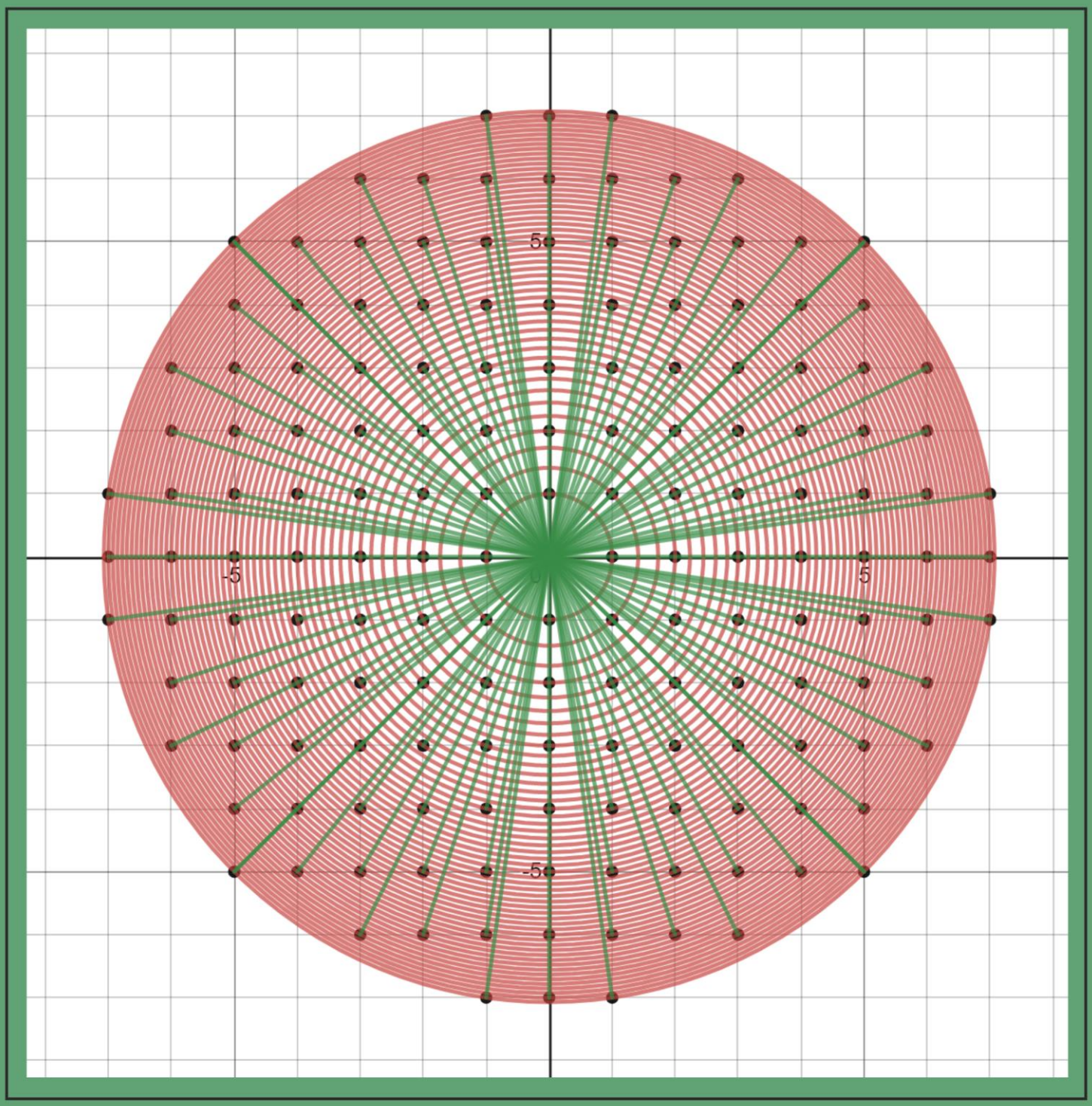
Kacper Błachut klasa I a
I Liceum Ogólnokształcące im. Seweryna Goszczńskiego
Plac Krasińskiego 1 34-400 Nowy Targ

Opiekun: mgr Joanna Trybuła

LICZBY PIERWSZE



Rok 2021

Dane kontaktowe

Szkoła:
I Liceum Ogólnokształcące w Nowym Targu, im. Seweryna Goszczyńskiego
Plac Krasińskiego 1 34-400 Nowy Targ
tel/fax: 18 266 29 55, 18 266 38 58
lo@goszczynski.nowytarg.pl

Opiekun:
mgr Joanna Trybuła
tel. 696 678 958 email: trybula.joanna@gmail.com

Autor:
Kacper Błachut tel. 601 22 32 31 email: kacper.blachut@gmail.com

Serdeczne podziękowania dla Pani ***mgr Joanny Trybuły***
za pomoc przy redagowaniu niniejszej pracy
oraz bezcenne uwagi przy korekcie

Zespół Szkół Ogólnokształcących Nr 1
w Nowym Targu
34-400 Nowy Targ, Pl.Krasińskiego 1
tel./fax (18) 2662955, 2663858
NIP 7351029206 Regon 000702848

Nowy Targ, 27.01.2021

INFORMACJA NAUCZYCIELA

Kacper jest uczniem klasy pierwszej o poszerzonym programie nauczania matematyki, fizyki i języka angielskiego. W Małopolskim Konkursie Prac Matematycznych startuje po raz piąty. W roku szkolnym 2016/2017 zajął II miejsce pisząc pracę nt Nieskończoności, w latach 2017/2018, 2018/2019 i 2019/2020 zajął I miejsca przedstawiając prace odpowiednio nt Matematyki Kostki Rubika, Funkcji i Matematyki Żonglowania.

Kacpra poznałam we wrześniu 2020 roku, gdy rozpoczął naukę w naszym liceum. Szybko dał się poznać jako osoba utalentowana matematycznie, pracowita i zdyscyplinowana. Przedstawiane przez Niego rozwiązania zadań cechuje wysoki stopień zrozumienia przerabianego materiału, precyzja, duża kultura matematyczna i przejrzystość zapisów.

Uczeń preferuje prace samodzielną, ale potrafi również zgodnie pracować w grupie, dzieląc się swoją wiedzą i doświadczeniem z innymi uczniami.

Kacper nieustannie poszerza swoją wiedzę i umiejętności matematyczne, rozwija swoje zainteresowania, czego dowodem jest niniejsza praca, napisana przez Niego samodzielnie. Treści w niej zawarte w sposób znaczny wykraczają poza podstawę programową nauczania matematyki w szkole ponadpodstawowej. Muszę przyznać, że przeczytałam ją z dużym zainteresowaniem, a pewne zagadnienia w niej zawarte okazały się być dla mnie nowością. Dlatego też praca z Kacprem daje dużo satysfakcji.

Joanna Trybuła

Joanna Trybuła
telefon: 696678958
e-mail: trybula.joanna@gmail.com

1. Wstęp

Już piąty raz przystępuję do pisania pracy na Małopolski Konkurs Prac Matematycznych organizowany przez Krakowskie Młodzieżowe Towarzystwo Przyjaciół Nauk i Sztuk im. H. Jordana. Poprzednie prace to: Nieskończoność, Matematyka kostki Rubika, Funkcje, Matematyka żonglowania. Zawsze starałem się znajdować takie tematy, które interesowały mnie ściśle z matematycznego punktu widzenia (Nieskończoność, Funkcje), czy znajdowałem matematykę, tam gdzie wydawałoby się jej nie ma (Matematyka kostki Rubika, Matematyka żonglowania).

Jak zwykle pierwszy krok był najtrudniejszy. Jaki temat wybrać? Czy wreszcie napiszę o ciągach (które fascynują mnie od czasu napisania pracy o nieskończoności)? A może wybrać całki (które z kolei byłyby kontynuacją pracy o funkcjach i pochodnych). A może dokończyć zeszłoroczną pracę o żonglowaniu (moje obliczenia utknęły tam w pewnym punkcie). A może zupełnie coś innego? Byłem w kropce.

Wtedy na jednym z obserwowanych przeze mnie kanałów matematycznych na YouTube (*3Blue1Brown*) znalazłem film o Riemannie. Zafascynowały mnie różne właściwości liczb pierwszych omówione w tym materiale. Oczywiście wiedziałem co to są liczby pierwsze. Znałem podstawowe ich właściwości. Wiedziałem, że są wykorzystywane do szyfrowania danych. Ale właściwie moja wiedza na ten temat była dość ograniczona. Im bardziej zagłębiałem się w temacie *prime numbers* tym więcej miałem pytań. Im więcej miałem pytań, tym mniej miałem wątpliwości, co będzie tematem mojej pracy. I tak oto powstało kolejne moje dzieło na temat matematyki.

W trakcie zbierania informacji na temat liczb pierwszych natknąłem się na kilka postaci matematyków, którzy właściwie całe swoje naukowe życie poświęcili walce z liczbami pierwszymi (a dokładniej z Hipotezą Riemanna) i … polegli. Przytoczę tu nazwisko **Louisa de Brangesa de Bourcii**, który już jako student w 1953r. obrał sobie za cel udowodnienie Hipotezy Riemanna. W 2004r. ogłosił (a więc po ponad półwieczu), że udowodnił HR. Niestety jego dowód okazał się błędny. Próbował dalej. Co jakiś czas na swojej stronie ogłaszał nowe dowody. Wszystkie one zawierały luki. W tej chwili już coraz mniej matematyków interesuje się tym co robi de Brangesa, który ostatnią wersję dowodu ogłosił bodajże w kwietniu 2017r. Inny wielki matematyk sir Michael Atiyah (laureat nagród Fields'a i Abela) w 2018r. stwierdził, że rozwiązał problem HR. Niestety on również się mylił. Przypomniała mi się wtedy moja praca sprzed czterech lat *Nieskończoność,* gdzie opisywałem jak wielu słynnych matematyków, którzy zajmowali się problemem nieskończoności, także poległo (w tym Hilbert). Niektórzy wręcz musieli się leczyć w szpitalach psychiatrycznych.

Liczby pierwsze mają zastosowanie w wielu dziedzinach naszego życia, a to poprzez wykorzystanie ich właściwości w szyfrowaniu. A w dzisiejszym zinformatyzowanym świecie szyfrujemy właściwie wszystko. Kodowane są płatne kanały w TV. Większość stron w Internecie przeglądamy już w protokole https (czyli szyfrowane). Gdy logujemy się do naszych

ulubionych mediów społecznościowych także wykorzystujemy kodowanie. Szyfrowane też są nasze połączenia z bankiem. Można tak wyliczać w nieskończoność.

W 1900r. wybitny matematyk David Hilbert przedstawił 23 ważne zagadnienia matematyczne z przełomu XIX i XX wieku. Sto lat później w 2000r Instytut Claya ogłosił zestaw siedmiu tak zwanych Problemów Milenijnych. Na obu listach, jako jedyna, znalazła się Hipoteza Riemanna – związana ściśle z liczbami pierwszymi. Już to pokazuje, jak ważna jest to sprawa. Za rozwiązanie któregoś z 7 Problemów Milenijnych Instytut Claya ufundował nagrodę w wysokości miliona dolarów.

Wiele zagadnień związanych z liczbami pierwszymi jest bardzo skomplikowana i trudna do zrozumienia. Nie chodzi tu wcale tylko o dowody twierdzeń, ale nawet o założenia i tezy. Dlatego w niniejszej pracy przedstawiam dowody tylko tych twierdzeń, które zrozumiałem (lub wydaje mi się, że rozumiem). Część twierdzeń tylko opisuję, bez formułowania dokładnych założeń i tez, gdyż są bardzo skomplikowane. Podaję tylko równania funkcji, nie wdając się w ich analizę. Tak na przykład jest z Hipotezą Riemanna, gdzie rozszerzenie analityczne dziedziny funkcji dzeta-Riemanna na liczby zespolone, które (na razie) wykracza poza moje rozumienie. Czasami wydaje mi się, że je rozumiem, a innym razem że nie.

W **Rozdziale 2** opisuję używane w dalszej części pojęcia matematyczne. Praca ta przeznaczona jest nie tylko dla entuzjastów matematyki, ale także – a może przede wszystkim - dla moich rówieśników, którzy mogą nie znać wszystkich symboli tu użytych.

Rozdział 3 to krótka historia liczb pierwszych. Od starożytności poprzez wieki średnie do dzisiaj. W sposób zwięzły pokazuję rozwój tej dziedziny matematyki.

W **Rozdziele 4** przedstawiam podstawowe definicje i twierdzenia dotyczące liczb pierwszych. Zawarte są w nim także dowody opisywanych twierdzeń.

Rozdział 5 poświęcony jest różnym rodzajom liczb pierwszych. W trakcie przygotowywania się do pisania niniejszej pracy, znalazłem mnóstwo bardzo różnych odmian liczb pierwszych. Tutaj opisuję tylko te, które – z niewiadomych dla mnie przyczyn – przykuły moją uwagę.

Własności liczb pierwszych opisuje wiele bardzo ciekawych twierdzeń. W **rozdziale 6** opisuję jedno z najciekawszych z nich: Twierdzenie Fermata o dwóch kwadratach. Mimo, że dowód tego twierdzenia jest bardzo skomplikowany, przytaczam go tutaj. Zrozumienie toku rozumowania przebiegu tego dowodu zajęło mi naprawdę dużo czasu i mam nadzieję, że moja interpretacja tego dowodu jest prostsza niż formułują ją zawodowi matematycy.

Rozdział 7 poświęciłem rozważaniom jak za pomocą liczb pierwszych można obliczyć liczbę π. W literaturze można znaleźć wiele różnych sposobów, za pomocą których można obliczyć liczbę π nie stosując jej zależności od koła. Spośród wielu z nich można także znaleźć takie, które wykorzystują liczby pierwsze.

Jeden z najgenialniejszych matematyków wszechczasów – Gauss – będąc mniej więcej w moim wieku zaczął zajmować się liczbami pierwszymi. **Rozdział 8** poświęcony jest właśnie jego wczesnymi zmaganiami z gęstością liczb pierwszych.

Rozdział 9 opisuje wkład polskiego matematyka Stanisław Ulama w badanie liczb pierwszych. Nie można zajmować się tym tematem i nie opisać tzw. Spirali Ulama.

W **rozdziale 10** opisuję wzór pentagonalny Eulera oraz jego wykorzystanie do znajdywania liczb pierwszych.

Jednym z najważniejszych w tej chwili zagadnień matematycznych współczesnego świata jest wspomniana Hipoteza Riemanna, ściśle związana z liczbami pierwszymi. **Rozdział 10** opisuje właśnie to zagadnienie.

Rozdział 11 to opis praktycznego zastosowania liczb pierwszych. Poświęcony jest kryptografii, na której opiera się właściwie cały zinformatyzowany świat.

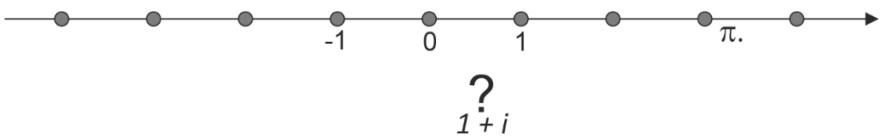
2. Używane pojęcia

Logarytm naturalny (ln) - logarytm o podstawie e, gdzie e = $\lim_{n\to\infty}\left(1+\frac{1}{n}\right)^n$

Liczby urojone – liczby postaci $a \cdot i$ gdzie $a \in \mathbb{R}$, a $i = \sqrt{-1}$

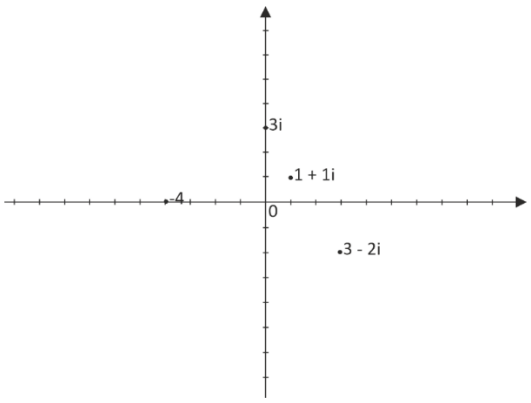
Liczby zespolone - liczby postaci $a + b \cdot i$ gdzie $a, b \in \mathbb{R}$, a $i = \sqrt{-1}$ (tzw. jednostka urojona).

Liczbę a nazywamy częścią rzeczywistą, a liczbę b częścią urojoną liczby zespolonej. Wprowadzając liczby zespolone matematycy *rozszerzyli* znany do tej pory największy zbiór liczb jakim są liczby rzeczywiste. O ile liczby rzeczywiste możemy przestawiać w sposób graficzny na osi liczbowej, o tyle liczb zespolonych tak przedstawić nie możemy. Bo gdzie można (na osi liczbowej) wstawić np. $1 + i$?



Rysunek 1 Gdzie na osi liczbowej przedstawić symbol i?

Dlatego wykorzystując układ kartezjański przyjęto, aby liczby zespolone przedstawiać na płaszczyźnie, gdzie na osi X zaznaczamy część rzeczywistą, a na osi Y część urojoną liczby zespolonej.



Rysunek 2 Przedstawienie liczb zespolonych na układzie współrzędnych

Liczby zespolone podobnie jak liczby rzeczywiste można dodawać, odejmować, mnożyć i dzielić (przy założeniu, że dzielnik jest różny od zera). Dodawanie i mnożenie jest banalne. Dodajemy (odejmujemy) osobno części rzeczywiste, osobno części urojone. Na przykład: $(1 + i) + (2 + 3i) = 3 + 4i, (3 + 2i) - (2 - i) = 1 + i$. Mnożenie liczb zespolonych przebiega podobnie jak mnożenie wyrażeń algebraicznych z niewiadomą (np. x), wykorzystujemy przy tym fakt, że $i^2 = -1$. Na przykład: $(5 + 2i) \cdot (3 - 7i) = 5 \cdot 3 - 5 \cdot 7i + 2i \cdot 3 - 2i \cdot 7i = 15 - 35i + 6i - 14i^2 = 15 - 29i + 14 = 29 - 29i$. Dzielenie liczby zespolonej jest trochę bardziej skomplikowane – wymaga użycia sprzężonej liczby zespolonej (tzn. takiej gdzie część urojona ma przeciwny znak): $\frac{1+8i}{2+3i} = \frac{1+8i}{2+3i} \cdot \frac{2-3i}{2-3i} = \frac{(1+8i)\cdot(2-3i)}{2^2-(3i)^2} = \frac{2-3i+16i-24i^2}{4-9i^2} = \frac{2+13i+24}{4+9} = \frac{26+13i}{13} = 2 + i$

W kolejnych krokach korzystałem z faktu, że liczba pomnożona przez 1 (w tym przypadku w postaci ułamka $\frac{2-3i}{2-3i}$) pozostaje tą samą liczbą, wykorzystywałem wzór skróconego mnożenia $(a+b)\cdot(a-b) = a^2 - b^2$ oraz fakt, że $i^2 = -1$.

Prześledźmy jakie dzielniki może mieć każda liczba zespolona $a + b\cdot i$ (gdzie $a, b \in \mathbb{R}$)? Oczywiście podobnie jak inne liczby: 1 oraz -1. Czy jeszcze jakieś inne? Zobaczmy co z liczbą *i*?

$$\frac{a+b\cdot i}{i} = \frac{a+b\cdot i}{i} \cdot \frac{-i}{-i} = \frac{-a\cdot i - b\cdot i^2}{-i^2} = \frac{b-a\cdot i}{1} = b - a\cdot i$$

Czyli i jest dzielnikiem dowolnej liczby zespolonej. Podobnie jest z liczbą $-i$:

$$\frac{a+b\cdot i}{-i} = \frac{a+b\cdot i}{-i} \cdot \frac{i}{i} = \frac{a\cdot i + b\cdot i^2}{-i^2} = \frac{-b+a\cdot i}{1} = -b + a\cdot i$$

Całka $\int f(x)\,dx$

W uproszczeniu całkowanie jest działaniem odwrotnym do różniczkowania. Żeby sprawnie liczyć całki, należy wcześniej dobrze opanować liczenie pochodnych. Całkę oznaczamy symbolem: ∫ Symbol ten pochodzi od łacińskiego słowa Summa (suma). Poniżej przytoczę definicje i twierdzenie za [9]

Definicja 2.1

Funkcję F nazywamy **funkcją pierwotną** funkcji f w przedziale I wtedy i tylko wtedy, gdy

$$\bigwedge_{x \in I} F'(x) = f(x)$$

Twierdzenie 2.1

Jeżeli funkcja F jest funkcją pierwotną f w przedziale I, to dowolna funkcja pierwotna ϕ funkcji f w tym przedziale ma postać $\phi(x) = F(x) + C$, gdzie $C \in \mathbb{R}$

Definicja 2.2

Całką nieoznaczoną funkcji f w pewnym przedziale I nazywamy zbiór jej wszystkich funkcji pierwotnych w tym przedziale. Całkę nieoznaczoną będziemy oznaczać symbolem $\int f(x)\,dx$ gdzie: f to funkcja podcałkowa, x to zmienna całkowania, a $f(x)dx$ to wyrażenie podcałkowe

Faktoryzacja

To inaczej rozkład liczby naturalnej (całkowitej) na czynniki pierwsze (liczby pierwsze). Faktoryzacja liczby naturalnej n polega na znalezieniu liczb pierwszych $m_1, m_2, \ldots m_k$, że $n = m_1 \cdot m_2 \cdot \ldots \cdot m_k$. Domyślnie chodzi nam o nietrywialny rozkład, tzn. $\forall\, i = 1..k$: $m_i \neq 1$ i $m_i \neq n$.

O ile mnożenie liczb całkowitych jest bardzo prostą czynnością, to nie są znane żadne szybkie metody faktoryzacji. Szybkie to znaczy działające w tzw. czasie wielomianowym względem ilości cyfr rozkładanej liczby. Wszystkie znane algorytmy działają w tzw. czasie wykładniczym względem liczby cyfr rozkładanej liczby.

Zaokrąglenia dolne i górne do liczb całkowitych

$\lceil x \rceil$ zaokrąglenie górne - jest to najmniejsza liczba całkowita większa lub równa x

Np. $\lceil 2{,}1 \rceil = 3$, $\lceil 2{,}9 \rceil = 3$

$\lfloor x \rfloor$ zaokrąglenie dolne - jest to największa liczba całkowita mniejsza lub równa x

Np. $\lfloor 2{,}1 \rfloor = 2$,$\lfloor 2{,}9 \rfloor = 2$

Arytmetyka modulo (modularna, zegarowa)

Jeżeli dzielimy dwie liczby całkowite otrzymujemy następujące równanie:

$$n = m \cdot Q + R$$

gdzie n to dzielna, m – dzielnik, Q – iloraz i R reszta z dzielenia.

