O znaczeniu równości

Dr Wiktor BARTOL

Czy język francuski jest bogatszy od polskiego? Tak postawione pytanie rodzi następne: co to właściwie znaczy ,,bogatszy"? Czy bogatszy jest ten język, który ma więcej słów, czy ten, w którym można więcej wyrazić, który pozwala dokładniej opisywać rzeczywistość? I wreszcie — dlaczego takie pytania pojawiają się w *Delcie*?

Nie będziemy mówili tu ani o języku francuskim, ani o języku polskim. Zajmiemy się językiem matematyki — bo przecież matematyka ma swój język i to nie jeden.

Upraszczając nieco można powiedzieć, że przedmiotem badań matematyka są pewne obiekty, operacje na nich określone i relacje wyrażające związki między nimi. I tak np. matematyk-analityk (czyli zajmujący się analizą matematyczną) bada przede wszystkim liczby rzeczywiste, funkcje określone na zbiorach liczb rzeczywistych (od funkcji najprostszych, takich jak dodawanie i mnożenie, aż po bardzo złożone — można je zresztą na ogół zdefiniować wychodząc od tych prostych), a także pewne relacje, jak np. relacja niewiększości. Po to by móc wygłaszać i zapisywać zdania mówiące o badanych obiektach trzeba mieć zatem nazwy dla interesujących nas funkcji i relacji, a także zmienne, które mogłyby reprezentować obiekty. Czasem możemy też potrzebować nazw dla pewnych konkretnych obiektów, które z jakichś powodów chcielibyśmy wyróżnić. Wróćmy do naszego analityka: musi on mieć nazwę dla funkcji dodawania (niech będzie nią ,,d") i nazwę dla mnożenia (wybierzmy symbol ,,m"), a także nazwę dla relacji niewiększości — użyjmy tu tradycyjnego symbolu ,,$\leqslant$". Pozwólmy mu też używać nazwy ,,0" do określenia tej jedynej liczby rzeczywistej, która dodana do jakiejkolwiek liczby nie zmienia jej wartości oraz nazwy ,,1" do określenia tej jedynej liczby rzeczywistej, przez którą można pomnożyć dowolną liczbę nie zmieniając jej wartości. Niech x, y, z itp. będą zmiennymi, reprezentującymi liczby rzeczywiste. Oznaczmy wreszcie zbiór wszystkich liczb rzeczywistych przez R. Dziedzinę badań matematyka-analityka możemy teraz opisać wymieniając nazwy interesujących go obiektów, funkcji i relacji: $(R, \mathrm{d}, \mathrm{m}, \leqslant, 0, 1)$ — rozumiejąc, że ,,d" i ,,m" są nazwami funkcji dwuargumentowych, ,,$\leqslant$" jest nazwą relacji dwuargumentowej (tzn. mówiącej o związku między dwoma obiektami), a ,,0" i ,,1" są nazwami pewnych wyróżnionych elementów zbioru R.

W tym momencie zajrzał mi przez ramię mój przyjaciel, matematyk-mnogościowiec (a więc zajmujący się teorią mnogości czyli teorią zbiorów). ,,Jak to — powiada — przecież opisałeś tu dziedzinę, którą ja się właśnie zajmuję. Interesują mnie własności podzbiorów zbioru liczb naturalnych. Zbiór wszystkich takich podzbiorów nazwałem R, nazwę ,,d" nadałem operacji sumowania zbiorów, nazwa ,,m" oznacza operację iloczynu (części wspólnej) zbiorów (są to więc nazwy funkcji dwuargumentowych), natomiast symbol ,,$\leqslant$" jest w moim systemie nazwą relacji zawierania. Użyłem też nazwy ,,0" dla zbioru pustego i nazwy ,,1" dla zbioru wszystkich liczb naturalnych — który jest przecież sam swoim podzbiorem."

