ROCZNIKI POLSKIEGO TOWARZYSTWA MATEMATYCZNEGO
Seria II: WIADOMOŚCI MATEMATYCZNE VII (1964)

Ministerstwo Nauki i Szkolnictwa Wyższego

L. J. MORDELL (Cambridge)

Równanie diofantyczne $y^2 = ax^3+bx^2+cx+d$ *

Jest rzeczą dobrze znaną, że najważniejsze osiągnięcia teorii liczb mają swój początek w zagadnieniach bardzo specjalnych. Tak na przykład wielkie twierdzenie Fermata, mówiące, że jeżeli n jest liczbą całkowitą większą od 2, to rozwiązania całkowite równania $x^n+y^n = z^n$ są określone warunkiem $xyz = 0$, doprowadziło do powstania teorii liczb algebraicznych, a więc znacznej części algebry współczesnej. Teoria form kwadratowych wielu zmiennych ma swój początek w wynikach Gaussa, Jacobiego i Eisensteina, dotyczących przedstawialności liczb naturalnych w postaci sumy $r \leqslant 10$ kwadratów oraz w twierdzeniu Legendre'a o rozwiązalności równania $ax^2+by^2+cz^2 = 0$.

Problem Waringa o przedstawialności liczby całkowitej w postaci sumy i-tych potęg liczb naturalnych był punktem wyjściowym prac Hilberta, Hardy'ego i Littlewooda oraz Winogradowa. Wynik Thuego o aproksymacji liczb algebraicznych liczbami wymiernymi doprowadził do mocniejszych twierdzeń Siegela i Dysona, a następnie do ostatecznego wyniku Rotha.

Badanie rozwiązań całkowitych równania

$$y^2 = ax^3+bx^2+cx+d, \tag{1}$$

gdzie a, b, c, d są liczbami całkowitymi doprowadziło mnie około 40 lat temu do twierdzenia o skończonej bazie dla punktów wymiernych na krzywej stopnia trzeciego. Po tym wyniku pojawiło się twierdzenie Weyla dotyczące krzywych genusu większego od 1, a następnie twierdzenie Siegela o skończoności liczby rozwiązań całkowitych ogólnego równania wielomianowego (z wyjątkiem pewnych przypadków szczególnych). Pomimo tych wyników wiele uwagi poświęca się obecnie zagadnieniom związanym z punktami wymiernymi na krzywej stopnia trzeciego. Równanie (1) uwydatnia dwie cechy charakterystyczne teorii liczb: jej wieczną młodość — to, że stosowanie starych metod i wyników prowadzi do

* Wykład wygłoszony 7 listopada 1963 roku w Instytucie Matematycznym PAN na seminarium z teorii liczb prowadzonym przez prof. W. Sierpińskiego.

nowych, oraz powstawanie istotnych i ważnych osiągnięć z prostych zagadnień i pomysłów. Równanie (1) pojawiło się najpierw w swej szczególnej postaci

$$y^2+k=x^3. \tag{2}$$

Przypadki szczególne $k=2, 4$ znane są z korespondencji Fermata z matematykami angielskimi. W jednym z listów postawił on następujący problem: *Udowodnić, że rozwiązania w liczbach całkowitych w przypadku* $k=2$ *są wyznaczone przez warunek* $x=3$, *a w przypadku* $k=4$, *przez warunek* $x=2, 5$. Postawienie tego problemu było kamieniem węgielnym w rozwoju teorii liczb. Wyniku tego nie potrafiono udowodnić; dopiero sto lat później Euler podał dowód, który okazał się błędny. Warto zanalizować dokładniej przypadek $k=2$.

Napiszmy równanie (2) dla $k=2$ w postaci

$$(y+\sqrt{-2})(y-\sqrt{-2})=x^3. \tag{3}$$

Z równości $YZ=X^3$, gdzie X, Y, Z są liczbami całkowitymi wymiernymi, łatwo jest wyprowadzić wniosek, że jeżeli Y, Z są liczbami względnie pierwszymi, to

$$Y=\pm B^3, \quad Z=\pm C^3, \quad X=\pm BC, \tag{4}$$

gdzie pierwsze dwa znaki są niezależne jeden od drugiego (mogą być one zresztą włączone do podstaw sześcianów). W przypadku, gdy Y i Z nie są względnie pierwsze, wzory (4) muszą być nieco zmodyfikowane.

