporządkowaną parę (x,y) liczb, gdzie

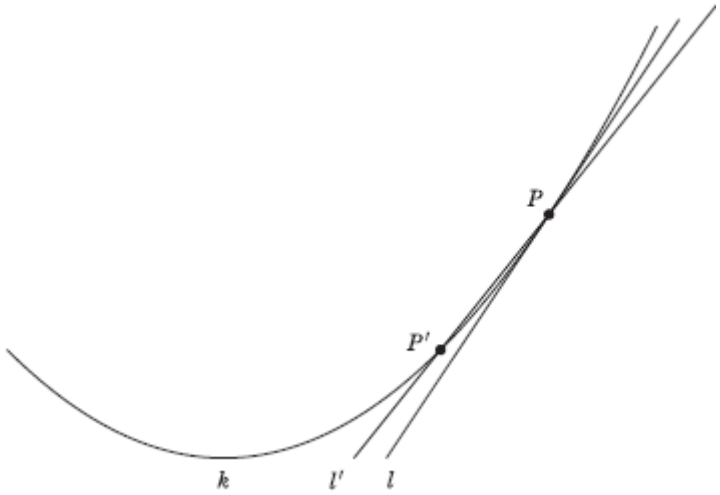
x = odległość pozioma do P od początku O;

y = odległość pionowa do P od początku O;

Krzywa algebraiczna to krzywa, której punkty spełniają równanie p(xy) = 0, gdzie p jest wielomianem. Na przykład punkty w odległości 1 od O spełniają $x^2 + y^2 = 1$, więc równanie okręgu jednostkowego to $x^2 + y^2 - 1 = 0$. Innym przykładem jest parabola $y = x^2$ lub $y - x^2 = 0$. Prowadzi to do klasyfikacji krzywych według stopnia wielomianu. Jeśli p jest stopnia 1 – czyli p(x,y) = ax +by - c – to p(x,y) = 0 jest równaniem linii. Jeśli p(x,y) jest stopnia 2, to Fermat i Kartezjusz odkryli (niezależnie), że p(x,y) = 0 jest jednym z przekrojów stożkowych badanych przez Greków. Oprócz przypadków zdegenerowanych, w których płaszczyzna przecinająca stożek spotyka go w punkcie lub liniach, są to elipsy, parabole i hiperbole. Typowe przykłady tych trzech typów pokazano na Rysunku 11.

Teraz styczna do krzywej k jest linią l, która spotyka k w punkcie „mnożenia”, w tym sensie, że istnieje wielokrotne rozwiązanie równania dla x, które wynika z podstawienia wyrażenia za y w l w równaniu p(x,y) = 0. Ten precyzyjny warunek algebraiczny uchwyca niejasną ideę, że l spotyka k w pojedynczym

izolowanym punkcie P, ale że każda linia l' przechodząca przez P i bliska l spotyka k w P i co najmniej jednym innym punkcie P' w pobliżu P. Rysunek 12 przedstawia tę sytuację.



„Podwójny punkt styku" P l jest granicą dwóch punktów styku P;P' l', gdy linia l' zbliża się do l. Na przykład linia y = 2x - 1 jest styczna do paraboli y = x^2 w punkcie x, ponieważ gdy podstawimy y = 2x - 1 w y = x^2 otrzymamy równanie 2x - 1 = x^2 lub x^2 - 2x + 1 = 0. Równanie to można zapisać jako $(x – 1)^2$ = 0, co pokazuje, że ma ono podwójne rozwiązanie x = 1.

Podobne obliczenia (choć naturalnie bardziej skomplikowane, jeśli krzywa ma wyższy stopień) pozwalają nam znaleźć styczne do dowolnej krzywej algebraicznej. Jednak znalezienie obszaru między krzywą algebraiczną a osią x jest trudniejszym problemem, nawet dla krzywej y = 1/x. To właśnie tutaj potrzeba rachunku różniczkowego staje się pilna.

Rachunek nieskończenie małych

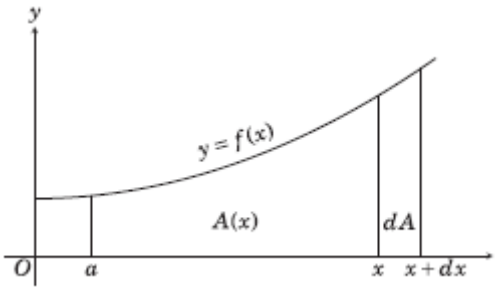
Algebraiczne kryterium styczności jest proste do sformułowania, choć może być trudne w użyciu. Bardziej poważnym zarzutem jest to, że dotyczy ono tylko krzywych algebraicznych, a niektóre krzywe fizycznie naturalne nie są algebraiczne. Jednym ze znanych przykładów jest łańcuch, kształt wiszącego łańcucha. Łańcuch wygląda raczej jak parabola, ale nią nie jest, więc potrzebna jest inna metoda, aby znaleźć jej styczne. W rzeczywistości nawet w przypadku krzywych algebraicznych pożądana jest prostsza metoda – taka, która znajduje nachylenie w dowolnym punkcie. Lepszą metodą jest rachunek różniczkowy, system, który powstał z algebry i geometrii hipotetycznych bytów zwanych nieskończenie małymi. Nieskończenie małe były używane przez wielu wczesnych wykładników rachunku różniczkowego, począwszy od lat trzydziestych XVII wieku i osiągnęły dojrzałość wraz z rachunkiem nieskończenie małych Leibniza w latach osiemdziesiątych XVII wieku. Leibniz wprowadził oznaczenia dx;dy;dz i tym podobne dla nieskończenie małych, które wydawały się być wielkościami mniejszymi od zwykłych liczb, ale niezerowymi. Na przykład nachylenie krzywej w punkcie (x,y) było postrzegane jako nachylenie dy/dx między (x,y) a punktem (x + dx, y + dy) na krzywej „nieskończenie bliskim" (x,y). Biorąc pod uwagę równanie y = f (x) dla krzywej, na ogół łatwo było obliczyć nachylenie dy/dx. Weźmy na przykład parabolę y= x^2. W tym przypadku,

$$dy = (x + dx)^2 - x^2 = 2x\ dx + (dx)^2,$$

więc

$$\frac{dy}{dx} = \frac{2x\ dx + (dx)^2}{dx} = 2x + dx.$$

Na tym etapie można swobodnie pominąć dx i wywnioskować, że nachylenie $y = x^2$ dla dowolnej wartości x wynosi 2x. (W szczególności, gdy x = 1 nachylenie wynosi 2, więc równanie stycznej w tym punkcie wynosi y = 2x - 1, jak stwierdzono w algebrze konwencjonalnej w poprzedniej sekcji). Podobne obliczenia z dx i dy łatwo dają nachylenie dowolnej krzywej algebraicznej, a zatem i styczną, w dowolnym punkcie krzywej. W szczególności nachylenie $y = x^n$ wynosi nx^{n-1}. Ale to tylko mała próbka magii infinitezymali. Umożliwiają one również obliczanie obszarów zakrzywionych, takich jak pole pod krzywą y = f (x). Aby to zrobić, należy rozpatrywać obszar jako funkcję A(x), biorąc obszar pomiędzy stałą wartością a i zmienną wartością x, jak na rysunku .



Nieskończenie mały wzrost dx w x powoduje nieskończenie mały wzrost dA w obszarze, co możemy zapisać

dA = f(x)dx

ponieważ dodatkowy pas dA powierzchni ma szerokość dx i wysokość różniącą się tylko nieskończenie małą od f(x). (Znów wybieramy wygodny moment, aby zaniedbać nieskończenie małe). Wnioskujemy, dzieląc obie strony przez dx, że

$$\frac{dA(x)}{dx} = f(x)$$

Innymi słowy, A(x) jest funkcją, której wykres ma nachylenie f(x). Zatem znalezienie obszarów pod krzywymi jest problemem odwrotnym do znalezienia nachyleń. Jeśli f(x) jest funkcją, którą już znaleźliśmy jako nachylenie dy/dx, to możemy wywnioskować, że funkcja pola A(x) jest taka sama jak y, przynajmniej w stałej. Na przykład, jeśli A(x) jest polem pod parabolą $y = x^2$ między 0 i x, to

$dA/dx = x^2$,

i możemy wywnioskować $A(x) = 1/3\ x^3$, ponieważ $dy/dx = x^2$, gdy $y = 1/3\ x^3$, a funkcje A(x) i y zgadzają się, gdy x = 0. To, w skrócie, jest rachunek nieskończenie małych Leibniza. Odwrotna relacja między problemami pola i stycznej nazywana jest jego podstawowym twierdzeniem. Newton odkrył podobny rachunek, choć bez wygodnej i sugestywnej notacji dx. Sztuka rachunku nieskończenie małych była zasadniczo algebrą i geometrią nieskończenie małych, powiązaną z pewnym dobrym osądem, kiedy „zaniedbywać" nieskończenie małe. (Na przykład, podziel przez dx, zanim je zaniedbasz!)

Nieskończenie małe: krytyka i unikanie

Nieskończenie małe były trochę jak liczby urojone. Wydawały się przeczyć przyjętym zasadom – tak jak liczby urojone przeczyły zasadzie, że kwadraty są dodatnie, nieskończenie małe przeczyły aksjomacie Archimedesa dla wielkości geometrycznych, podanemu w rozdziale 2.4 – mimo to umożliwiały obliczenia, które w innym przypadku byłyby trudne lub niemożliwe. Dla matematyków XVII i XVIII wieku był to zazwyczaj wystarczający powód, aby je zaakceptować. Ale w inny sposób nieskończenie małe nie były jak liczby urojone. Chociaż prawdą jest, że liczby urojone nie są częścią zwykłego systemu liczbowego, system ten można łatwo rozszerzyć, aby je uwzględnić – na przykład definiując liczby urojone jako uporządkowane pary zwykłych liczb, jak zrobił to Hamilton w 1835 roku. Rozszerzenie systemu liczbowego o nieskończenie małe nie jest wcale tak łatwe ani wygodne. Nie wiedziano nawet, że jest to możliwe aż do XX wieku, długo po tym, jak matematycy zdecydowali, że lepiej unikać nieskończenie małych, tak jak

Euklides i Archimedes unikali nieskończoności.

Kiedy rachunek nieskończenie małych był w powijakach, było dość łatwo, choć żmudnie, zastąpić nieskończenie małe argumenty metodą wyczerpania. Ale gdy nieskończenie małe algebra i geometria rosły w siłę, a wiara w ich poprawność rosła, stało się niewdzięcznym zadaniem przepisywanie argumentów w rygorystyczny, starożytny sposób. Już w 1659 roku Huygens napisał:

Matematycy nigdy nie będą mieli wystarczająco dużo czasu, aby przeczytać wszystkie odkrycia w geometrii (ilość, która rośnie z dnia na dzień i wydaje się prawdopodobne, że w tej naukowej epoce rozwinie się do ogromnych rozmiarów), jeśli będą one nadal przedstawiane w rygorystycznej formie zgodnie ze sposobem starożytnych

Filozofowie słusznie wyśmiewali koncepcję nieskończenie małych – Berkeley nazywał je „duchami minionych wielkości" – z powodu luźnego i czasami niespójnego sposobu, w jaki byli wykorzystywani przez matematyków. Ale matematycy nie zrezygnowali całkowicie z nieskończenie małych, dopóki nie zostali do tego zmuszeni z przyczyn matematycznych. A kiedy to zrobili, było to częścią ogólnej rewolucji w matematyce, która zastąpiła geometrię arytmetyką w podstawach matematyki. Zobaczymy, jak do tego doszło w następnych dwóch sekcjach.

Analiza zespolona

Rachunek różniczkowy i całkowy oraz liczby urojone, choć oba miały wątpliwe pochodzenie, utworzyły potężny sojusz w XVIII i XIX wieku. Najbardziej niezwykłymi sojusznikami były funkcje kołowe i wykładnicze, które badano oddzielnie w XVII wieku i stwierdzono, że można je wyrazić za pomocą szeregów nieskończonych. Około 1670 roku Newton odkrył szereg dla funkcji wykładniczej,

$$e^x = 1 + \frac{x}{1!} + \frac{x^2}{2!} + \frac{x^3}{3!} + \dots,$$

i on i inni odkryli na nowo szereg sinusów i cosinusów, które zostały już odkryte w Indiach:

$$\cos x = 1 - \frac{x^2}{2!} + \frac{x^4}{4!} - \frac{x^6}{6!} + \dots,$$

$$\sin x = \frac{x}{1!} - \frac{x^3}{3!} + \frac{x^5}{5!} - \frac{x^7}{7!} \dots.$$

Zastąpienie x przez ix w szeregu dla ex daje cudowny wzór

$$e^{ix} = \cos x + i \sin x,$$

odkryte przez Eulera (1748). Wzór ten nie tylko pozwala na wyrażenie sinusa i cosinusa za pomocą wykładników, a mianowicie

$$\cos x = \frac{1}{2}(e^{ix} + e^{-ix}), \quad \sin x = \frac{1}{2i}(e^{ix} - e^{-ix}),$$

ale zwraca również uwagę na ich „hiperboliczne" odpowiedniki:

$$\cosh x = \frac{1}{2}(e^{x} + e^{-x}), \ \sinh x = \frac{1}{2}(e^{x} - e^{-x}).$$

Ta analogia może pomóc wyjaśnić szalone przypuszczenie Lamberta (1766, §82). Lambert wprowadził funkcje hiperboliczne i bez wątpienia był świadomy wzorów trygonometrii sferycznej, które obejmują sinus i cosinus. Mógł wtedy zgadnąć, że analogiczne wzory obejmujące sinus hiperboliczny i cosinus opisują trygonometrię na „sferze o promieniu urojonym". Jeśli tak, może to wyjaśnić jego przypuszczenie z 1766 r., że geometria nieeuklidesowa może obowiązywać na sferze urojonej.

Zagadnienia filozoficzne

Zanim omówimy rachunek różniczkowy i całkowy, należy wspomnieć o ważnym produkcie ubocznym książki Kartezjusza (1637) o geometrii algebraicznej – mnożeniu długości. Jak wspomnieliśmy, istnieje naturalna suma długości, która sama w sobie jest długością, ale Euklides i jego następcy uznali iloczyn dwóch długości za pudełko, co poważnie ograniczyło jakąkolwiek algebrę długości. Istnieje jednak inny sposób zdefiniowania iloczynu długości, oparty na proporcjonalności trójkątów podobnych (rysunek 14). Aby to zrobić, wybieramy długość jednostkową, oznaczoną 1 i konstruujemy trójkąty podobne zawierające długości a i b w pokazanych pozycjach. Następnie wynika z proporcjonalności, że tak jak przedłużenie długości 1 jest długością b, przedłużenie długości a jest długością ab. Kartezjusz wykorzystał ten fakt do zdefiniowania iloczynu długości, nadając w ten sposób długościom strukturę algebraiczną podobną do struktury liczb. Wskazał również na prostą konstrukcję geometryczną (również obejmującą podobne trójkąty) dla pierwiastka kwadratowego z dowolnej długości. W ten sposób Kartezjusz (1637) zmiótł niektóre ze starożytnych problemów stawianych przez istnienie niewymiernych długości:

długości zachowują się jak liczby, a rozsądne było, aby pierwiastek kwadratowy z dowolnej liczby był liczbą. Ale gdy tylko usunięto jedną trudność filozoficzną, rachunek różniczkowy i całkowy stworzył kolejną. W rzeczywistości wynalezienie rachunku różniczkowego i całkowego zakłóciło filozofię matematyki być może bardziej niż jakiekolwiek wydarzenie od czasu odkrycia niewymiernych wielkości. Filozofowie słusznie kwestionowali koncepcję nieskończenie małych, ale matematycy początkowo ignorowali ich krytykę i nadal wierzyli, że mogą uzyskać wyniki rachunku różniczkowego i całkowego za pomocą metod Euklidesa (choć rzadko to faktycznie robili). Powstał impas, który nie został przełamany aż do roku 1800, kiedy matematycy musieli przyznać, że Euklides nie był, mimo wszystko, odpowiednim fundamentem dla matematyki.