Wobec takiego podobieństwa zaczęliśmy poszukiwać własności, które pozwoliłyby odróżnić jedną dziedzinę od drugiej. Oczywiście mogliśmy wyrazić taką własność tylko w języku zawierającym symbole ,R" ,,d", ,,m" itd.; umówiliśmy się też, że w każdym języku — a więc i w naszym — będzie występował symbol ,,=", który będziemy interpretować (czyli rozumieć) zawsze jako równość. Na początek wybraliśmy własność następującą:

$$\mathrm{m}(x, y) = \mathrm{m}(y, x).$$

,,Moja dziedzina ma tę własność — powiedział mnogościowiec — jeśli wezmę iloczyn dwóch dowolnych podzbiorów x i y, to otrzymany zbiór będzie równy iloczynowi podzbiorów y i x." Przekonaliśmy się też, że jeśli rozumieć ,,m" jako mnożenie liczb rzeczywistych, to również w dziedzinie liczb rzeczywistych występuje powyższa własność — jest to przecież przemienność mnożenia. Szukaliśmy więc dalej:

$$x \leqslant \mathrm{d}(x, y).$$

Sprawdziliśmy najpierw dziedzinę mojego przyjaciela, stwierdzając, że ma ona powyższą własność: rzeczywiście, każdy zbiór jest zawarty w swojej sumie z dowolnym zbiorem y. Czy powyższa nierówność będzie prawdziwa dla dowolnych liczb rzeczywistych x i y (gdy ,,d" jest dodawaniem)?

Łatwo zauważyliśmy, że nie: jeśli x jest liczbą 2, a y liczbą -10, to nierówność nie będzie prawdziwa. Tak więc znaleźliśmy własność odróżniającą dwie rozpatrywane przez nas dziedziny, opisane tym samym językiem.
Zachęceni tym sukcesem, zastanawialiśmy się dalej: czy można znaleźć taką rozróżniającą własność dającą się zapisać przy pomocy najprostszej relacji, jaką jest równość? Okazało się, że tak. Taką własnością jest np.

$$x = \mathrm{d}(x, x).$$

Jest ona spełniona przez dowolny podzbiór zbioru liczb naturalnych (gdy „d" interpretujemy jako sumę zbiorów), natomiast nie każda liczba rzeczywista ją spełnia (gdy „d" oznacza sumę liczb rzeczywistych).

Wypisując wszystkie te własności dziedziny mojego przyjaciela mnogościowca, które dadzą się wyrazić przy pomocy relacji równości jako jedynej relacji, wyznaczylibyśmy pewną klasę dziedzin: tych, którym te własności przysługują. I tak np. do tej klasy należałaby dziedzina składająca się z rodziny wszystkich podzbiorów dowolnego zbioru A oraz operacji sumy i iloczynu zbiorów z wyróżnionym zbiorem pustym i całym zbiorem A (co może ciekawsze, do tej klasy należałyby także wszystkie tzw. algebry Boole'a; to stwierdzenie nie jest już jednak całkiem trywialne).
Jak widzieliśmy wyżej, nie należy do wyznaczonej klasy zbiór liczb rzeczywistych z dodawaniem i mnożeniem, zerem i jedynką.

Jak wyglądają klasy dziedzin — w przypadku, gdy jedyną rozpatrywaną relacją jest równość, dziedziny nazywamy algebrami — a więc, jak wyglądają klasy algebr, które dadzą się w podobny sposób opisać przy pomocy pewnego zbioru równości? Mówiąc inaczej: jakie klasy algebr można scharakteryzować przy pomocy równości? A *scharakteryzować* — to znaczy znaleźć taki zbiór równości E, że algebra należy do charakteryzowanej klasy wtedy i tylko wtedy, gdy jest w niej prawdziwa każda równość ze zbioru E.

Odpowiedź na powyższe pytania została udzielona w latach trzydziestych XX wieku przez amerykańskiego matematyka Garretta Birkhoffa, który zapoczątkował badania abstrakcyjnych struktur składających się ze zbioru i operacji określonych na tym zbiorze. Algebra uniwersalna (tak został nazwany ten kierunek badawczy) zajmuje się takimi własnościami struktur, które nie zależą od tego, czym są ich elementy i jak konkretnie są na nich określone operacje.