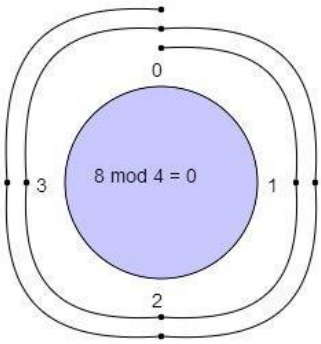
Niekiedy interesuje nas tylko i wyłącznie reszta z dzielenia. Resztę taką oznaczamy operatorem $modulo\ (mod)$.

$$n\ mod\ m = R$$

Na przykład $23\ mod\ 5 = 3$ bo $23 = 4 \cdot 5 + 3$

Zauważmy, że jeżeli zwiększamy n dzieląc przez m to wynik modulo zwiększa się nam od 0 do $n - 1$ i dalej cyklicznie. Na przykład: $0\ mod\ 4 = 0, 1\ mod\ 4 = 1, 2\ mod\ 4 = 2, 3\ mod\ 4 = 3, 4\ mod\ 4 = 0, 5\ mod\ 4 = 1, 6\ mod\ 4 = 2$ itd.

Dla takiego działania możemy skonstruować zegar, aby na przykład wyliczyć $8\ mod\ 4$:



Rysunek 3 Graficzna interpretacja arytmetyki modularnej

Dla dzielnika równego 4 budujemy tarczę zegara z liczbami 0, 1, 2, 3. Startujemy od 0 i idziemy w prawo 8 kroków dookoła. Otrzymujemy liczby: 1, 2, 3, 0, 1, 2, 3, 0. Skończyliśmy na 0, a więc $8\ mod\ 4$=0.

Dla działania mod spełnione są poniższe równania:

dodawanie $(a + b)\ mod\ m = (a\ mod\ m + b\ mod\ m) mod\ m$

odejmowanie $(a - b)\ mod\ m = (a\ mod\ m - b\ mod\ m) mod\ m$

mnożenie $(a \cdot b)\ mod\ m = (a\ mod\ m \cdot b\ mod\ m)\ mod\ m$

potęgowanie $a^b\ mod\ m = ((a\ mod\ m)^b)\ mod\ m$

Tabela ASCII

Podstawowa zawiera kody, które przypisują poszczególnym liczbom z zakresu od 0 do 255 litery alfabetu angielskiego (duże i małe), cyfry, znaki interpunkcyjne, symbole specjalne, symbole matematyczne, proste znaki graficzne.

000d	00h		(nul)	016d	10h	▶	(dle)	032d	20h	sp	048d	30h	0	064d	40h	@	080d	50h	P	096d	60h	`	112d	70h	p
001d	01h	☺	(soh)	017d	11h	◀	(dc1)	033d	21h	!	049d	31h	1	065d	41h	A	081d	51h	Q	097d	61h	a	113d	71h	q
002d	02h	●	(stx)	018d	12h	↕	(dc2)	034d	22h	"	050d	32h	2	066d	42h	B	082d	52h	R	098d	62h	b	114d	72h	r
003d	03h	♥	(etx)	019d	13h	‼	(dc3)	035d	23h	#	051d	33h	3	067d	43h	C	083d	53h	S	099d	63h	c	115d	73h	s
004d	04h	♦	(eot)	020d	14h	¶	(dc4)	036d	24h	$	052d	34h	4	068d	44h	D	084d	54h	T	100d	64h	d	116d	74h	t
005d	05h	♣	(enq)	021d	15h	§	(nak)	037d	25h	%	053d	35h	5	069d	45h	E	085d	55h	U	101d	65h	e	117d	75h	u
006d	06h	♠	(ack)	022d	16h	▬	(syn)	038d	26h	&	054d	36h	6	070d	46h	F	086d	56h	V	102d	66h	f	118d	76h	v
007d	07h	•	(bel)	023d	17h	↨	(etb)	039d	27h	'	055d	37h	7	071d	47h	G	087d	57h	W	103d	67h	g	119d	77h	w
008d	08h	◘	(bs)	024d	18h	↑	(can)	040d	28h	(	056d	38h	8	072d	48h	H	088d	58h	X	104d	68h	h	120d	78h	x
009d	09h		(tab)	025d	19h	↓	(em)	041d	29h	)	057d	39h	9	073d	49h	I	089d	59h	Y	105d	69h	i	121d	79h	y
010d	0Ah		(lf)	026d	1Ah		(eof)	042d	2Ah	*	058d	3Ah	:	074d	4Ah	J	090d	5Ah	Z	106d	6Ah	j	122d	7Ah	z
011d	0Bh	♂	(vt)	027d	1Bh	←	(esc)	043d	2Bh	+	059d	3Bh	;	075d	4Bh	K	091d	5Bh	[	107d	6Bh	k	123d	7Bh	{
012d	0Ch	♀	(np)	028d	1Ch	∟	(fs)	044d	2Ch	,	060d	3Ch	<	076d	4Ch	L	092d	5Ch	\	108d	6Ch	l	124d	7Ch	|
013d	0Dh		(cr)	029d	1Dh	↔	(gs)	045d	2Dh	-	061d	3Dh	=	077d	4Dh	M	093d	5Dh	]	109d	6Dh	m	125d	7Dh	}
014d	0Eh	♫	(so)	030d	1Eh	▲	(rs)	046d	2Eh	.	062d	3Eh	>	078d	4Eh	N	094d	5Eh	^	110d	6Eh	n	126d	7Eh	~
015d	0Fh	☼	(si)	031d	1Fh	▼	(us)	047d	2Fh	/	063d	3Fh	?	079d	4Fh	O	095d	5Fh	_	111d	6Fh	o	127d	7Fh	⌂

Rysunek 4 Tablica ASCII

Dowód nie wprost

to forma dowodu (matematycznego) opartego na równowartości $(p \Rightarrow q) \Leftrightarrow (\neg q \Rightarrow \neg p)$, tzn. jeśli zaprzeczając tezie dojdziemy do sprzeczności z założeniem to implikacja wyjściowa jest prawdziwa.

Przykład

Twierdzenie: Jeżeli $x = 3$, to $x^2 = 9$.

Hipoteza: $x^2 \neq 9$

Dowód: $x^2 \neq 9 \Leftrightarrow x \neq 3 \wedge x \neq -3$, co jest sprzeczne z założeniem

Tak więc przypuszczenie o fałszywości tezy doprowadziło nas do sprzeczności z założeniem, co na mocy przytoczonej równoważności oznacza, że wyjściowa implikacja jest prawdziwa.

Sumy i iloczyny

$\sum_{i=1}^{k} a_i = a_1 + a_2 + \cdots + a_k$ – suma k-elementów

$\prod_{i=1}^{k} a_i = a_1 \cdot a_2 \ldots a_k$ iloczyn k – elementów

Symbol „#”

Symbol „#” w literaturze angielskiej nazywamy *primorial* i jest to prosta analogia do nazwy *factorial*, która oznacza silnię. *Primorial* to silnia, w której mnożymy kolejne liczby pierwsze (w przeciwieństwie do silni, gdzie mnożymy kolejne liczby naturalne)

Przykład : $20\# = 2 \cdot 3 \cdot 5 \cdot 7 \cdot 11 \cdot 13 \cdot 17 \cdot 19 = 9699690$

3. Trochę historii

Historia liczb naturalnych sięga tak daleko wstecz, że właściwie nie jesteśmy jednoznacznie stwierdzić, kiedy pojawiły się po raz pierwszy (ani gdzie). Były z nami od zawsze. Ludzie pierwotni znali takie pojęcia jak *jeden*, *dwa*, *wiele* (a także *nic*). Liczyli w ten sposób na przykład liczbę upolowanych zwierząt. Rozwój człowieka pociągnął za sobą wzrost umiejętności liczenia. Niejednokrotnie zdobycze trzeba było podzielić na całe plemię lub jego część. Istotnym problemem stało się jak uczciwie podzielić trofea (np. 6 saren, 13 zajęcy). Niekiedy trofea można było podzielić na równe części tylko w taki sposób, że każda część będzie zawierała jeden element (np. $7 = 1 + 1 + 1 + 1 + 1 + 1 + 1$). Tak po raz pierwszy ludzkość spotkała się z liczbami pierwszymi. Historycy wskazują, że niewielkie liczby pierwsze znane były ludziom zamieszkującym tereny obecnego Konga w Afryce już ok. 20 tysięcy lat temu.

Liczby pierwsze wzbudziły zainteresowanie Euklidesa w IV w p.n.e. . Poświęcił im księgę VII w słynnych Elementach. W księdze tej można znaleźć wiele interesujących twierdzeń (i ich dowodów) na temat liczb pierwszych – między innymi Twierdzenie o nieskończoności zbioru liczb pierwszych. Później niejaki Eratostenes z Cyreny – grecki matematyk żyjący w latach 276-194r. p.n.e. opracował algorytm (zwany dzisiaj sitem Eratostenesa) wykorzystywany do „przeczesywania" zbioru liczb naturalnych w poszukiwaniu liczb pierwszych.

Przez wieki badań nad liczbami pierwszymi próbowano znaleźć matematyczne wzory, które wyznaczałyby tylko liczby pierwsze.

W XVII wieku Fermat sformułował hipotezę (zwaną Małym Twierdzeniem Fermata). Matematyk ów twierdził, że liczby postaci

$$f(n) = 2^{2^n} + 1 \text{ gdzie } n \in \mathbb{N}$$

są liczbami pierwszymi. Nie są to oczywiście kolejne (wszystkie) liczby pierwsze, ale Fermat przypuszczał, że wzór powyższy daje tylko liczby pierwsze. Rzeczywiście, jeżeli zbadamy je dla początkowych *n* mamy

$$f(1) = 5, f(2) = 17, f(3) = 257, f(4) = 65537$$

Wszystkie one są liczbami pierwszymi, ale już $f(5) = 4294967297 = 641 \cdot 6700417$. Czyli wzór już dla *n* = 5 okazuje się niepoprawny!.

Próbowano dalej. Matematycy zaproponowali poniższy wzór:

$$f(n) = n^2 - n + 41 \text{ gdzie } n \in \mathbb{N}$$

Rzeczywiście daje on dużo liczb pierwszych (np. dla *n* = 1…40), ale już z samej jego konstrukcji widać, że $f(41) = 41^2 - 41 + 41 = 41^2$, a to nie liczba pierwsza.

Inna próba stworzenia wzoru na liczby pierwsze to

$$f(n) = n^2 - 79n + 1601 \text{ gdzie } n \in \mathbb{N}$$

Daje on liczby pierwsze przy wartościach *n = 1 ..79,* przy *n = 80* zawodzi.

W XVI w Marin Mersenne zaproponował inną postać liczby pierwszej: $M_n = 2^n - 1$. Liczby takie nazywamy dzisiaj liczbami Mersenne'a. Niestety on również się mylił. Na przykład 11 z kolei liczba Mersenne'a równa się 2047 = 23 × 89.

To pokazuje, że w przypadku liczb pierwszych polegli nawet najwybitniejsi matematycy.

W rozwoju naszej wiedzy na temat liczb pierwszych nie brakuje także Polaków. W 1963r. słynny matematyk Stanisław Ulam wskazał graficzną metodę pokazywania pewnych niewyjaśnionych do dziś prawidłowości w rozmieszczeniu liczb pierwszych na osi liczb naturalnych, tzw. Spiralę Ulama.

Na przestrzeni dziejów powstało wiele twierdzeń dotyczących liczb pierwszych. Niektóre z nich wydają się oczywiste, inne są prawie jak *science fiction*. Wiele z nich udowodniono, niektóre zaś czekają na dowód do dzisiaj.

4. Zacznijmy od początku

Liczby naturalne stworzył dobry Bóg. Reszta jest dziełem człowieka jak powiedział Leopold Kronecker. Czy rzeczywiście? Według mnie Kronecker się mylił. Bóg stworzył liczby pierwsze, a liczby naturalne są efektem wtórnym!

Każdy, kto choć trochę interesuje się matematyką (a nawet ci, którzy od królowej nauk stronią) wiedzą, co to są liczby pierwsze. Liczby naturalne (czyli 1, 2, 3, 4, …) można podzielić na trzy podzbiory

- Liczbę 1
- Liczby pierwsze
- Liczby złożone

Liczby 1 (jeden) nie muszę tłumaczyć.

Liczby pierwsze to takie liczby, które dzielą się przez samą siebie i przez 1 np. 2, 3, 5, 7, 173, 5483.

Liczby złożone to liczby, które można rozłożyć *na iloczyn czynników pierwszych* np. $6 = 2 \cdot 3$, $24 = 2 \cdot 2 \cdot 3 \cdot 3$, $1386 = 2 \cdot 3 \cdot 3 \cdot 7 \cdot 11$

A dokładniej

Definicja 4.1 – liczby pierwsze i złożone

Liczba pierwsza – liczba naturalna większa od 1, która ma dokładnie dwa dzielniki naturalne: jedynkę i siebie samą. Zbiór wszystkich liczb pierwszych oznacza się symbolem $\mathbb{P}$.

Liczba złożona – to liczba naturalna większa od 1, która nie jest liczbą pierwszą, tzn. taka, która posiada więcej niż dwa dzielniki.

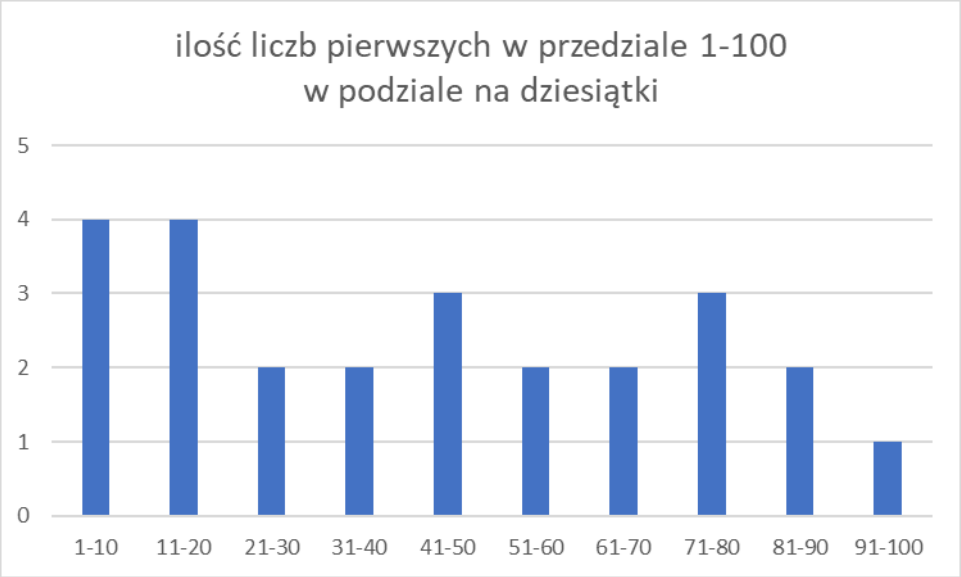
Liczby złożone można przedstawić jako iloczyn liczb pierwszych (o tym później). Można więc powiedzieć, że wszystkie liczby naturalne (poza 1) są albo liczbami pierwszymi, albo są iloczynem liczb pierwszych. Czyli liczby pierwsze są jakby elementem podstawowym, takim atomem wśród liczb (jeżeli przyjmiemy model Bohra). Dlatego myślę, że Kronecker się mylił. Bóg dał nam liczby pierwsze, a reszta to wymysł człowieka!

Co może być ciekawego w liczbach pierwszych? Nuda! Nic bardziej mylnego. Liczby pierwsze to jedna z najbardziej fascynujących dziedzin matematyki. I nie tylko ja tak twierdzę!

Przyjrzyjmy się na początek liczbom pierwszym w kilku pierwszych dziesiątkach liczb naturalnych, a właściwie ich liczebności:

dziesiątka	1-10	11-20	21-30	31-40	41-50	51-60	61-70	71-80	81-90	91-100
l. pierwsze	2 3 5 7	11 13 17 19	23 29	31 37	41 43 47	53 59	61 67	71 73 79	83 89	97
ilość lp	4	4	2	2	3	2	2	3	2	1

Rysunek 5 Tabela z ilością liczb pierwszych w kolejnych dziesiątkach liczb naturalnych



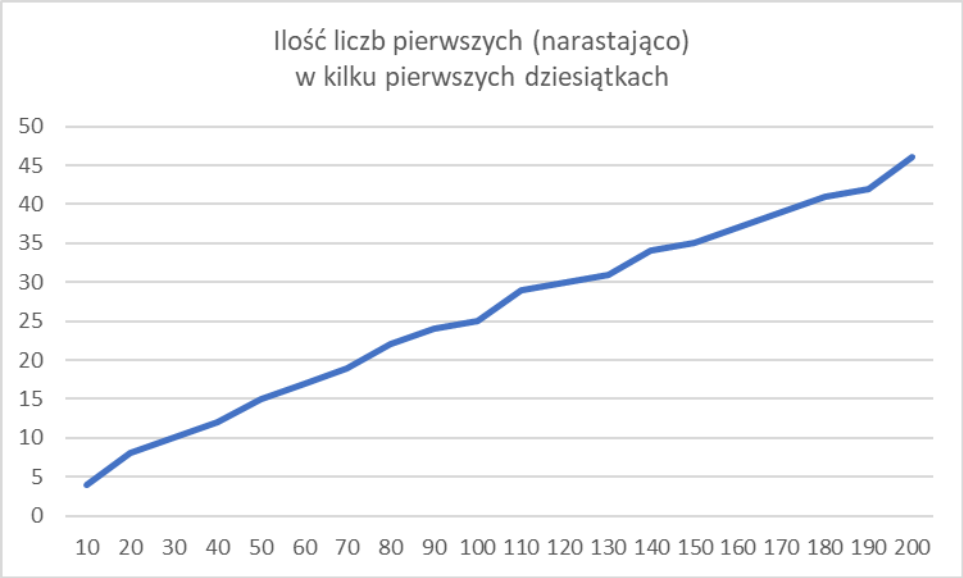
Rysunek 6 Ilość liczb pierwszych w przedziale 1-100 w podziale na dziesiątki

Nie widać tu żadnej regularności ilości liczb pierwszych w kolejnych dziesiątkach, żadnego schematu. A może za mały przedział? Spróbujmy rozszerzyć zakres do 1000, a przedziały do „setek”:

Rysunek 7 ilość liczb pierwszych w przedziale 1-1000 w podziale na setki

Dalej nie widać żadnej regularności! Moglibyśmy tak rozszerzać zakresy i nie znajdziemy żadnego schematu rozmieszczenia liczb pierwszych!

A może to złe podejście? Spróbujmy zobaczyć jak wzrasta ilość liczb pierwszych w pewnych zakresach.

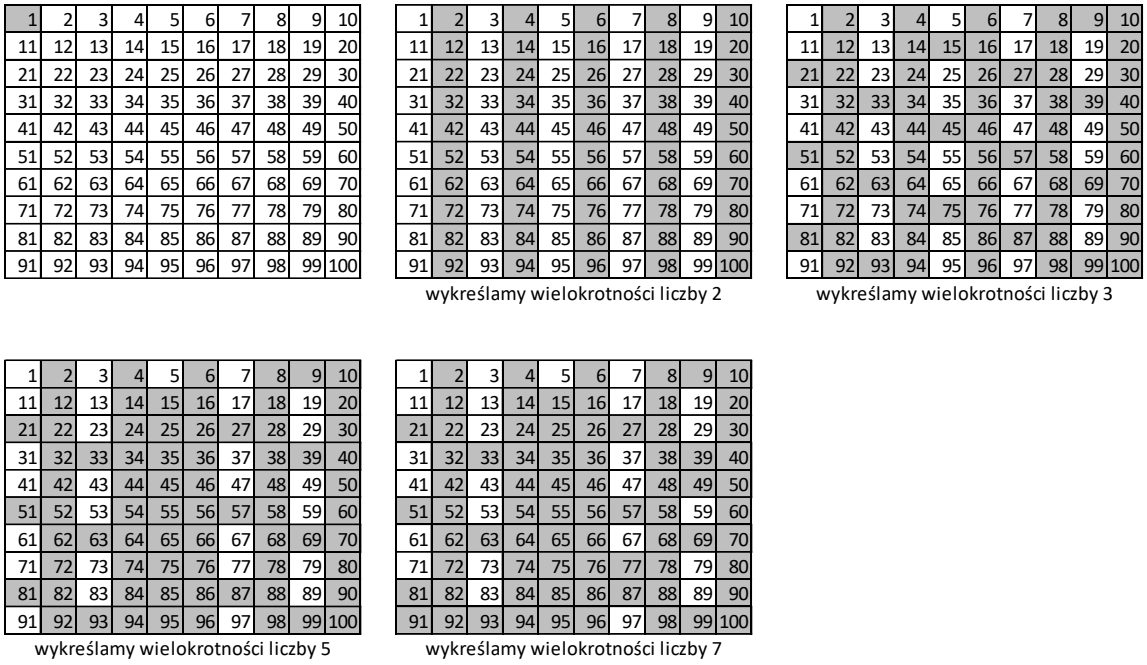


Rysunek 8 Ilość liczb pierwszych (narastająco) w kilku pierwszych dziesiątkach

Dalej nie widać żadnej regularności. Dlaczego? Po prostu nie ma żadnego schematu rozmieszczenia liczb pierwszych wśród liczb naturalnych (a przynajmniej nikt do tej pory nie znalazł takiego schematu). Dwie kolejne liczby pierwsze (nawet te naprawdę wielkie) mogą się różnić o dwa (np.$2996863034895 \cdot 2^{1290000} - 1$ oraz $2996863034895 \cdot 2^{1290000} + 1$), a z drugiej strony przerwa pomiędzy kolejnymi liczbami pierwszymi może być ogromna. Czy rzeczywiście? Posłużmy się pojęciem silni: $n! = 1 \cdot 2 \cdot 3 \cdot \ldots n$. Liczba $n!$ z definicji jest podzielna przez 2, 3 … n. Więc $n!$ na pewno nie jest liczbą pierwszą. Idźmy dalej: o liczbie $n! + 1$ nie jesteśmy w stanie powiedzieć, czy jest liczbą pierwszą, czy nie, ale już liczba $n!+2$ jest na pewno podzielna przez 2, bo $n!+2 = 1 \cdot 2 \cdot 3 \cdot \ldots n + 2 = 2 \cdot (1 \cdot 3 \cdot 4 \cdot \ldots n + 1)$, liczba $n!+3$ jest podzielna przez 3 i tak dalej, a więc znaleźliśmy n-1 kolejnych liczb, które nie są liczbami pierwszymi. A skoro n może być dowolnie duże, to „dziura” pomiędzy kolejnymi liczbami pierwszymi może być także dowolnie duża.