Intuicja i logika. Intuicja odgrywała wiodącą rolę w rachunku różniczkowym (i nadal odgrywa), gdzie „nieskończenie bliskie punkty na krzywej" i „nieskończenie cienkie paski" były używane do tworzenia równań do obliczania stycznych i powierzchni. Jednak gdy równania zostały znalezione, siła symboliki algebraicznej (w algebrze nieskończenie małych) przeważyła. Matematycy wierzyli w rachunek różniczkowy ze względu na jego niesamowity sukces w rozwiązywaniu problemów z geometrii i

mechaniki. Uważali również, że stoją na solidnym gruncie, wierząc, że wszystkie ich wyniki można uzyskać rygorystycznie metodą wyczerpania.

Znaczenie i istnienie. Jednak, jak zauważył Berkeley, istnienie nieskończenie małych było wysoce wątpliwe, więc co wyjaśniało ich sukces? (Przed Berkeleyem Hobbes ostro krytykował rachunek różniczkowy i całkowy oraz stosowanie algebry w geometrii. Zniszczył jednak swoją wiarygodność wśród matematyków, proponując nie do utrzymania teorię okręgu – twierdząc, że zawiera on tylko skończenie wiele punktów – i twierdząc, że w ten sposób rozwiązuje starożytny problem „kwadratury okręgu".)

Dyskretne i ciągłe. Z drugiej strony, nieskończenie małe liczby stanowiły nowy pomost (choć chwiejny) między dyskretnym a ciągłym. Obraz kontinuum był niejasny i być może sprzeczny, ale można było poprawnie obliczyć długości i pola. Więc może nieskończenie małe liczby mogłyby pomóc wyjaśnić naturę kontinuum? W dłuższej perspektywie okazało się, że problemy podnoszone przez nieskończenie małe liczby to problemy dotyczące natury kontinuum, a w szczególności jego natury jako szczególnego rodzaju nieskończonego zbioru. W kolejnych dwóch sekcjach zobaczymy, jak rozwinął się ten problem i jego filozoficzne implikacje.

Funkcje ciągłe i liczby rzeczywiste

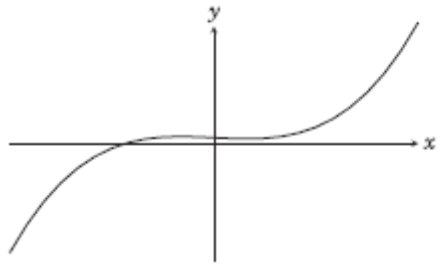
PODGLĄD

Rachunek różniczkowy i całkowy mógł zajmować się wieloma konkretnymi funkcjami, ale ogólne pojęcie funkcji pozostało niejasne. Około 1800 roku konieczne stało się udowodnienie pewnych ogólnych własności funkcji ciągłych – co zaskakujące, w celu udowodnienia podstawowego twierdzenia algebry. Po kilku wątpliwych próbach udowodnienia tego twierdzenia w XVIII wieku Gauss zaproponował kilka dowodów. Najbardziej przekonujący był jeden z nich autorstwa Gaussa (1816), który sprowadził twierdzenie do szczególnego przypadku wielomianów nieparzystych stopni. Taki wielomian p(x) przyjmuje wartości, które zmieniają się „ciągle" od dodatnich (dla dużego dodatniego x) do ujemnych (dla dużego ujemnego x). Wydaje się wtedy oczywiste, że wielomian przyjmuje gdzieś wartość 0, co jest wymagane w przypadku podstawowego twierdzenia. Bolzano (1817) położył palec na kluczowym założeniu dowodu Gaussa, twierdzeniu o wartości pośredniej dla funkcji ciągłych: jeśli f (x) jest funkcją, która zmienia się w sposób ciągły od ujemnej do dodatniej, gdy x się zmienia, to f (x) = 0 dla pewnej wartości x. Bolzano był w stanie podać zadowalającą definicję ciągłości, ale aby udowodnić twierdzenie o wartości pośredniej, musiał założyć własność liczb rzeczywistych, własność najmniejszej górnej granicy: jeśli S jest ograniczonym zbiorem liczb rzeczywistych, to S ma najmniejszą górną granicę. Nie było możliwe udowodnienie własności najmniejszej górnej granicy z geometrycznej koncepcji liczby rzeczywistej Euklidesa, która do tej pory była uważana za wystarczającą. W końcu trzeba było zmierzyć się z niedokończonym zadaniem pogodzenia dyskretności z ciągłością

Podstawowe twierdzenie algebry

Widzieliśmy , że istnieje wzór na rozwiązywanie równań sześciennych, opublikowany po raz pierwszy przez Cardano w 1545 r. W tej samej książce Cardano opublikował również wzór na rozwiązywanie równań kwartalnych (czwartego stopnia). Oba wzory wyrażają rozwiązanie za pomocą współczynników równania, operacji wymiernych +; -; x; ÷; oraz pierwiastków kwadratowych i sześciennych. Wzbudziło to nadzieję na podobne rozwiązania równań wyższego stopnia – rozwiązania przez pierwiastki (n-tego stopnia), jak je nazywano. Nadzieja ta została ostatecznie rozwiana w latach dwudziestych XIX wieku przez Abela i Galois, którzy wykazali, że takie wzory nie istnieją dla równań piątego stopnia i wyższych. Ale zanim rozwiązanie przez pierwiastki zostało wykluczone, Gauss zasugerował już inne podejście do równań wielomianowych: zadowalać się udowodnieniem istnienia rozwiązania, zamiast znajdować je

we wzorze. Około 1800 roku Gauss przedstawił kilka dowodów w tym stylu, których kluczowym składnikiem było odwołanie się do własności funkcji ciągłych. Uproszczenie możliwe dzięki dowodowi istnienia można zilustrować w przypadku równania sześciennego, powiedzmy $x^3 - x + 2 = 0$. Jeśli przyjrzymy się wykresowi $y = x^3 - x + 2$



, od razu stanie się jasne, dlaczego $x^3 - x + 2 = 0$ dla pewnej wartości x. Wykres przechodzi nieprzerwanie od wartości ujemnych (dla dużych ujemnych wartości x) do wartości dodatnich (dla dużych dodatnich wartości x), a zatem gdzieś przyjmuje wartość 0. Ten sam argument dotyczy dowolnego wielomianu nieparzystego stopnia. W 1816 roku Gauss rozszerzył go na wielomiany dowolnego stopnia za pomocą pewnej pomysłowej algebry, która redukuje rozwiązanie równania stopnia 2n do rozwiązania równania kwadratowego (gdzie rozwiązanie może być zespolone) i równania stopnia n. Wielokrotne dzielenie stopnia na pół prowadzi do równania o nieparzystym stopniu, a zatem istnieje rozwiązanie. Tak więc argument Gaussa (1816) redukuje istnienie rozwiązań dowolnego równania wielomianowego – podstawowego twierdzenia algebry, jak się je obecnie nazywa – do pewnej algebry plus pozornie oczywistego faktu dotyczącego funkcji ciągłych. Gauss chętnie zakładał fakty dotyczące ciągłości, ale w rzeczywistości stał na skraju nowego świata rzeczywistej analizy, w którym w centrum uwagi znajdowała się natura i właściwości funkcji ciągłych.

Twierdzenie o wartości pośredniej

Co ciekawe, dowód Gaussa (1816) został natychmiast zauważony, a jego kluczowa idea dotycząca funkcji ciągłych została sformułowana jako twierdzenie przez Bolzano (1817). To właśnie nazywamy obecnie twierdzeniem o wartości pośredniej.

Twierdzenie o wartości pośredniej. Jeśli f jest funkcją ciągłą, zdefiniowaną dla x, dla którego $a \le x \le b$, i $f(a) < 0$ i $f(b) > 0$, to $f(c) = 0$ dla pewnego c między a i b.

Aby sformułować twierdzenie, Bolzano musiał zdefiniować, co oznacza „ciągły”, i stworzył zasadniczo nowoczesną definicję: f jest ciągła w x, jeśli dla każdego $\varepsilon > 0$ istnieje $\delta > 0$ takie, że $|x - x'| < \delta$ implikuje $|f(x) - f(x')| < \varepsilon$. (Mówiąc mniej formalnie: możemy sprawić, by f(x’) było tak bliskie f(x), jak chcemy, wybierając x’ wystarczająco blisko x.) Wtedy f jest po prostu ciągłe dla x od a do b, jeśli f jest ciągłe w x, gdy $a \le x \le b$. Nawet gdy jest to sformułowane nieformalnie, nie jest jasne, czy ta definicja mówi, co ma oznaczać „ciągły”. Wolelibyśmy powiedzieć „wykres y = f(x) jest nieprzerwany” lub coś w tym stylu – w rzeczywistości coś w rodzaju twierdzenia o wartościach pośrednich. Nadzwyczajny wgląd Bolzano doprowadził go do definicji, która pozwala, aby niejasna globalna własność nieprzerwalności wynikała z dokładnej lokalnej własności ciągłości w punkcie. Bolzano z pewnością miał właściwą definicję ciągłości, aby udowodnić twierdzenie o wartościach pośrednich, ale jego dowód zależał od nieudowodnionej własności liczb. Była to zasada najmniejszej górnej granicy, stwierdzająca, że każdy ograniczony zbiór liczb ma najmniejszą górną granicę. Bolzano mógł podać jedynie mgliste uzasadnienie tej zasady, ponieważ miał jedynie mglistą, geometryczną koncepcję linii liczbowej. Nie

było możliwe pójście dalej, dopóki koncepcja liczby nie została zdefiniowana w sposób czysto arytmetyczny.

Definicja liczb rzeczywistych

Kilka dekad po Bolzano, którego praca nie wzbudziła większego zainteresowania, Dedekind doświadczył podobnego niezadowolenia z intuicyjnych argumentów w rachunku różniczkowym i całkowym. W swojej broszurze z 1872 r. Continuity and Irrational Numbers napisał:

Jako profesor w Szkole Politechnicznej w Zurychu po raz pierwszy poczułem się zobowiązany do wygłoszenia wykładu na temat elementów rachunku różniczkowego i bardziej niż kiedykolwiek wcześniej odczułem brak prawdziwie naukowych podstaw dla arytmetyki. Dyskutując na temat pojęcia zbliżania się zmiennej wielkości do ustalonej wartości granicznej, a zwłaszcza dowodząc twierdzenia, że wielkość, która rośnie w sposób ciągły, ale nie poza wszelkie granice, musi z pewnością zbliżać się do wartości granicznej, uciekałem się do argumentów geometrycznych… dokładniejsze badanie przekonało mnie, że twierdzenie to lub jakiekolwiek inne równoważne mu twierdzenie można w jakiś sposób uznać za wystarczającą podstawę do analizy nieskończenie małej. Pozostało mi tylko odkryć jego prawdziwe pochodzenie w elementach arytmetyki i w ten sposób jednocześnie zabezpieczyć rzeczywistą definicję istoty ciągłości. Udało mi się to 24 listopada 1858 r.

Przez „ciągłość" Dedekind rozumie to, co obecnie nazywamy spójnością liczb rzeczywistych; nieformalnie, że „nie mają luk". Pod tym względem kontrastują z liczbami wymiernymi, które mają lukę w pozycji √2 i wielu innych miejscach. Czym dokładnie jest luka w liczbach wymiernych? Jest to podział liczb wymiernych na dwa zbiory L i U, tak że każdy element U jest większy niż każdy element L; L nie ma największego elementu, a U nie ma najmniejszego elementu. Luka w √2, na przykład, ma górny zbiór U składający się ze wszystkich dodatnich liczb wymiernych z kwadratem większym niż 2, a dolny zbiór L składający się ze wszystkich pozostałych liczb wymiernych. Ponieważ √2 nie jest wymierne, L nie ma największego elementu, a U nie ma najmniejszego. Ogólnie rzecz biorąc, każda liczba niewymierna odpowiada luce w liczbach wymiernych, którą Dedekind nazwał cięciem. Ta idea jest wyraźnie podobna do traktowania wielkości niewymiernych w Księdze V Elementów Euklidesa , z tym wyjątkiem, że nieskończone zbiory nie są już unikane. Śmiałym, ale prostym pomysłem Dedekinda było używanie nieskończonych zbiorów jako obiektów matematycznych. „Luka w liczbach wymiernych" jest wówczas sensownym obiektem matematycznym – parą zbiorów L; U z powyższymi własnościami – którą możemy przyjąć, aby zdefiniować liczbę niewymierną. Tak więc liczby wymierne i niewymierne razem tworzą system liczbowy bez luk, który obecnie nazywamy systemem liczb rzeczywistych $\mathbb{R}$. Opierając liczby rzeczywiste na liczbach wymiernych w ten sposób, Dedekind odkrył, że ich spójność ma swoje „prawdziwe źródło w elementach arytmetyki". Co więcej, łatwo jest pokazać, że algebraiczne własności liczb wymiernych (takich jak $a + b = b + a$ i $ab = ba$) przenoszą się na liczby rzeczywiste w sposób naturalny. A stosując koncepcję zbioru, Dedekind odkrył, że równie łatwo jest udowodnić zasadę najmniejszej górnej granicy Bolzano, a zatem zapewnić arytmetyczne podstawy dla analizy rzeczywistej. Aby udowodnić zasadę najmniejszej górnej granicy, najpierw przedstawiamy każdą liczbę rzeczywistą x za pomocą zbioru L_x liczb wymiernych, który jest ograniczony z góry i „zamknięty z dołu": to znaczy z własnością, że jeśli $r \in L_x$ i $s < r$, to $s < L_x$. Dla liczby niewymiernej x, L_x jest L w parze L; U, która definiuje x; dla liczby wymiernej x przyjmujemy, że L_x jest zbiorem liczb wymiernych $\leq x$. Ta reprezentacja ma wygodną własność, że uporządkowanie liczb rzeczywistych odpowiada zawieraniu zbioru: mianowicie, $x \leq y$ wtedy i tylko wtedy, gdy $L_x \subseteq L_y$. Wtedy, jeśli mamy ograniczony zbiór liczb rzeczywistych x, zbiory L_x są również ograniczone, a zatem ich suma L również. Wynika z tego, że liczba rzeczywista l określona przez zbiór L jest najmniejszą górną granicą liczb x.