Aby móc tu przedstawić ową odpowiedź Birkhoffa, musimy uściślić lub wprowadzić kilka pojęć właśnie z zakresu algebry uniwersalnej. Każdą strukturę postaci $(A, f_1, \ldots, f_n)$, gdzie A jest pewnym niepustym zbiorem, a $f_1, \ldots, f_n$ są operacjami na zbiorze A (a więc funkcjami, które elementom zbioru A — a mogą to być funkcje niekoniecznie jednej zmiennej — przyporządkowują znów elementy tego zbioru), nazwiemy algebrą. Napisałem wyżej, że operacje w algebrze mogą być funkcjami „niekoniecznie jednej zmiennej"; mogą to być funkcje o więcej niż jednej (i wtedy sprawa jest jasna) bądź mniej niż jednej zmiennej (argumencie). Ale czy może istnieć funkcja bez zmiennych? Otóż tak — i można takie funkcje dokładnie opisać na gruncie teorii mnogości. Tu zauważmy, że wartość funkcji „bez zmiennych" musi być stała, bo przecież od niczego nie może zależeć. Każda taka funkcja wybiera zatem jakiś element ze zbioru A, możemy ją więc z tym elementem utożsamić. Okazuje się zatem, że w ten sposób zawarliśmy w naszym pojęciu algebry także to, że mogą w niej występować jakieś elementy wyróżnione.

Będziemy mówili, że algebry $(A, f_1, \ldots, f_n)$ i $(B, g_1, \ldots, g_m)$ są podobne, jeśli mają tyle samo operacji (czyli $n = m$) i ponadto operacje jednej z nich mają tyle samo argumentów, co odpowiednie operacje drugiej; mówiąc dokładniej, dla każdego $i = 1, \ldots, n$, operacje f_i i g_i mają tyle samo argumentów, są funkcjami tej samej liczby zmiennych. I tak np. algebra $(R, +, \cdot, 0, 1)$ liczb rzeczywistych z operacjami dodawania, mnożenia, zerem i jedynką jest podobna do algebry $(P(N), \cup, \cap, \emptyset, N)$, gdzie $P(N)$ oznacza zbiór wszystkich podzbiorów zbioru liczb naturalnych N, $\cup$ i $\cap$ są odpowiednio operacjami sumy i iloczynu zbiorów, a zbiór pusty $\emptyset$ i zbiór wszystkich liczb naturalnych N są wyróżnionymi elementami zbioru $P(N)$.
Właśnie dlatego, że algebry te są podobne, mogliśmy używać do ich opisu tego samego języka, przedstawionego na początku artykułu (pomijamy tu relację $\leqslant$).

Niech więc $\mathbf{A} = (A, f_1, \ldots, f_n)$ i $\mathbf{B} = (B, g_1, \ldots, g_n)$ będą algebrami podobnymi. Przekształcenie $h: A \to B$ nazwiemy *homomorfizmem* pierwszej algebry w drugą, jeśli dla każdej operacji f_i oraz dowolnych elementów $a_1, \ldots, a_{k_i} \in A$ (gdzie k_i jest liczbą argumentów operacji f_i — a więc i g_i) zachodzi równość

$$h(f_i(a_1, \ldots, a_{k_i})) = g_i(h(a_1), \ldots, h(a_{k_i})).$$

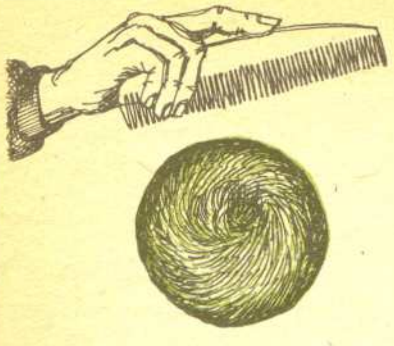
Rozwiązanie zadania M 300. Podstawiając $a = x$, $b = -y$, $c = -z$ w tożsamości

$$a^3 + b^3 + c^3 - 3abc = (a + b + c)(a^2 + b^2 + c^2 - ab - bc - ac)$$

otrzymamy z pierwszego równania

$$0 = x^3 - y^3 - z^3 - 3xyz = (x - y - z)(x^2 + y^2 + z^2 + xy - yz + zx) = (x - y - z)(x^2 + (y - z)^2 + xy + yz + xz),$$

skąd, wobec dodatniości x, y, z, wynika, że $x - y - z = 0$, czyli $x = y + z$. Z drugiego równania mamy teraz $x^2 = 2x$, skąd otrzymujemy, że $x = 2$ i $y = z = 1$.