Powstaje teraz pytanie, czy z równania (3) wynika, że

$$y+\sqrt{-2}=(a+b\sqrt{-2})^3, \quad y-\sqrt{-2}=(a-b\sqrt{-2})^3, \quad x=a^2+2b^2, \tag{5}$$

gdzie a, b są liczbami wymiernymi całkowitymi. Jeżeli tak jest, to z równości (5) otrzymujemy

$$1=b(3a^2-2b^2), \tag{6}$$

skąd, oczywiście, $b=1$, $a=\pm 1$, $x=3$.

Trudno byłoby wywnioskować coś z równania (6), gdyby nie przedstawiało ono szczególnej krzywej stopnia trzeciego względem a i b. Zawsze jednakże możemy założyć, że zachodzą wzory (5), a to może nas doprowadzić do znalezienia rozwiązań w liczbach całkowitych. Istotnym pytaniem jest, czy równości (5) dają wszystkie rozwiązania całkowite. W tym celu musimy poznać arytmetykę ciała kwadratowego $R(\sqrt{-2})$. Łatwo jest wykazać, że dogodnie jest zdefiniować nowe liczby całkowite jako liczby postaci $a+b\sqrt{-2}$, gdzie a i b są liczbami całkowitymi wymiernymi i że jedynymi jednościami, tzn. dzielnikami liczby 1, są 1 oraz -1. Najważniejszą własnością tego ciała jest to, że zachodzi w nim algorytm

Euklidesa, a więc i twierdzenie o jednoznaczności rozkładu na czynniki pierwsze. Stąd, jak poprzednio, z równości $YZ = X^3$ i $(Y, Z) = 1$ wynika, że $Y = B^3$, $Z = C^3$, $X = BC$. W równości (3) x nie jest parzyste, a więc $(y+\sqrt{-2}, y-\sqrt{-2})$ jako dzielnik $2\sqrt{-2}$ jest równy 1. Stąd wynikają równości (5).

Dowody dla $k = -2, -4$ mogły być podane przez Fermata bez korzystania z liczb zespolonych i algebraicznej teorii liczb. Fermat twierdził jednak, że posiada godną podziwu metodę badania takich równań. Z pewnością nie mógł zajmować się on równaniem w postaci ogólnej (2), ani w szczególności równaniem $y^2 = x^3+17$, które ma osiem rozwiązań, w których największa wartość x jest równa 5234. Landau i Ostrowski udowodnili, że równanie (2) ma skończoną ilość rozwiązań w liczbach całkowitych (dla $k \neq 0$). Inna metoda udowodnienia, że w pewnych przypadkach równanie (2) nie ma rozwiązań w liczbach całkowitych pochodzi od V. A. Lebesgue'a.

Dla uproszczenia rozpatrzmy równanie

$$y^2 = x^3+7. \tag{7}$$

Oczywiście mamy tu $x \not\equiv 0 \pmod 2$, $x \not\equiv -1 \pmod 4$, a więc $x \equiv 1 \pmod 4$. Napiszmy równanie (7) w postaci

$$y^2+1 = (x+2)(x^2-2x+4).$$

Widzimy, że $x+2$ jest dodatnie i $x+2 \equiv 3 \pmod 4$, a więc nie może być dzielnikiem liczby y^2+1.

Obydwa pomysły badania rozwiązań w liczbach całkowitych równania (2) będą stosowane w dalszym ciągu niniejszego artykułu.

Fermat stawiając wspomniany problem napisał również, że łatwo byłoby znaleźć nieskończenie wiele rozwiązań w liczbach wymiernych równania $y^2+2 = x^3$ stosując metodę Bacheta. Tak więc podstawiając $x = 3+X$ otrzymujemy

$$y^2 = 25+27X+9X^2+X^3.$$

Podstawiając $y = 5+\frac{27}{10}X$ otrzymujemy $X = \frac{129}{100}$. Postępowanie takie może być kontynuowane i możemy oczekiwać, że znajdziemy nieskończenie wiele rozwiązań w liczbach wymiernych.