Kontrintuicyjne krzywe

Wypędzając intuicję geometryczną z podstaw analizy, Bolzano i Dedekind umożliwili badanie i dokładną analizę obiektów kontrintuicyjnych. Dla niektórych matematyków był to szokujący rozwój sytuacji i „odwrócili się z przerażeniem i obrzydzeniem od tej strasznej zarazy" (parafrazując list Hermite'a do Stieltjesa z 1893 r.). Ale dla innych było to mile widziane, oferując nie tylko rygor, ale także sposób na rzeczywiste wyostrzenie intuicji i zobaczenie rzeczy, których naiwna intuicja nie potrafiła. Stary formalizm nieskończenie małych nie był wystarczająco ostry, ponieważ sugeruje, że ciągła funkcja y = f(x) to taka, dla której nieskończenie mała zmiana dx w x powoduje nieskończenie małą zmianę dy w y. Notacja ta prowadzi nas następnie do przekonania, że dy=dx istnieje; to znaczy, że wykres ciągłej funkcji y = f(x) ma nachylenie w każdym punkcie. Rzeczywiście, Newton i inni twórcy rachunku różniczkowego i całkowego wierzyli w to samo. Ale się mylili: Bolzano i inni matematycy XIX wieku odkryli, że istnieją ciągłe krzywe bez stycznej w żadnym punkcie i inne kontrintuicyjne właściwości. Piękny przykład krzywej bez stycznych, uzyskanej jako granica ciągu krzywych wielokątnych, podał von Koch (1904). Myślę, że jest „intuicyjne", że ten ciąg ma ciągłą krzywą graniczną. Ale można też „zobaczyć", że nie ma stycznych. Z jego konstrukcji jasno wynika, że krzywą graniczną można podzielić na cztery części (odpowiadające czterem odcinkom linii na drugim rysunku), z których każda wygląda dokładnie tak samo jak cała krzywa po powiększeniu przez 3. Ale gdyby krzywa miała styczną w dowolnym punkcie, otoczenie tego punktu stałoby się prostsze po powiększeniu. Wiele innych przykładów dziwacznych obiektów, które intuicja może pojąć, można znaleźć w książce Vilenkina In Search of Infinity (1995). Pokazują one, że nasza geometryczna intuicja jest zdolna do znacznie więcej, niż sądzili matematycy przed XIX wiekiem. Niemniej jednak w XIX wieku pojawiły się inne wyzwania dla intuicji, jak zobaczymy w następnej sekcji.

Zagadnienia filozoficzne

Dowód twierdzenia o wartościach pośrednich to konstrukcja nieskończona, która obejmuje punkt c, gdzie f(c) = 0, poprzez wielokrotne dzielenie na pół przedziału, w którym powinno znajdować się c. Nie różni się to aż tak bardzo od klasycznego dowodu przez wyczerpanie (zakładając istnienie najmniejszych górnych granic), z wyjątkiem tego, że zależy od „danej" funkcji f. W czasach Bolzano nie było jasne, co to oznacza, co wywołało podejrzenie, że twierdzenie o wartościach pośrednich jest czysto twierdzeniem o istnieniu, w którym obiekt c jest rzekomo istniejący, ale nie skonstruowany.

W rzeczywistości konstrukcja c nie stanowi problemu, gdy f jest wielomianem, jak w podstawowym twierdzeniu algebry. Jednak później pojawiło się więcej twierdzeń o istnieniu dotyczących funkcji ciągłych, które stały się kością niezgody z „konstruktywistami" pod koniec XIX wieku. W każdym razie bardziej problematyczną częścią dowodu Bolzano jest zdefiniowanie liczb rzeczywistych w taki sposób, aby zagwarantować najmniejszą własność górnej granicy, a także ich strukturę algebraiczną: krótko mówiąc, arytmetyzacja linii. Doprowadziło to do wybuchu nieprzewidzianych problemów filozoficznych, jak zobaczymy w dwóch następnych sekcjach. Ale w bezpośrednim następstwie Bolzano i Dedekinda pojawiło się już kilka problemów.

Intuicja i logika. Wbrew intuicji, że algebra jest dyskretna, podstawowe twierdzenie algebry wydaje się obejmować ciągłość. A intuicja dotycząca krzywych ciągłych (i rachunku różniczkowego w ogóle) wymaga głębszego fundamentu w arytmetycznej teorii liczb rzeczywistych.

Znaczenie i istnienie. Co oznacza udowodnienie istnienia, nie podając wzoru na obiekt, o którym twierdzi się, że istnieje? Czym dokładnie są liczby rzeczywiste i co wyjaśnia ich zupełność; to znaczy ich domknięcie w różnych nieskończonych operacjach? (Definicja Dedekinda daje jedną odpowiedź; czy istnieją alternatywy?)

Dyskretne i ciągłe. W szczególności, jak najlepiej możemy – unikając wątpliwych środków nieskończenie małych – zdefiniować ciągłe (liczby rzeczywiste) w kategoriach dyskretnych (liczb naturalnych)?

Od geometrii nieeuklidesowej do arytmetyki

PODGLĄD

Tymczasem Euklidesowi brakowało czegoś innego. Przez setki, jeśli nie tysiące lat, aksjomat równoległości był uważany za niepotrzebną skazę na systemie Euklidesa. Wielu matematyków próbowało udowodnić go za pomocą innych aksjomatów Euklidesa, ale do 1800 roku nadzieje zaczęły słabnąć. Niektórzy zaczęli rozważać coś wcześniej nie do pomyślenia: geometrię nieeuklidesową, w której aksjomat równoległości był fałszywy. Ta dziwna, ale cudowna geometria była początkowo badana tylko hipotetycznie. Ale Beltrami (1868) znalazł jej modele, pokazujące, że geometria nieeuklidesowa jest równie spójna jak euklidesowa. Dzięki temu odkryciu geometria euklidesowa straciła swoją uprzywilejowaną pozycję u podstaw matematyki (i jako domniemana geometria przestrzeni fizycznej). Nadszedł czas na zbudowanie nowego fundamentu matematyki, a arytmetyka była gotowa zająć miejsce geometrii. Dedekind znalazł arytmetyczną definicję osi liczb rzeczywistych $\mathbb{R}$, na której można było zbudować nowy fundament geometrii i analizy: rzeczywiste przestrzenie wektorowe Grassmanna.

Aksjomat równoległości

Jak widzieliśmy w poprzedniej sekcji, matematycy w pierwszej połowie XIX wieku stracili wiarę w intuicję geometryczną na poziomie mikro, gdzie nie udało się jej uchwycić lokalnej natury linii. W mniej więcej tym samym okresie nastąpiła również utrata wiary w intuicję geometryczną na poziomie makro, gdzie nie udało się jej uchwycić globalnej natury linii – w szczególności zachowania równoległości. Sekcja wcześniejsza pokazała, gdzie leży problem: w aksjomacie równoległości Euklidesa, który implicite twierdzi, że istnieją i niepowtarzalność równoległości. Aksjomat równoległości był bardzo użytecznym aksjomatem, używanym w dowodzie wielu sygnaturowych twierdzeń geometrii Euklidesa, takich jak twierdzenie Pitagorasa. W rzeczywistości nawet istnienie kwadratów zależy od aksjomatu równoległości, ponieważ do udowodnienia, że suma kątów trójkąta wynosi π, a zatem, że suma kątów czworokąta wynosi 2π, potrzebne są niepowtarzalne równoległości. Inną ważną konsekwencją aksjomatu równoległości, która dobrze współgra z naszym doświadczeniem, jest to, że figury dowolnej wielkości mogą mieć ten sam kształt. Dla wielu takie twierdzenia były bardziej akceptowalne niż sam aksjomat równoległości i znaleziono dowody, że implikują aksjomat równoległości. Ostateczną nadzieją było to, że inne aksjomaty Euklidesa – które są z pewnością bardziej prawdopodobne niż aksjomat równoległości – mogą zostać uznane za implikujące aksjomat równoległości, tak aby można było usunąć tę „skazę" na Euklidesie. Inne aksjomaty Euklidesa mówią bezsporne rzeczy, takie jak: istnieje jedyna prosta przechodząca przez dowolne dwa punkty, proste są nieskończone, a jeśli dwa trójkąty są zgodne co do dwóch boków i kąta zawartego, to są zgodne co do wszystkich boków i wszystkich kątów. Jeśli weźmie się inne aksjomaty i doda aksjomat, że równoległe nie są jedyne, nadzieja była taka, że zostanie odkryta sprzeczność. Pierwszym, który zajął się nieuchwytną sprzecznością, był Saccheri (1733) w swojej książce Euclid Vindicated from Every Blemish. Nie było zaskoczeniem, że odkrył, że aksjomaty te implikują, że suma kątów trójkąta jest mniejsza niż π, a suma kątów czworokąta jest mniejsza niż 2π. Odkrył również, co było jeszcze bardziej zaskakujące, że dwie linie mogą być asymptotyczne; to znaczy, że mogą się one dowolnie zbliżać do siebie, nie spotykając się. Saccheri uznał to za „odrażające dla natury linii prostych", ale mimo to nie było to sprzecznością. W rzeczywistości było to dokładne spojrzenie na świat nieeuklidesowy.

Geometria nieeuklidesowa

W pierwszych dekadach XIX wieku matematycy zaczęli badać hipotetyczny świat nieeuklidesowy mniej sceptycznie. Sam Gauss zdawał sobie sprawę z implikacji niejednoznacznych paraleli, ale bał się je opublikować. W latach dwudziestych XIX wieku Bolyai i Lobachevsky samodzielnie opracowali i opublikowali podstawowe twierdzenia. W latach trzydziestych XIX wieku w pracy Mindinga pojawiła się również wskazówka, że wzory trygonometrii nieeuklidesowej mają sens na niektórych powierzchniach w kształcie siodła. Te wzory, które są podobne do wzorów trygonometrii sferycznej, z tą różnicą, że sinus i cosinus są zastąpione ich hiperbolicznymi analogami, są prawdziwe, jeśli „linie" zostaną uznane za geodezyjne – krzywe o najkrótszej długości – na powierzchni. Jednak odkrycie Mindinga nie było pełną realizacją geometrii nieeuklidesowej, ponieważ jego powierzchnie były niekompletne – nie dopuszczały nieskończonych „linii" w każdym kierunku. Trudności modelowania geometrii nieeuklidesowej zostały ostatecznie przezwyciężone przez Beltramiego (1868) poprzez złagodzenie definicji „odległości". W rzeczywistości Beltrami znalazł kilka realizacji lub modeli geometrii nieeuklidesowej, w których „linie" były całkiem eleganckie i naturalne. Być może najłatwiejszy do zrozumienia jest model dysku konforemnego pokazany na rysunku 17. Na tym rysunku „płaszczyzna" jest wnętrzem dysku, więc jego „punkty" są punktami wewnątrz granicy dysku. Jego „linie" są łukami kołowymi prostopadłymi do granicy dysku, a „kąty" są rzeczywistymi kątami między „liniami". (Dlatego model ten nazywa się konforemnym, co oznacza, że wiernie przedstawia kąty.) Koncepcja odległości jest za pośrednictwem pewnego wzoru, którego nie podam, ale można uzyskać ogólne wrażenie „odległości" w następujący sposób. Na rysunku przedstawiono wiele „trójkątów", z których każdy ma ten sam kształt, ponieważ ma kąty $\pi/2;\pi/3;\pi/7$ (tak więc, jak wszystkie trójkąty nieeuklidesowe, mają sumę kątów $< \pi$). Ponieważ figury o tym samym kształcie mają ten sam rozmiar w geometrii nieeuklidesowej, wszystkie te trójkąty mają ten sam nieeuklidesowy rozmiar. W szczególności widać, że płaszczyzna nieeuklidesowa i jej „linie" są nieskończone, ponieważ każda linia przechodzi przez nieskończenie wiele trójkątów na swojej drodze do granicy dysku. Można również zgadnąć, że „linie" są krzywymi o najkrótszej „długości", ponieważ wydaje się, że „linia" daje najkrótszą ścieżkę między dowolnymi dwoma punktami, jeśli mierzy się ją liczbą trójkątów, przez które przechodzi.

Wpływ geometrii nieeuklidesowej

Geometria nieeuklidesowa ostatecznie zabiła wszelkie szanse na wydedukowanie aksjomatu równoległości z innych aksjomatów Euklidesa, ponieważ inne aksjomaty obowiązują w modelu Beltramiego, ale aksjomat równoległości nie. Oczywiście istnieją modele, w których aksjomat równoległości obowiązuje, więc widzimy, że aksjomat jest niezależny od innych aksjomatów i można go zastąpić nieeuklidesowym aksjomatem równoległości, stwierdzającym istnienie wielu równoległości, bez obawy o sprzeczność. Bezpośrednim skutkiem tego odkrycia – pierwszego

dowodu niezależności w matematyce – była zwiększona nieufność do geometrii Euklidesa jako podstawy matematyki i zwiększone poparcie dla arytmetyki. W rzeczywistości, jak zobaczymy w następnej sekcji, arytmetyczne podejście do geometrii oparte na rzeczywistych przestrzeniach wektorowych było gotowe do wdrożenia, chociaż niewielu matematyków było gotowych je zrozumieć. Z pewnością arytmetyzacja analizy zyskiwała poparcie. Dzięki wpływowi Weierstrassa, który zaproponował arytmetyczną definicję liczb rzeczywistych niezależnie od Dedekinda w 1864 r., dowody arytmetyczne podstawowych twierdzeń o funkcjach ciągłych (a zatem i podstawowego twierdzenia algebry) zaczęły krążyć w latach 70. XIX wieku. Innym niezwykłym odkryciem, dokonanym przez Poincarégo i Kleina w latach 80. XIX wieku, było to, że geometria nieeuklidesowa była już obecna w matematyce. Zauważyli, że w zachowaniu funkcji na jednostkowym dysku liczb zespolonych występuje okresowość nieeuklidesowa i że zjawisko to zostało zaobserwowane wcześniej przez Gaussa, Riemanna i Schwarza, nie zdając sobie sprawy z jego znaczenia geometrycznego. Schwarz (1872) stworzył nawet

diagramdysku, który nie jest niczym innym, jak teselacją modelu dysku konforemnego przez przystające trójkąty nieeuklidesowe! Odkrycia te ugruntowały pozycję geometrii nieeuklidesowej w matematyce, a także w arytmetyce liczb zespolonych. Poza geometrią niezależność aksjomatu równoległości zapowiadała inne dowody niezależności w matematyce. W przypadkach, gdy nie wiadomo, że zdanie σ jest konsekwencją aksjomatów P , można mieć nadzieję na udowodnienie niezależności σ od P poprzez znalezienie dwóch modeli P : jednego spełniającego σ i drugiego spełniającego negację σ. Przykłady podamy później. Niezależność aksjomatu równoległości była również zwiastunem matematyki odwrotnej, gałęzi, w której poszukuje się „właściwego aksjomatu", aby udowodnić dane twierdzenie. Aksjomat równoległości jest właściwym aksjomatem do udowodnienia, powiedzmy, twierdzenia Pitagorasa, ponieważ oba są równoważne – biorąc pod uwagę inne aksjomaty Euklidesa – ale żadnego z nich nie można udowodnić na podstawie samych tych aksjomatów. Rzeczywiście, aksjomat równoległości jest właściwym aksjomatem do udowodnienia całej masy twierdzeń, które okazały się równoważne aksjomatowi równoległości, zanim poznano jego niezależność. Jak zobaczymy w rozdziale 9, zjawisko niezależności stało się wielką historią w matematyce XX wieku, zmieniając sposób, w jaki myślimy o prawdzie i dowodzie.