Co jeszcze możemy powiedzieć o liczbach pierwszych, tak *na pierwszy rzut oka*? Otóż jedyna parzysta liczba pierwsza to 2 (dwa), wszystkie pozostałe liczby parzyste dzielą się przez dwa, a więc z definicji nie są liczbami pierwszymi (mają co najmniej 3 dzielniki). Idąc dalej tym tokiem rozumowania: 3 jest liczbą pierwszą, ale żadna jej wielokrotność nie jest liczbą pierwszą; 5 jest liczbą pierwszą – żadna wielokrotność 5 nią nie jest … I tak dochodzimy do sita Eratostenesa. Jest to sposób znajdowania liczb pierwszych wymyślony przez greckiego matematyka i astronoma Eratostenesa w III w p.n.e. Polega on na tym, że wypisujemy wszystkie liczby naturalne z obranego zakresu (tutaj od 1 do 100), następnie w pierwszym kroku wykreślamy wszystkie liczby podzielne przez 2 (oprócz samej liczby 2). Dalej wykreślamy

wszystkie liczby podzielne przez 3 (oprócz samej trójki). Wielokrotności liczby 4 nie musimy wykreślać, gdyż zostały one skreślone przy okazji wielokrotności liczby 2. Następnie wykreślamy wielokrotności liczby 5. Liczbę sześć mamy już wykreśloną, więc przechodzimy do liczby 7 i wykreślamy jej wielokrotności. Uogólniając, w sicie Eratostenesa wykreślamy wszystkie wielokrotności liczb niewykreślonych do tej pory (licząc od najmniejszych). Czyli wykreślamy wielokrotności liczb pierwszych i na końcu pozostają nam w tabeli tylko liczby pierwsze.



Rysunek 9 Sito Eratostenesa

Od razu widać, że ta metoda jest dobra, ale tylko dla małych zakresów liczb pierwszych. Znajdowanie wielkich liczb pierwszych za pomocą sita Eratostenesa jest dość problematyczne. Czy możemy sobie uprościć „przechodzenie" przez sito? Czy musimy liczyć, do końca, czy też możemy przerwać w pewnym momencie? Zauważmy, że przy liczbach z zakresu od 1 do 100 – po dojściu do liczby 7 mieliśmy już wykreślone wszystkie liczby złożone. Otóż zauważmy na początek pewną zależność:

Twierdzenie 4.1 Każda liczba złożona dzieli się przez liczbę pierwszą mniejszą od siebie

$$\boldsymbol{n = a \cdot p}$$

gdzie n jest liczbą złożoną,
p jest liczbą pierwszą oraz $p < n$,
a jest liczbą naturalną

Dowód

Ponieważ n jest liczbą złożoną, więc na pewno można ją przedstawić jako iloczyn dwóch liczb naturalnych (nie koniecznie pierwszych) $n = a_1 \cdot a_2$ takich, że $a_1 < n$, $a_2 < n$. Jeżeli a_1 lub a_2 jest liczbą pierwszą, to mamy sytuację z twierdzenia 4.1. Jeżeli a_1 i a_2 są złożone to np. a_1 możemy przedstawić jako iloczyn liczb naturalnych mniejszych od a_1. Wtedy $a_1 = b_1 \cdot b_2$. Jeżeli b_1 lub b_2 jest

liczbą pierwszą (np. b_1), to wtedy $n = a_1 \cdot a_2 = b_1 \cdot b_2 \cdot a_2 = b_1 \cdot (b_2 \cdot a_2)$ i mamy sytuację z twierdzenia 4.1. Dla dowolnej liczby naturalnej n jeżeli będziemy ją dzielić na coraz mniejsze czynniki w końcu pojawi się nam liczba pierwsza. A więc twierdzenie 4.1 jest prawdziwe.

Czy zależność przedstawioną w twierdzeniu 4.1 możemy jeszcze bardziej uszczegółowić? Okazuje się, że tak:

Twierdzenie 4.2 Każda liczba złożona dzieli się przez liczbę pierwszą mniejszą lub równą od pierwiastka z siebie

$$\boldsymbol{n = a \cdot p}$$

gdzie n jest liczbą złożoną,
p jest liczbą pierwszą oraz $p \leq \sqrt{n}$,
a jest liczbą naturalną

Dowód

Jeżeli n jest liczbą złożoną, to może być przedstawiona jako iloczyn dwóch liczb naturalnych większych od jeden i mniejszych n. Możliwe są dwa przypadki:

(1) $n = \sqrt{n} \cdot \sqrt{n}$
(2) $n = a \cdot p$

W sytuacji (1) mamy spełnioną tezę twierdzenia 4.2 (gdy $\sqrt{n}$ nie jest liczbą pierwszą, to można go rozłożyć zgodnie z lematem 4.1 na iloczyn liczby naturalnej i liczby pierwszej mniejszej od $\sqrt{n}$). Natomiast w sytuacji (2) $n = a \cdot p$ albo

(3) $p \leq \sqrt{n}$ oraz $\sqrt{n} < a$.
(4) $a < \sqrt{n}$ oraz $\sqrt{n} < p$ albo na odwrót

Gdyby obie liczby $a, p < \sqrt{n}$, to ich iloczyn byłby mniejszy od n. I odwrotnie, gdyby obie liczby $a, p > \sqrt{n}$, to ich iloczyn byłby większy niż n. W sytuacji (3) mamy spełnioną tezę. Natomiast jeżeli zachodzi sytuacja (4), to zgodnie z twierdzeniem 4.2 liczbę naturalną a $(a < \sqrt{n})$ możemy przedstawić jako iloczyn liczby pierwszej q oraz liczby naturalnej b. Skoro $a < \sqrt{n}$, to każdy z jej dzielników musi być mniejszy od $\sqrt{n}$, a więc także liczba pierwsza q. Co kończy nam dowód.

Uzbrojeni w taką wiedzę powróćmy w takim razie do sita Eratostenesa. Wypisaliśmy w nim wszystkie liczby naturalne od 1 do 100. Następnie wykreślaliśmy wszystkie wielokrotności liczb pierwszych (narastająco). Zgodnie z twierdzeniem 4.2 wystarczy wykreślić wszystkie wielokrotności liczb pierwszych 2, 3, 5, 7 aby pozostałe liczby były liczbami pierwszymi. Dlaczego? Otóż $\sqrt{100} = 10$. A więc każdą liczbę złożoną mniejszą lub równą od 100 możemy przedstawić jako wielokrotność liczby pierwszej mniejszej od $\sqrt{100} = 10$. A są to liczby pierwsze: 2, 3, 5, 7.Widzimy już, że liczby pierwsze nie poddają się pewnym regułom (nie można znaleźć schematu rozmieszczenia liczb pierwszych wśród liczb naturalnych), a z drugiej

strony mają ciekawe właściwości. Dojrzeliśmy już chyba, aby zadać sobie fundamentalne pytanie: ile właściwie jest tych liczb pierwszych? Czy jest ich nieskończenie wiele, czy wręcz przeciwnie jest ich pewna określona ilość? Odpowiedź na to pytanie dał nam już Euklides. Liczb pierwszych jest nieskończenie wiele, a dokładniej:

Twierdzenie 4.3

Żaden skończony zbiór nie zawiera wszystkich liczb pierwszych

Dowód

Niech P będzie niepustym skończonym zbiorem liczb pierwszych

$P = \{p_1, p_2, p_3, \dots, p_k\}$, gdzie p_i - liczby pierwsze

Weźmy liczbę N równą iloczynowi liczb p_i; $(i = 1..k)$

$$N = p_1 \cdot p_2 \cdot p_3 \cdot \dots \cdot p_k$$

Wówczas na pewno N jest podzielne przez p_1. Wtedy liczba $N + 1$ nie jest podzielna przez p_1.

(gdyby liczby N oraz $N + 1$ były podzielne przez p_1 wtedy ich różnica byłaby podzielna przez p_1: $N + 1 = b \cdot p_1$ oraz $N = a \cdot p_1$ $(a, b \in \mathbb{N})$ => $N + 1 - N = b \cdot p_1 - a \cdot p_1$ => $1 = (b - a) \cdot p_1$. Liczba p_1 jest liczbą pierwszą, a więc jest większa od 1. Czyli powyższa równość nie może być prawdziwa. Sprzeczność ta dowodzi, że jeżeli N jest podzielna przez p_1 to $N + 1$ nie może być podzielna przez a $\cdot p_1$)

W analogiczny sposób można udowodnić, że N + 1 nie jest podzielne przez żadną z liczb $p_1, p_2, \dots p_k$. Zatem liczba N + 1 nie jest podzielna przez żadną liczbę pierwszą mniejszą od siebie zawartą w zbiorze P. Na podstawie twierdzenia 4.1 liczbę N + 1 możemy określić na dwa sposoby:

(1) Liczba N + 1 jest liczbą pierwszą, której nie ma w skończonym zbiorze P
(2) Liczba N + 1 jest liczbą złożoną mającą dzielnik, który nie jest zawarty w zbiorze P

Zarówno w sytuacji (1) jak i w sytuacji (2) możemy powiedzieć, że dowolny skończony zbiór liczb pierwszych nie zawiera wszystkich liczb pierwszych. Co to oznacza? Że liczb pierwszych musi być nieskończenie wiele.

W Internecie wiele jest niepoprawnych dowodów na wyżej wymienione twierdzenie. Część z nich zakłada, że jeżeli wybierzemy pewną ilość najmniejszych liczb pierwszych, to ich iloczyn powiększony o 1 jest liczbą pierwszą. Nie jest to prawda. Liczba ta nie dzieli się oczywiście przez żadną z tych wybranych liczb pierwszych, ale może dzielić się przez jakąś inną liczbę pierwszą. Na przykład: 2 · 3 · 5 · 7 · 11 · 13 + 1 = 30 031 = 59 · 509. Z dowodami matematycznymi w Internecie trzeba uważać (jak zresztą ze wszystkim co spotykamy w sieci)!

Poniżej przytaczamy fundamentalne twierdzenie matematyki zwane Podstawowym Twierdzeniem Arytmetyki. Brzmi ono następująco:

Twierdzenie 4.4

Każda liczba naturalna n może być zapisana jako iloczyn liczb pierwszych na dokładnie jeden sposób.

Dowód

Postawmy hipotezę, że istnieją liczby naturalne, które można rozłożyć na iloczyn liczb pierwszych na co najmniej dwa sposoby. Jeżeli istnieją takie liczby, to istnieje również najmniejsza taka liczba naturalna. Nazwijmy ją *n*. Wtedy:

$$n = p_1 \cdot p_2 \cdot p_3 \ldots \cdot p_i \quad \text{oraz} \quad n = q_1 \cdot q_2 \cdot q_3 \cdot \ldots q_j \tag{1}$$

gdzie p_i i q_i oznaczają liczby pierwsze.

Poukładajmy p i q tak aby

$$p_1 \leq p_2 \leq \ldots \cdot \leq p_i \text{ oraz } q_1 \leq q_2 \leq \ldots \leq q_j. \tag{2}$$

Możemy tak zrobić, gdyż mnożenie jest przemienne

Zauważmy, że $p_1 \neq q_1$. Gdyby $p_1 = q_1$, wtedy oba równania (1) moglibyśmy podzielić przez p_1 i otrzymalibyśmy liczbę $\frac{n}{p_1} < n$, która ma istotnie dwa różne rozkłady na liczby pierwsze, co byłoby sprzeczne z założeniem, że *n* jest najmniejszą taką liczbą. Zatem $p_1 \neq q_1$. W takiej sytuacji możemy mieć dwie sytuacje: albo $p_1 < q_1$ albo $p_1 > q_1$. Załóżmy, że $p_1 < q_1$ (gdyby było odwrotnie w dalszej części dowodu po prostu zamieniamy miejscami p i q).

Tworzymy liczbę naturalną n' w następujący sposób:

$$n' = n - p_1 \cdot q_2 \cdot q_3 \cdot \ldots \cdot q_j \tag{3}$$

n' jest liczbą naturalną gdyż założyliśmy, że $p_1 < q_1$, a więc powyższa różnica jest większa od zera.

Teraz podstawiając oba równania (1) do równania (3) otrzymujemy dwa kolejne równania:

$$n' = p_1 \cdot p_2 \cdot \ldots p_i - p_1 \cdot q_2 \cdot q_3 \cdot \ldots \cdot q_j = p_1 \cdot (p_2 \cdot p_3 \cdot \ldots p_i - q_2 \cdot q_3 \cdot q_j) \tag{4}$$

$$n' = q_1 \cdot q_2 \cdot q_3 \cdot \ldots q_j - p_1 \cdot q_2 \cdot q_3 \cdot \ldots \cdot q_j = (q_1 - p_1) \cdot (q_2 \cdot q_3 \cdot \ldots \cdot q_j) \tag{5}$$

Zauważmy, że z równania (5) oraz z faktu, że $p_1 < q_1$ wynika że $n' < n$. A jeżeli tak, to rozkład *n'* musi być jedyny (zakładaliśmy, że n jest najmniejszą liczbą, która ma dwa różne rozkłady). Zatem równania (4) i (5) przedstawiają dwa takie same rozkłady (jeżeli nie bierzemy pod uwagę kolejności czynników). Z (4) wynika, że p_1 jest dzielnikiem *n'*, jeżeli zatem weźmiemy pod uwagę równanie (5) to p_1 jest dzielnikiem

$q_2 \cdot q_3 \cdot \ldots \cdot q_j$ lub (a)

$q_1 - p_1$. (b)

Weźmy pod uwagę sytuację (a), jeżeli p_1 byłoby dzielnikiem $q_2 \cdot q_3 \cdot \ldots \cdot q_j$. Wtedy można by zapisać następujące równanie:

$$n'' = p_1 \cdot k = q_2 \cdot q_3 \cdot \ldots \cdot q_j \quad (6)$$

Z nierówności (2) wiemy, że każde *q* jest większe od p_1, a zatem równanie (6) przedstawia rozkład liczby *n"* na dwa różne czynniki. Z konstrukcji liczby *n"* wynika, że *n" < n,* co jest sprzeczne z założeniem, że n jest najmniejszą liczbą mającą dwa różne rozkłady na liczby pierwsze.

Teraz weźmy pod uwagę sytuację (b): p_1 jest dzielnikiem $q_1 - p_1$ wtedy dla pewnego naturalnego k mamy:

$$q_1 - p_1 = p_1 \cdot k \Rightarrow q_1 = p_1 + p_1 \cdot k \Rightarrow q_1 = p_1 \cdot (k+1) \quad (7)$$

Z równań (7) wynika, że p_1 jest dzielnikiem q_1, co jest sprzeczne z założeniem, że q_1 jest liczbą pierwszą.

Zaprzeczenie tezy doprowadziło nas do sprzeczności z założeniem, a więc twierdzenie jest prawdziwe.

Już po tych kilku twierdzeniach i ich dowodach widzimy, że liczby pierwsze są tematem, który może fascynować. Zadziwiającym jest fakt, że z jednej strony nie możemy stwierdzić, gdzie na osi liczbowej pojawi się nam liczba pierwsza, a z drugiej strony znamy wiele własności tych liczb. Te pokazane w tym rozdziale to dopiero wierzchołek góry lodowej.

I jeszcze jedna uwaga: często spotykamy na różnych forach się z pytaniem *Czy liczba **jeden** jest liczba pierwszą? Dlaczego liczba **jeden** nie jest liczbą pierwszą?.* Jak wynika z definicji 4.1 liczba pierwsza to taka, która jest większa od 1 oraz dzieli się przez jeden oraz samą siebie (czyli ma dwa dzielniki). Liczba jeden po pierwsze nie jest większa od 1 i do tego ma tylko jeden dzielnik (liczbę jeden). Gdyby jedynkę zakwalifikować do liczb pierwszych to nieprawdziwe byłoby Podstawowe Twierdzenie Arytmetyki (twierdzenie 4.4). Przecież np. liczbę 12 można byłoby zapisać jako 12 = 2 · 2 · 3 oraz 12 = 1 · 2 · 3 · 3 (ogólnie $n = p_1 \cdot p_2 \cdot p_3 \cdot \ldots \cdot p_k$ oraz $n = 1 \cdot p_1 \cdot p_2 \cdot p_3 \cdot \ldots \cdot p_k$). Zatem liczbę 1 wyklucza się ze zbioru liczb pierwszych celem zapewnienia jednoznaczności twierdzeń w tej teorii.

Innym bardzo interesującym twierdzeniem dotyczącym liczb pierwszych jest ciekawe spostrzeżenie pewnej ich zależności od cyfry 6.

Twierdzenie 4.5

Każda liczba pierwsza (różna od 2 i 3) jest o 1 większa bądź mniejsza od wielokrotności liczby 6, czyli

$$\boldsymbol{p = 6n \pm 1,}$$

gdzie p jest dowolną liczbą pierwszą (poza 2 i 3), a n to liczba naturalna.

Dowód tego twierdzenia jest bardzo prosty:

- $6n + 0 = 6n$, ale $6n$ jest podzielne przez 6, więc nie może to być liczba pierwsza
- $6n + 1$ może być liczbą pierwszą, ponieważ nie jest podzielne przez 2 ani 3
- $6n + 2 = 2(3n + 1)$, ale $2(3n + 1)$ jest podzielne przez 2, więc nie może to być liczba pierwsza
- $6n + 3 = 3(2n + 1)$, ale $3(2n + 1)$ jest podzielne przez 3, więc nie może to być liczba pierwsza
- $6n + 4 = 2(3n + 2)$, ale $2(3n + 2)$ jest podzielne przez 2, więc nie może to być liczba pierwsza
- $6n + 5 = 6(n + 1) - 1$ może być liczbą pierwszą, ponieważ nie jest podzielne przez 2 i nie jest podzielne przez 3

Jednak twierdzenie odwrotne nie jest poprawne. Nie każda liczba o postaci $6n \pm 1$ jest liczbą pierwszą. Weźmy na przykład liczby 25 i 35, które oczywiście nie są pierwsze, a spełniają równania $25 = 6 \cdot 4 + 1$, $35 = 6 \cdot 6 - 1$.

Na podstawie powyższego twierdzenia można udowodnić jeszcze jedno. Tym razem ma ono związek z kwadratami liczb:

Twierdzenie 4.6

Kwadrat każdej liczby pierwszej (różnej od 2 i 3) jest o 1 większy bądź mniejszy od wielokrotności 24, czyli

$$\boldsymbol{p^2 = 24n \pm 1,}$$

gdzie p jest dowolną liczbą pierwszą (poza 2 i 3), a n liczbą naturalną.

Dowód wyżej wymienionego twierdzenia jest już trochę trudniejszy.

5. Rodzaje liczb pierwszych

Przez wieki badań nad liczbami pierwszymi matematycy znajdowali pewne specyficzne, charakterystyczne liczby pierwsze posiadające określone własności. I to o nich jest właśnie ten rozdział. Rodzajów liczb pierwszych jest dużo, w tym rozdziale przedstawię kilka z nich, które mnie zainteresowały bardziej od innych (z nieznanych mi powodów).

Liczby bliźniacze

To takie pary liczb pierwszych, których różnica wynosi 2. Już Pitagorejczycy, którzy cenili ład i harmonię wśród liczb interesowali się liczbami bliźniaczymi.

Przykłady liczb bliźniaczych: z początku osi liczb naturalnych: (3, 5), (5, 7), (11, 13), (17, 19), i te naprawdę wielkie: $(1639494 \cdot 2^{4423} - 1,\ 1639494 \cdot 2^{4423} + 1)$

Zauważmy, że liczba 5 jest bliźniacza zarówno do 3 jak i do 7.

W połowie XX w odkryto taką oto zależność: dla $n \geq 2$ liczby *n* i *n* + 2 są bliźniacze jeżeli

$$4((n-1)! + 1) + n \equiv 0(\mathrm{mod} n(n+2))$$

Natomiast początkiem XX w udowodniono, że suma odwrotności liczb bliźniaczych

$$\frac{1}{3} + \frac{1}{5} + \frac{1}{7} + \frac{1}{11} + \frac{1}{13} + \frac{1}{17} + \frac{1}{19} + \cdots$$

jest zbieżna. Wiemy, że szereg ten nie przekroczy nigdy pewnej stałej zwanej stałą Bruna. Dzisiaj znamy ją z dokładnością do 12 cyfr po przecinku γ = 1,902160583104 …

Największe do tej pory znalezione pary liczb bliźniaczych to (za wikipedia.org sierpień 2019):

$(2996863034895 \cdot 2^{1290000} - 1, 2996863034895 \cdot 2^{1290000} + 1).$

Nie udowodniono do tej pory skończoności, bądź nieskończoności zbioru par bliźniaczych.

Liczby czworacze

Istnieją także „czwórki" kolejnych liczb pierwszych, które tworzą pary liczb bliźniaczych. Jeżeli taką czwórkę tworzą liczby pierwsze $p, p+2, p+6$ i $p+8$, to grupy takie nazywamy liczbami czworaczymi.

Dlaczego $p, p+2, p+6$ i $p+8$, a nie $p, p+2, p+4$ i $p+6$?

$p \neq 2 \ \wedge \ p \neq 3$. Wtedy liczba p może przybrać dwie postaci:

(1) $p = 3k + 1$
(2) $p = 3k + 2$

jeżeli p jest postaci $3k + 1$ to $p + 2 = 3k + 1 + 2 = 3k + 3 = 3(k+1)$, więc $p + 2$ nie jest liczbą pierwszą. Zatem p musi być postaci $3k + 2$. Wtedy $p + 2 = 3k + 2 + 2 = 3k + 4$, więc $p + 2$ może być liczbą pierwszą.

Wiemy już że p musi być postaci $3k + 2$ aby zarówno p jak i $p + 2$ mogły być liczbami pierwszymi. A co z $p + 4$?

$p + 4 = 3k + 2 + 4 = 3k + 6 = 3(k + 2)$, a więc $p + 4$ nie jest liczbą pierwszą.

A co z $p + 6$?

$p + 6 = 3k + 2 + 6 = 3k + 8$, a więc $p + 6$ może być liczbą pierwszą.

Podobnie z $p + 8$

$p + 8 = 3k + 2 + 8 = 3k + 10$, a więc $p + 8$ może być liczbą pierwszą.

Dlatego istnieją „czwórki" liczb pierwszych postaci $p, p + 2, p + 6$ i $p + 8$, a nie istnieją liczby czworacze postaci $p, p + 2, p + 4$ i $p + 6$.