Szukanie rozwiązań równania w liczbach wymiernych może być czasami zagadnieniem trywialnym, czasami zaś problemem bardzo trudnym. Tak na przykład, wszystkie rozwiązania w liczbach wymiernych równania

$$y^2 = ax^3+bx^2z+cxz^2+dz^3$$

otrzymujemy natychmiast podstawiając $x = py$, $z = qy$, gdzie p i q są dowolnymi liczbami wymiernymi. Znalezienie wszystkich rozwiązań

w liczbach całkowitych stanowi poważny problem teorii liczb. Częściowe jego rozwiązanie było podane przez Lagrange'a. W roku 1912 udowodniłem, że wszystkie rozwiązania w liczbach całkowitych spełniające pewne proste warunki, np. $(x, z) = 1$, są dane skończoną liczbą wyrażeń postaci

$$x = f(p, q), \quad z = g(p, q), \quad y = h(p, q),$$

gdzie f i g są wielomianami stopnia czwartego względem p i q, h — stopnia szóstego, a p i q przybierają wszystkie wartości całkowite. W tym czasie nie znałem twierdzenia Thuego, ogłoszonego w roku 1909, a mówiącego, że równanie diofantyczne $g(p, q) = 1$ ma, z wyjątkiem pewnych trywialnych przypadków, skończoną liczbę rozwiązań w liczbach całkowitych. Wynika stąd, że równanie (1) ma tylko skończoną ilość rozwiązań w liczbach całkowitych, jeżeli prawa jego strona nie ma czynnika będącego kwadratem wielomianu liniowego względem x. Kilka lat później znalazłem inny dowód tego faktu. Próbowałem wówczas wykazać, że równanie $y^2 = ax^4 + bx^3 + cx^2 + dx + e$ ma tylko skończoną ilość rozwiązań w liczbach całkowitych. Wysiłki moje nie zostały uwieńczone powodzeniem, lecz za to znalazłem rozwiązanie znacznie ważniejszego problemu. Punkt wyjścia stanowiła metoda Bacheta znajdowania rozwiązań równania $y^2 + 2 = x^3$ w liczbach wymiernych.

Niech będzie dane równanie stopnia trzeciego $f(x, y) = 0$, gdzie f jest wielomianem nieprzywiedlnym o współczynnikach wymiernych. Przypuśćmy, że krzywa $f(x, y) = 0$ nie ma punktu podwójnego, a więc jest genusu 1. Punkt leżący na krzywej będziemy nazywali *wymiernym*, jeżeli obydwie jego współrzędne są wymierne. Przypuśćmy, że znamy jeden punkt wymierny P_1. Styczna w punkcie P_1 przecina krzywą w punkcie wymiernym P_{11} na ogół różnym od P_1 (P_{11} pokrywa się z P_1 tylko wtedy, gdy P_1 jest punktem przegięcia). W podobny sposób otrzymamy z punktu P_{11} punkt $P_{11,11}$. Postępowanie to da na ogół nieskończoną ilość punktów wymiernych, ale w pewnych przypadkach szczególnych może prowadzić tylko do skończonej ich liczby. Przypuśćmy teraz, że znamy jeszcze inny punkt wymierny, P_2. Wówczas prosta P_1P_2 przecina krzywą w innym punkcie wymiernym P_{12}. Jest rzeczą oczywistą, że jeżeli dany jest zbiór punktów wymiernych $P_1, P_2, \ldots, P_n$, to możemy zastosować opisaną metodę siecznych i stycznych do tego zbioru i zbioru punktów otrzymywanych tą metodą. Około czterdziestu lat temu udowodniłem twierdzenie nazywane obecnie *twierdzeniem Mordella o skończonej bazie*, mówiące, że *wszystkie punkty wymierne mogą być otrzymane ze skończonej ich liczby przez zastosowanie metody siecznych i stycznych*. Okazało się, że twierdzenie to odgrywa podstawową rolę w teorii równań diofantycznych stopnia trzeciego. Stwierdzono, że nie jest łatwą rzeczą znaleźć wszystkie rozwiązania w liczbach całkowitych równania (1), chociaż liczba ich jest skończona. Zostało to zrobione w przypadku pewnych

równań, które pojawiły się przy próbach znalezienia takich równań, które miałyby określoną liczbę $N = 1, 2, 3, \ldots$ rozwiązań w liczbach wymiernych. Czasami z równaniami postaci (1) wiążą się interesujące pytania, na przykład problem kwadratowej piramidy: udowodnić, że $x = 1, 24$ są jedynymi rozwiązaniami w liczbach naturalnych równania

$$6y^2 = x(x+1)(2x+1).$$

Zostało to udowodnione przez G. N. Watsona. Wiadomości arytmetyczne wymagane do rozwiązania tego problemu nie są trudne. Równanie to prowadzi do równań postaci

$$x = pa^2, \quad x+1 = qb^2, \quad 2x+1 = rc^2,$$

gdzie p, q, r są czynnikami liczby 6. Ponieważ $qb^2 - pa^2 = 1$ itd., korzystamy tylko z teorii ciał kwadratowych.