Arytmetyzacja geometrii

W mało znanej publikacji zatytułowanej Die Ausdehnungslehre (teoria rozszerzeń) Grassmann (1844) wprowadził rewolucyjne podejście do geometrii, oparte na tym, co obecnie nazywamy przestrzeniami wektorowymi rzeczywistymi. Niestety styl Grassmanna był nieprzenikniony dla jego współczesnych, nawet gdy opublikował uproszczoną wersję w 1847 r., kładąc nacisk na rolę iloczynu skalarnego. Pierwszym, który docenił jego pracę, był Peano (1888), który przedstawił system aksjomatów dla przestrzeni wektorowych rzeczywistych, jeden z pierwszych nowoczesnych systemów aksjomatów. Co ciekawe, Grassmann przypisał zalążek pomysłu szkicowemu rękopisowi Leibniza z 1679 r. zatytułowanemu Characteristica Geometrica. Jeśli Grassmann naprawdę dostrzegł przestrzenie wektorowe w Characteristica, był jedynym, który to zrobił (inni uważali, że Characteristica zapowiada zupełnie inny przedmiot, jakim jest topologia kombinatoryczna). Przynajmniej Grassmann zasługuje na uznanie za rozpoznanie koncepcji przestrzeni wektorowej i za to, że zauważył, co należy dodać, aby stała się ona prawdziwie geometryczna. Dzisiaj przestrzenie wektorowe widzimy wszędzie w matematyce, więc większość matematyków zna podstawowe idee. Istnieje zbiór V obiektów zwanych wektorami, które można mnożyć przez liczby rzeczywiste i dodawać. Obejmują one wektor zerowy 0 i dla każdego wektora a wektor a zwany przeciwieństwem a. Wektory są rządzone przez następujące aksjomaty. Pierwsza grupa mówi, że suma wektorów þ zachowuje się jak zwykła suma; to znaczy dla dowolnego u; v; w w V:

$$\begin{aligned} \boldsymbol{u}+\boldsymbol{v} &= \boldsymbol{v}+\boldsymbol{u}, \\ \boldsymbol{u}+(\boldsymbol{v}+\boldsymbol{w}) &= (\boldsymbol{u}+\boldsymbol{v})+\boldsymbol{w} \\ \boldsymbol{u}+\mathbf{0} &= \boldsymbol{u}, \\ \boldsymbol{u}+(-\boldsymbol{u}) &= \mathbf{0}. \end{aligned}$$

Druga grupa twierdzi, że wielokrotności wektorów zachowują się jak wielokrotności liczb, czyli dla dowolnych a;b w $\mathbb{R}$ i u;v w V,

$$\begin{aligned} a(b\boldsymbol{u}) &= (ab)\boldsymbol{u}, \\ 1\boldsymbol{u} &= \boldsymbol{u}, \\ a(\boldsymbol{u}+\boldsymbol{v}) &= a\boldsymbol{u}+a\boldsymbol{v}, \\ (a+b)\boldsymbol{u} &= a\boldsymbol{u}+b\boldsymbol{u}. \end{aligned}$$

Z tych aksjomatów jasno wynika, że $\mathbb{R}$ samo w sobie jest przestrzenią wektorową rzeczywistą, ale ciekawszym przykładem jest V ¼ $\mathbb{R}$2, z sumami i wielokrotnościami wektorów ðx;yÞ określonymi dla każdego x,y,a ∈ $\mathbb{R}$ przez

$$(x_1,y_1) + (x_2,y_2) = (x_1 + x_2, y_1 + y_2)$$
$$a(x,y) = (ax, ay).$$

$\mathbb{R}^2$ ma znaczną treść geometryczną. Można zdefiniować linie, linie równoległe i względne pojęcie długości wzdłuż danej linii. Na przykład można powiedzieć, że wielokrotności tv wektora niezerowego v tworzą linię przechodzącą przez 0, a 2v jest dwa razy dalej od 0 niż v. Jednak nie ma pojęcia odległości między dowolnymi punktami. Aby uzyskać naturalne pojęcie odległości, Grassmann wprowadził iloczyn skalarny:

$$(x_1,y_1) \cdot (x_2,y_2) = x_1x_2 + y_1y_2.$$

Z tej definicji wynika, że jeśli v ¼ ðx;yÞ, to

$$\boldsymbol{v} \cdot \boldsymbol{v} = x^2 + y^2 = |v|^2$$

gdzie |v| jest odległością |v| od początku 0 podaną przez twierdzenie Pitagorasa. Jak zauważył Grassmann (1847), jego iloczyn skalarny jest zasadniczo równoważny twierdzeniu Pitagorasa. Pojęcie kąta jest również nieodłączne w iloczynie skalarnym, ponieważ

$$\boldsymbol{u} \cdot \boldsymbol{v} = |\boldsymbol{u}||\mathrm{v}|\cos\theta,$$

gdzie θ jest kątem między liniami od 0 do u i v.

Geometria wektorowa

Poprzednia sekcja pokazuje, że płaszczyzna Kartezjusza , w której punkty są parami uporządkowanymi (x,y) liczb rzeczywistych x,y, a odległość jest podana przez twierdzenie Pitagorasa, ma te same składniki co przestrzeń wektorowa $\mathbb{R}^2$ z iloczynem skalarnym Grassmanna. Różnica polega na tym, że algebra Kartezjusza, która obejmuje dowolne wielomiany w x i y, idzie znacznie dalej niż Euklidesa. Algebra Grassmanna przestrzeni wektorowej $\mathbb{R}^2$, obejmująca tylko sumy i wielokrotności rzeczywiste wektorów oraz iloczyn skalarny, jest na odpowiednim poziomie, aby uchwycić geometrię Euklidesa. Z tego powodu przestrzeń wektorowa $\mathbb{R}^2$ z iloczynem skalarnym Grassmanna nazywana jest płaszczyzną euklidesową. Geometria wektorowa Grassmanna uogólnia się również do dowolnej liczby wymiarów bez dodatkowego wysiłku. Poprzez pracę z n-krotkami (x_1,x_2, . . . ,x_n) liczb rzeczywistych, można wykonywać geometrię w dowolnej liczbie wymiarów bez potrzeby wizualizacji, łamiąc w ten sposób barierę wymiarów, która powstrzymywała Greków. W innym kierunku, przydatne jest uogólnienie definicji „odległości" poprzez uogólnienie pojęcia iloczynu skalarnego. Znanym przykładem jest przestrzeń Minkowskiego zdefiniowana przez Minkowskiego (1908) jako otoczenie dla szczególnej teorii względności Einsteina. Jest to przestrzeń $\mathbb{R}^4$ 4-krotek (t,x,y,z) z odległością wyprowadzoną z iloczynu skalarnego

$$(t_1,x_1,y_1,z_1) \cdot (t_2,x_2,y_2,z_2) = -t_1t_2 + x_1x_2 + y_1y_2 + z_1z_2.$$

Zatem odległość Minkowskiego (t,x,y,z) od początku układu współrzędnych ma kwadrat równy $-t^2 + x^2 + y^2 + z^2$, który czasami jest ujemny. Przestrzeń Minkowskiego jest naturalną przestrzenią dla fizyki, gdzie t oznacza czas, a x,y,z są zmiennymi dla trzech wymiarów przestrzeni. Ale matematycznie równie

interesujące jest spojrzenie na przestrzeń trójwymiarową uzyskaną przez pominięcie współrzędnej z. W tej przestrzeni istnieje „sfera o promieniu √-1”, składająca się z punktów v =(t,x,y) z $|v|^2 = 1$, czyli

$$-t^2 + x^2 + y^2 = -1, \text{ or } t^2 - x^2 - y^2 = 1$$

To nie jest nic innego, jak hiperboloida. Pod względem odległości Minkowskiego, geometria na tej hiperboloidzie nie jest niczym innym, jak nieeuklidesową geometrią Beltramiego! Wskazaliśmy, jak trójkąty w modelu dysku konforemnego odpowiadają trójkątom na hiperboloidzie, które są równe w sensie odległości Minkowskiego. Przestrzeń Minkowskiego nadaje treść szalonemu pomysłowi Lamberta z 1766 roku , że geometria nieeuklidesowa może obowiązywać na sferze o promieniu urojonym. Przestrzenie euklidesowa i Minkowskiego pokazują, że $\mathbb{R}$ jest naturalnym i wygodnym fundamentem dla geometrii, zarówno euklidesowej, jak i nieeuklidesowej. Następne pytanie brzmi: czym jest fundament dla $\mathbb{R}$?

Zagadnienia filozoficzne

Kluczowym krokiem w modelowaniu geometrii nieeuklidesowej przez Beltramiego była jego decyzja o uogólnieniu pojęcia długości. Ośmieliła go do tego lektura Riemanna (1854), przełomowego dzieła z zakresu podstaw geometrii, które po raz pierwszy opublikowano w 1868 roku. W nowym, radykalnym podejściu do geometrii (obecnie znanego jako geometria riemannowska) Riemann uogólnił arytmetyczną geometrię Kartezjusza z płaszczyzny na przestrzenie zakrzywione dowolnego wymiaru n.

Punkty przestrzeni, zamiast być reprezentowane przez uporządkowane pary liczb, były reprezentowane przez uporządkowane n-krotki (x_1, x_2, . . . , x_n). A odległość, zamiast być dana przez wzór „pitagorejski”, była uzyskiwana przez rachunek różniczkowy i całkowy z „nieskończenie małego” wzoru pitagorejskiego. Własność „nieskończenie małej pitagorejskiej” oznacza, że geometria przestrzeni (objawiająca się na przykład sumą kątów trójkąta) zbliża się do geometrii euklidesowej w małych obszarach, ale może się od niej różnić w dużych obszarach. Przykładem jest geometria kuli, gdzie małe trójkąty mają sumę kątów bliską π, ale duże trójkąty mają sumę kątów znacznie większą niż π. Kula jest przykładem przestrzeni o stałej dodatniej krzywiźnie. Riemann krótko przyjrzał się przestrzeniom o stałej ujemnej krzywiźnie, ale to Beltrami zauważył, że takie przestrzenie wykazują geometrię nieeuklidesową i że ich „linie” (krzywe o najkrótszej długości) można modelować po prostu za pomocą łuków kołowych. Tak więc geometria Riemanna jest gruntowną arytmetyzacją geometrii, wskazującą drogę do arytmetycznej bazy dla euklidesowej, nieeuklidesowej i wszelkiego rodzaju geometrii zakrzywionych. Rewolucja dokonana przez Riemanna i Beltramiego podniosła kilka kwestii dotyczących natury geometrii i jej miejsca w matematyce.

Intuicja i logika. Modele pokazują, że geometrie euklidesowa i nieeuklidesowa są równie solidne, a liczby rzeczywiste stanowią podstawę dla obu (a także dla rachunku różniczkowego i całkowego oraz fizyki matematycznej). Pozostaje znaleźć dobrą podstawę dla liczb rzeczywistych; miejmy nadzieję, opartą na arytmetyce liczb naturalnych. Jak wiemy, Dedekind znalazł jeden sposób, aby to zrobić.

Znaczenie i istnienie. Ale liczby rzeczywiste obejmują więcej niż arytmetykę; mianowicie pewne założenie o nieskończoności. Jak bardzo konieczne i uzasadnione jest założenie?

Dyskretne i ciągłe. Zniwelowanie luki między dyskretnym a ciągłym jest zasadniczo problemem zdefiniowania liczb rzeczywistych w kategoriach liczb naturalnych. Stąd możliwość zniwelowania luki (i arytmetyzowania geometrii) zależy od tego, co uzasadnione jest założenie o nieskończoności.

Teoria mnogości i jej paradoksy

PODGLĄD

Definicja liczb rzeczywistych Dedekinda zapoczątkowała nową erę w myśli matematycznej, w której zbiory nieskończone były postrzegane jako obiekty matematyczne. Było to niemile widziane przez wielu matematyków, którzy przyjęli starożytny pogląd, że rzeczywista nieskończoność jest niedopuszczalna. Rzeczywiście, w niektórych kręgach opór wobec rzeczywistej nieskończoności trwa do dziś. Jeszcze bardziej niemile widziane było odkrycie Cantora (1874), że liczby rzeczywiste tworzą nieprzeliczalny zbiór – taki, którego nie można udoskonalić jako nieskończoności jedynie potencjalnej, a nie rzeczywistej. Ale nieprzeliczalność nie była ostatnią kroplą. Cantor (1891) znalazł proste uogólnienie swojego argumentu, które pokazuje, że nie ma największego zbioru. Rzeczywiście, dla dowolnego zbioru S, podzbiory S są liczniejsze niż elementy S. Wynika z tego, że nie istnieje „zbiór wszystkich zbiorów", a zatem nie każda własność jest realizowana przez zbiór. W szczególności nie istnieje zbiór, który realizuje własność „X jest zbiorem". Nawet Dedekind był zaniepokojony tym rozwojem sytuacji.

Przed Cantorem

Wspomnieliśmy, że Grecy byli podejrzliwi wobec nieskończoności i wyjaśniliśmy, jak unikali jej tak dalece, jak to możliwe. Niemniej jednak Grecy z pewnością stosowali procesy nieskończone, i to bardzo pomysłowo. Ich unikanie w praktyce oznaczało rozważanie dowolnych skończonych części nieskończonej całości lub procesu – „potencjalnej" a nie „rzeczywistej" nieskończoności. Rzeczywistej nieskończoności unikano ze względu na jej pozornie paradoksalne właściwości, takie jak pozorne naruszenie zasady, że „całość jest większa od części". Ten paradoks powrócił wraz z odrodzeniem greckiej nauki w czasach średniowiecza i renesansu. Na przykład Galileusz wskazał na korespondencję, która pozornie pokazuje, że cały zbiór liczb całkowitych dodatnich nie jest większy od jego części składającej się z kwadratów. Do końca XIX wieku tego rodzaju relacja – korespondencja jeden do jednego między częścią a całością jakiegoś zbioru – była uważana za paradoksalną. Jednak ostatecznie korespondencja jeden do jednego została uznana za właściwy sposób porównywania zbiorów nieskończonych. Rzeczywiście, Dedekind (1888) przyjął, że definiującą własnością zbioru nieskończonego jest to, że jego całość można umieścić w jednoznacznej korespondencji z częścią, zamieniając w ten sposób paradoks w definicję. Dwa zbiory, które można umieścić w jednoznacznej korespondencji ze sobą, nazywane są równolicznymi lub tej samej kardynalności. Wiele zbiorów jest równolicznych ze zbiorem $\mathbb{N}$ liczb całkowitych dodatnich. Na przykład:

1 Zbiór $\mathbb{Z}$ wszystkich liczb całkowitych, które są równe $\mathbb{N}$ poprzez korespondencję

1	2	3	4	5	6	7	...
↕	↕	↕	↕	↕	↕	↕	
0	1	−1	2	−2	3	−3	

2 $\mathbb{Q}^+$ dodatnich wymierników, poprzez korespondencję

1	2	3	4	5	6	7	8	9	...
↕	↕	↕	↕	↕	↕	↕	↕	↕	
1/1	2/1	1/2	3/1	1/3	4/1	3/2	2/3	1/4	

gdzie w dolnym wierszu wymienione są poszczególne ułamki m/n w grupach: najpierw te, dla których m + n = 2, następnie te, dla których m + n = 3, te, dla których m + n = 4 i tak dalej.