Liczby „szóstkowe" (ang. *sexy primes*)

Sexy prime to ciąg (para, trójka, itd…) liczb pierwszych, które różnią się od siebie o sześć (stąd nazwa). $(n, n + 6)$ $(n, n + 6, n + 12)$ $(n, n + 6, n + 12, n + 18)$

Największa znana „para" liczb szóstkowych to para liczb postaci $(n, n + 6)$ gdzie :
$n = (48011837012 \cdot ((53238 \cdot 7879\#)2 - 1) + 2310) \cdot 53238 \cdot 7879\#/385 + 1$

Największa znana trójka liczb szóstkowych to trójka liczb postaci (n , n+6 , n+12) gdzie :
$n = (84055657369 \cdot 205881 \cdot 4001\# \cdot (205881 \cdot 4001\# + 1) + 210) \cdot (205881 \cdot 4001\# - 1) / 35 + 1$

n ma 5132 cyfr
(znaczenie symbolu # *primorial* opisałem w rozdziale 2)

Liczby lustrzane

To takie pary liczb pierwszych, z których jedna powstaje przez zapisanie cyfr w odwrotnej kolejności. Przykład: 13 i 31, 17 i 71, 37 i 73, 79 i 97,107 i 701

Liczby względnie pierwsze

Liczby n_1 i n_2 są względnie pierwsze jeżeli ich największy wspólny dzielnik wynosi 1

Truncated primes (left-, right-)

Są to liczby pierwsze, które po skreśleniu kolejnych cyfr z prawej (lewej) strony pozostają liczbami pierwszymi

Np. 7393913, 739391, 73939, 7393, 739, 73, 7.

Liczby pierwsze izolowane

Liczba pierwsza p jest izolowana, jeśli najbliższa liczba pierwsza różni się od niej co najmniej o 4. Na przykład 89, 157, 173.

Liczby pierwsze palindromiczne

To liczby pierwsze, które nie zmieniają się, gdy ich cyfry zapiszemy w odwrotnej kolejności. Przykłady: 11, 101, 131, 191, 929.

Liczby Mersenne'a

Są to liczby postaci $M_n = 2^n - 1$, gdzie $n \in \mathbb{N}$. Nie wszystkie liczby Mersenne'a to liczby pierwsze, np. $M_{11} = 2047 = 23 \cdot 89$. Jednak są to bardzo dobre kandydatki do liczb pierwszych. Przez całe stulecia największymi znanymi liczbami pierwszymi były właśnie liczby Mersenne'a. W 1588 odkryto, że $M_{19} = 524287$ jest liczbą pierwszą i przez wiele lat była to największa znana taka liczba. Euler dwieście lat później udowodnił, że $M_{31} = 2147483647$ jest również liczbą pierwszą, ona także przez długi czas była rekordzistką. W czasach maszyn komputerów udowodniono także, że liczby $M_{521}, M_{607}, M_{2203}, M_{2281} \ldots . M_{13466917}, M_{57885161}$ są liczbami pierwszymi.

Liczby Mersenne'a są bardzo popularne wśród badaczy liczb pierwszych. Może to wynikać z faktu, że liczby te w rozwinięciu dwójkowym złożone są z samych jedynek (np. $M_{19} = 524287 = 1111111111111111111_2$). Dlatego komputerom jest dość prosto sprawdzać ich pierwszość. Jest jeszcze jeden powód popularności liczb Mersenne'a. Otóż istnieje twierdzenie związane z tymi liczbami, które otwiera pewną furtkę do poszukiwania naprawdę ogromnych liczb pierwszych

Twierdzenie 5.1

Jeśli liczba M_n jest liczbą pierwszą, to liczba n także jest pierwsza.

Twierdzenie odwrotne nie jest prawdziwe (np. $M_{11} = 2047 = 23 \cdot 89$). Wystarczy więc zajmować się liczbami Mersenne'a o podstawach liczb pierwszych i badać czy są one pierwsze.

W 1996r. powstał projekt Great Internet Mersenne Prime Search (GIMPS - mersenne.org) oparty na obliczeniach rozproszonych (wielka ilość komputerów udostępnia swoją moc obliczeniową). Od momentu powstania do chwili obecnej przy pomocy projektu GIMPS odkryto 17 liczb pierwszych Mersenne'a, w tym największą do tej pory $M_{82589933}$ składającą się z 24862048 cyfr.

Dzięki tak wielkiej popularności liczb Mersenne'a znamy dość dużo jej własności:

- jeżeli n jest liczbą złożoną, to M_n jest liczbą złożoną
- jeżeli liczba $n - 3$ jest podzielna przez 4 i n jest liczbą pierwszą, to $2n + 1$ dzieli M_n dokładnie wtedy, gdy $2n + 1$ jest liczbą pierwszą;
- dzielnik pierwszy liczby M_p, gdzie p jest liczbą pierwszą, jest postaci $2kp + 1$, gdzie k jest liczbą całkowitą dodatnią;
- liczba cyfr M_n liczby wyraża się wzorem: $D(n) = E(ln(2n - 1) + 1)$, gdzie: $E(x)$ oznacza część całkowitą liczby x,

Liczby Fermata

Liczbami Fermata nazywamy liczby postaci $F_n = 2^{2^n} + 1$, gdzie $n \in \mathbb{N} \cup \{0\}$. Fermat sprawdził, że dla $n \in \{0, 1 \dots 4\}$ wzór powyższy daje liczby pierwsze: $F_0 = 3$, $F_1 = 5$, $F_2 = 17, F_3 = 257, F_4 = 65537$. Niestety Euler wykazał, że $F_5 = 4294967297 = 641 \cdot 6700417$. Przy okazji udowodnił ciekawą właściwość liczb Fermata: każdy ewentualny dzielnik pierwszy (*p*) liczby F_n musi mieć postać $p = 2^{n+2} \cdot k + 1, gdzie\ k \in \mathbb{N}$. Sprawdźmy to dla podanej przed chwilą liczby F_5: $n = 5 \Rightarrow p = 2^{5+2}k + 1 = 2^7 \cdot k + 1$ i mamy dla $k = 10, p = 641$.

Dzisiaj wiemy, że $F_6, F_7, F_8, F_9, F_{10}, F_{11}$ są liczbami złożonymi. Co więcej, znamy ich rozkłady. Wiemy, że kolejna liczba Fermata F_{12} jest liczbą złożoną, ale nie znamy jeszcze jej rozkładu – podobnie jest z następnymi dwunastoma liczbami. Jak widać z tych przykładów liczby Fermata nie są zbyt dobrymi kandydatami do liczb pierwszych (w przeciwieństwie do liczb Mersenne'a).

Liczby pierwsze Gaussa

Zacznijmy od liczb Gaussa. Otóż są to liczby zespolone $z = a + b \cdot i$, gdzie $a, b \in \mathbb{Z}$ ($\mathbb{Z}$ - zbiór liczb całkowitych). Spróbujmy teraz zdefiniować liczby pierwsze Gaussa. Otóż możemy spróbować zacząć podobnie jak ze zwykłymi liczbami pierwszymi - są to liczby Gaussa podzielne tylko przez siebie i przez 1. Widzimy już od razu błąd. Każda liczba Gaussa ma przecież 8 dzielników (z, $-z$, zi, $-zi$, 1, -1, i, $-i$), a nie 2 jak w przypadku liczb naturalnych. Możemy więc zmienić naszą definicję uwzględniając te dzielniki.

Liczba pierwsza Gaussa z to taka liczba, która dzieli się tylko przez z, $-z$, zi, $-zi$, 1, -1, i, $-i$.

Zwróćmy uwagę, że liczby naturalne należą do zbioru liczb zespolonych, a w szczególności do liczb Gaussa (część urojona jest zerowa). Ale naturalne liczby pierwsze nie muszą być liczbami pierwszymi Gaussa. Na przykład 5 jest liczbą pierwszą, ale nie jest liczbą pierwszą Gaussa, bo $5 = (2 + i) \cdot (2 - i)$.

6. Twierdzenie Fermata o Dwóch Kwadratach

Słynny francuski matematyk Pierre de Fermat także analizował właściwości liczb pierwszych. Wynikiem tych badań jest m.in. tak zwane Twierdzenie Fermata o dwóch kwadratach.

Twierdzenie 6.1 O dwóch kwadratach

1. Każdą liczbę pierwszą o postaci $4k + 1$ można przedstawić dokładnie na jeden sposób jako $x^2 + y^2$
2. Żadnej liczby pierwszej o postaci $4k + 3$ w ten sposób przedstawić się nie da

gdzie $k \in \mathbb{N} \cup \{0\}; \quad x, y \in \mathbb{N}$

Liczba 2 jest tutaj specjalna, ponieważ nie bierzemy jej pod uwagę w dowodzie.

Dowód:

Załóżmy, że p to liczba pierwsza o postaci $4k + 1$, chcemy teraz udowodnić, że $p = x^2 + y^2$, $x, y \in \mathbb{N} \cup \{0\}$. Liczba p musi być nieparzysta (wszystkie liczby pierwsze poza 2 są nieparzyste), więc $x^2 + y^2$ musi być nieparzyste. Oznacza to, że jeden z tych kwadratów musi być parzysty, a drugi nie. Załóżmy, że y jest parzyste (gdyby x było parzyste, dowód wygląda tak samo, tylko zamieniamy oznaczenia) - czyli jest o postaci $2k$ => $p = x^2 + (2k)^2$ i tutaj dobrze by było zmienić minimalnie naszą notację. Otóż literę k wykorzystujemy we wzorze $4k + 1$, więc zamieńmy k ze wzoru $p = x^2 + (2k)^2$ na y, ponieważ już go nie używamy. Rozpisując powyższy wzór otrzymujemy: $p = x^2 + (2y)^2 = x^2 + 4y^2$. To jest nadal trudne do udowodnienia równanie. Spróbujmy więc uprościć je trochę. Wiemy, że $y^2 = y \cdot y$. Zamieńmy więc jeden y na z, czyli: $p = x^2 + 4yz$. Teraz, jeżeli znajdziemy wszystkie rozwiązania tego równania to możemy próbować udowodnić, że przy jednym z nich $y = z$. Zacznijmy więc szukać rozwiązań. Wiemy, że $p = 4k + 1$, więc $4k + 1 = x^2 + 4yz$.

Możemy teraz zauważyć już dwa trywialne rozwiązania. Otóż jeżeli $x = 1, y = 1$ i $z = k$ (bądź $y = k, z = 1$) to otrzymujemy: $x^2 + 4yz = 1^2 + 4 \cdot 1 \cdot k = 4k + 1$ (albo $x^2 + 4yz = 1^2 + 4 \cdot k \cdot 1 = 4k + 1$). Ale czy są jeszcze jakieś inne rozwiązania? Oczywiście, że tak! Jest ich nawet wiele. Można łatwo udowodnić, że jeżeli nasze pierwotne twierdzenie jest prawdziwe, to ilość rozwiązań musi być nieparzysta. Dla każdego rozwiązania istnieje rozwiązanie przeciwne, gdzie y i z są zamienione miejscami. Poza jednym rozwiązaniem – tym gdzie $y = z$. Jest to jednak dowód niepełny, ponieważ uznajemy, że twierdzenie jest prawdziwe. Jednak nadal może nam się przydać. Musimy go jednak odwrócić, czyli: jeżeli ilość rozwiązań jest nieparzysta to jednym z nich musi być przypadek, gdzie $y = z$.

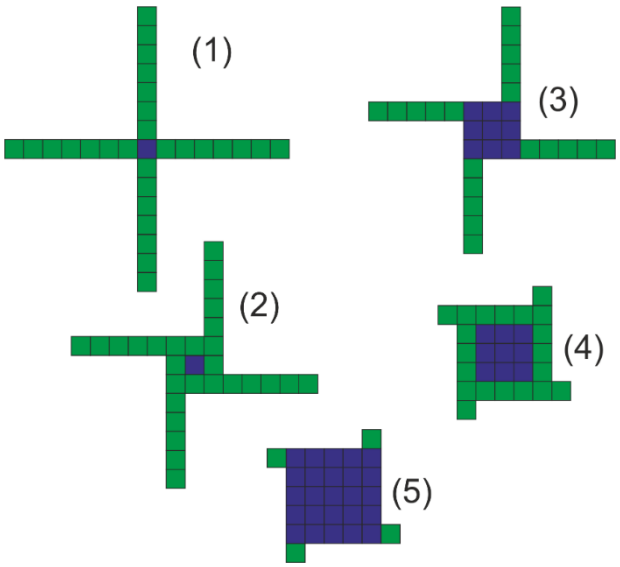
Oznacza to, że jeżeli udowodnimy w inny sposób, że ilość rozwiązań równania $4k + 1 = x^2 + 4yz; \ p = 4k + 1, p \in \mathbb{P}$ jest nieparzysta to udowodniliśmy twierdzenie Fermata. <A>

Spróbujmy geometrycznie przedstawić nasz wzór: $p = x^2 + 4yz$. Przedstawmy x^2 jako kwadrat o boku x, natomiast $4yz$ jako 4 prostokąty o bokach y i z. Aby dowód był bardziej

zrozumiały rozpatrzmy najpierw sytuację dla jednego konkretnego przypadku. Np. dla liczby pierwszej 29. Jest ona postaci $4k + 1$. Można ją rozłożyć na kilka sposobów:

	x	y	z	
29 =	$1^2 + 4 \cdot$	$1 \cdot$	7	(1)
29 =	$1^2 + 4 \cdot$	$7 \cdot$	1	(2)
29 =	$3^2 + 4 \cdot$	$1 \cdot$	5	(3)
29 =	$3^2 + 4 \cdot$	$5 \cdot$	1	(4)
29 =	$5^2 + 4 \cdot$	$1 \cdot$	1	(5)

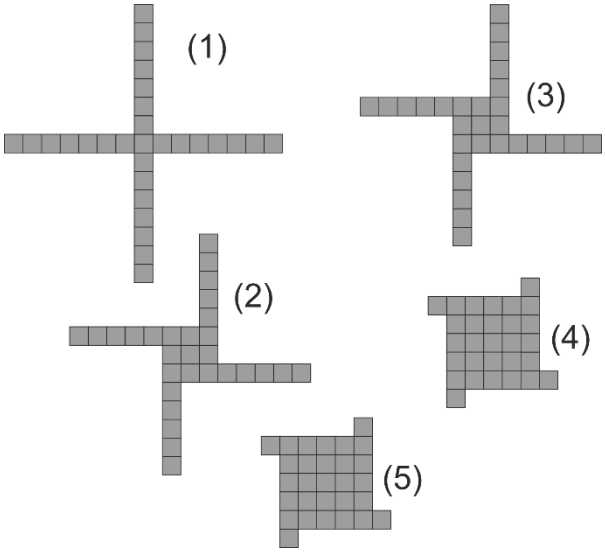
Wiemy, że kwadrat ma 4 boki, więc możemy nasze 4 prostokąty „przykleić" do jego boków w taki sposób:



Rysunek 10 Rozkład liczby 29 (pokolorowany)

Wszystkie te rozwiązania występują parami (poza ostatnim). Równania (1) i (2) oraz (3) i (4) to są te same równania, tylko zamieniamy miejscami y i z. Na razie rozkładaliśmy liczbę 29. Ale jeżeli prześledzimy konstrukcję przekształcania rozkładu liczby na jej graficzną reprezentację to zobaczymy, że każdy rozkład możemy tak przedstawić.

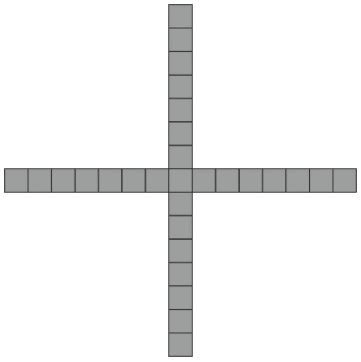
Musimy teraz udowodnić, że ilość takich figur o polu p jest nieparzysta. Zauważmy, że jeżeli popatrzymy na każdą z tych figur to możemy ją podzielić na 2 różne rozwiązania.



Rysunek 11 Rozkład liczby 29 (niepokolorowany)

Zauważmy, że jeżeli na Rysunku 10 zrezygnujemy z kolorów to będziemy mieć parami identyczne figury (2) i (3) oraz (4) i (5). Znowu oprócz jednego tym razem (1)

Dowód na to, że każdą figurę można tak przedstawić jest dosyć skomplikowany, więc nie będę go tutaj przedstawiał. Ale jeżeli każda figura (poza jedną) reprezentuje 2 rozwiązania to znaczy, że ilość rozwiązań musi być nieparzysta, gdyż jedna figura reprezentuje tylko jedno rozwiązanie. Wygląda ona jak plus:



Rysunek 12 Pojedyncze graficzne rozwiązanie

Czy zawsze taka figura powstanie nam z rozkładu liczby? Tak, wynika to z rozumowania przedstawionego kilka paragrafów wyżej – jest to jeden z trywialnych rozkładów $(x = 1, y = 1, z = k)$.

Udowodniliśmy więc, że ilość rozwiązań jest nieparzysta. Czyli (patrz <A>) jedno z nich jest o postaci $y = z$.

Czyli udowodniliśmy twierdzenie…

No prawie udowodniliśmy. Nie pokazaliśmy jeszcze, że istnieje tylko jedno rozwiązanie równania $p = x^2 + y^2$. Tak naprawdę, to wiemy tylko, że ilość rozwiązań jest nieparzysta. Oznacza to, że mogą być np. 3 rozwiązania. Więc teraz następnym krokiem jest udowodnienie, że każde 2 rozwiązania są sobie równoważne. Jest to dowód dosyć skomplikowany i zawierający dużo symboli.

Lemat 6.1

Jeżeli mamy dwie liczby, które można przedstawić jako sumę dwóch kwadratów: $a^2 + b^2$ i $c^2 + d^2$ to ich iloczyn można przedstawić jako suma dwóch kwadratów na dwa różne sposoby:

$$(a^2 + b^2)(c^2 + d^2) = (ac - bd)^2 + (ad + cb)^2$$

$$(a^2 + b^2)(c^2 + d^2) = (ac + bd)^2 + (ad - cb)^2$$

Załóżmy, że $a^2 + b^2$ i $c^2 + d^2$, gdzie $a, b, c, d \in \mathbb{N} \cup \{0\}$, to dwa różne sposoby na przedstawienie liczby pierwszej p o postaci $4k + 1$. Oznacza to, że $p^2 = (a^2 + b^2)(c^2 + d^2)$ również można przedstawić na dwa sposoby jako sumę dwóch kwadratów:

$$p^2 = (ac - bd)^2 + (ad + cb)^2$$

$$p^2 = (ac + bd)^2 + (ad - cb)^2$$

Teraz zaczynamy przekształcać p:

$$p = a^2 + b^2 \qquad p = c^2 + d^2$$

$$p - a^2 = b^2 \qquad p - c^2 = d^2$$

$$d^2(p - a^2) = b^2 d^2 \qquad b^2(p - c^2) = b^2 d^2$$

Widzimy, że prawe strony równań są sobie równe. Oznacza to, że lewe strony też są równe:

$$d^2(p - a^2) = b^2(p - c^2)$$

$$pd^2 - a^2 d^2 = pb^2 - c^2 b^2$$

$$pd^2 - pb^2 = a^2 d^2 - c^2 d^2$$

$$p(d^2 - b^2) = (ad)^2 - (cd)^2$$

$$p(d^2 - b^2) = (ad - cb)(ad + cb)$$

p jest liczbą pierwszą, więc musi dzielić albo $(ad - cb)$ albo $(ad + cb)$. Załóżmy pierwsze, że $p|(ad - cb)$. Oznacza to, że $(ad - cb)$ musi być wielokrotnością liczby p. Innym wzorem, w którym mieliśmy $(ad - cb)$ jest $p^2 = (ac + bd)^2 + (ad - cb)^2$. Gdyby $(ad - cb)$ było równe $p, 2p, 3p$ lub więcej to prawa strona równania byłaby większa niż lewa. Oznacza to, że $(ad - cb)$ musi być równe 0.

Wracając do wzoru $p(d^2 - b^2) = (ad - cb)(ad + cb)$ możemy podstawić wartość 0 pod $(ad - cb)$:

$$p(d^2 - b^2) = (ad - cb)(ad + cb)$$

$$p(d^2 - b^2) = 0(ad + cb)$$

$$p(d^2 - b^2) = 0$$

$$d^2 - b^2 = 0$$

$$d^2 = b^2$$

Wracając teraz do samego początku:

$$p = a^2 + b^2 = c^2 + d^2$$

$$a^2 + b^2 = c^2 + d^2$$

$$a^2 + b^2 = c^2 + b^2$$

$$a^2 = c^2$$

Oznacza to, że obydwa sposoby na przedstawienie p jako sumy dwóch kwadratów są sobie równoważne.

Ale został nam jeszcze drugi przypadek, gdzie $p|(ad + cb)$. Teraz używamy go przy wzorze $p^2 = (ac - bd)^2 + (ad + cb)^2$. Wiemy, że $(ad + cb)$ nie może być równe 0 (ponieważ p jest

większe od 0), więc musi być równe p (gdyby było większe to prawa strona równania będzie większa niż lewa). Więc:

$$p^2 = (ac - bd)^2 + (ad + cb)^2$$

$$p^2 = (ac - bd)^2 + p^2$$

$$0 = (ac - bd)^2$$

$$0 = ac - bd$$

$$ac = bd$$

Znowu wracając do początku: $p = a^2 + b^2$ możemy zauważyć, że największym wspólnym dzielnikiem a i b musi być 1. Gdyby ten dzielnik był większy to p miało by również ten dzielnik, co oznaczałoby, że p nie jest pierwsze. Oznacza to, że na podstawie wzoru $ac = bd$ można powiedzieć, że $a|d$ i $b|c$. Ale p jest również równe $c^2 + d^2$, a wiemy już, że c jest wielokrotnością a oraz d jest wielokrotnością b. Gdyby $c > a$ i $d > b$ to $a^2 + b^2 < c^2 + d^2$, więc $p < c^2 + d^2$. Natomiast gdyby $c < a$ (lub $d < b$) to $p = d^2$ (lub $p = c^2$) a wiemy, że p jest liczbą pierwszą, co oznacza, że p nie może być kwadratem.

Widzimy teraz, że oby dwie opcje są sprzeczne z założeniem, więc oba rozwiązania są sobie równoważne.

Podsumowując. Na początku, aby uprościć dowód zastosowaliśmy sprytną „sztuczkę" z zamianą jednego y z $4y^2$ na z. Potem graficznie przedstawiliśmy rozwiązania nowego równania. Wykazaliśmy, że rozwiązań musi być nieparzysta ilość – co oznaczało, że przynajmniej jednym z nich było rozwiązanie gdzie $y = z$. A następnie udowodniliśmy, że takie rozwiązanie jest tylko jedno.