Mówiąc prościej: jeśli najpierw wykonamy na elementach $a_1, \ldots, a_{k_i}$ algebry **A** operację f_i, a wynik tego działania przeniesiemy przy pomocy funkcji h do algebry **B**, to otrzymamy to samo, co przy przeniesieniu do algebry **B** elementów $a_1, \ldots, a_{k_i}$ i wykonaniu na ich obrazach operacji g_i. O takiej funkcji h algebraicy mówią krótko, że „zachowuje operacje".

Jeśli przy tym funkcja h przeprowadza zbiór A na zbiór B (czyli każdy element zbioru B jest obrazem — ze względu na h — pewnego elementu zbioru A), to algebrę **B** nazwiemy *obrazem homomorficznym* algebry **A**. Możemy np. łatwo sprawdzić, że algebra $(R, +)$ jest obrazem homomorficznym algebry $(R^+, \cdot)$ — gdzie R^+ jest zbiorem wszystkich dodatnich liczb rzeczywistych — ze względu na homomorfizm $h: R^+ \to R$ taki, że $h(a) = \log a$ dla każdej liczby rzeczywistej dodatniej a.

Zatrzymajmy się na chwilę przy algebrze $(R^+, \cdot)$. Zauważmy, że jeśli wykonamy mnożenie dowolnych dwóch liczb rzeczywistych z przedziału $(0,1]$, to otrzymamy znów liczbę należącą do tego przedziału; możemy więc powiedzieć, że przedział $(0,1]$ jest zamknięty ze względu na jedyną operację algebry $(R^+, \cdot)$; nie jest tak np. z przedziałem $(1,2)$ ani z podzbiorem $\{1/2, 1/3, 1/4\}$. Każdy podzbiór zbioru R^+, który ma taką własność, nazwiemy podalgebrą algebry $(R^+, \cdot)$. Ogólniej, gdy $\mathbf{A} = (A, f_1, \ldots, f_n)$ jest dowolną algebrą, a B jest podzbiorem zbioru A takim, że wykonanie którejkolwiek z operacji $f_1, \ldots, f_n$ na elementach ze zbioru B daje w wyniku element zbioru B, to B nazwiemy *podalgebrą* algebry **A**. Ściślej mówiąc, podalgebrą będzie algebra $\mathbf{B} = (B, f_1, \ldots, f_n)$ z operacjami zredukowanymi do zbioru B, ale przecież wskazaliśmy już operacje mówiąc o algebrze **A**, więc tu możemy je pominąć. Możemy sprawdzić, że oprócz przedziału $(0,1]$ podalgebrami algebry $(R^+, \cdot)$ są np. przedział $\left(0, \frac{1}{2}\right)$, zbiór liczb całkowitych dodatnich, zbiór liczb wymiernych dodatnich, ale nie zbiór wszystkich liczb pierwszych.

Potrafimy już otrzymać z danej algebry **A** nowe algebry przekształcając **A** przez homomorfizm albo wybierając z niej podzbiory zamknięte ze względu na wszystkie operacje algebry **A**. Musimy teraz omówić konstrukcję nieco bardziej skomplikowaną.

Niech $\mathbf{A} = (A, f_1, \ldots, f_n)$ i $\mathbf{B} = (B, g_1, \ldots, g_n)$ będą algebrami podobnymi. Utwórzmy ze wszystkich elementów zbioru A i zbioru B pary tak, aby w każdej parze na pierwszym miejscu był element zbioru A, a na drugim — element zbioru B. Taką parę utworzoną z elementu $a \in A$ i elementu $b \in B$ oznaczmy symbolem $\langle a, b\rangle$, a $A \times B$ niech będzie zbiorem wszystkich par tego typu. I tak np. gdy $A = B = R$, to zbiorem wszystkich takich par (czyli zbiorem $R \times R$) jest zbiór wszystkich par liczb rzeczywistych, który możemy utożsamić z płaszczyzną euklidesową. Zdefiniujemy w zbiorze $A \times B$ operacje $F_1, \ldots, F_n$ tak, aby otrzymać algebrę $(A \times B, F_1, \ldots, F_n)$ podobną do algebr A i B. Przypuśćmy dla uproszczenia, że f_i i g_i są operacjami dwuargumentowymi — zatem F_i też musi być operacją dwuargumentową. Niech $\langle a, b\rangle$ i $\langle a_1, b_1\rangle$ będą parami ze zbioru $A \times B$; wówczas $F_i(\langle a, b\rangle, \langle a_1, b_1\rangle) = \langle f_i(a, a_1), g_i(b, b_1)\rangle$. Operacja F_i działa więc „na współrzędnych". Jej wartością dla danych dwóch par $\langle a, b\rangle$ i $\langle a_1, b_1\rangle$ jest para, w której na pierwszym miejscu jest wynik działania operacji f_i pierwszej algebry **A** na „pierwszych współrzędnych" tych par, a na drugim miejscu jest wynik działania operacji g_i drugiej algebry **B** na „drugich współrzędnych". Jeśli więc np. $\mathbf{A} = \mathbf{B} = (R, +)$, to w zbiorze $R \times R$ definiujemy „dodawanie" tak: $\langle x, y\rangle + \langle x_1, y_1\rangle = \langle x + x_1, y + y_1\rangle$. Algebrę zbudowaną w powyższy sposób z algebr **A** i **B** nazywamy *produktem* algebr **A** i **B** i oznaczamy przez $\mathbf{A} \times \mathbf{B}$.