3 Zbiór $\mathbb{Q}$ wszystkich liczb wymiernych można następnie przedstawić jako równoliczny z $\mathbb{N}$ za pomocą tej samej sztuczki, której użyto do wyliczenia $\mathbb{Z}$: najpierw wypisz 0, a następnie naprzemiennie wypisz dodatnią i ujemną formę każdej dodatniej liczby wymiernej.

Te wyniki pokazują, że $\mathbb{Z}$, $\mathbb{Q}$þ i $\mathbb{Q}$ można postrzegać jako „potencjalne" nieskończoności, takie jak $\mathbb{N}$. Jeszcze silniejszy wynik w tym samym kierunku znalazł Dedekind w 1874 r.: zbiór wszystkich liczb algebraicznych (rozwiązania równań wielomianowych ze współczynnikami całkowitymi) jest równoliczny z $\mathbb{N}$. Więc on również można postrzegać jako „potencjalną" nieskończoność. Dzięki tym wynikom można było odrzucić większość starożytnych obaw dotyczących rzeczywistej nieskończoności – ponieważ każda nieskończoność wydawała się być jedynie „potencjalna" – ale tuż za rogiem czekała wielka niespodzianka.

Argument przekątny Cantora

Cantor (1874) udowodnił, że liczby rzeczywiste nie są równoliczne z liczbami całkowitymi dodatnimi. Oznacza to, że każde sparowanie liczb całkowitych dodatnich z liczbami rzeczywistymi;

1	2	3	4	5	6	7	. . .,
↕	↕	↕	↕	↕	↕	↕	
x_1	x_2	x_3	x_4	x_5	x_6	x_7	

nie obejmuje wszystkich liczb rzeczywistych. W rzeczywistości, biorąc pod uwagę dowolną listę $x_1,x_2,x_3,x_4,$ . . . liczb rzeczywistych, możemy wyraźnie opisać liczbę rzeczywistą x, której nie ma na liście. Dowód z 1874 r. nie był łatwy do naśladowania – szczególnie dla społeczności matematycznej zupełnie nieprzygotowanej na niego – ale Cantor (1891) podał inny dowód, który uzyskuje „świadka" x z maksymalną jasnością. Jest to słynny (lub, dla niektórych, niesławny) argument przekątny. Istnieje wiele sposobów, w jaki można podać liczby rzeczywiste $x_1,x_2,x_3,x_4,$ . . .., ale dla ścisłości załóżmy, że są one podane jako nieskończone liczby dziesiętne. Zignorujemy również cyfry przed przecinkiem dziesiętnym, więc możemy sobie wyobrazić liczby wyświetlane jak na rysunku , pokazując tylko cyfry po przecinku dziesiętnym:

x_1	**1**	1	1	1	. . .
x_2	0	**2**	0	1	. . .
x_3	7	7	**7**	7	. . .
x_4	0	0	0	**0**	. . .
⋮					
x	**2**	**1**	**1**	**1**	. . .

Rysunek pokazuje również pierwsze kilka cyfr x. Zapewniamy, że x ≠ każde x_n jest różne na n-tym miejscu po przecinku (i nie używamy cyfr 0 i 9 w x, ponieważ liczby z tymi cyframi mogą być takie same, nawet jeśli ich cyfry są różne – na przykład 0 · 4999 . . . = 0 · 5000 . . .). Konkretnie definiujemy x przez

$$n\text{th digit of } x = \begin{cases} 2 & \text{if } n\text{th digit of } x_n \text{ is } 1 \\ 1 & \text{if } n\text{th digit of } x_n \text{ is not } 1. \end{cases}$$

Zatem x różni się od wszystkich liczb $x_1,x_2,x_3,x_4,$ . . ., stąd podana lista nie obejmuje wszystkich liczb rzeczywistych. Z tego powodu mówimy, że zbiór wszystkich liczb rzeczywistych jest nieprzeliczalny – zbiór przeliczalny to taki, którego elementy można sparować z liczbami całkowitymi dodatnimi.

Ponieważ tylko zbiory przeliczalne można uznać za „potencjalnie” nieskończone, zbiór $\mathbb{R}$ liczb rzeczywistych jest nieuchronnie rzeczywistą nieskończonością. Argument Cantora nazywa się „przekątnym”, ponieważ obejmuje on tylko cyfry na przekątnej tabeli, pokazane pogrubioną czcionką na rysunku. Dlatego musimy jedynie zbadać skończoną ilość każdego rozwinięcia dziesiętnego – mianowicie pierwsze n cyfr x_n – aby obliczyć n-tą cyfrę x. Wspominam o tym, aby rozwiać powszechne błędne przekonanie, że argument przekątny jedynie dowodzi istnienia liczby x, która nie znajduje się na podanej liście $x_1,x_2,x_3,x_4, \ldots$ W rzeczywistości pokazuje to, że x jest tak samo konstruowalne jak liczby $x_1,x_2,x_3,x_4, \ldots$ same w sobie.

Wyższe nieskończoności

Istotą argumentu przekątnego jest stwierdzenie: biorąc pod uwagę liczbę rzeczywistą x_n sparowaną z każdą liczbą naturalną n, możemy zdefiniować liczbę rzeczywistą x ≠ dla każdego x_n zgodnie z regułą

$$n\text{th digit of } x = \begin{cases} 2 & \text{if } n\text{th digit of } x_n \text{ is } 1 \\ 1 & \text{if } n\text{th digit of } x_n \text{ is not } 1. \end{cases}$$

ponieważ daje to $x \neq x_n$ na n-tym miejscu po przecinku. Podobny argument pokazuje, że nie może istnieć lista $S_1,S_2,S_3,S_4, \ldots$ wszystkich podzbiorów $\mathbb{N}$. Ponieważ dla dowolnej takiej listy możemy zdefiniować zbiór S ≠ każdego S_n według reguły

$$n \in S \Leftrightarrow n \notin S_n.$$

ponieważ ta reguła odróżnia S od S_n w odniesieniu do liczby n. W swojej pracy z 1891 r. Cantor doprowadził ten tok myślenia do końca, pokazując, że każdy zbiór X ma więcej podzbiorów niż elementów. Aby zrozumieć dlaczego, załóżmy, że każdy element x zbioru X jest sparowany z podzbiorem S_x zbioru S. Ale wtedy możemy zdefiniować podzbiór S zbioru X różny od każdego Sx w odniesieniu do elementu x. Mianowicie niech

$$x \in S \Leftrightarrow x \notin S_x$$

Zatem zbiór podzbiorów Sx sparowany z elementami x zbioru X nie obejmuje wszystkich podzbiorów zbioru X. Nieco alarmujące w tym wyniku jest to, że pokazuje on, że nie istnieje zbiór wszystkich zbiorów. Gdyby istniał taki zbiór, powiedzmy X, zbiór jego podzbiorów byłby zbiorem większym niż zbiór wszystkich zbiorów, co jest wyraźnie sprzeczne. Cantor wkrótce zauważył tę konsekwencję argumentu diagonalnego, ale pozostał spokojny. (Dedekind był jednak zaskoczony tym rozwojem sytuacji i w rezultacie opóźnił publikację drugiego wydania swojej książki Dedekind [1888]). Historia pokazałaby, że naturalne jest, aby zbiór wszystkich zbiorów nie był zbiorem, podobnie jak naturalne jest, aby zbiór wszystkich dodatnich liczb całkowitych nie był liczbą całkowitą dodatnią. Jednak w tamtym czasie panował niepokój z powodu wcześniejszego przekonania, że każda własność powinna być realizowana przez zbiór. Po odkryciu Cantora stało się jasne, że istnieją wyjątki od tego przekonania, takie jak własność bycia zbiorem.

Następstwa argumentu przekątnego

Argument przekątny był kluczowym odkryciem w podstawach matematyki. W jednym kierunku prowadził, pozornie nieubłaganie, do wyższych sfer nieskończoności – i ku nowym rodzajom paradoksów. W innym kierunku doprowadził do odkrycia ograniczeń w matematyce formalnej: niezupełności i nierozwiązywalności algorytmicznej, jak zobaczymy w dwóch następnych sekcjach. I wywołał również reakcję nowego rodzaju sceptyków, których sceptycyzm był skierowany nie tylko na

nieskończoność, ale na samą logikę. Odkrycie Cantora, że nie istnieje największy zbiór i że nie każda własność jest realizowana przez zbiór, doprowadziło do rozwoju teorii mnogości, systematycznego badania zbiorów nieskończonych w ogólności i $\mathbb{R}$ w szczególności. Nieprzeliczalność $\mathbb{R}$ była zaledwie pierwszym z wielu mylących odkryć dotyczących $\mathbb{R}$, pokazującym, że oś liczbowa, którą myśleliśmy, że znamy, jest bardziej skomplikowana, niż ktokolwiek sobie wyobrażał. W rzeczywistości, aby odpowiedzieć na wiele pytań dotyczących $\mathbb{R}$, musimy poczynić założenia dotyczące całego wszechświata zbiorów. Z drugiej strony, prosta i obliczeniowa natura argumentu przekątnego sprawia, że można go stosować do najbardziej przyziemnych „potencjalnych" rodzajów nieskończoności. Na przykład pokazuje, że nie istnieje obliczalna lista wszystkich obliczalnych liczb. Ponieważ jeśli $x_1,x_2,x_3,x_4,$ . . . jest obliczalną listą obliczalnych liczb, to liczba przekątna x jest również obliczalna – i nie znajduje się na liście. Ten wynik ma głębokie implikacje dla tego, co można obliczyć, we wszystkich dziedzinach matematyki. Nic dziwnego, że argument przekątny i jego implikacje nie były mile widziane przez wszystkich matematyków. Ci, którzy już podejrzliwie podchodzili do nieskończoności, tacy jak berliński kolega Cantora, Kronecker, stali się jeszcze bardziej zagorzali w swoim sprzeciwie wobec zbiorów takich jak $\mathbb{R}$. Weierstrass był bardziej przychylny, ale przekonał Cantora, by złagodził aspekt nieprzeliczalności dowodu i zamiast tego położył nacisk na pozytywny wynik: nowy i elementarny dowód istnienia liczb niealgebraicznych – tak zwanych liczb transcendentalnych. (Wynika to bezpośrednio z nieprzeliczalności $\mathbb{R}$ i wyniku Dedekinda, że zbiór liczb algebraicznych jest przeliczalny). W kolejnych latach idee Cantora stopniowo zyskiwały poparcie większości, dzięki wsparciu wybitnych matematyków, takich jak Hilbert. Pozostał pewien sprzeciw, ale stał się bardziej niuansowy, ponieważ stopniowo zrozumiano, jak wiele matematyki trzeba poświęcić, jeśli odrzuci się różne rodzaje nieskończoności. Wśród skrajnych odrzucających najbardziej prominentni są konstruktywiści, którzy akceptują dowody istnienia tylko wtedy, gdy dostarczają one konstrukcji obiektu, o którym twierdzi się, że istnieje. Ta postawa miała pozytywny wpływ nawet na matematyków, którzy jej nie podzielają. Na przykład obecnie potrafimy konstruować wiele obiektów, takich jak rozwiązania równań wielomianowych, których istnienie po raz pierwszy udowodniono za pomocą argumentów niekonstruktywnych.

Zagadnienia filozoficzne

Warto wspomnieć, że pierwszy dowód niepoliczalności $\mathbb{R}$ Cantora był również argumentem diagonalnym, ale w mniej przejrzystej formie. Miał tę samą istotę: skonstruowanie liczby rzeczywistej x krok po kroku i uczynienie x różnym od liczby rzeczywistej xn w kroku n. Jednak konstrukcja była specyficzna dla liczb rzeczywistych, więc spektakularne uogólnienie na dowolne zbiory znalezione przez Cantora (1891) było nowe. Istnieje inny dowód niepoliczalności, specyficzny dla $\mathbb{R}$, który jest moim zdaniem bardziej odkrywczy. Działa on w następujący sposób. Biorąc pod uwagę liczby rzeczywiste $x_1,x_2,x_3,x_4,$ . . ., pokrywamy każdy x_n przedziałem o długości 2^{-n}. Wówczas całkowita długość pokrytej osi liczbowej wynosi co najwyżej

$$\frac{1}{2}+\frac{1}{4}+\frac{1}{8}+\ldots=1,$$

co z pewnością nie jest całą linią. Tak więc punkty $x_1,x_2,x_3,$. . . nie mogą obejmować wszystkich liczb rzeczywistych. Dowód ten można udoskonalić, aby uzyskać konkretną liczbę x, która nie jest objęta. Wybieramy pierwszą (binarną) cyfrę x, aby uniknąć pierwszego przedziału, następnie drugą cyfrę, aby uniknąć drugiego przedziału, i tak dalej – na którym etapie staje się jasne, że ta konstrukcja jest kolejnym argumentem diagonalnym.

Istnieje w rzeczywistości inna droga do niepoliczalności, poprzez teorię liczb porządkowych Cantora. Choć jest to interesujące, podnosi ona kolejny zestaw kwestii, na których omówienie nie mamy tu miejsca. Samo istnienie niepoliczalnych zbiorów podnosi wystarczająco dużo kwestii

Intuicja i logika. Jak wielkie zrozumienie nieskończoności jest możliwe dla człowieka? Dedekind (1888) podjął śmiałą próbę udowodnienia, że istnieją zbiory nieskończone (po podobnej próbie Bolzano 1851). Opiera się ona na fakcie, że tylko zbiór nieskończony może być równoliczny z właściwym podzbiorem samego siebie:

Twierdzenie. Istnieją nieskończone systemy.

Dowód. Moje własne królestwo myśli, tj. całość S wszystkich rzeczy, które mogą być obiektami mojej myśli, jest nieskończone. Bo jeśli s oznacza element S, to myśl jest s', że s może być obiektem mojej myśli, samo będąc elementem S. Jeśli potraktujemy to jako transformację $\phi(s)$ elementów s, to transformacja ϕ z S, w ten sposób określona, ma własność, że transformacja S' jest częścią S; a S' jest z pewnością właściwą częścią S, ponieważ istnieją w S elementy, które nie są zawarte w S'.