Co kończy dowód.

7. *Liczba* π a liczby pierwsze

Czy liczby pierwsze, które są tematem mojej pracy mają coś wspólnego z liczbą π? Wydawałoby się, nie. Przecież π (3,1415 …) jest wykorzystywana do obliczania obwodu i pola koła. Ma zastosowanie w trygonometrii. A liczby pierwsze? Co to, to nie! Na pewno nie ma „wspólnego mianownika" pomiędzy π, a liczbami pierwszymi! Cóż, w pewnym sensie to prawda, ale… Istnieje pewien wzór na obliczanie π, który w swoim dowodzie wykorzystuje liczby pierwsze. Nieskończona suma odwrotności liczb nieparzystych, gdzie na co drugim miejscu jest znak ujemny wynosi „ćwierć π":

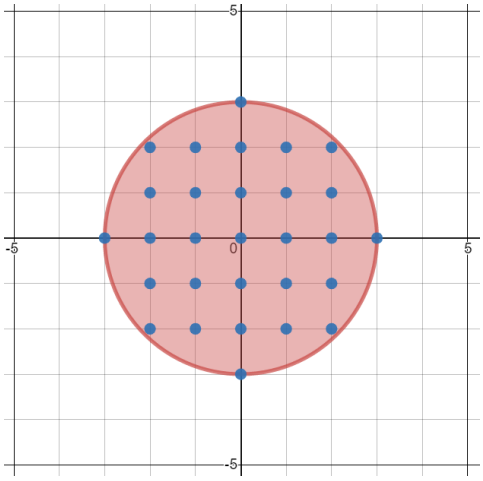
$$1-\frac{1}{3}+\frac{1}{5}-\frac{1}{7}+\frac{1}{9}-\frac{1}{11}+\frac{1}{13}-\cdots=\frac{\pi}{4}$$

Powyższe równanie można udowodnić na dwa różne sposoby. Tutaj wykorzystamy ten, który „pokazuje gdzie jest koło" (bo przecież wszędzie gdzie π musi być ukryte jakieś koło).

Zacznijmy więc od tego, w jaki sposób można policzyć π. Są dwie łatwe metody. Pierwsza to przekształcenie wzoru na pole $P=\pi r^2$, natomiast druga to przekształcenie wzoru na obwód $Obw=2\pi r$. Skupmy się na pierwszym wzorze. Można go przekształcić do postaci: $\pi=\frac{P}{r^2}$. Teraz wystarczy, że znajdziemy inny sposób (nie wykorzystujący π) na obliczenie pola koła (P) o określonym promieniu (r).

Twierdzenie 7.1

Pole koła to w przybliżeniu liczba punktów kratowych znajdujących się w nim.



Rysunek 13 Punkty kratowe

Powyższe twierdzenie wynika pośrednio ze wzoru Picka na obliczanie pola powierzchni wielokąta prostego:

$$P=W+\frac{1}{2}B-1,$$

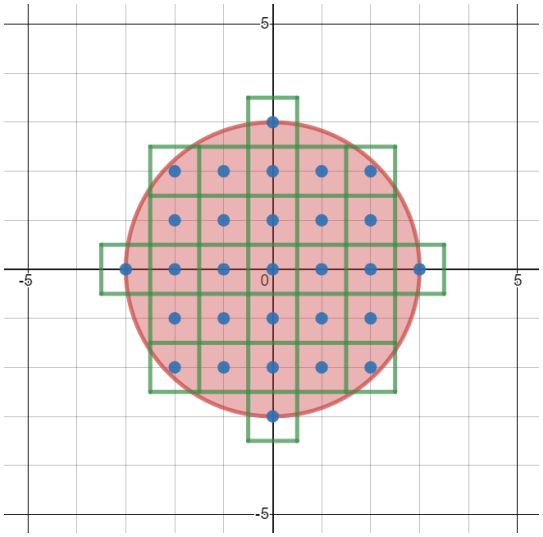
gdzie W to liczba punktów kratowych leżących wewnątrz wielokąta, B – oznacza liczbę punktów kratowych leżących na brzegu wielokąta. Dowód znajdziemy w [10]. Możemy przyjąć, że koło to wielokąt foremny z nieskończoną liczbą boków.

Punkty kratowe, to nic innego jak punkty w układzie współrzędnych, o obu współrzędnych całkowitych.

Dla kół o dużym promieniu wzór Picka możemy uprościć do postaci: $P=W+B$ (dowód można znaleźć w [10])

Popatrzmy na rysunek 13. W tym przypadku mamy koło o promieniu 3. Ilość punktów kratowych znajdujących się w nim to 29. Wykorzystując nasz wzór na π otrzymujemy:

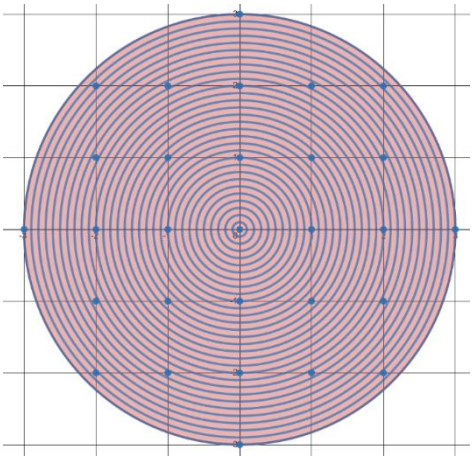
$$\pi=\frac{P}{r^2}=\frac{29}{3^2}=3{,}222\ldots$$



Aby lepiej wyobrazić sobie dlaczego ilość punktów kratowych jest przybliżona do wartości pola można wyobrazić sobie kwadraty o polu 1 wokół każdego z tych punktów. Im większy jest promień koła tym dokładniej ten sposób będzie wyznaczał jego pole.

Rysunek 14 Punkty kratowe określające pole koła

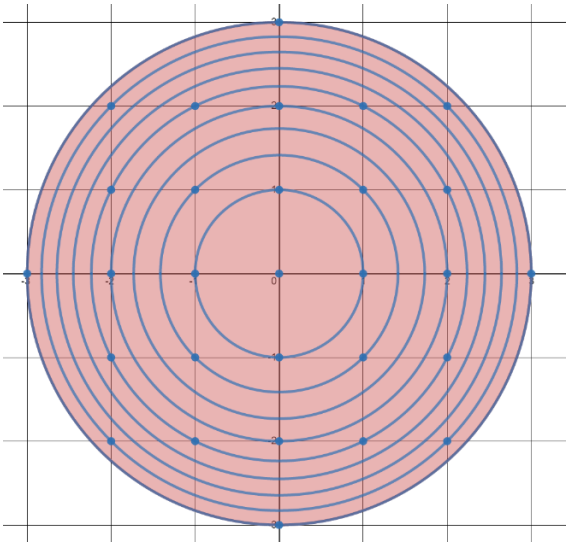
Teraz pozostaje pytanie, w jaki sposób policzyć ilość punktów kratowych. Otóż można to zrobić w taki sposób, że bierzemy szereg współśrodkowych okręgów o rosnącym promieniu R od 0 do r. Przyjmijmy, że R rośnie o jakąś stałą. Zliczamy ilość punktów kratowych na każdym z tych okręgów. Na rysunku 15 znajduje się przykład z okręgami, gdzie różnice promieni wynoszą $0{,}1$.



Rysunek 15 Okręgi współśrodkowe, o różnicy promienia 0,1

Nie są to oczywiście wszystkie okręgi. Aby policzyć wszystkie punkty kratowe należałoby narysować jeden okrąg dla każdej liczby rzeczywistej między 0, a r. Co oczywiście nie jest możliwe.

Zauważmy, że punkty kratowe mają współrzędne postaci (a, b), gdzie a i b są liczbami całkowitymi. Zatem ich odległość od środka układu wynosi: $R = \sqrt{a^2 + b^2}$, skoro a i b są całkowite, to $a^2 + b^2$ będzie liczbą naturalną. Oznacza to, że R może przyjmować wartości równe tylko pierwiastkom z liczb naturalnych. Dzięki temu spostrzeżeniu możemy ograniczyć ilość okręgów z nieskończoności do r^2 (gdzie r – promień koła).



Rysunek 16 Okręgi współśrodkowe o promienia będącymi pierwiastkami liczb naturalnych

W jaki sposób możemy policzyć ilość punktów kratowych na danym okręgu? Do tego jest nam teraz potrzebna trochę bardziej zaawansowana matematyka. Po pierwsze musimy zamienić punkty całkowite (a, b) na liczby zespolone. Czyli na przykład punkt $(2,1)$ zamieniamy w liczbę $2 + 1i$. Jest pewna ciekawa własność liczb zespolonych: $(a + bi)(a - bi) = a^2 + b^2$. Zauważmy, że $a^2 + b^2$ to dokładnie kwadrat naszej liczby R. Czyli możemy przekształcić nasze pytanie: ile jest liczb zespolonych takich, że iloczyn danej liczby zespolonej i jej sprzężenia daje n, gdzie n jest liczbą naturalną, będącą zarazem R^2. I tu zaczyna się robić już ciekawie. Możemy zauważyć, że mając jakąś liczbę n (np. 70), można uzyskać ilość wszystkich par liczb zespolonych $(a + bi, a - bi)$ w następujący sposób:

1. Dzielimy liczbę na czynniki pierwsze $(70 = 2 \cdot 5 \cdot 7)$
2. Znajdujemy pary liczb zespolonych dla każdej liczby pierwszej (np. $5 = (2 + i) \cdot (2 - i)$, gdyż $(2 + i)(2 - i) = 2^2 + 1^2 = 4 + 1 = 5$)
3. Mnożymy ze sobą po jednej liczbie zespolonej z każdej pary dla danej liczby n (daje nam to wszystkie możliwe kombinacje)
4. Mnożymy ilość kombinacji przez 4 (każdą parę liczb zespolonych można pomnożyć przez $1, -1, i, -i$ i nadal ich iloczyn będzie równy n)

Spróbujmy więc dokładniej opisać ten proces.

Punkt 1. Jest łatwy do zrozumienia, nie potrzeba nam do niego na razie żadnych wzorów.

Natomiast już w **punkcie 2.** pojawia się problem. Istnieje prosty dowód na to, że jeżeli liczbę pierwszą możemy rozłożyć na iloczyn liczb zespolonych całkowitych, to rozkłada się ona w dokładnie jeden sposób (z pominięciem par symetrycznych). Jednak skąd możemy wiedzieć, czy dana liczba pierwsza jest rozkładalna w sposób przedstawiony w tym punkcie. Otóż do tego przyda nam się Twierdzenie Fermata o Dwóch Kwadratach (patrz rozdział nr 6). Mówi ono o tym, że liczbę pierwszą można przedstawić jako sumę dwóch kwadratów tylko i wyłącznie wtedy, gdy można ją zapisać jako $4k + 1$. Oznacza to, że liczby pierwsze o postaci

$4k+3$ nie można tak przedstawić. A wiemy przecież, że zapisanie liczby jako sumy dwóch kwadratów jest równoznaczne ze znalezieniem pary liczb zespolonych, których szukamy.

Teraz **punkt 3.** W jaki sposób możemy pomnożyć ze sobą te liczby, tak aby powstały nam poprawne pary dla liczby n. Możemy to zrobić posługując się tabelą. Otóż w pierwszej kolumnie wpisujemy wszystkie liczby pierwsze powstałe w punkcie 1. W drugiej i trzeciej kolumnie wpisujemy pary liczb zespolonych odpowiadających danej liczbie pierwszej. Oto tabela dla liczby $5525 = 5 \cdot 5 \cdot 13 \cdot 17$.

Liczba	Czynnik 1.	Czynnik 2.
5	$2+i$	$2-i$
5	$2+i$	$2-i$
13	$3+2i$	$3-2i$
17	$4+i$	$4-i$

Rysunek 17 Rozkład 5525 na czynniki pierwsze Gaussa

Gdy pomnożymy ze sobą osobno liczby zespolone z drugiej kolumny oraz te z trzeciej kolumny to powstanie nam para liczb zespolonych reprezentująca liczbę n – 5525.

$$(2+i)\cdot(2+i)\cdot(3+2i)\cdot(4+i) = -14+73i$$

$$(2-i)\cdot(2-i)\cdot(3-2i)\cdot(4-i) = -14-73i$$

Zauważmy, że mogliśmy tę tabelę skonstruować na kilka sposobów. Ponieważ w pierwszym wierszu zamiast $2+i, 2-i$ mogliśmy wpisać $2-i, 2+i$. Gdy teraz pomnożymy ze sobą odpowiednie kolumny powstanie nam nowa para. Ile jest w takim razie możliwości na jakie możemy stworzyć tabelę? Na pierwszy rzut oka możemy stwierdzić, że dla każdej liczby pierwszej mamy dwie możliwości, więc ilość kombinacji to $2^{ilość\ liczb\ pierwszych}$. Jest to jednak wzór błędny. Zauważmy, że mając dwie piątki (dwie pierwsze linie na rysunku 17), nie możemy z nich stworzyć 4 kombinacji, ale tylko 3 (w drugiej kolumnie możemy wpisać: $2+i$, $2+i$ lub $2+i$, $2-i$ lub $2-i$, $2-i$). I tak samo z innymi liczbami pierwszymi. Oznacza to, że poprawniejszy byłby wzór: $(w_1+1)(w_2+1)(w_3+1)$ … gdzie w_1 to kolejne wykładniki liczb pierwszych. Na przykładzie powyżej: $5525 = 5^2 \cdot 13^1 \cdot 17^1$, czyli: $w_1 = 2, w_2 = 1, w_3 = 1$, używając powyższego wzoru otrzymujemy:

$$(w_1+1)(w_2+1)(w_3+1) = (2+1)(1+1)(1+1) = 3 \cdot 2 \cdot 2 = 12$$

Czyli istnieje 12 par liczb postaci $a+bi$ i $a-bi$, takich, że ich iloczyn jest równy 5525.

Należy jednak zauważyć, że powyższy przykład (5525) miał czynniki pierwsze tylko postaci $4k+1$. Co by się stało, gdybyśmy mieli liczbę postaci $4k+3$?

Liczba	Czynnik 1.	Czynnik 2.
3	?	?
5	$2+i$	$2-i$
5	$2+i$	$2+i$
13	$3+2i$	$3+2i$

Rysunek 18 Rozkład 975 na czynniki pierwsze Gaussa

Otóż od razu pojawia się kłopot. Nie mamy gdzie wpisać trójki. Jeżeli umieścimy ją w jednej z dwóch kolumn, to ich iloczyny nie będą o postaci $a+bi, a-bi$. Problem byłby rozwiązany, gdyby były dwie trójki, albo cztery, albo jakakolwiek parzysta liczba trójek. Ponieważ wtedy moglibyśmy je po równo rozłożyć między te dwie kolumny (nie zmienia się wtedy ilość kombinacji). Natomiast jeżeli ich ilość jest nieparzysta, to nie ma żadnego sposobu na rozłożenie czynników (jest to równoważne z tym, że nie ma żadnych par $a+bi, a-bi$, których iloczyn to n).

Zostaje nam teraz tylko liczba 2. Nie jest ona ani postaci $4k+1$ ani $4k+3$. Dzieli się ona na $1+i, 1-i$. Należy jednak zauważyć, że $(1-i)*i = 1+i$ – po pomnożeniu jednej z nich przez i wychodzi nam druga. Oznacza to, że przy wstawianiu tej pary do tabelki mamy tylko jedną możliwość (a nie dwie), ponieważ zmiana czynników miejscami jest równoważne z mnożeniem przez i lub $-i$. Nic to nie zmieni, gdyż tę samą czynność wykonujemy w pkt. 4.

Punkt 4 jest łatwy do zrozumienia. Każdą z liczb wyjściowych z punktu 3 możemy „obrócić" w układzie kartezjańskim. Jest to równoznaczne z mnożeniem danej liczby zespolonej przez $1, -1, i$ lub$-i$. Są 4 sposoby na pomnożenie, więc mnożymy wynik przez 4.

Mamy więc teoretycznie łatwy sposób na obliczenie ilości punktów kratowych znajdujących się na okręgu o promieniu R (czyli $\sqrt{n}$). Ważne jest tu słowo „teoretycznie". Są dwa zasadnicze problemy z naszym sposobem. Pierwszy problem jest taki, że musimy daną liczbę n podzielić na czynniki pierwsze, a jak się okazuje, nie jest to wcale takie proste. Drugim problemem jest fakt, że algorytm zawarty w punkcie 3 jest trudny do opisania (np. za pomocą funkcji). Spróbujmy te problemy teraz rozwiązać.

Zacznijmy od drugiego. Otóż jednym ze sposobów na usystematyzowanie algorytmu jest zapisać właściwe kroki:

1. Dzielimy czynniki pierwsze na 3 grupy: $4k+1, 4k+3, 2$
2. Sprawdzamy wykładniki dla każdego czynnika postaci $4k+3$. Jeżeli chociaż jeden jest nieparzysty to wynikiem jest 0.
3. Jeżeli przy kroku 2 nie wyszło nam 0, to wynikiem jest iloczyn kolejnych sum wykładników czynników pierwszych i jeden (czyli $(w_1+1)(w_2+1)(w_3+1)\dots$)
4. Czynnik pierwszy 2 nic nie zmienia.

Krok drugi możemy rozłożyć na taki wzór (prawda – 1, fałsz – 0):

$$(czy\ w_1\ jest\ parzyste?)(czy\ w_2\ jest\ parzyste?)(czy\ w_3\ jest\ parzyste?)$$
$$(czy\ w_4\ jest\ parzyste?)\dots$$

Gdy wynik w/w wzoru pomnożymy z wynikiem kroku 3 będziemy mieli jeden wzór na cały algorytm. Tylko jak teraz bardziej matematycznie przedstawić pytanie „$czy\ w_1\ jest\ parzyste?$" (nie używając logiki)? Można to zrobić następująco: podnosimy p_1(liczbę pierwszą o wykładniku w_1) kolejno do potęg: $0, 1, 2, \dots, w_1$. Następnie tworzymy nową funkcję $f(x)$, która daje wartość 1 dla co drugiej potęgi; i (-1) dla pozostałych. Możemy powiedzieć, że $(czy\ w_1\ jest\ parzyste?) = f(p_1^0) + f(p_1^1) + f(p_1^2) + \dots + f(p_1^{w_1})$. Dla uproszczenia powiedzmy, że $f(1) = 1$. W tym przypadku nasze p jest postaci $4k+3$, co

oznacza że $f(4k+3)$ musi być równe (-1), ponieważ gdyby w_1 było równe 1, to w tym miejscu zatrzyma się suma, a wiemy że w_1 jest nieparzyste, czyli wynik musi być 0 ($f(1)=1$, więc kolejna musi być (-1)). A co jeżeli $w_1=2$? Wtedy mamy: $(4k+3)^2=16k^2+24k+9=4(4k^2+6k+2)+1=4m+1$, oznacza to, że dla wszystkich liczb postaci $4k+1$ funkcja f przyjmuje wartość 1 (podobny argument jak poprzednio). Dalej mamy $w_1=3$, czyli $(4m+1)\,(4k+3)=16mk+12m+4k+3=4(4mk+3m+k)+3=4j+3$, i już ustaliliśmy, że liczby tej postaci dają (-1), co się zgadza z nowym wzorem. Kolejne wartości w_1 powtarzają się: raz $4k+3$, a następnie $4k+1$.

Spróbujmy teraz tę funkcję wykorzystać do kroku 3. Okazuje się to nadzwyczajnie proste. Otóż rozpisujemy je w dokładnie taki sam sposób, jak przy kroku 2: $f(p_1^0)+\ f(p_1^1)+f(p_1^2)+\cdots+f\left(p_1^{w_1}\right)$. Funkcję f mamy już zdefiniowaną, okazuje się, że poprawnie również i dla tej sytuacji. Otóż: $(4k+1)^2=16k^2+8k+1=4(4k^2+2k)+1=4m+1$, czyli wszystkie potęgi $4k+1$ również będą postaci $4k+1$, a $f(4k+1)=1$, więc $f(p_1^0)+\ f(p_1^1)+f(p_1^2)+\cdots+f\left(p_1^{w_1}\right)=1+1+1+\cdots+1=w_1+1$ (taki wynik chcieliśmy otrzymać).

Zostaje jeszcze krok 4. Nieważne jaki jest wykładnik liczby 2, zawsze wynik musi być 1 (ponieważ chcemy ze sobą wszystko na końcu pomnożyć, a mnożenie przez 1 nie zmienia wartości liczby). Używając jeszcze raz (żeby wszystkie liczby były traktowane w ten sam sposób) wzoru $f(p_1^0)+\ f(p_1^1)+f(p_1^2)+\cdots+f\left(p_1^{w_1}\right)$, wiemy, że $f(1)=1$, więc wszystkie pozostałe czynniki muszą być równe 0. Czyli $f(2^k)=0$, dla uproszczenia możemy pod tę grupę zakwalifikować wszystkie liczby parzyste, czyli $f(2k)=0$.