Nietrudno taką konstrukcję uogólnić i mówić o produkcie trzech, czterech itd. algebr. Myślę, że potraficie to zrobić bez kłopotu. Dodajmy jeszcze, że można zbudować produkt nie tylko skończonej rodziny algebr, ale także nieskończonej rodziny algebr podobnych. W takim przypadku jednak konstrukcja wymaga bardziej zawiłego opisu, pomińmy ją zatem.

Jesteśmy już teraz w stanie przytoczyć obiecane twierdzenie Birkhoffa, mówiące o klasach charakteryzowanych przez równości (jak mówią matematycy — *definiowalnych równościowo*). Otóż:

Klasa $\mathscr{K}$ algebr podobnych jest definiowalna równościowo wtedy i tylko wtedy, gdy jest zamknięta ze względu na obrazy homomorficzne, podalgebry i produkty.

Co to znaczy, że klasa jest *zamknięta* ze względu na jakąś konstrukcję? Znaczy to, że jeśli wykonamy taką konstrukcję na algebrach należących do tej klasy, to otrzymamy algebrę, która znów do tej klasy należy. W szczególności, klasa $\mathscr{K}$ jest zamknięta ze względu na obrazy homomorficzne, jeśli każdy obraz homomorficzny algebry należącej do $\mathscr{K}$ też należy do klasy $\mathscr{K}$; podobnie $\mathscr{K}$ jest zamknięta ze względu na podalgebry, jeśli każda podalgebra algebry należącej do $\mathscr{K}$ też do $\mathscr{K}$ należy. Wreszcie $\mathscr{K}$ jest zamknięta ze względu na produkty, jeśli produkt dowolnej rodziny algebr z klasy $\mathscr{K}$ jest znów algebrą należącą do $\mathscr{K}$.

Twierdzenie Birkhoffa mówi więc, że przy pomocy równości można opisać te i tylko te własności algebr, które są zachowywane przy przekształceniach homomorficznych, które przechodzą z algebry na podalgebrę i które nie giną przy budowaniu produktu z algebr mających te własności.

Nie będziemy przytaczać pełnego dowodu tego twierdzenia. Zauważmy tylko, że jeśli pewne równości są prawdziwe w algebrze A, to tym bardziej jest tak w każdej podalgebrze algebry A, bo elementy tej podalgebry są przecież także elementami algebry A. Równości te są także prawdziwe w każdym obrazie homomorficznym A, gdyż homomorfizm przyporządkowuje oczywiście — tak jak każda funkcja — równym elementom elementy równe, zachowując przy tym operacje. Wreszcie, jeśli pewne równości są prawdziwe na każdej ,,współrzędnej" produktu algebr, to prawdziwe są także i w całym produkcie, w którym operacje, jak pamiętamy, określone są właśnie ,,na współrzędnych". Widać z powyższego rozumowania, że prawdziwość równości zachowywana jest przez podalgebry, obrazy homomorficzne i produkty, a zatem każda klasa definiowalna równościowo jest zamknięta ze względu na te konstrukcje.