Spoilsports bez wątpienia zaprotestują, że w swoim życiu Dedekind miał tylko skończoną liczbę myśli, więc coś musi być nie tak z tym „dowodem". Ale kto wie, jaka jest naprawdę dziedzina myśli Dedekinda? Argument Dedekinda miał kilku wybitnych zwolenników, takich jak Russell (1903, 357). Bolzano i Dedekind rozważali tylko zbiory przeliczalne, ale moglibyśmy również dążyć do zbiorów nieprzeliczalnych. Czy nasza pozorna intuicja ciągłości jest intuicją nieprzeliczalności i/lub rzeczywistej nieskończoności?

Znaczenie i istnienie. Przed Cantorem pytanie o nieskończoność było proste: czy akceptujemy rzeczywistą nieskończoność, czy nie? Po Cantorze pytanie stało się bardziej złożone: ile nieskończoności akceptujemy? Gdy okaże się, że nieskończoność ma nieskończenie wiele możliwych poziomów, możliwych jest wiele różnych poziomów akceptacji. Niektórzy matematycy akceptowali tylko przeliczalnie nieskończone zbiory (potencjalna nieskończoność), inni akceptowali $\mathbb{R}$, ale nie wszystkie podzbiory $\mathbb{R}$ itd. Francuscy matematycy Borel, Baire, Lebesgue i Hadamard przeprowadzili na ten temat ożywioną debatę w 1905 r., którą można przeczytać w Ewald (1996). W odpowiedzi na pogląd Borela, że pewne procedury, takie jak dokonywanie nieprzeliczalnie wielu wyborów, są „poza matematyką", Hadamard odparł:

Od wynalezienia rachunku nieskończenie małego do chwili obecnej, wydaje mi się, że zasadniczy postęp w matematyce wynikał z sukcesywnego anektowania pojęć, które dla Greków, renesansowych geometrów lub poprzedników Riemanna były „poza matematyką", ponieważ nie dało się ich opisać.

Dyskretne i ciągłe. Nieprzeliczalność $\mathbb{R}$ jednoznacznie pokazuje, że matematyka obejmująca $\mathbb{R}$ – taka jak analiza – dotyka teorii mnogości. Co najmniej projekt arytmetyzowania geometrii i analizy musi wyjść poza czystą arytmetykę: musi zawierać pewne założenia dotyczące zbiorów nieskończonych. W następnej sekcji przyjrzymy się standardowemu systemowi aksjomatów arytmetyki i rzucimy okiem na aksjomaty dotyczące zbiorów, które leżą poza nim.

Systemy formalne

PRZEGLĄD

Nagłe pojawienie się teorii mnogości i jej paradoksów nie było jedynym znaczącym filozoficznie wydarzeniem w matematyce końca XIX wieku. Innym było pojawienie się logiki formalnej i, bardziej ogólnie, formalnych systemów matematyki. Począwszy od pojawienia się Analizy matematycznej logiki Boole'a w 1847 roku, logika i matematyka zostały przełożone na symbolikę, w której dedukcje były zasadniczo obliczeniami. Sto lat temu znane były systemy gotowe do formalizacji dla arytmetyki (Peano 1889), geometrii (Hilbert 1899) i teorii mnogości (Zermelo 1908). Istniał również formalny system dla logiki, za sprawą Fregego (1879). Systemy formalne ożywiły stare marzenie Leibniza: rachunek rozumowy, dzięki któremu prawdziwość dowolnego twierdzenia można było ustalić za pomocą obliczeń. Prawdą jest, że twierdzenia formalnego systemu Σ można uzyskać przez obliczenia, więc jeśli τ jest twierdzeniem Σ, ostatecznie zaobserwujemy ten fakt, systematycznie generując twierdzenia. Nie jest jednak jasne, czy

1 wszystkie prawdy wyrażalne w Σ są twierdzeniami, (Zupełność)

2 jakaś reguła decyduje dla każdego τ, czy τ jest twierdzeniem, (Rozstrzygalność)

3 jeśli Σ udowadnia τ, to Σ nie udowadnia negacji τ. (Spójność)

Niemniej jednak są to wszystkie pytania o wyniki obliczeń skończonych, takie jak pytania teorii liczb. Można więc mieć nadzieję na ich rozstrzygnięcie za pomocą środków elementarnych (które Hilbert nazwał „skończonymi”). W szczególności, za pomocą programu rozstrzygania pytania o spójność dla formalnej teorii mnogości, Hilbert miał nadzieję rozwiać wszelkie wątpliwości co do stosowania nieskończoności w matematyce.

Hilbert

W latach 90. XIX wieku Hilbert ożywił geometrię Euklidesa, dogłębnie studiując aksjomaty geometryczne i ich związek z algebrą i liczbami rzeczywistymi. Opierając się na pracach niektórych swoich poprzedników, takich jak von Staudt i Pasch,

1 wypełnił luki w systemie Euklidesa, wyraźnie stwierdzając aksjomaty, których Euklides używał nieświadomie,

2 pogrupował aksjomaty na różne typy koncepcyjne: incydencja, rząd, przystawanie, przecięcie okręgu,

3 wyprowadził algebraiczne własności sumy i iloczynu z aksjomatów incydencji, a ich własności porządku z aksjomatów porządku,

4 dodał dwa aksjomaty niepotrzebne w geometrii, ale potrzebne do wyprowadzenia własności linii liczb rzeczywistych: aksjomat Archimedesa (stwierdzający, że nie ma nieskończenie małych) i aksjomat zupełności (stwierdzający, że nie ma luk, w sensie Dedekinda).

Z uwzględnienia aksjomatów Archimedesa i zupełności wynika, że Hilbert interesował się liczbami rzeczywistymi tak samo, jak geometrią. W swojej książce z 1899 r. Grundlagen der Geometrie (Podstawy geometrii) zamieścił również pomysłową nową konstrukcję $\mathbb{R}$ z aksjomatów geometrii nieeuklidesowej (uzyskaną przez zastąpienie aksjomatu równoległego). Wyprowadzenie $\mathbb{R}$ z aksjomatów geometrycznych nie było często stosowane w matematyce, ale wniosło ciekawy wkład do filozofii. Hartry Field (1980) w swojej Science without Numbers zaadaptował wynik Hilberta, aby wyprowadzić strukturę liczb rzeczywistych z aksjomatów dotyczących przestrzeni fizycznej. Użył tego wyprowadzenia, aby twierdzić, że $\mathbb{R}$ jest zbędny w nauce, a zatem nikt nie musi zakładać jego istnienia. Jednak dla większości matematyków założenie, że struktura $\mathbb{R}$ istnieje (w przestrzeni fizycznej lub w

dowolnym abstrakcyjnym królestwie) jest tym samym, co założenie istnienia samego $\mathbb{R}$. A założenie, że przestrzeń fizyczna jest archimedesowa, zupełna i euklidesowa – tylko po to, aby uniknąć założenia istnienia $\mathbb{R}$ – wydaje się przeczyć fizyce tak samo, jak i matematyce.

Systemy Peano i Zermelo

Najbardziej powszechnym sposobem dojścia do $\mathbb{R}$ w matematyce jest użycie aksjomatów dla liczb naturalnych i zbiorów. Aksjomaty dla liczb naturalnych zostały podane przez Dedekinda (1888) i Peano (1889) w systemie zwanym obecnie arytmetyką Peano, lub PA. Peano uznał ideę Grassmanna (1861), że kluczowym pojęciem arytmetyki jest indukcja. „Indukcja" nie jest tutaj ogólnym pomysłem wyciągania ogólnego wniosku z konkretnych przypadków, ale żelazną gwarancją, że ogólny wniosek jest poprawny. Ten rodzaj indukcji jest często nazywany indukcją całkowitą, ponieważ dla liczb naturalnych obejmuje wszystkie przypadki. To, co umożliwia udowodnienie ogólnego twierdzenia σ(n) dotyczącego dowolnej liczby naturalnej n, to fakt, że n można osiągnąć od 0 przez wielokrotne dodawanie 1. Dlatego jeśli możemy udowodnić, że σ(0) jest prawdziwe, a σ(m+1) jest prawdziwe zawsze, gdy σ(m) jest prawdziwe, to możemy być pewni, że σ(n) jest prawdziwe dla każdej liczby naturalnej n. Peano wbudował tę ideę w swój system aksjomatów na dwa sposoby.

• Po uwzględnieniu nazw 0 i S dla liczby początkowej i funkcji następczej S(m) = m + 1, z odpowiednimi własnościami, podał indukcyjne definicje sumy i iloczynu:

$$l + 0 = l,\ l + S(m) = S(l + m),$$
$$l \cdot 0 = 0,\ l \cdot S(m) = l \cdot m + l.$$

Z tych definicji wynika, że funkcje + i · są zdefiniowane dla wszystkich liczb naturalnych. Na przykład pierwsze równanie w definicji + definiuje l + m dla m = 0; drugie definiuje l + S(m) po zdefiniowaniu l + m, a zatem definiuje l + n dla wszystkich liczb naturalnych n, przez indukcję zupełną. Podobnie, druga para równań definiuje l · n dla wszystkich n, zakładając, że + jest już zdefiniowane.

• Definicje + i · umożliwiają wyprowadzenie wszystkich faktów szczegółowych dotyczących sumy i iloczynu dla liczb 0S(0),SS(0), . . . poprzez podstawienie w równaniach definiujących. Ale aby udowodnić ogólne fakty, takie jak l + m = m + l, Peano podaje aksjomat indukcyjny dla każdej własności σ: Jeśli σ(0) i jeśli σ(m)) => σ(S(m)) dla wszystkich m, to σ(n) obowiązuje dla każdego n.

(Równoważnym aksjomatem indukcyjnym jest to, że każdy zbiór liczb naturalnych ma najmniejszy element lub że malejący ciąg liczb naturalnych jest skończony. W tej ostatniej formie indukcja sięga czasów Euklidesa.) Zermelo podał aksjomaty dla teorii mnogości w 1908 roku. Pomijamy szczegóły, ale są one podobne w duchu do aksjomatów Dedekinda lub Peano, jak przyznał Zermelo. Twierdzą oni, że istnieje zbiór początkowy Ø (zbiór pusty, który można postrzegać jako 0), operacje budowania kolejnych zbiorów (które między innymi pozwalają na budowanie następników 0) oraz aksjomat nieskończoności stwierdzający istnienie zbioru obejmującego 0 i wszystkich jego następców. Istnieje również aksjomat fundamentu, który jest podobny do indukcji. Rzeczywiście, jeśli pominiemy aksjomat nieskończoności, teoria mnogości Zermelo ma zasadniczo tę samą treść, co aksjomaty Peano. Tak więc teoria mnogości w pewnym sensie jest „teorią liczb þ nieskończoność". Teoria mnogości jest niezwykle potężnym systemem, zdolnym do pokrycia praktycznie całej matematyki głównego nurtu. Dzieje się tak, ponieważ ma zasady konstrukcji zbiorów – takie jak tworzenie zbioru P(X) wszystkich podzbiorów zbioru X – które powodują wybuchowy wzrost, gdy tylko istnieje nieskończony zbiór X. Począwszy od Hilberta w latach 30. XX wieku, pojawiło się zainteresowanie systemami z łagodniejszymi zasadami konstrukcji zbiorów, dostosowanymi do analizy. Okazuje się, że w tych systemach możemy mierzyć „siłę" różnych twierdzeń analizy za pomocą zasad konstrukcji zbiorów potrzebnych do ich

udowodnienia. (Zdumiewająco często zdarza się, że możemy znaleźć „właściwe aksjomaty konstrukcji zbiorów”, aby udowodnić twierdzenia analizy, podobnie jak znalezienie aksjomatu równoległości jako „właściwego aksjomatu”, aby udowodnić wiele twierdzeń geometrii.

System logiki Fregego

Do tej pory nie byliśmy do końca pewni, w jaki sposób aksjomaty mogą być „formalizowane”, aby zrealizować marzenie Leibniza o znalezieniu prawdy poprzez obliczenia. Brakującym składnikiem jest logika formalna. Pierwsze kroki podjął Boole (1847), w tym, co później stało się znane jako algebra Boole'a lub logika zdań. Boole zauważył, że spójniki „lub” i „i” działają na zdania podobnie jak suma i iloczyn. Spełniają one pewne podstawowe tożsamości, takie jak a + b = b + a i ab = ba, z których ogólne tożsamości między zdaniami złożonymi mogą być udowodnione przez algebrę. Jednak logika zdań nie jest wystarczająco ekspresywna dla matematyki, gdzie wewnętrzna struktura zdania jest ważna. Zazwyczaj zdanie matematyczne zawiera: Zmienne obejmujące pewną domenę jednostek, takie jak liczby. Symbole predykatów oznaczające własności lub relacje w domenie. Symbole logiczne, które obejmują nie tylko spójniki, takie jak „i”, „lub” i „nie”, ale także kwantyfikatory „dla wszystkich x” i „istnieje x”, stosowane do dowolnej zmiennej x. Gdy te elementy językowe są uwzględnione, mamy język logiki predykatów. W przypadku logiki predykatów nie jest wcale jasne, jak udowodnić ważne propozycje – to znaczy te prawdziwe dla wszystkich dziedzin i wszystkich interpretacji symboli predykatów – ale, co zadziwiające, jest to możliwe. Frege w 1879 r. podał zbiór aksjomatów i reguł wnioskowania zdolnych do ich wygenerowania. Nie będziemy tutaj wymieniać wszystkich aksjomatów i reguł Fregego. Przykłady jego aksjomatów, zapisanych za pomocą spójników => (jeśli … wtedy) i ¬ (nie) oraz $\forall x$ (dla wszystkich x), to

- $a \Rightarrow (b \Rightarrow a)$,
- $a \Rightarrow (\neg\neg a)$,
- $\forall x P(x) \Rightarrow P(c)$, gdzie c jest literą nie występującą w P(x).

Przykłady jego reguł wnioskowania to

- modus ponens: z B i B => A wnioskuj A
- ∀ wprowadzenie: z A(x) wnioskuj $\forall x$A(x).

Frege najwyraźniej uważał, że jego aksjomaty i reguły wnioskowania wystarczą do udowodnienia każdej ważnej propozycji. Gödel (1930) udowodnił, że Frege miał rację. To twierdzenie o zupełności dla logiki predykatów było pierwszym z kilku zadziwiających wkładów Gödla do logiki matematycznej (i filozofii matematyki). Dowód Gödla twierdzenia o zupełności przy okazji udowodnił dwie inne ważne właściwości logiki predykatów.

Modelowanie zdań spójnych. Jeśli Σ jest zbiorem zdań, których logiczne konsekwencje nie zawierają sprzeczności, to istnieje model Σ . Oznacza to, że istnieje dziedzina D indywiduów i interpretacji symboli predykatów Σ na D, w ramach których każde zdanie w P jest prawdziwe.