Spróbujmy połączyć ze sobą powyższe obliczenia, aby uzyskać wzór na ilość punktów kratowych:

$$\left(f(p_1^0)+\ f(p_1^1)+\cdots+f\left(p_1^{w_1}\right)\right)\left(f(p_2^0)+\ f(p_2^1)+\cdots+f\left(p_2^{w_2}\right)\right)\left(f(p_3^0)+\ f(p_3^1)+\cdots+f\left(p_3^{w_3}\right)\right)\ldots$$

Należy zauważyć, że zdefiniowana powyżej funkcja f jest multiplikatywna, czyli $f(a)*f(b)=f(a*b)$. Otóż znamy już wzór funkcji f, jest to:

$$f(x)=\begin{cases} 1\ dla\ x=4k+1 \\ 0\ dla\ x=2k \\ -1\ dla\ x=4k+3 \end{cases}$$

Wystarczy teraz, że pomnożymy ze sobą każdy możliwy przypadek i udowodnimy, że jest to funkcja multiplikatywna. Przypadków tych jest 6:

1. $f(4k_1+1)*f(4k_2+1)=f\big((4k_1+1)(4k_2+1)\big)=f(16k_1k_2+4k_1+4k_2+1)=f(4(4k_1k_2+k_1k_2)+1)=f(4m+1)=\mathbf{1}$
   $f(4k_1+1)*f(4k_2+1)=1*1=\mathbf{1}$
2. $f(4k_1+1)*f(2k_2)=f\big((4k_1+1)(2k_2)\big)=f(8k_1k_2+2k_2)=f\big(2(4k_1k_2+k_2)\big)=f(2m)=\mathbf{0}$
   $f(4k_1+1)*f(2k_2)=1*0=\mathbf{0}$

3. $f(4k_1+1)*f(4k_2+3)=f\big((4k_1+1)(4k_2+3)\big)=f(16k_1k_2+12k_1+4k_2+3)=f(4(4k_1k_2+3k_1+k_2)+3)=f(4m+3)=\mathbf{-1}$
   $f(4k_1+1)*f(4k_2+3)=1*(-1)=\mathbf{-1}$
4. $f(2k_1)*f(2k_2)=f\big((2k_1)(2k_2)\big)=f(4k_1k_2)=f(2(2k_1k_2)=f(2m)=\mathbf{0}$
   $f(2k_1)*f(2k_2)=0*0=\mathbf{0}$
5. $f(2k_1)*f(4k_2+3)=f\big((2k_1)(4k_2+3)\big)=f(8k_1k_2+6k_1)=f\big(2(4k_1k_2+3k_2)\big)=f(2m)=\mathbf{0}$
   $f(2k_1)*f(4k+3)=0*(-1)=\mathbf{0}$
6. $f(4k_1+3)*f(4k_2+3)=f\big((4k_1+3)(4k_2+3)\big)=f(16k_1k_2+12k_1+12k_2+9)=f(4(4k_1k_2+3k_1+3k_2+2)+1)=f(4m+1)=\mathbf{1}$
   $f(4k_1+3)*f(4k_2+3)=(-1)*(-1)=\mathbf{1}$

Wiedząc, że funkcja ta jest multiplikatywna możemy rozłożyć nasz wzór na ilość punktów kratowych:

$$\Big(f(p_1^0)+\ f(p_1^1)+\cdots+f\big(p_1^{w_1}\big)\Big)\Big(f(p_2^0)+\ f(p_2^1)+\cdots+f\big(p_2^{w_2}\big)\Big)\Big(f(p_3^0)+\ f(p_3^1)+\cdots+f\big(p_3^{w_3}\big)\Big)\ldots$$

Otóż możemy wymnożyć ze sobą wszystkie nawiasy. Może się wydawać, ze jest to niepotrzebne komplikowanie wzoru, jednak takie wymnożenie jest bardzo przydatne. Zauważmy, że mnożymy ze sobą na wszystkie sposoby potęgi czynników pierwszych liczby n. Oznacza to, że tworzymy dzielniki liczby n. Czyli powstaje nam nowy wzór postaci:

$$f(d_1)+f(d_2)+f(d_3)+f(d_4)+\cdots$$

Gdzie d to dzielniki liczby n. Zauważmy, że przy okazji zniknął nam problem pierwszy. Nie musimy już dzielić n na czynniki pierwsze.

Wróćmy teraz więc do uproszczonego wzory Picka z początku rozdziału $P=W+B$. Suma punktów kratowych na okręgach o promieniach R jest równa polu koła o promieniu r. Oznaczmy funkcję dającą ilość punktów kratowych na danym okręgu jako $g(n)$, gdzie n to R^2. Czyli wzór na pole koła ma postać

$$P=\sum_{n=1}^{r^2} g(n)$$

Pojawia się jednak kolejny kłopot. Otóż skąd mamy wiedzieć jakie są czynniki pierwsze n. Spróbujmy rozpisać $g(n)$ dla kilku pierwszych n.

$$g(1)=f(1)$$
$$g(2)=f(1)+f(2)$$
$$g(3)=f(1)+f(3)$$
$$g(4)=f(1)+f(2)+f(4)$$

$$g(5) = f(1) + f(5)$$
$$g(6) = f(1) + f(2) + f(3) + f(6)$$
$$g(7) = f(1) + f(7)$$
$$g(8) = f(1) + f(2) + f(4) + f(8)$$
$$g(9) = f(1) + f(3) + f(9)$$
$$g(10) = f(1) + f(2) + f(5) + f(10)$$
$$g(11) = f(1) + f(11)$$

Może się wydawać, że tylko skomplikowaliśmy sprawę. Zamiast rozkładu liczb pierwszych musimy teraz znać rozkład dzielników liczb! Okazuje się jednak, że jest to niepotrzebne. Spróbujmy ułożyć powyższe równania w kolumny:

$$\begin{aligned}
g(1) &= f(1) \\
g(2) &= f(1) + f(2) \\
g(3) &= f(1) + f(3) \\
g(4) &= f(1) + f(2) + f(4) \\
g(5) &= f(1) + f(5) \\
g(6) &= f(1) + f(2) + f(3) + f(6) \\
g(7) &= f(1) + f(7) \\
g(8) &= f(1) + f(2) + f(4) + f(8) \\
g(9) &= f(1) + f(3) + f(9) \\
g(10) &= f(1) + f(2) + f(5) + f(10) \\
g(11) &= f(1) + f(11)
\end{aligned}$$

Możemy teraz zauważyć, że każda liczba ma jako dzielnik 1, co druga ma dzielnik 2, co trzecia 3, co czwarta 4, itd. Oznacza to, że r^2 okręgów będzie miało we wzorze $f(1)$, średnio $\frac{r^2}{2}$ okręgów będzie miało $f(2)$, średnio $\frac{r^2}{3}$ okręgów będzie miało $f(3)$, itd. Musimy jeszcze ten wzór pomnożyć przez 4 (z punktu 4). Oznacza to, że wzór na ilość punktów kratowych w kole o promieniu r (suma ilości punktów na okręgach), a zarazem na pole, wygląda następująco:

$$P = 4\left(\frac{r^2}{1} * f(1) + \frac{r^2}{2} * f(2) + \frac{r^2}{3} * f(3) + \frac{r^2}{4} * f(4) + \frac{r^2}{5} * f(5) + \cdots\right)$$
$$= 4r^2\left(\frac{f(1)}{1} + \frac{f(2)}{2} + \frac{f(3)}{3} + \frac{f(4)}{4} + \frac{f(5)}{5} + \cdots\right)$$

Wiemy także, że innym wzorem na pole koła jest $P = \pi r^2$, więc:

$$\pi r^2 = 4r^2\left(f(1) + \frac{f(2)}{2} + \frac{f(3)}{3} + \frac{f(4)}{4} + \frac{f(5)}{5} + \cdots\right)$$

$$\pi = 4\left(f(1) + \frac{f(2)}{2} + \frac{f(3)}{3} + \frac{f(4)}{4} + \frac{f(5)}{5} + \cdots\right)$$

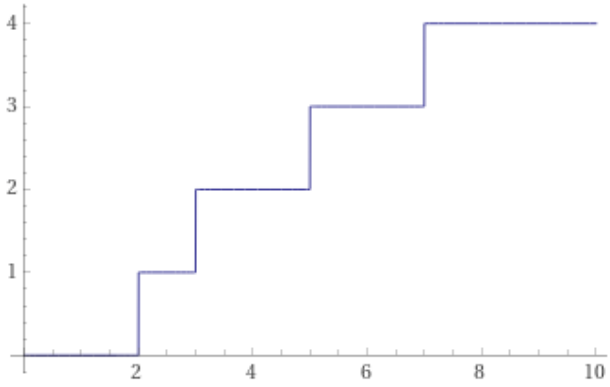
$$\frac{\pi}{4} = \left(f(1) + \frac{f(2)}{2} + \frac{f(3)}{3} + \frac{f(4)}{4} + \frac{f(5)}{5} + \cdots\right) = \left(1 + \frac{0}{2} + \frac{-1}{3} + \frac{0}{4} + \frac{1}{5} + \cdots\right)$$

$$\frac{\pi}{4} = 1 - \frac{1}{3} + \frac{1}{5} - \frac{1}{7} + \frac{1}{9} - \cdots$$

Czyli wyznaczyliśmy wartość liczby π, a do obliczeń wykorzystaliśmy liczby pierwsze.

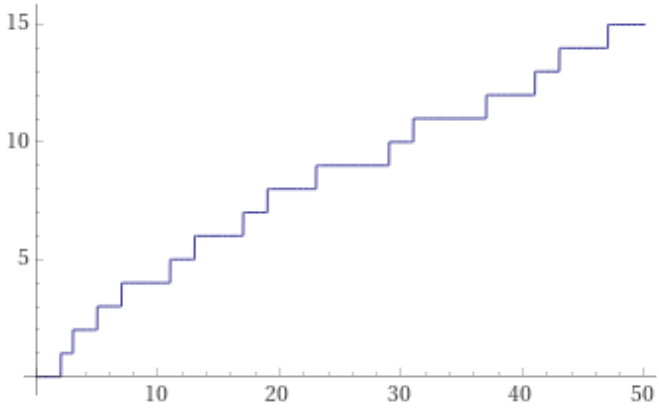
8. Gęstość liczb pierwszych

W rozdziale 4 poruszyliśmy temat rozmieszczenia liczb pierwszych na osi liczbowej. Tutaj rozwiniemy ten temat. Nie będziemy zajmować się pojedynczymi liczbami pierwszymi. Spróbujemy odpowiedzieć na pytanie: jak gęsto są one rozłożone? Matematycy wprowadzili funkcję $\pi(n)$ dającą w wyniku ilość liczb pierwszych w przedziale od 1...n ($n \in \mathbb{N}$). I tutaj taka mała dygresja: funkcja π nie ma nic wspólnego ze stałą π = 3,14.. (której wartość obliczaliśmy w poprzednim rozdziale), co więcej funkcja $\pi(n)$ (a właściwie jej nazwa) została wprowadzona wcześniej niż oznaczenie stałej. Zobaczmy jak wygląda wykres funkcji $\pi(n)$ gdy $n \in \{1, 2, \dots, 10\}$.



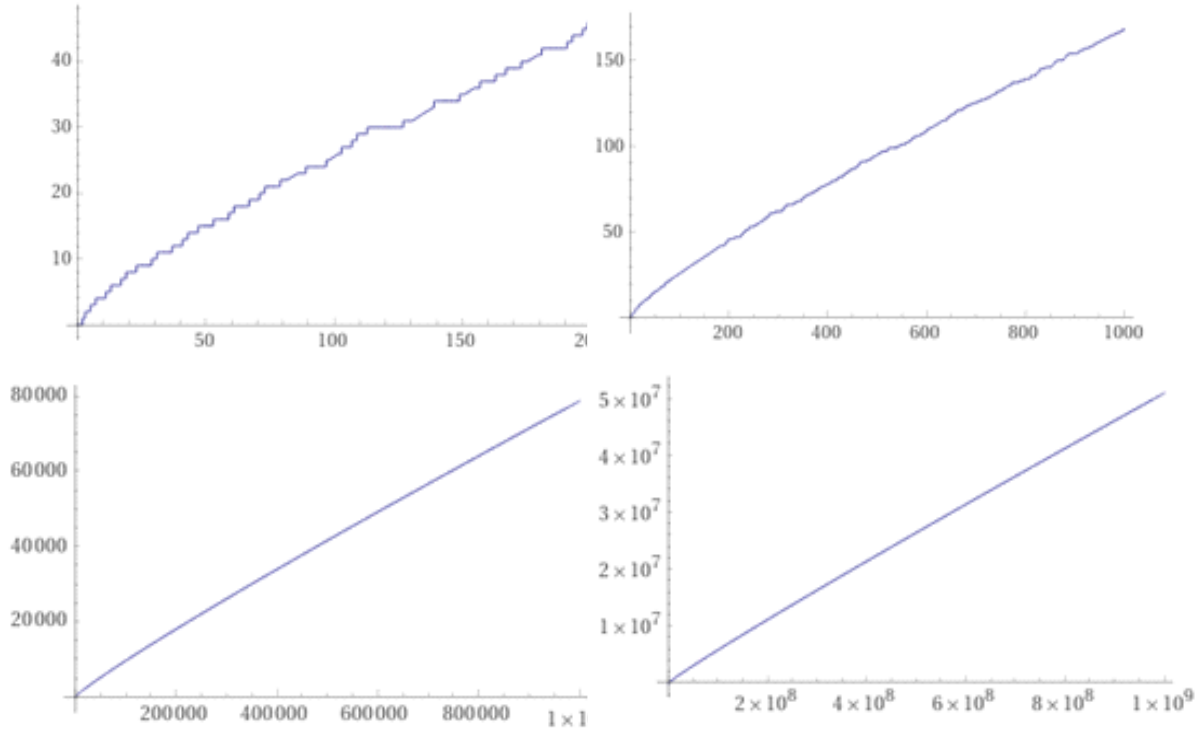
Rysunek 19 Wykres ilości liczb pierwszych w przedziale od 1 do 10

Na wykresie możemy zauważyć linie pionowe, co oczywiście kłóci się z definicją funkcji. Jednak program, który został użyty – *Wolfram Alpha* – w ten sposób wizualizuje tę funkcję. W przedziale [0,2) funkcja π przyjmuje oczywiście wartość 0 (brak liczb pierwszych). W przedziale [2,3) jej wartość wynosi 1. A potem schodkowo rośnie w górę. Właśnie wykres tej funkcji przypomina schodki i to dość nieregularne - nieregularne jest przecież rozmieszczenie liczb pierwszych. Oddalmy się trochę od wykresu (tzn. zobaczmy większy jego wycinek).



Rysunek 20 Wykres ilości liczb pierwszych w przedziale od 1 do 50

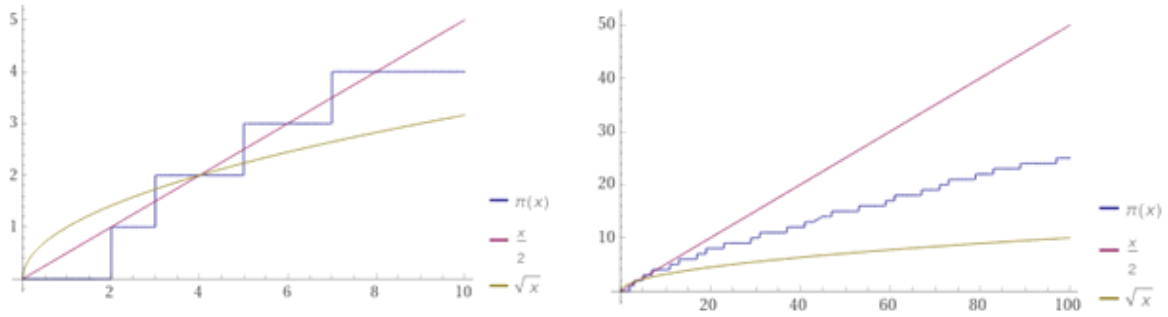
Tutaj również widzimy „schodki”, co się stanie jeżeli będziemy się jeszcze bardziej „oddalać”?



Rysunek 21Wykres ilości liczb pierwszych w kolejnych przedziałach

Przy skali obejmującej pierwszy milion liczb naturalnych już nie widać żadnych schodków, oczywiście przy miliardzie wykres będzie jeszcze „gładszy".

Spróbujmy nałożyć na siebie wykresy $\pi(n)$, $\frac{n}{2}$, oraz $\sqrt{n}$ aby oszacować w przybliżeniu ilość liczb pierwszych w pewnych przedziałach



Rysunek 22 Wykresy ilości liczb pierwszych oraz funkcji $\frac{n}{2}$ i $\sqrt{n}$

Widać, że dla początkowych liczb naturalnych wykresy te przecinają się, ale później tendencja jest stała: Liczb pierwszych w przedziale $[1, n]$ jest więcej niż $\sqrt{n}$, ale mniej niż $\frac{n}{2}$.

Wprowadźmy teraz pojęcie gęstości zbioru liczb pierwszych w przedziale $[1, n]$ jako stosunek ilości liczb pierwszych do ilości liczb naturalnych w tym przedziale $\frac{\pi(n)}{n}$.

Zobaczmy jakie wartości przyjmuje tak zdefiniowana funkcja.

n	$\pi(n)$	$\frac{\pi(n)}{n}$
10	4	0,4000
100	25	0,2500
1 000	168	0,1680
10 000	1 229	0,1229
100 000	9 592	0,0959
1 000 000	78 498	0,0785
10 000 000	664 579	0,0665
100 000 000	5 761 455	0,0576
1 000 000 000	50 847 534	0,0508
10 000 000 000	455 052 511	0,0455
100 000 000 000	4 118 054 813	0,0412
1 000 000 000 000	37 607 912 018	0,0376
10 000 000 000 000	346 065 536 839	0,0346
100 000 000 000 000	3 204 941 750 802	0,0320
1 000 000 000 000 000	29 844 570 422 669	0,0298
10 000 000 000 000 000	279 238 341 033 925	0,0279
100 000 000 000 000 000	2 623 557 157 654 230	0,0262
1 000 000 000 000 000 000	24 739 954 287 740 800	0,0247
10 000 000 000 000 000 000	234 057 667 276 344 000	0,0234
100 000 000 000 000 000 000	2 220 819 602 560 910 000	0,0222
1 000 000 000 000 000 000 000	21 127 269 486 018 700 000	0,0211

Rysunek 23 Wartości funkcji $\pi(n), \frac{\pi(n)}{n}$

Zwróćmy uwagę, że wraz ze zwiększaniem się zakresu *n* gęstość zbioru liczb pierwszych maleje. Oznacza to, że im większe jest *n* tym „rzadziej" rozłożone są liczby pierwsze. Nie znaczy to oczywiście, że w okolicach miliarda nie ma liczb bliźniaczych, różniących się od siebie tylko o 2. Mówimy tu o podejściu statystycznym dla całego przedziału. Spróbujmy do tej tabeli dodać jeszcze dwie kolumny: jedną, mówiącą o tym jaką część wszystkich liczb naturalnych nie większych od danej liczby stanowią liczby pierwsze, a drugą mówiącą jak zmienia się poprzednia wartość:

n	$\pi(n)$	$\frac{\pi(n)}{n}$	$\frac{n}{\pi(n)}$	Δ
10	4	0,4000	2,500	2,500
100	25	0,2500	4,000	1,500
1 000	168	0,1680	5,952	1,952
10 000	1 229	0,1229	8,137	2,184
100 000	9 592	0,0959	10,425	2,289
1 000 000	78 498	0,0785	12,739	2,314
10 000 000	664 579	0,0665	15,047	2,308
100 000 000	5 761 455	0,0576	17,357	2,310
1 000 000 000	50 847 534	0,0508	19,667	2,310
10 000 000 000	455 052 511	0,0455	21,975	2,309
100 000 000 000	4 118 054 813	0,0412	24,283	2,308
1 000 000 000 000	37 607 912 018	0,0376	26,590	2,307
10 000 000 000 000	346 065 536 839	0,0346	28,896	2,306
100 000 000 000 000	3 204 941 750 802	0,0320	31,202	2,306
1 000 000 000 000 000	29 844 570 422 669	0,0298	33,507	2,305
10 000 000 000 000 000	279 238 341 033 925	0,0279	35,812	2,305
100 000 000 000 000 000	2 623 557 157 654 230	0,0262	38,116	2,304
1 000 000 000 000 000 000	24 739 954 287 740 800	0,0247	40,420	2,304
10 000 000 000 000 000 000	234 057 667 276 344 000	0,0234	42,725	2,304
100 000 000 000 000 000 000	2 220 819 602 560 910 000	0,0222	45,028	2,304
1 000 000 000 000 000 000 000	21 127 269 486 018 700 000	0,0211	47,332	2,304

Rysunek 24 Wartości funkcji $\pi(n), \frac{\pi(n)}{n}, \frac{n}{\pi(n)}$

Tutaj widzimy jak na dłoni, że wartość w ostatniej kolumnie wzrasta mniej więcej stale (ok. 2,3). W każdym wierszu zakres badanych danych wzrasta nam dziesięciokrotnie. Czy coś nam to mówi? Oczywiście mamy zależność logarytmiczną i to o podstawie ok. 2,3. I co? Przecież to podstawa logarytmu naturalnego (przybliżona oczywiście)!