Innym zagadnieniem jest udowodnienie, że 6 i 210 są jedynymi liczbami naturalnymi, które są iloczynami zarówno dwóch, jak i trzech kolejnych liczb całkowitych $6 = 2\cdot 3 = 1\cdot 2\cdot 3$, $210 = 14\cdot 15 = 5\cdot 6\cdot 7$. Problem ten prowadzi do równania

$$y(y+1) = x(x+1)(x+2).$$

Podstawiając $2y+1 = Y$, $2x+2 = X$ otrzymujemy

$$2Y^2 = X^3 - 4X + 2. \tag{8}$$

Należy teraz wykazać, że wszystkimi wartościami całkowitymi X, spełniającymi równanie (8) są $X = 0, \pm 2, 4, 12$. Niech θ będzie pierwiastkiem równania $t^3 - 4t + 2 = 0$. Wówczas równanie (8) można napisać w postaci

$$2Y^2 = (X-\theta)(X^2+\theta X+\theta^2-4). \tag{9}$$

Podobnie, jak równanie $y^2+2 = x^3$ wymagało badania ciała kwadratowego $R(\sqrt{-2})$, równanie (8) wymaga badania własności arytmetycznych ciała sześciennego $R(\theta)$. Można udowodnić, że

1° liczbami całkowitymi tego ciała są liczby $a+b\theta+c\theta^2$, gdzie a, b, c są liczbami całkowitymi wymiernymi,

2° jednościami, tzn. dzielnikami liczby 1, są liczby $\pm\varepsilon^l\eta^m$, gdzie $\varepsilon = \theta-1$, $\eta = 2\theta-1$, a l, m przybierają wszystkie wartości całkowite.

3° w ciele $R(\theta)$ zachodzi twierdzenie o jednoznaczności rozkładu na czynniki.

Z równania (9) wynika więc, że

$$X-\theta = \pm d\varepsilon^l\eta^m(a+b\theta+c\theta^2)^2, \tag{10}$$

gdzie d zależy tylko od wspólnego dzielnika liczb $X-\theta$ i $X^2+\theta X+\theta^2-4$; mamy tu $d = \theta(4\theta-3)^n$, gdzie n jest pewną liczbą całkowitą nieujemną.

Ponieważ parzyste potęgi ε, η oraz $4\theta-3$ mogą być zawarte w wyrażeniu $(a+b\theta+c\theta^2)^2$, możemy ograniczyć się tylko do rozpatrzenia przypadków $l, m, n = 0, 1$. Niemożliwe jest, by $n = 1$, gdyż biorąc normy obydwu stron równania (10) mielibyśmy wówczas

$$X^3-4X+2 = \pm 2\cdot 37Z^2,$$

gdzie Z jest liczbą całkowitą wymierną. Tak więc

$$X = 2X_1 \quad \text{i} \quad 4X_1^3-4X_1+1 = \pm 37Z^2.$$

Ponieważ $x_1^3-x_1 \equiv 0 \pmod 2$, otrzymujemy kongruencję mod 8 niemożliwą. Pozostają więc do rozpatrzenia cztery przypadki:

$$\langle l, m\rangle = \langle 0, 0\rangle, \langle 1, 0\rangle, \langle 0, 1\rangle, \langle 1, 1\rangle.$$

Tak więc z równania (10), po uprzednim wyrażeniu θ^3, θ^4 przez θ^2, θ, 1 i przyrównaniu wyrazów z θ, θ^2 po obydwu stronach, otrzymujemy cztery układy równań kwadratowych względem a, b, c. W przypadku ogólnym zagadnienia takie są trudne do rozstrzygnięcia, w naszym jednak przypadku tak nie jest. Gdy $l = 1$, otrzymujemy równanie sprzeczne mod 4. Przypadek $l = 0$, $m = 0$ jest łatwy, gdyż prowadzi do równań

$$b(a+4c) = c^2, \qquad (a+4c)^2+4b^2-4bc = 1.$$

Jeżeli $b = 0$, to $c = 0$, $a = \pm 1$, jeżeli zaś $b \neq 0$, to

$$\frac{c^4}{b^2}+4b^2-4bc = 1.$$

Lewa strona traktowana jako funkcja c przyjmuje najmniejszą wartość dla $c = b$, skąd otrzymujemy $b = \pm 1$, $a = \mp 3$. Przypadek $l = 0$, $m = 1$ także daje się łatwo rozwiązać.