Zwartość. Jeśli każdy skończony podzbiór Σ' z Σ ma model, to Σ również. (Wynika to z modelowania, ponieważ jeśli każdy skończony podzbiór Σ' ma model, to z Σ' nie wynika żadna sprzeczność. Ale wtedy nie wynika żadna sprzeczność z Σ , ponieważ każdy dowód jest skończony i stąd obejmuje tylko skończony podzbiór Σ' z Σ .)

Zupełność i niezupełność

Zupełność logiki predykatów oznacza, że wszystkie twierdzenia matematyczne są dowodliwe w następującym względnym sensie. Jeśli τ wynika z aksjomatów $\alpha_1, \ldots, \alpha_k$, wówczas implikacja ($\alpha_1$ i . . . i α_k) => τ jest ważna, a zatem dowodliwa w logice predykatów. (Przywodzi to na myśl początkowe słowa Russella [1903]: czysta matematyka jest klasą, jeśli wszystkie propozycje w postaci „p implikuje q".) Może jednak istnieć pewien obszar matematyki, którego nie da się całkowicie zaksjomatizować, w tym sensie, że żaden zbiór aksjomatów, który możemy zapisać, nie będzie logicznie implikował wszystkich prawd tego obszaru.

Zagadnienia filozoficzne

Kiedy po raz pierwszy wprowadzono systemy formalne, wiele pytań na ich temat dotyczyło znaczenia wzorów i metod stosowanych do wyprowadzania twierdzeń, takich jak indukcja.

Intuicja i logika. Poincaré w 1894 roku dokonał mocnego kontrastu między indukcją w nauce a indukcją w matematyce:

Indukcja stosowana w naukach fizycznych jest zawsze niepewna, ponieważ opiera się na wierze w ogólny porządek wszechświata i porządek poza nami. Indukcja matematyczna... wręcz przeciwnie, narzuca się koniecznie, ponieważ jest tylko potwierdzeniem właściwości samego umysłu.

Opisując indukcję jako właściwość umysłu, twierdzi, że indukcja należy do intuicji, a nie logiki. Jakkolwiek by było, indukcję można z pewnością wyrazić jako właściwość liczb. Biorąc pod uwagę redukcję matematyki do arytmetyki plus pewnej teorii mnogości, pozostaje znaleźć odpowiednie aksjomaty dla arytmetyki i zbiorów. W przypadku arytmetyki indukcja wydaje się zarówno intuicyjna, jak i kluczowa – ale czy jest wystarczająca? W przypadku zbiorów nie jest oczywiste, które aksjomaty są intuicyjne lub wystarczające. Jak pewni jesteśmy, że przyjęte aksjomaty są spójne?

Znaczenie i istnienie. Czy istnienie zbiorów nieskończonych jest udowodnione? Bolzano (1851) i Dedekind (1888) uważali tak, jak widzieliśmy w rozdziale 7.5, ale Zermelo przyjął to jako aksjomat. Z drugiej strony, nie wydaje się, aby istniała również praktyczna teoria skończonego wszechświata matematycznego. Jeśli pominiemy aksjomat nieskończoności Zermela z jego teorii mnogości (lub dodamy jego negację), otrzymamy teorię zbiorów skończonych, ale nadal istnieje nieskończenie wiele zbiorów skończonych. Nic więc dziwnego, że nie możemy znaleźć żadnego skończonego modelu teorii zbiorów skończonych. Inne pytanie o istnienie postawił Hilbert w 1900 roku. Cytuję z angielskiego tłumaczenia w Hilbercie (1902):

Jeśli do pojęcia zostaną przypisane sprzeczne atrybuty, to twierdzę, że matematycznie pojęcie to nie istnieje. Tak więc na przykład liczba rzeczywista, której kwadrat wynosi -1, nie istnieje matematycznie. Ale jeśli można udowodnić, że atrybuty przypisane pojęciu nigdy nie mogą prowadzić do sprzeczności przez zastosowanie skończonej liczby procesów logicznych, twierdzę, że matematyczne istnienie pojęcia (na przykład liczby lub funkcji, która spełnia pewne warunki) jest w ten sposób udowodnione.

Raczej satysfakcjonujące uzasadnienie twierdzenia Hilberta daje modelowanie zdań spójnych, które wynika z dowodu Gödla twierdzenia o zupełności dla logiki predykatów. Jeśli „atrybuty przypisane pojęciu" można wyrazić w logice predykatów (co zwykle ma miejsce w matematyce), to pojęcie ma model, który może służyć do ustalenia jego istnienia.

Dyskretny i ciągły. Wadą modelu zdań spójnych uzyskanego z dowodu Gödla o zupełności jest to, że niekoniecznie jest to zamierzony model. W szczególności, jeśli zbiór zdań jest policzalny – co dla wszystkich praktycznych celów musi być – to model jest również policzalny. W szczególności każdy

spójny zbiór zdań o kontinuum ma przeliczalny model, co z pewnością nie jest zamierzone! Ten wynik, uzyskany przez Skolema (1922) i znany jako paradoks Skolema, nie jest w rzeczywistości sprzeczny. Oznacza on jedynie, że w modelu nie ma funkcji, która łączyłaby członków jego „kontinuum" z liczbami naturalnymi. Ale z pewnością pokazuje, jak trudno jest w pełni opisać kontinuum.

Niektórzy matematycy sprzeciwiali się formalizacji, nie tylko dlatego, że nie wierzyli w formalizację niektórych obiektów, ale również dlatego, że formalizacja wyklucza intuicję, którą uważali za istotny składnik matematyki. Ten ostatni zarzut mógłby zostać zasadniczo obalony, gdyby znaleziono kompletny formalny system matematyki – ale tak się nie stanie, jak zobaczymy w następnej sekcji. W każdym razie zarzut ten nie miał znaczenia.

Nawet najwięksi zwolennicy formalizacji, tacy jak Hilbert, cenili rolę intuicji w matematyce. Ich powód formalizacji był inny: tworzenie dowodów, co do których wszyscy zgadzają się, że są dowodami. Wszyscy zgodzą się, że dane twierdzenie wynika z danych aksjomatów zgodnie z danymi regułami, ponieważ jest to stwierdzenie dotyczące skończonego obliczenia, które każdy może zweryfikować. Interpretacja ciągów symboli w dowodzie nie ma znaczenia. Celem formalizacji jest raczej uczynienie z kwestii spójności kwestii obliczeń; mianowicie, czy sprzeczny ciąg, taki jak 0 = 1, wynika z aksjomatów na podstawie podanych reguł. Gdy na to pytanie zostanie udzielona twierdząca odpowiedź dla formalnego systemu F, wówczas nawet konstruktywiści będą musieli przyznać, że F jest nieszkodliwy, nawet jeśli nie uznają wszystkich jego twierdzeń za sensowne. Był to program Hilberta, za pomocą którego miał nadzieję udowodnić, że rozumowanie o nieskończoności jest nieszkodliwe. Program nie okazał się taki, jakiego Hilbert się spodziewał, jak zobaczymy w następnej sekcji, ale doprowadził do niezwykłych osiągnięć w filozofii matematyki.

Nierozwiązywalność i niezupełność

PRZEGLĄD

Formalne systemy omówione w poprzedniej sekcji są częściową realizacją marzenia Leibniza o rachunku argumentacyjnym – formalnym systemie do ustalania prawdy – ale tylko częściową. W tej sekcji przyjrzymy się drugiej stronie historii: ograniczeniom formalnych systemów. Ograniczenia te wynikają z oszałamiającej propozycji znanej jako teza Churcha: istnieje kompletna i matematycznie precyzyjna definicja obliczeń (a wraz z nią precyzyjne definicje algorytmu, problemu rozwiązywalnego i problemu nierozwiązywalnego). Możliwość wszechstronnej definicji obliczeń została po raz pierwszy podejrzewana przez Posta w latach 20. XX wieku, a najbardziej przekonująco argumentowana przez Turinga (1936). Z natury tezę Churcha można w zasadzie sfalsyfikować, ale żadna falsyfikacja nigdy nie wyszła na jaw. Tak więc tezę Churcha uważa się obecnie za równie dobrze potwierdzoną, jak jakiekolwiek prawo natury. W każdym razie możemy zbudować teorię obliczeń na tezie Churcha, a teoria ta prowadzi do następujących niezwykłych wyników:

1 Istnieją algorytmicznie nierozwiązywalne problemy w matematyce.

2 Każdy wystarczająco silny formalny system arytmetyki jest niekompletny lub niespójny.

3 Spójność wystarczająco silnego formalnego systemu arytmetyki nie jest udowodniona w ramach systemu.

Te wyniki, jak pokażemy w zarysie poniżej, wszystkie wynikają z prostych wariacji argumentu diagonalnego Cantora.

Obliczalność

W XVII wieku, gdy Leibniz marzył o rozstrzyganiu prawdy za pomocą obliczeń, pojęcie obliczeń miało raczej ograniczone znaczenie. Sam Leibniz zaprojektował maszynę liczącą zdolną do wykonywania obliczeń arytmetycznych na liczbach i bez wątpienia przyjąłby, że algebra również jest obliczeniami. Do 1850 roku Boole doszedł do tego, że wykonywał logikę zdań za pomocą obliczeń algebraicznych. Jednak ogólna definicja obliczeń musiała poczekać na rozwój formalnych systemów matematyki, około 1900 roku. Dopiero wtedy stało się jasne, jak szeroka musiała być definicja obliczeń, aby marzenie Leibniza się spełniło. Najbardziej wpływowym formalnym systemem na początku XX wieku były Principia Mathematica Whiteheada i Russella (1910). Principia twierdziły, że pokazują, w jaki sposób wszystkie twierdzenia matematyki można wygenerować z poszczególnych aksjomatów za pomocą pewnych reguł. Reguły były takie, że w zasadzie były mechaniczne, a zatem można je było stosować bez zastanowienia do ciągów symboli – eliminując wszelkie możliwości ludzkiego błędu lub stronniczości. Na początku lat 20. XX wieku reguły zostały przeanalizowane i uproszczone przez Posta, aż do momentu, gdy zostały zredukowane do formy

gw -> Wg'

dla skończonej liczby par ciągów symboli (g,g'). Reguła gW→Wg' mówi, że w dowolnym ciągu zaczynającym się od g, g można usunąć z lewej strony i następnie dołączyć g' z prawej. Post nazwał taki system reguł systemem normalnym. Początkowo Post myślał, że prostota systemów normalnych pozwoli mu zdecydować, czy dany ciąg symboli jest twierdzeniem Principia, czy nie. Jednak później nie był w stanie przewidzieć zachowania bardzo prostych systemów normalnych i zdał sobie sprawę, że sytuacja była odwrotna do tego, na co liczył: każde obliczenie może być symulowane przez system normalny i nie ma ogólnej reguły rozstrzygającej, czy dany system normalny generuje dany ciąg. Post dostrzegł przyszłość logiki matematycznej, która miała się rozwinąć w ciągu następnych piętnastu lat (jego opis można znaleźć w Post 1941). Jednakże powstrzymała go bezprecedensowa trudność: jak można być pewnym, że koncepcja układu normalnego (lub jakakolwiek inna definicja) w pełni oddaje koncepcję obliczeń? Zobaczył, że to twierdzenie (później znane jako teza Churcha lub teza Churcha-Turinga) jest czymś w rodzaju prawa natury – wymagającego ciągłej weryfikacji i zagrożonego możliwą falsyfikacją.

Nierozwiązywalność

Podczas gdy praca Posta pozostała nieopublikowana i nieznana, niezależne próby zdefiniowania pojęcia obliczeń podjęli Church (1936) i Turing (1936). Podejście Turinga (obecnie znane jako koncepcja maszyny Turinga) było niezwykle przekonujące, będąc zasadniczo idealizacją ludzkiego „komputera" pracującego z ołówkiem i papierem:

• Zamiast papieru maszyna Turinga ma nieskończoną taśmę podzieloną na kwadraty. Każdy kwadrat może pomieścić jeden ze skończonego alfabetu symboli, w tym puste miejsce.

• Zamiast człowieka z ołówkiem maszyna Turinga ma głowicę odczytująco-zapisującą, która może przyjąć jeden ze skończonej liczby stanów wewnętrznych (takich jak stany mentalne). Głowica skanuje jeden kwadrat na raz i, w zależności od stanu wewnętrznego q_i i zeskanowanego symbolu S_j, zastępuje S_j symbolem S_k, przesuwa się o jeden kwadrat w lewo lub w prawo i wchodzi w stan q_l.

Każda maszyna Turinga jest zatem opisana przez skończenie wiele kwintetów, w postaci $q_i S_j S_k R q_l$ lub $q_i S_j S_k L q_l$. Obliczenie maszyny jest określone przez dane wejściowe, składające się ze skończonej sekwencji oznaczonych kwadratów taśmy, początkowo zeskanowanego kwadratu i stanu początkowego. Przy odrobinie praktyki łatwo jest zrealizować obliczenia ołówkiem i papierem za pomocą obliczeń maszyny Turinga. Sekret polega na wyobrażeniu sobie, jak można by wykonać

obliczenia, gdyby pozwolono nam oglądać tylko jeden symbol na raz. Najbardziej ambitnym obliczeniem, o którym należy pomyśleć, jest obliczenie uniwersalnej maszyny Turinga: maszyny U, która może przyjąć opis dowolnej maszyny Turinga T i dowolne dane wejściowe I, i symulować obliczenie T na I. Pierwszą trudnością do pokonania jest to, że U, jak każda inna maszyna Turinga, ma tylko skończony alfabet symboli i skończenie wiele stanów wewnętrznych. Jeśli U ma w ogóle odczytać opis T, ten opis musi być zapisany w ustalonym, skończonym alfabecie. Na przykład, można użyć alfabetu q;S;L;R;0g i użyć symbolu pierwszego 0, aby zapisać qi jako q00 ... (q z i liczbami pierwszymi) i zapisać Sj jako S00 ... (S z j liczbami pierwszymi). Tak więc pojedynczy symbol w opisie T musi zostać zastąpiony ciągiem symboli, koniecznie rozłożonym na sekwencji kwadratów taśmy U. Naturalnie, sprawia to, że symulacja T przez U jest dość powolna (podobnie jak potrzeba, aby U ciągle „odwoływał się" do opisu T w celu wykonania każdego kroku obliczeń T). Niemniej jednak, w zasadzie widać, dlaczego istnieje uniwersalna maszyna Turinga i że U może symulować to, co T robi na danym wejściu, krok po kroku. (Dziś uniwersalne maszyny Turinga są wszechobecne; wspólny język programowania jest równoważny uniwersalnej maszynie Turinga. Wadą tego faktu, jak zaraz udowodnimy, jest to, że trudno przewidzieć, co dany program zrobi na podstawie danych wejściowych.) Teraz jedną rzeczą jest wykonywanie instrukcji; inną jest przewidywanie, dokąd one doprowadzą. Na przykład możemy obliczyć dowolną liczbę miejsc dziesiętnych liczby π, ale obecnie nie wiemy, czy kiedykolwiek wystąpi 1000 kolejnych 7. W przypadku maszyn Turinga ostateczny wynik obliczeń jest niepewny, co możemy zobaczyć w następującym przykładzie:

Problem samobadania. Mając daną maszynę Turinga T, zdecyduj, czy T, mająca swój własny opis des (T) jako dane wejściowe, kiedykolwiek zatrzyma się na pustym kwadracie. Używamy tutaj terminu „problem" w znaczeniu nieskończonego zbioru pytań, w tym przypadku następujących, dla każdej maszyny Turinga T:

Q_T: Czy T, na danych wejściowych des (T), kiedykolwiek zatrzyma się na pustym kwadracie?