A więc można zapisać $\pi(n) \approx \frac{n}{\ln(n)}$. Zobaczmy jak wyglądają powyższe zależności w tabeli:

n	$\pi(n)$	$\frac{n}{\ln(n)}$	błąd bezwzględny
10	4	4,34	0,34
100	25	21,71	3,29
1 000	168	144,76	23,24
10 000	1 229	1 085,74	143,26
100 000	9 592	8 685,89	906,11
1 000 000	78 498	72 382,41	6 115,59
10 000 000	664 579	620 420,69	44 158,31
100 000 000	5 761 455	5 428 681,02	332 773,98
1 000 000 000	50 847 534	48 254 942,43	2 592 591,57
10 000 000 000	455 052 511	434 294 481,90	20 758 029,10
100 000 000 000	4 118 054 813	3 948 131 653,67	169 923 159,33
1 000 000 000 000	37 607 912 018	36 191 206 825,27	1 416 705 192,73
10 000 000 000 000	346 065 536 839	334 072 678 387,12	11 992 858 451,88
100 000 000 000 000	3 204 941 750 802	3 102 103 442 166,08	102 838 308 635,92
1 000 000 000 000 000	29 844 570 422 669	28 952 965 460 216,80	891 604 962 452,21
10 000 000 000 000 000	279 238 341 033 925	271 434 051 189 532,00	7 804 289 844 392,62
100 000 000 000 000 000	2 623 557 157 654 230	2 554 673 422 960 300,00	68 883 734 693 925,00
1 000 000 000 000 000 000	24 739 954 287 740 800	24 127 471 216 847 300,00	612 483 070 893 476,00
10 000 000 000 000 000 000	234 057 667 276 344 000	228 576 043 106 975 000,00	5 481 624 169 369 380,00
100 000 000 000 000 000 000	2 220 819 602 560 910 000	2 171 472 409 516 260 000,00	49 347 193 044 651 000,00
1 000 000 000 000 000 000 000	21 127 269 486 018 700 000	20 680 689 614 440 600 000,00	446 579 871 578 137 000,00

Rysunek 25 Wartości funkcji $\pi(n), \frac{n}{ln\,(n)}$ oraz błąd bezwzględny

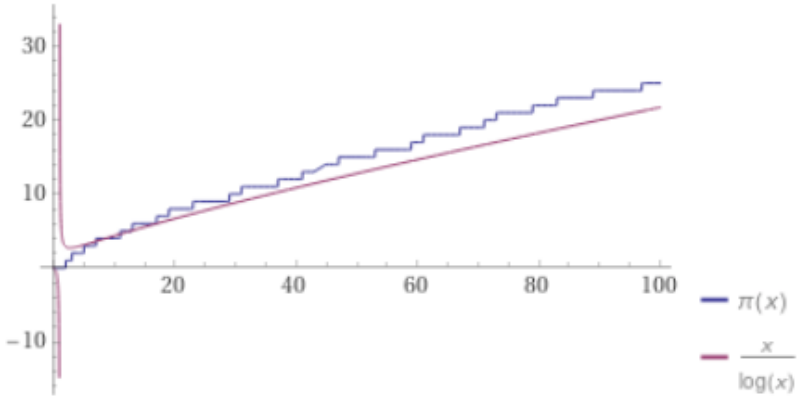
Wydaje się, że chyba jednak gdzieś wkradł się błąd. Przecież błędy bezwzględne wraz ze wzrostem *n* rosną do niebotycznych rozmiarów. Ale, właśnie jest jedno „ale". W ostatniej kolumnie przedstawiono błędy bezwzględne, to znaczy o jaką wartość różnią się $\pi(n)$ oraz $\frac{n}{\ln(n)}$. Dodajmy do tabeli jeszcze dwie następne kolumny zawierające: błąd względny oraz dokładność przybliżenia, aby zobaczyć, że jednak wszystko jest w porządku.

n	$\pi(n)$	$\frac{n}{\ln(n)}$	błąd bezwzględny	błąd względny	dokładność
1,00E+01	4	4,34	0,34	3,4294%	96,5706%
1,00E+02	25	21,71	3,29	3,2853%	96,7147%
1,00E+03	168	144,76	23,24	2,3235%	97,6765%
1,00E+04	1 229	1 085,74	143,26	1,4326%	98,5674%
1,00E+05	9 592	8 685,89	906,11	0,9061%	99,0939%
1,00E+06	78 498	72 382,41	6 115,59	0,6116%	99,3884%
1,00E+07	664 579	620 420,69	44 158,31	0,4416%	99,5584%
1,00E+08	5 761 455	5 428 681,02	332 773,98	0,3328%	99,6672%
1,00E+09	50 847 534	48 254 942,43	2 592 591,57	0,2593%	99,7407%
1,00E+10	455 052 511	434 294 481,90	20 758 029,10	0,2076%	99,7924%
1,00E+11	4 118 054 813	3 948 131 653,67	169 923 159,33	0,1699%	99,8301%
1,00E+12	37 607 912 018	36 191 206 825,27	1 416 705 192,73	0,1417%	99,8583%
1,00E+13	346 065 536 839	334 072 678 387,12	11 992 858 451,88	0,1199%	99,8801%
1,00E+14	3 204 941 750 802	3 102 103 442 166,08	102 838 308 635,92	0,1028%	99,8972%
1,00E+15	29 844 570 422 669	28 952 965 460 216,80	891 604 962 452,21	0,0892%	99,9108%
1,00E+16	279 238 341 033 925	271 434 051 189 532,00	7 804 289 844 392,62	0,0780%	99,9220%
1,00E+17	2 623 557 157 654 230	2 554 673 422 960 300,00	68 883 734 693 925,00	0,0689%	99,9311%
1,00E+18	24 739 954 287 740 800	24 127 471 216 847 300,00	612 483 070 893 476,00	0,0612%	99,9388%
1,00E+19	234 057 667 276 344 000	228 576 043 106 975 000,00	5 481 624 169 369 380,00	0,0548%	99,9452%
1,00E+20	2 220 819 602 560 910 000	2 171 472 409 516 260 000,00	49 347 193 044 651 000,00	0,0493%	99,9507%
1,00E+21	21 127 269 486 018 700 000	20 680 689 614 440 600 000,00	446 579 871 578 137 000,00	0,0447%	99,9553%

Rysunek 26 Wartości funkcji $\pi(n), \frac{n}{ln\,(n)}$ błąd bezwzględny oraz błąd względny i dokładność

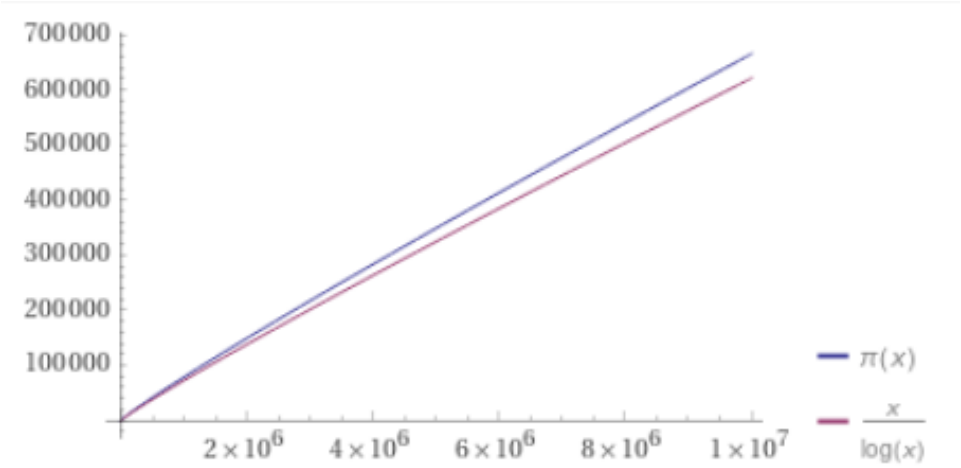
Pierwsza kolumna została zmodyfikowana, aby tabela była bardziej czytelna. Widzimy teraz, że błąd względny wraz ze wzrostem n maleje, a dokładność rośnie prawie do 100%.

Zobaczmy powyższe zależności jeszcze na wykresie, w zakresie n: 1-100:



Rysunek 27 Wykresy funkcji $\pi(n), \frac{n}{ln\,(n)}$ w przedziale 1-100

Widać „schodkową naturę” funkcji $\pi(n)$ i zbliżający się wykres $\frac{n}{\ln(n)}$. Natomiast jeżeli „odejdziemy” trochę dalej – czyli zwiększymy zakres *n* zobaczymy, ze wykresy obu funkcji są bardzo zbliżone.



Rysunek 28 Wykres funkcji $\pi(n), \frac{n}{ln(n)}$ w przedziale 1-10 000 000

Wykres 28 pokazuje również, że funkcja $\frac{n}{\ln(n)}$ dość dobrze przybliża ilość liczb pierwszych nieprzekraczających n.

Nie jest to oczywiście dowód matematyczny. Nieraz okazywało się, zwłaszcza w temacie liczb pierwszych, że twierdzenia „padały” dopiero przy naprawdę bardzo ogromnych liczbach.

Posługując się takimi (lub bardzo podobnymi obliczeniami) młody Gauss – jeden z najwybitniejszych matematyków wszechczasów, sformułował podstawowe twierdzenie o rozmieszczeniu liczb pierwszych wśród liczb naturalnych.

Twierdzenie 8.1 o liczbach pierwszych

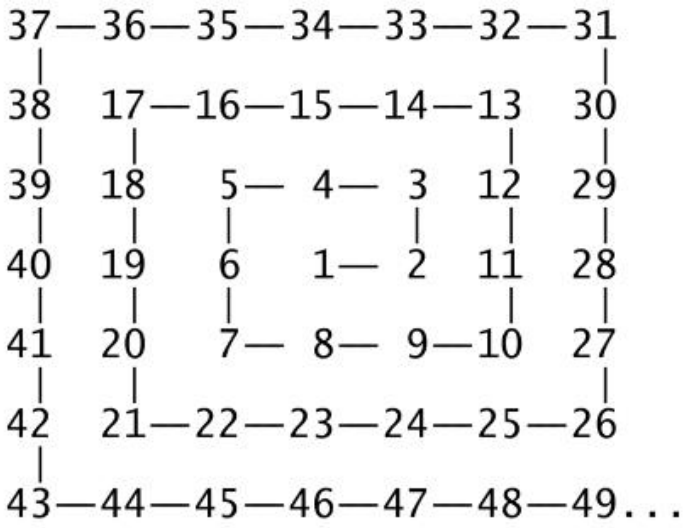
Niech $\pi(n)$ będzie ilością liczb pierwszych nieprzekraczających n, wtedy

$$\lim_{n\to\infty} \frac{\pi(n)}{\frac{n}{\ln(n)}} = 1$$

Czyli jeżeli zwiększamy n, to nasze przybliżenie jest coraz bardziej dokładne.

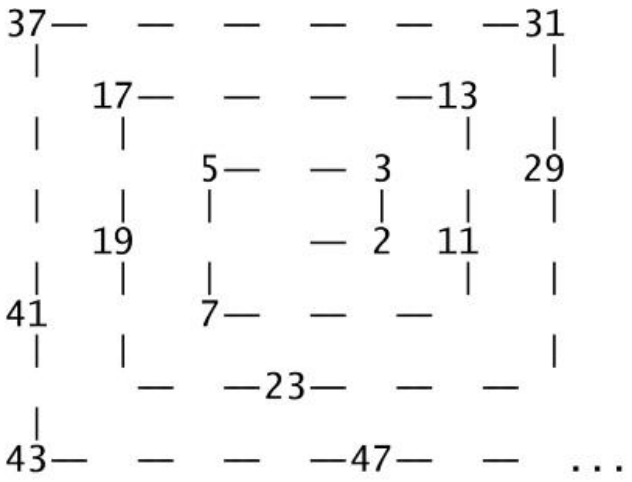
9. Spirala Ulama

Nie można napisać pracy na temat liczb pierwszych pomijając osiągnięcia polskich matematyków w tym zakresie. Najsłynniejszą postacią wśród naukowców zajmujących się liczbami pierwszymi jest Stanisław Ulam. Legenda głosi, że podczas pewnego długiego wykładu w 1963r. Ulam z nudów zaczął rysować spiralę kwadratową poczynając od 1:



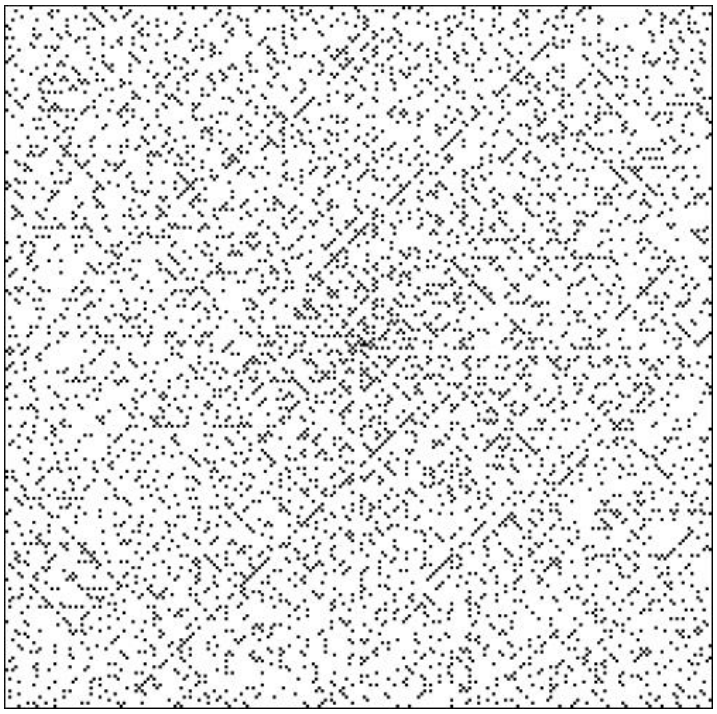
Rysunek 29 Liczby wpisane w spiral e kwadratową

Następnie wykreślił wszystkie liczby poza liczbami pierwszymi.



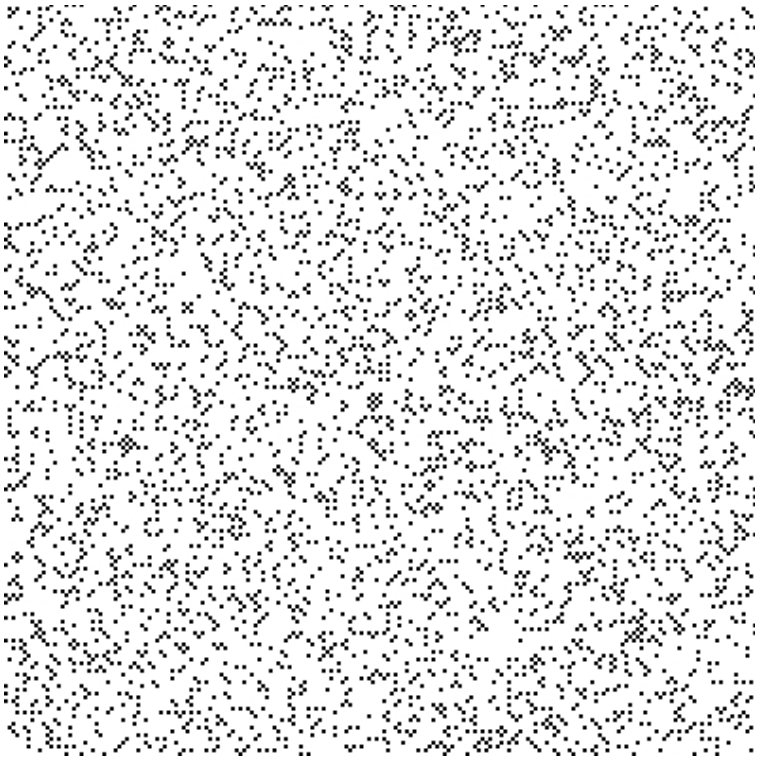
Rysunek 30 Spirala Ulama

Już na spirali 7x7 pokazanej powyżej widzimy zalążki pewnych, wydawałoby się, prawidłowości w rozmieszczeniu liczb pierwszych. Na niektórych przekątnych jest więcej liczb pierwszych niż na innych. Lepiej widać te prawidłowości, gdy zwiększymy zakres spirali np. 200x200 i zamiast liczb będziemy wstawiać w miejsce liczb pierwszych ciemne kropki. Białe (puste) miejsca reprezentują tutaj liczby złożone.



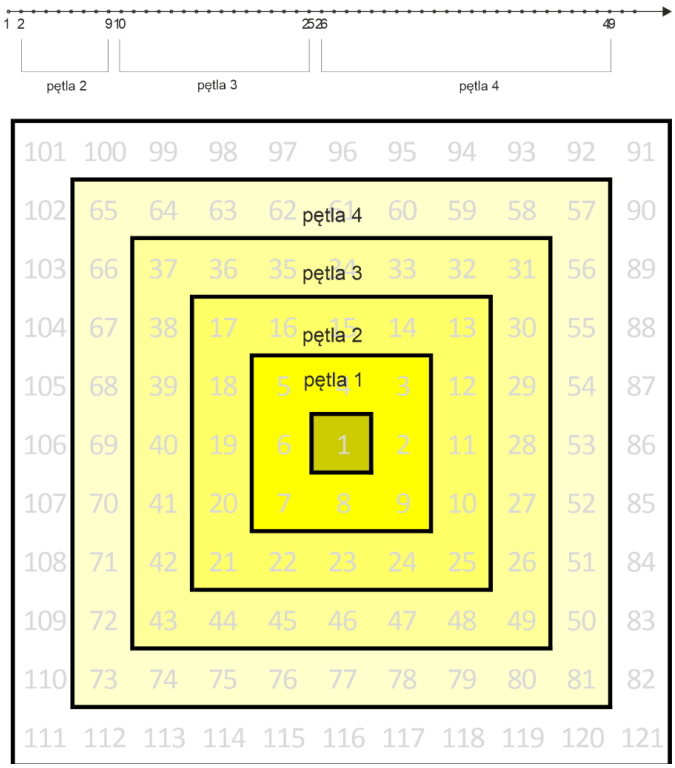
Rysunek 31 Spirala Ulama 200x200

Jeżeli przyjrzymy się powyższemu rysunkowi zobaczymy, że liczby pierwsze pojawiają się na nim w sposób, który wydaje się nam „jakoś” uporządkowany. Poniżej przedstawiono dla porównania losowe rozmieszczenie liczb nieparzystych w podobnym zakresie 200x200:



Rysunek 32 Losowy rozkład liczb w zakresie 200x200

Zastanówmy się teraz jak można wyjaśnić uporządkowany (częściowo) świat liczb pierwszych w spirali Ulama. Otóż konstrukcja spirali polega na „zawijaniu” osi liczbowej. Z jednowymiarowej osi X przechodzimy na dwuwymiarową płaszczyznę XY.



Rysunek 33 Sposób przełożenia osi liczbowej na spiralę

W spirali Ulama, jak wcześniej pokazaliśmy, liczby pierwsze pojawiają się zadziwiająco często na pewnych prostych. Najczęściej są to proste diagonalne (skośne), ale czasami mogą nas też interesować proste poziome lub pionowe (na których nie występują w ogóle liczby pierwsze, bądź występują niezmiernie rzadko). Jak wyznaczyć równanie takiej prostej, tzn. równanie funkcji, której wartości należałyby do w/w prostej? Spirala Ulama to przestrzeń dwuwymiarowa, więc narzuca się pomysł z funkcją kwadratową. Ale jak go skonstruować?

Na początek mała dygresja: Każda funkcja kwadratowa $f(n) = an^2 + bn + c$ da się przedstawić w formie rekurencyjnej $f(n+1) = f(n) + dn + e$ o ile znamy „wartość początkową” takiej funkcji $f(1)$.

$$f(n+1) = f(n) + dn + e \equiv an^2 + bn + c$$

$$a\cdot(n+1)^2 + b\cdot(n+1) + c \equiv an^2 + bn + c + dn + e$$

$$an^2 + 2an + a + bn + b + c \equiv an^2 + bn + c + dn + e$$

$2an + a + b \equiv dn + e \Rightarrow 2a = d$ oraz $a + b = e \Rightarrow a = \frac{d}{2}$ oraz $b = e - a = e - \frac{d}{2}$

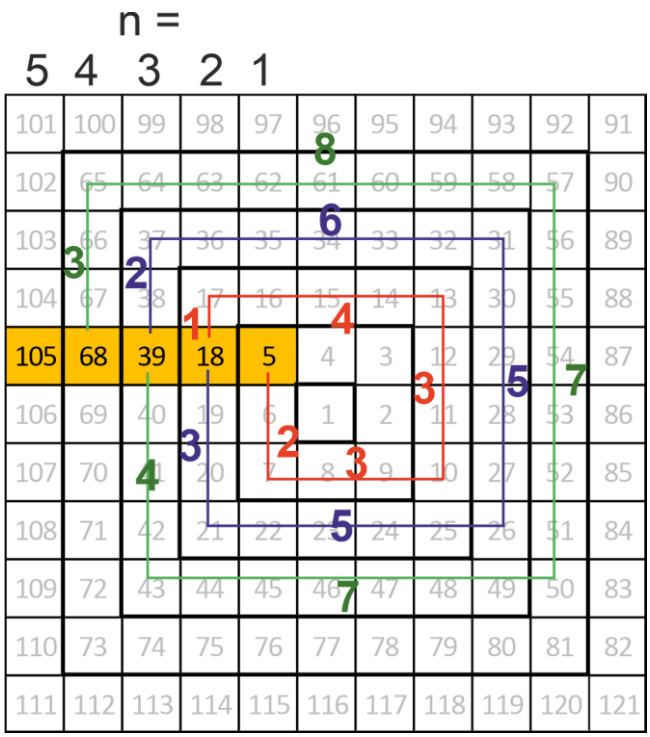
Jeżeli znamy $f(1)$ to

$f(1) = a\cdot 1^2 + b\cdot 1 + c = a + b + c \Rightarrow c = f(1) - a - b = f(1) - a - (e - a) = f(1) - e$

A więc przy założeniu, że znamy wzór rekurencyjny $f(n+1) = f(n) + dn + e$ znając ponadto wartość $f(1)$ możemy wyprowadzić wzór na funkcję kwadratową postaci:

$f(n) = an^2 + bn + c$ gdzie: $a = \frac{d}{2},\ b = e - \frac{d}{2}, c = f(1) - e$ (1)

Spróbujmy wykorzystać rekurencyjną metodę tworzenia funkcji kwadratowej do sporządzenia wzoru. Popatrzmy na rysunek 34



Rysunek 34 Graficzna reprezentacja wyliczeń funkcji kwadratowej dla liczb 5, 18, 39, 68, 105

$f(1) = 5$

$f\ (2) = f(1) + 2 + 3 + 3 + 4 + 1$ kolor czerwony

$f(3) = f(2) + 3 + 5 + 5 + 6 + 2$ kolor niebieski

$f(4) = f(3) + 4 + 7 + 7 + 8 + 3$ kolor zielony

…

$$f(n+1) = f(n) + (n+1) + (2n+1) + (2n+1) + (2n+2) + n$$

$$f(n+1) = f(n) + 8n + 5$$

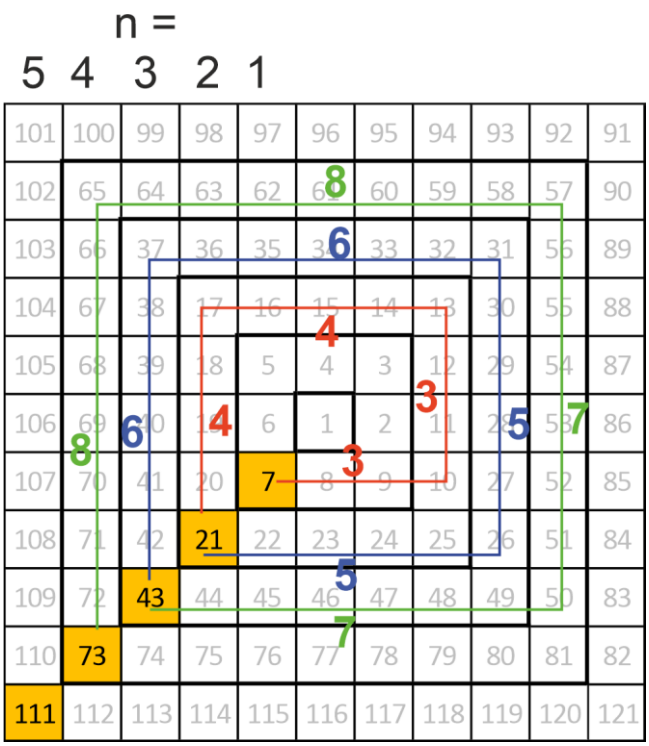
Podstawiając do wzoru (1) mamy $d = 8, e = 5, f(1) = 5$ czyli:

$$f(n) = 4n^2 + (5-4)n + (5-5)$$

$$\boldsymbol{f(n) = 4n^2 + n}$$

Czyli linia pozioma zaznaczona na rysunku 34 odpowiada przedstawionej powyżej funkcji kwadratowej.