Zauważyłem niedawno, że metoda, którą stosowaliśmy do równania (7), pozwala znaleźć równania postaci (1) z dokładnie N rozwiązaniami w x ($N = 1, 2$).

Rozważmy równanie

$$y^2+d^2 = 2Ax^3+(6-2A-2C+8B)x^2+2Cx, \tag{11}$$

gdzie $A > 0$, B, C są dowolnymi liczbami całkowitymi i d jest liczbą nieparzystą niepodzielną przez żadną liczbę pierwszą $\equiv 3 \pmod 4$. Mamy oczywiście $x \not\equiv 0 \pmod 2$, gdyż w przeciwnym przypadku $y^2 \equiv -1 \pmod 4$, a także $x \not\equiv 1 \pmod 4$, gdyż dla $x \equiv 1 \pmod 4$ mielibyśmy

$$y^2+1 \equiv 2A+(6-2A-2C+8B)+2C \equiv 6 \pmod 8.$$

Tak więc $x \equiv 3 \pmod 4$, ale x nie może być dodatnie, gdyż liczba dodatnia $\equiv 3 \pmod 4$ nie dzieli y^2+d^2. Tak więc jedynymi możliwymi

wartościami x są liczby ujemne $\equiv 3 (\mathrm{mod}\, 4)$. Muszą one także spełniać nierówność

$$2Ax^2+(6-2A-2C+8B)x+2C<0,$$

skąd łatwo oszacować x.

Jeżeli teraz przyjmiemy $d=1$ i zażądamy, by $x=-1$, $y=\pm 1$ było rozwiązaniem, otrzymamy $C=1-A+2B$ i równanie (11) możemy napisać w postaci

$$y^2+1 = 2Ax^3+(4+4B)x^2+2(1-A+2B)x, \quad A>0. \tag{12}$$

Wartości $x=-1$, $y=\pm 1$ będą jedynymi rozwiązaniami tego równania, jeżeli $3A>1+B$. Następna możliwa wartość x będzie -5, gdyż $x \not\equiv 0, 1, 2 (\mathrm{mod}\, 4)$. Udowodnimy, że dla $x \leqslant -5$ mamy

$$2Ax^3+(4+4B)x^2+2(1-A+2B)x-1<0,$$

czyli

$$(x+1)[2Ax^2+(4-2A+4B)x-2]<-1.$$

Aby to zachodziło, wystarcza by

$$2Ax^2+(4-2A+4B)x-2>0 \quad \text{lub} \quad 10A-(4-2A+4B)>0.$$

Jeżeli w równaniu (12) przyjmiemy $B=3A+1$, $A>0$:

$$y^2 = 2Ax^3+(12A+8)x^2+(10A+6)x-1, \tag{13}$$

to będzie ono miało rozwiązanie $x=-5$, $y=\pm 13$. Udowodnimy, że $x=-1, -5$ wyznaczają wszystkie rozwiązania w liczbach całkowitych. Wystarczy w tym celu wykazać, że jeżeli $x \leqslant -9$, to

$$2Ax^3+(12A+8)x^2+(10A+6)x-1<0,$$

czyli

$$(x+1)[2Ax^2+(10A+8)x-2]<-1.$$

Aby zachodziła ostatnia nierówność wystarczy by

$$Ax^2+(5A+4)x-1>0 \quad \text{lub} \quad 9A>5A+4, \quad \text{a więc} \quad A>1.$$

Gdy $A=1$, otrzymujemy równanie

$$y^2 = 2x^3+20x^2+16x-1.$$

Ponieważ $x=-9$ nie jest rozwiązaniem tego równania, wynik pozostaje prawdziwy dla $A=1$.

Tak więc jedynymi rozwiązaniami równania (13) są $x = -1, -5$. Można wykazać, że wyniki tego samego typu zachodzą dla równań postaci

$$y^2 \pm kd^2 = ax^3 + bx^2 + cx, \quad \text{gdzie} \quad k = 2, 3, 5.$$

Łatwo można otrzymać wyniki ogólniejsze, dotyczące równań $y^2 = f(x)$, gdzie $f(x)$ jest wielomianem o współczynnikach całkowitych, pisząc równanie w postaci

$$y^2 \pm kd^2 = a_0 x^n + a_1 x^{n-1} + \ldots + a_{n-1} x.$$

Warto zauważyć, że w roku 1926 Siegel w anonimowym artykule podpisanym X udowodnił, że w ogólności równanie $y^2 = f(x)$ ma tylko skończoną liczbę rozwiązań.