Pokażemy, że żadna maszyna Turinga S nie może poprawnie odpowiedzieć na wszystkie pytania Q_T . Można założyć, że S otrzymuje pytanie Q_T w formie des (T), ponieważ Q_T można zrekonstruować z des (T). Można również założyć, że S odpowiada „nie", zatrzymując się na pustym kwadracie, i „tak", zatrzymując się na niepustym kwadracie. Ale wtedy S nie może udzielić prawidłowej odpowiedzi na pytanie Q_S. Jeśli S, przy danym Q_S, zatrzymuje się na pustym kwadracie, to odpowiedź na Q_S brzmi „tak", więc S nie powinien zatrzymać się na pustym kwadracie. A jeśli S nie zatrzymuje się na pustym kwadracie, to odpowiedź na pytanie Q_S brzmi „nie", a S musi zatrzymać się na pustym kwadracie. Ta sprzeczność pokazuje, że problem samobadania nie może zostać rozwiązany przez maszynę Turinga. A zatem, jeśli teza Churcha-Turinga jest poprawna, tego problemu nie można rozwiązać żadnymi obliczeniami. Jak mówimy, problem jest algorytmicznie nierozwiązywalny lub po prostu nierozwiązywalny. Ponieważ problem samobadania jest raczej ewidentnie samoniszczący, można by mieć nadzieję, że nierozwiązywalność jest aberracją, a nie czymś, co dzieje się naturalnie. Tak nie jest, ponieważ maszyny Turinga mogą być „symulowane" przez różne naturalne systemy w matematyce i logice. W rzeczywistości Church i Turing natychmiast zauważyli, że logika predykatów może „symulować" maszyny Turinga, w wyniku czego problem decydowania o ważności w logice predykatów – tzw. Entscheidungsproblem – jest nierozwiązywalny.

Niezupełność

Pojęcie obliczalności, które jakimś cudem wydaje się być zupełne i absolutne, stoi w kontraście do pojęcia dowodzenia, które okazuje się niezupełne i względne. Łącznikiem między nimi jest nieobliczalność, która wynika z obliczalności przez argument przekątny. Najwygodniejszą formą argumentu przekątnego jest ta, której użył Cantor (1891), aby udowodnić, że dowolny zbiór ma więcej

podzbiorów niż elementów. Podążając za Postem, stosujemy ten argument do zbiorów powiązanych z maszynami Turinga, zwanych zbiorami obliczalnie przeliczalnymi. Zbiór liczb naturalnych jest nazywany obliczalnie przeliczalnym, jeśli istnieje maszyna Turinga, która wymienia jego elementy. Sposób tworzenia listy nie jest ważny, o ile dowolna maszyna Turinga ma powiązany z nią zbiór obliczalnie przeliczalny (być może zbiór pusty) i możemy zaobserwować, kiedy dana maszyna T wymienia daną liczbę n. Odwołamy się do tezy Churcha-Turinga, aby stwierdzić, że każdy zbiór, który jest intuicyjnie wyliczalny, jest wyliczalny przez maszynę Turinga. $\mathbb{N}$ jest obliczalnie wyliczalny, podobnie jak wiele jego podzbiorów, takich jak zbiór liczb pierwszych. Co więcej, możemy obliczalnie wyliczyć wszystkie opisy maszyny Turinga, wymieniając je w kolejności leksykograficznej. Niech Wn będzie obliczalnie wyliczalnym zbiorem wymienionym przez n-tą maszynę Turinga. Wynika z tego argument diagonalny, że zbiór

$$D = \{n : n \notin W_n\}$$

nie jest obliczalnie wyliczalny, ponieważ różni się od każdego W_n względem elementu n. (Co ciekawe, jego dopełnienie $\mathbb{N} - D = \{n : n \in W_n\}$ jest obliczalnie wyliczalne. Możemy wypisać n, które pojawiają się w $\mathbb{N}$ - D, pozwalając pierwszym k maszynom Turinga działać przez k kroków, dla k = 1, 2, 3, . . ., i umieszczając n na liście, jeśli kiedykolwiek zostanie wypisane przez n-tą maszynę.) Teraz formalny system F, zgodnie z najszerszą możliwą definicją, ma obliczalnie wyliczalny zbiór twierdzeń, więc możemy obliczalnie wyliczyć twierdzenia F, które mają postać n $\notin$ D. Ponieważ D nie jest obliczalnie wyliczalny, twierdzenia te (jeśli są poprawne) nie obejmują wszystkich prawdziwych stwierdzeń postaci n $\notin$ D . Ten fakt wskazuje na pewne niepowodzenie systemu F, ale musimy wydobyć znaczenie warunku „jeśli jest poprawny". Prowadzi to do ciągu coraz silniejszych twierdzeń o niezupełności.

1 Załóżmy, że F jest niesprzeczny i że zbiór n, dla którego F dowodzi „n $\notin W_n$" jest zbiorem obliczalnie przeliczalnym W_m. Zakładamy również, że F jest wystarczająco silny, aby symulować maszyny Turinga, a zatem udowodnić n $\in W_n$, kiedykolwiek jest to prawdą. Co możemy teraz powiedzieć o zdaniu „m = $\notin W_m$"? Jeśli F dowodzi „m = $\notin W_m$", to m $\in$ Wm , zgodnie z definicją Wm. Ale założyliśmy, że F może udowodnić wszystkie takie prawdziwe stwierdzenia, więc F dowodzi również „m $\in$ Wm", zaprzeczając spójności. Zatem F nie dowodzi „m = $\notin W_m$", więc m = $\notin W_m$, ponownie zgodnie z definicją W_m. To sprawia, że „m = $\notin W_m$" jest konkretnym prawdziwym zdaniem, którego F nie potrafi udowodnić. Zauważ, że zdanie „m = $\notin W_m$" zasadniczo mówi „nie jestem udowodnialny", ponieważ m = $\notin W_m$ oznacza, że „m = $\notin W_m$" nie jest udowodnialne, zgodnie z definicją W_m.

2 Przekładając działanie maszyn Turinga na arytmetykę – co jest możliwe, choć dalekie od oczywistości – możemy pokazać, że arytmetyczna PA Peano jest „wystarczająco silna" w powyższym sensie. Tak więc język PA ma zdanie, które jest równoważne z m = $\notin W_m$, a zatem nie do udowodnienia w PA, jeśli PA jest niesprzeczne. Innymi słowy, jeśli PA jest niesprzeczne, to zdanie „m = $\notin W_m$" nie jest do udowodnienia w PA. Lub, równoważnie, jeśli PA jest niesprzeczne, tom = $\notin W_m$.

3 Podobna „arytmetyzacja" dowodów w PA pozwala nam wyrazić niesprzeczność PA za pomocą zdania w języku PA, Con(PA), i wywnioskować z niego, że m = $\notin W_m$. Tak więc „Con(PA)) m = $\notin W_m$" jest twierdzeniem PA.

Z tego wynika, że Con(PA) nie może być udowodnione w PA, w przeciwnym razie modus ponens dałoby dowód „m = $\notin W_m$", a wiemy, że „m = $\notin W_m$" nie ma dowodu w PA. Zatem PA, jeśli jest spójne, nie może udowodnić swojej własnej spójności (i nie może tego zrobić żaden spójny system, który obejmuje PA, ponieważ podobny argument będzie miał zastosowanie). To błyskotliwy tok myślenia – jeden z najbardziej oszałamiających w matematyce. Pozycja 1 przypomina wyniki uzyskane przez Posta w

latach 20. XX wieku. Pozycje 2 i 3 są zasadniczo pierwszym i drugim twierdzeniem o niezupełności Gödla (1931), chociaż Gödel znalazł pierwsze twierdzenie inaczej, bezpośrednio konstruując zdanie, które mówi „nie jestem udowodnialny". Co ciekawe, Gödel najpierw udowodnił niezupełność dla systemu wyższego poziomu Principia Mathematica, ponieważ nie zdawał sobie sprawy, że obliczenia lub dowody można arytmetyzować. Do arytmetyzowania zachęcił go von Neumann, który również zasługuje na uznanie za drugie twierdzenie o niezupełności. Von Neumann wskazał na niemożność udowodnienia spójności w liście do Gödla, zanim sam Gödel ogłosił wynik.

Niezupełność teorii mnogości

Argument przekątny jest łatwym i skutecznym sposobem na wykazanie niezupełności powszechnych systemów aksjomatów dla matematyki, takich jak arytmetyka Peano. Jednak jak dotąd nie udało się wykazać nieudowodnialności (w PA lub jakimkolwiek silniejszym systemie) żadnych znanych, nieudowodnionych zdań, takich jak hipoteza Goldbacha, hipoteza liczb pierwszych bliźniaczych lub hipoteza Riemanna. W rzeczywistości wszystkie nieudowodnione zdania znalezione do tej pory zostały opracowane przez logików. Sytuacja jest inna w teorii mnogości, gdzie kilka twierdzeń otwartych od czasów Cantora okazało się później nieudowodnialnych na podstawie standardowych aksjomatów teorii mnogości. Dwa z najciekawszych to aksjomat wyboru i hipoteza continuum. Aksjomat wyboru stwierdza, że dowolny zbiór X niepustych zbiorów x ma funkcję wyboru: funkcję f taką, że $f(x) \in x$ dla każdego $x \in x$. Ten aksjomat jest używany niejawnie (a na początku był często używany nieświadomie), gdy zbiór jest definiowany za pomocą nieskończonego ciągu wyborów. Na przykład, aby udowodnić, że nieskończony zbiór X zawiera nieskończony ciąg elementów $x_1,x_2,x_3, \ldots$ chcemy powiedzieć: wybierz x_1 z X, następnie x_2 z $X - \{x_1\}$, następnie x_3 z $X - \{x_1,x_2\}$, i tak dalej. Ale nawet ta prosta definicja nie może być uzasadniona standardowymi aksjomatami teorii mnogości. Cohen (1963a,b) pokazał to, konstruując model standardowych aksjomatów, w którym istnieje nieskończony zbiór (liczb rzeczywistych) nie zawierający nieskończonego ciągu. Hipoteza continuum powstała z odkrycia Cantora, że $\mathbb{R}$ jest zbiorem większym niż $\mathbb{N}$. W swojej naiwnej formie hipoteza stwierdza, że $\mathbb{R}$ jest kolejnym największym zbiorem po $\mathbb{N}$. (Cantor podał również bardziej wyrafinowaną formę, obejmującą liczby porządkowe, której nie mam tu miejsca wyjaśnić). Cohen (1963) wykazał, że hipoteza continuum nie jest udowadnialna ze standardowych aksjomatów, poprzez skonstruowanie modelu tych aksjomatów, w którym hipoteza jest fałszywa. Należy również wspomnieć, że aksjomatu wyboru i hipotezy continuum nie można obalić ze standardowych aksjomatów. Pokazał to Gödel (1938), gdy skonstruował model, w którym oba stwierdzenia są prawdziwe. Tak więc dwa z najbardziej naturalnych pytań dotyczących zbiorów nieskończonych nie mogą zostać rozstrzygnięte przez naturalne aksjomaty. Sytuacja jest trochę podobna do sytuacji aksjomatu równoległego w stosunku do innych aksjomatów Euklidesa (lub Hilberta). Możemy swobodnie dodać albo aksjomat, albo jego negację i nie powstanie żadna sprzeczność, zakładając, że inne aksjomaty są spójne. Różnica polega na tym, że – o ile nam wiadomo – w przypadku teorii mnogości nie istnieją żadne oczywiste naturalne modele dla alternatyw – nie ma teorii mnogości „Cantora" i „nie-Cantora".

Zagadnienia filozoficzne

Biorąc pod uwagę rozwój od lat 30. do 60. XX wieku, możemy zobaczyć, jak daleko rozwinęła się filozofia matematyki od czasów starożytnej Grecji:

1 Pojęcie obliczalności, nieuznawane przez Greków, zostało precyzyjnie zdefiniowane i obecnie rządzi dużymi obszarami matematyki. Wyznacza ono granice koncepcji dowodu i rozstrzyga, czy wiele problemów jest rozwiązywalnych, czy nie.

2 Pojęcie dowodu w logice zostało opisane w całości (w przypadku logiki predykatów, najbardziej istotne dla matematyki): istnieje obliczenie, które generuje wszystkie logicznie poprawne formuły. Nie

ma jednak algorytmu, który decyduje, biorąc pod uwagę dowolną formułę, czy formuła ta jest udowodnialna (a zatem ważna), czy nie.

3 Pojęcie dowodu w matematyce jest niekompletne, a zatem nie jest absolutną prawdą matematyczną; w rzeczywistości dla dowolnego spójnego systemu F zawierającego pewną ilość arytmetyki istnieją prawdziwe stwierdzenia F, których nie da się udowodnić za pomocą F. Z jasnej strony, z zupełności logiki wynika, że możemy znaleźć wszystkie konsekwencje danego zestawu aksjomatów.

4 Twierdzenie Hilberta, że istnienie matematyczne jest tym samym, co spójność, zostało uznane za możliwe do utrzymania, ponieważ dowód Gödla na zupełność logiki predykatów daje model dla dowolnego spójnego zestawu aksjomatów (jak wspomniano w rozdziale 8.3). Jednak obliczalny zestaw aksjomatów nie musi mieć obliczalnego modelu, więc ten rodzaj istnienia nie jest akceptowalny dla konstruktywistów.

5 Spójność matematyki musi być kwestią intuicji; w tym sensie, że żaden rozsądnie silny (i spójny) system nie zawiera dowodu swojej własnej spójności. Program Hilberta do udowadniania spójności wydaje się być nie do naprawienia uszkodzony przez ten wynik.

6 Znaczenie i istnienie nieskończoności stało się bardziej skomplikowanym pytaniem. Kontinuum, które akceptuje większość matematyków, można pojąć tylko jako rzeczywistą nieskończoność, ponieważ jest niepoliczalne. Ale istnieje nieskończenie wiele rzeczywistych nieskończoności. Nie jest jasne, gdzie przeprowadzić granicę między akceptowalnym a nieakceptowalnym w rzeczywistych nieskończonościach.

7 Pomimo swoich trudności, kontinuum stało się podstawą analizy, stąd większość geometrii i większość fizyki matematycznej. Te dziedziny matematyki zostały zarytmetyzowano; to znaczy, w oparciu o aksjomaty dla liczb naturalnych plus pewne aksjomaty dla zbiorów nieskończonych.

8 Kontinuum nadal nie jest w pełni zrozumiane, ponieważ hipoteza kontinuum nie jest ustalona przez standardowe aksjomaty teorii mnogości. Z drugiej strony, nie znaleziono żadnego „konstruktywnego” analogu kontinuum, który byłby akceptowalny dla większości matematyków.