Dowód tego, że każda klasa $\mathscr{K}$ zamknięta ze względu na obrazy homomorficzne, podalgebry i produkty jest definiowalna równościowo nie da się tu przedstawić w sposób podobnie prosty. Polega on na pokazaniu, że każda algebra, w której prawdziwe są wszystkie równości prawdziwe jednocześnie we wszystkich algebrach z klasy $\mathscr{K}$ — jest obrazem homomorficznym pewnej algebry z $\mathscr{K}$ (do jej zbudowania używa się produktów i podalgebr), a więc z założenia do klasy $\mathscr{K}$ też należy. Wynika stąd, że zbiór tych równości charakteryzuje jednoznacznie klasę $\mathscr{K}$.

Klasą definiowalną równościowo jest np. klasa wszystkich algebr Boole'a i klasa wszystkich grup, rozumianych jako algebry z operacją składania, odwracania i z wyróżnionym elementem neutralnym względem składania (jednością). Nie można natomiast scharakteryzować przy pomocy równości klasy wszystkich algebr $(A, f_1, \ldots, f_n)$ takich, że zbiór A ma dokładnie 2 elementy: produkt dwóch takich algebr będzie miał przecież 4 elementy, nie będzie więc do tej klasy należał. Nie jest to dziwne — własność bycia zbiorem 2-elementowym nie jest własnością algebraiczną, nie dotyczy działania operacji na zbiorze, podczas gdy równość mówi zawsze o tym, że stosując w różny sposób operacje algebry otrzymujemy zawsze to samo. Nie jest też definiowalna równościowo klasa wszystkich grup, rozumianych jako algebry z jedną tylko operacją składania. Dla przykładu, grupą jest algebra $(R^+, \cdot)$, bo istnieje w niej element neutralny względem mnożenia (jest nim liczba 1), a dla każdego elementu istnieje element odwrotny $\left(\text{dla liczby } r \in R \text{ jest nim liczba } \frac{1}{r}\right)$. Algebra ta ma jednak podalgebrę — a mianowicie zbiór liczb całkowitych dodatnich — która grupą nie jest.

Klasy definiowalne równościowo mają wiele interesujących własności. Jeśli $\mathscr{K}$ jest taką klasą, to istnieje zbiór równości E taki, że algebra A należy do klasy $\mathscr{K}$ wtedy i tylko wtedy, gdy spełnia wszystkie równości ze zbioru E. Mogą być jednak spełnione w algebrze A należącej do $\mathscr{K}$ także równości spoza zbioru E, których nie spełniają inne algebry z $\mathscr{K}$. Dla przykładu, w klasie wszystkich grup istnieje grupa jednoelementowa, spełniająca równość $x = y$ (równość ta mówi, że wszystkie elementy tej grupy są równe, a więc ma ona tylko jeden element); oczywiście żadna grupa, która ma więcej elementów, tej równości nie spełnia. Zawsze jednak istnieją w klasie $\mathscr{K}$ definiowalnej równościowo algebry (i to dowolnie duże), które nie spełniają żadnej innej równości poza równościami ze zbioru E — są to tzw. algebry wolne. Równości spełnione w tych algebrach są więc równościami spełnionymi w całej klasie $\mathscr{K}$. Co więcej, z jednej tylko takiej algebry wolnej można otrzymać wszystkie pozostałe algebry z klasy $\mathscr{K}$ stosując przekształcenia homomorficzne, wybierając podalgebry albo budując produkty.

Okazuje się, że język używający jednego tylko ,,czasownika", jednej relacji — równości — jest dostatecznie bogaty, by mówić o wielu własnościach algebr, by opisać wiele znanych klas algebr, a także — by sprawić wiele satysfakcji matematykom.

Rozwiązanie zadania F 115. Różnica potencjałów między punktami A i B przed zwarciem jest równa

$$U_{AB} = \frac{\mathscr{E}_1 r_2 - \mathscr{E}_2 r_1}{r_1 + r_2}.$$

Przez bezoporowy przewodnik zwierający popłynie więc prąd, narastający aż do wartości (różnica prądów zwarcia źródeł):

$$I_{AB} = \frac{\mathscr{E}_1}{r_1} - \frac{\mathscr{E}_2}{r_2}.$$

Różnica potencjałów U_{AB} zmaleje wtedy do zera, jednak prąd będzie nadal płynął dzięki bezoporowości przewodnika.
Prąd między punktami A i B nie popłynie (niezależnie od oporności przewodnika zwierającego), jeśli przed zwarciem $U_{AB} = 0$, tzn. $\mathscr{E}_1 r_2 = \mathscr{E}_2 r_1$.