Zobaczmy jeszcze linię skośną przedstawioną na rysunku 35:



Rysunek 35 Graficzna reprezentacja wyliczeń funkcji kwadratowej dla liczb 7, 21, 43, 73, 111

Podobnie jak poprzednio korzystamy z rekurencji

$$f(1) = 7$$

$f\ (2) = f(1) + 3 + 3 + 4 + 4$ kolor czerwony

$f(3) = \ f(2) + 5 + 5 + 6 + 6$ kolor niebieski

$f(4) = f(3) + 7 + 7 + 8 + 8$ kolor zielony

…

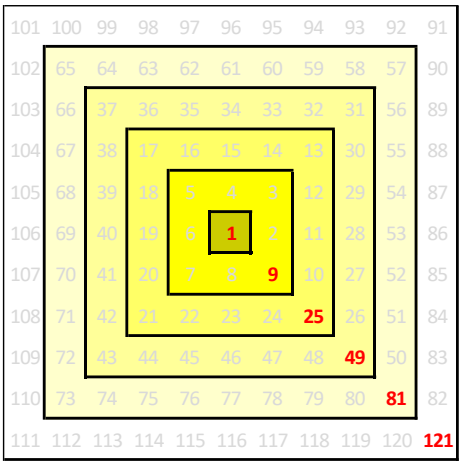
$$f(n+1) = f(n) + (2n+1) + (2n+1) + (2n+2) + (2n+2)$$

$$f(n+1) = f(n) + 8n + 6$$

Podstawiając do wzoru (1) mamy $d\ = \ 8, e = 6, f(1) = 7$ czyli:

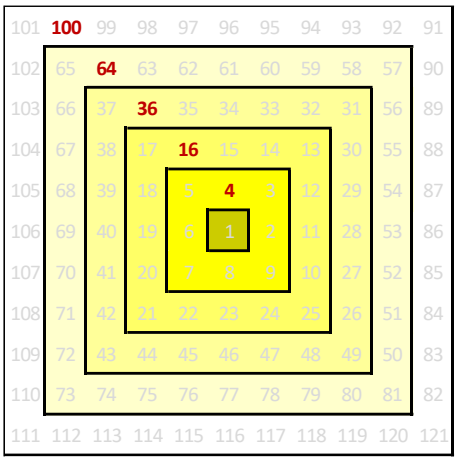
$$f(n) = 4n^2 + (6-4)n + (7-6)$$

$$\boldsymbol{f(n) = 4n^2 + 2n + 1}$$



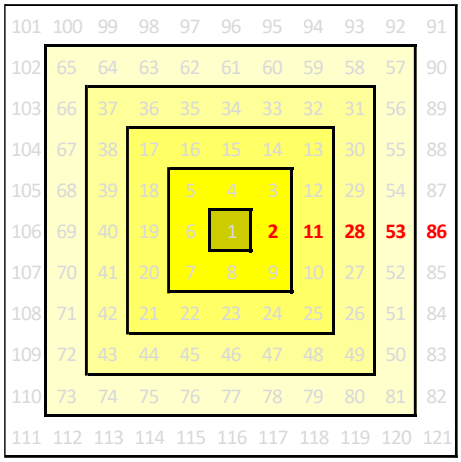
Rysunek 37 Funkcja kwadratowa dla liczb 1, 9, 25 , 49, 81, 121

$$f(n) = 4n^2 - 4n + 1$$



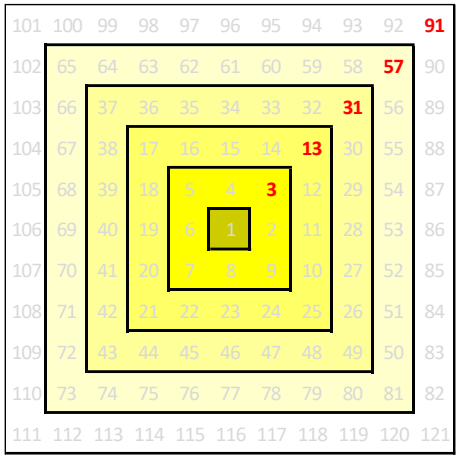
Rysunek 36 Funkcja kwadratowa dla liczb 4, 16, 36, 64, 100

$$f(n) = 4n^2$$



Rysunek 38 Funkcja kwadratowa dla liczb 2, 11, 28, 53, 86

$$f(n) = 4n^2 - 3n + 1$$



Rysunek 39 Funkcja kwadratowa dla liczb 3, 13, 31, 57, 91

$$f(n) = 4n^2 - 10n + 7$$

Wszystkie proste (pionowe, poziome i skośne) na spirali Ulama mają postać $f(n) = 4n^2 + bn + c$. Można udowodnić, że dla b parzystego proste są skośne, natomiast dla b nieparzystego linie są poziome, bądź pionowe. Nie wszystkie linie muszą zaczynać się od środka spirali, przy większych wartościach b i c linie zaczynają się dalej od środka. Co więcej nie wszystkie początkowe wartości funkcji leżą na prostej.

101	100	99	98	97	96	95	94	93	92	91
102	65	64	63	62	61	60	59	58	57	90
103	66	37	36	35	34	33	32	31	56	89
104	67	38	17	16	15	14	13	30	55	88
105	68	39	18	5	4	3	12	29	54	87
106	69	40	19	6	1	2	11	28	53	86
107	70	41	20	7	8	9	10	27	52	85
108	71	42	21	22	23	24	25	26	51	84
109	72	43	44	45	46	47	48	49	50	83
110	73	74	75	76	77	78	79	80	81	82
111	112	113	114	115	116	117	118	119	120	121

$$f(n) = 4n^2 + 5$$

Rysunek 40 Nie wszystkie wartości początkowe funkcji leżą na prostej

Dlaczego tyle czasu poświęciłem opisując równania prostych na spirali Ulama? Okazuje się, że niektóre funkcje kwadratowe generują bardzo dużo liczb pierwszych (*ang. prime-rich quadratic polynomials)*.

Jedną z takich funkcji kwadratowych jest funkcja $f(n) = n^2 + n + 41$, gdzie $n \in \mathbb{N}$, o której wspomniano już w rozdziale 3. Funkcja ta generuje liczby pierwsze „nienaturalnie" często.

0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19
41	43	47	53	61	71	83	97	113	131	151	173	197	223	251	281	313	347	383	421
20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39
461	503	547	593	641	691	743	797	853	911	971	1033	1097	1163	1231	1301	1373	1447	1523	1601
40	41	42	43	44	45	46	47	48	49	50	51	52	53	54	55	56	57	58	59
1681	1763	1847	1933	2021	2111	2203	2297	2393	2491	2591	2693	2797	2903	3011	3121	3233	3347	3463	3581
60	61	62	63	64	65	66	67	68	69	70	71	72	73	74	75	76	77	78	79
3701	3823	3947	4073	4201	4331	4463	4597	4733	4871	5011	5153	5297	5443	5591	5741	5893	6047	6203	6361
80	81	82	83	84	85	86	87	88	89	90	91	92	93	94	95	96	97	98	99
6521	6683	6847	7013	7181	7351	7523	7697	7873	8051	8231	8413	8597	8783	8971	9161	9353	9547	9743	9941

Rysunek 41 Liczby pierwsze generowane przez wielomian $n^2 + n + 41$ w zakresie 0-99

W pierwszej setce liczb naturalnych (i zera) tylko dla 14 wartości n funkcja daje wartości złożone. Dla pozostałych 86 są to liczby pierwsze. A więc 86% wartości funkcji kwadratowej $f(n) = n^2 + n + 41$ dla $0 \leq n \leq 99$ daje w wyniku liczbę pierwszą.

A co dzieje się gdy rozszerzymy zakres? Na rysunku 42 przedstawiono wyniki badanej funkcji dla n z zakresu od 0 do 999 (dla przejrzystości pokazujemy już w tej tabeli tylko liczbę n, a nie jak powyżej n i $f(n)$). I co się okazuje? Na 1 000 wyników 581 jest liczbami pierwszymi, co stanowi 58% całości. Wynik dużo mniejszy niż poprzednio. Ale gdy porównamy go z ilością liczb pierwszych mniejszych od 1 000 (czyli 168) jest ich dalej zaskakująco dużo.

0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19
20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39
40	41	42	43	44	45	46	47	48	49	50	51	52	53	54	55	56	57	58	59
60	61	62	63	64	65	66	67	68	69	70	71	72	73	74	75	76	77	78	79
80	81	82	83	84	85	86	87	88	89	90	91	92	93	94	95	96	97	98	99
100	101	102	103	104	105	106	107	108	109	110	111	112	113	114	115	116	117	118	119
120	121	122	123	124	125	126	127	128	129	130	131	132	133	134	135	136	137	138	139
140	141	142	143	144	145	146	147	148	149	150	151	152	153	154	155	156	157	158	159
160	161	162	163	164	165	166	167	168	169	170	171	172	173	174	175	176	177	178	179
180	181	182	183	184	185	186	187	188	189	190	191	192	193	194	195	196	197	198	199
200	201	202	203	204	205	206	207	208	209	210	211	212	213	214	215	216	217	218	219
220	221	222	223	224	225	226	227	228	229	230	231	232	233	234	235	236	237	238	239
240	241	242	243	244	245	246	247	248	249	250	251	252	253	254	255	256	257	258	259
260	261	262	263	264	265	266	267	268	269	270	271	272	273	274	275	276	277	278	279
280	281	282	283	284	285	286	287	288	289	290	291	292	293	294	295	296	297	298	299
300	301	302	303	304	305	306	307	308	309	310	311	312	313	314	315	316	317	318	319
320	321	322	323	324	325	326	327	328	329	330	331	332	333	334	335	336	337	338	339
340	341	342	343	344	345	346	347	348	349	350	351	352	353	354	355	356	357	358	359
360	361	362	363	364	365	366	367	368	369	370	371	372	373	374	375	376	377	378	379
380	381	382	383	384	385	386	387	388	389	390	391	392	393	394	395	396	397	398	399
400	401	402	403	404	405	406	407	408	409	410	411	412	413	414	415	416	417	418	419
420	421	422	423	424	425	426	427	428	429	430	431	432	433	434	435	436	437	438	439
440	441	442	443	444	445	446	447	448	449	450	451	452	453	454	455	456	457	458	459
460	461	462	463	464	465	466	467	468	469	470	471	472	473	474	475	476	477	478	479
480	481	482	483	484	485	486	487	488	489	490	491	492	493	494	495	496	497	498	499
500	501	502	503	504	505	506	507	508	509	510	511	512	513	514	515	516	517	518	519
520	521	522	523	524	525	526	527	528	529	530	531	532	533	534	535	536	537	538	539
540	541	542	543	544	545	546	547	548	549	550	551	552	553	554	555	556	557	558	559
560	561	562	563	564	565	566	567	568	569	570	571	572	573	574	575	576	577	578	579
580	581	582	583	584	585	586	587	588	589	590	591	592	593	594	595	596	597	598	599
600	601	602	603	604	605	606	607	608	609	610	611	612	613	614	615	616	617	618	619
620	621	622	623	624	625	626	627	628	629	630	631	632	633	634	635	636	637	638	639
640	641	642	643	644	645	646	647	648	649	650	651	652	653	654	655	656	657	658	659
660	661	662	663	664	665	666	667	668	669	670	671	672	673	674	675	676	677	678	679
680	681	682	683	684	685	686	687	688	689	690	691	692	693	694	695	696	697	698	699
700	701	702	703	704	705	706	707	708	709	710	711	712	713	714	715	716	717	718	719
720	721	722	723	724	725	726	727	728	729	730	731	732	733	734	735	736	737	738	739
740	741	742	743	744	745	746	747	748	749	750	751	752	753	754	755	756	757	758	759
760	761	762	763	764	765	766	767	768	769	770	771	772	773	774	775	776	777	778	779
780	781	782	783	784	785	786	787	788	789	790	791	792	793	794	795	796	797	798	799
800	801	802	803	804	805	806	807	808	809	810	811	812	813	814	815	816	817	818	819
820	821	822	823	824	825	826	827	828	829	830	831	832	833	834	835	836	837	838	839
840	841	842	843	844	845	846	847	848	849	850	851	852	853	854	855	856	857	858	859
860	861	862	863	864	865	866	867	868	869	870	871	872	873	874	875	876	877	878	879
880	881	882	883	884	885	886	887	888	889	890	891	892	893	894	895	896	897	898	899
900	901	902	903	904	905	906	907	908	909	910	911	912	913	914	915	916	917	918	919
920	921	922	923	924	925	926	927	928	929	930	931	932	933	934	935	936	937	938	939
940	941	942	943	944	945	946	947	948	949	950	951	952	953	954	955	956	957	958	959
960	961	962	963	964	965	966	967	968	969	970	971	972	973	974	975	976	977	978	979
980	981	982	983	984	985	986	987	988	989	990	991	992	993	994	995	996	997	998	999

Rysunek 42 Liczby pierwsze generowane przez wielomian $n^2 + n + 41$ w zakresie 0-999

Funkcję $f(n) = n^2 + n + 41$ możemy oczywiście zapisać w postaci: $4f(n) = 4n^2 + 4n + 164$, a taka funkcja tworzy prostą skośną w spirali Ulama.

Podsumowując: linie skośne, które tak ładnie widać na spirali Ulama, to właśnie linie generowane, przez funkcje zwane z angielskiego *prime-rich quadratic polynomials*. I przeciwnie niektóre funkcje dają tylko i wyłącznie liczby parzyste (a więc nie będące liczbami pierwszymi /z wyjątkiem dwójki/) np. $4n^2 + 2n + 4$. Te proste odpowiedzialne są znowu za „puste aleje" na spirali Ulama.

10. Wzór Pentagonalny Eulera

Liczenie podziałów liczb (w sensie sumowania – zobacz tabela 1 - poniżej) jest operacją bardzo skomplikowaną. Problem ten zawsze stanowił duże wyzwanie dla matematyków. Już Euler zaproponował pewny algorytm podziału liczb. Sam sposób liczenia Eulera nie zawierał liczb pierwszych, ale jeżeli nieznacznie zmienimy definicję ciągu wprowadzonego przez tego matematyka, to powstanie nam szereg ilości dzielników kolejnych liczb. A to już jeden krok do znalezienia liczb pierwszych. Wystarczy, że w ciągu znajdziemy liczbę dwa. Wszak liczby pierwsze mają tylko dwa dzielniki.

Zaczniemy od czegoś prostszego. Załóżmy, że chcemy policzyć na ile sposobów można przedstawić liczbę n jako sumę liczb naturalnych. Na przykład dla $n = 5$ mamy takie możliwości:

Tabela 1

$$
\begin{aligned}
&5\\
&4+1\\
&3+2\\
&2+3\\
&1+4\\
&3+1+1\\
&2+2+1\\
&2+1+2\\
&1+3+1\\
&1+2+2\\
&1+1+3\\
&2+1+1+1\\
&1+2+1+1\\
&1+1+2+1\\
&1+1+1+2\\
&1+1+1+1+1
\end{aligned}
$$

Jest ich 16. Ale czy możemy znaleźć wzór ogólny?

Zaczynając od postaci $1+1+\cdots+1+1$ możemy każdy z $+$-sów „usunąć” poprzez dodawanie liczb po jego dwóch stronach. Ponadto każdy z tych znaków możemy albo usunąć (poprzez dodanie liczb po jego stronach) albo nie, co oznacza, że każdy ma jakby dwa stany. Na początku znaków jest zawsze $n-1$, więc wzór na ilość podziałów to 2^{n-1}.

Jest to dosyć łatwy wzór. Ale można zauważyć pewną rzecz. W naszym przykładzie dla $n = 5$ wiele podziałów się powtarza, tylko że liczby są w innej kolejności, a przecież operacja dodawania jest przemienna ($a+b=b+a$) więc takie podziały są tożsame. Takie grupy powtarzających się podziałów zaznaczono poniżej w identycznych kolorach:

$$5$$
$$4+1$$
$$3+2$$
$$2+3$$
$$1+4$$
$$3+1+1$$
$$2+2+1$$
$$2+1+2$$
$$1+3+1$$
$$1+2+2$$
$$1+1+3$$
$$2+1+1+1$$
$$1+2+1+1$$
$$1+1+2+1$$
$$1+1+1+2$$
$$1+1+1+1+1$$

Gdyby te wszystkie powtarzające się podziały zredukować do jednej wyglądałoby to tak:

$$5$$
$$4+1$$
$$3+2$$
$$3+1+1$$
$$2+2+1$$
$$2+1+1+1$$
$$1+1+1+1+1$$

Zostało tylko 7 możliwości! Okazuje się, że taki sposób na podział jest o wiele ciekawszy.

Możemy teraz stworzyć ciąg (s), gdzie s_n to ilość możliwości na jakie możemy przedstawić liczbę n w powyższy sposób. Początek tego ciągu wygląda następująco:

$$1, 2, 3, 5, 7, 11, 15, 22, 30, 42, 56, 77, 101, 135, 176, 231, 297, 385, 490, 627 \ldots \quad <1>$$

Możemy zauważyć, że w/w ciąg ma dziwne własności. Na początku $(1, 2, 3, 5)$ wygląda jakby to był ciąg Fibbonaciego (bez pierwszej jedynki). Następnie $(1, 2, 3, 5, 7, 11)$ jakby to były liczby pierwsze (oraz 1). Niestety po 11 wszystko zaczyna się psuć i trudno znaleźć więcej zależności.

Euler był zainteresowany w/w ciągiem i postanowił znaleźć jego wzór ogólny. Poszedł tropem Fibbonaciego. Uznał, że oprócz dodawania dwóch ostatnich liczb, co będzie jeżeli odejmiemy jakąś dalszą? Dla przypomnienia gdyby to był ciąg Fibbonaciego to wzór wyglądałby następująco:

$$s_n = s_{n-1} + s_{n-2}$$

Euler zaproponował, żeby przed pierwszą jedynką dodać jeszcze jedną (tak aby bardziej przypominał on ciąg Fibbonaciego: $s_0 = 1, s_1 = 1$) oraz, żeby niektóre wyrazy ciągu występujące wcześniej dodawać i odejmować.

Zacznijmy od zwykłego wzoru: $s_n = s_{n-1} + s_{n-2}$. Daje on nam (na czerwono zaznaczono liczby, które nie zgadzają się z oryginalnym ciągiem (s)):

$$1, 1, 2, 3, 5, 8, 13, 21, 34, 55, 89, 144, 233, 377, 610, 987, 1597, 2584, 4181, 6765, \ldots$$

Teraz podążajmy tak jak Euler. Pierwszy wyraz, który nie zgadza się z <1> to 8 na piątym miejscu (zaczynamy od zera) w ciągu. Tam powinna być 7. Aby otrzymać poprawny wyraz musimy odjąć 1. Taką wartość w ciągu mamy na miejscu zerowym. czyli: $s_n = s_{n-1} + s_{n-2} - s_{n-5}$. Wymieniony wzór daje nam ciąg postaci:

$$1, 1, 2, 3, 5, 7, 11, 16, 24, 35, 52, \ldots$$

16 jest większe niż 15 (poprawna w ciągu <1>), więc musimy odjąć jeszcze raz: $s_n = s_{n-1} + s_{n-2} - s_{n-5} - s_{n-7}$. Ciąg zgadza się trochę lepiej (ale nadal psuje się po pewnym czasie). Następnie więc dodajemy lub odejmujemy kolejne zmienne wartości dopóki nie otrzymamy poprawnego wzru. Eulerowi udało się udowodnić, że takie podejście faktycznie zadziała. Udowodnił także, że znak liczb, które dodajemy zmienia się co 2, czyli:

$$s_n = s_{n-m_1} + s_{n-m_2} - s_{n-m_3} - s_{n-m_4} + s_{n-m_5} + s_{n-m_6} - s_{n-m_7} - \cdots$$

Teraz pozostaje pytanie jak znaleźć liczby $m_1, m_2, m_3, \ldots$? Otóż możemy stworzyć kolejny ciąg! Nazwijmy go (m):

$$1, 2, 5, 7, 12, 15, 22, 26, 35, 40, 51, 57, 70, 77, 92, 100, 117, 126, 145, 155, 176, 187, \ldots \quad <2>$$

Wygląda on równie dziwnie jak nasz oryginalny ciąg <1>. Ma on także ciekawe własności. Na przykład jeżeli na podstawie ciągu (m) stworzymy następny szereg poprzez odejmowanie od siebie dwóch kolejnych wyrazów otrzymujemy:

$$1, 3, 2, 5, 3, 7, 4, 9, 5, 11, 6, 13, 7, 15, 9, 17, \ldots$$

Teraz patrząc na co drugi element ciągu możemy zobaczyć albo:

$$1, 2, 3, 4, 5, 6, 7, 8, 9, \ldots$$

Albo:

$$3, 5, 7, 9, 11, 13, 15, 17, \ldots$$

Widzimy od razu, że wyrazami pierwszego ciągu są liczby naturalne, natomiast drugiego są liczby nieparzyste (z pominięciem 1). Dzięki tym dwóm prostym ciągom możemy ciągle dopisywać kolejne wyrazy do ciągu (m). Trudno jednak na ich podstawie stworzyć ładny wzór. Popatrzmy więc jeszcze raz na ciąg (m): $1, 2, 5, 7, 12, 15, 22, 26, 35, 40, 51, 57, 70, 77, 92, 100, 117, \ldots$ Euler zauważył tu pewną własność. Otóż co druga liczba (zaczynając od 1) jest liczbą pięciokątną (pentagonalną). Ale co to właściwie jest liczba pięciokątna? Otóż jest to ilość punktów jakie są potrzebne, aby skonstruować pięciokąt foremny o boku n. Wzór opisujący liczby pięciokątne to:

$$P(n) = \frac{3n^2 - n}{2}$$

Sprawdźmy, czy się zgadza. Kolejne liczby pentagonalne to:

$$1, 5, 12, 22, 35, 51, 70, 92, 117, 145, 176, 210, 247, \ldots$$

Przypatrzmy się na nasz ciąg (m):

$$1, 2, 5, 7, 12, 15, 22, 26, 35, 40, 51, 57, 70, 77, 92, 100, 117, 126, 145, 155, 176, 187, \ldots$$

Zgadza się są to kolejne liczby pentagonalne! Ale co z pozostałymi? Euler dalej podążał szlakiem liczb pentagonalnych. Postanowił popatrzeć jakie wartości przyjmuje $P(n)$ dla ujemnych n. Okazuje się, że i tutaj mu się udało, gdyż otrzymał następujący ciąg:

$$2, 5, 15, 26, 40, 57, 77, 100, 126, 155, 187, 222, 260, \ldots$$

Tworzy on pozostałe wyrazy w ciągu (m). Oznacza to, że wzór ogólny ciągu (m) ma postać:

$$m_n = P\left((-1)^{n+1}\left\lceil\frac{n}{2}\right\rceil\right)$$

Jest to wzór ciągu zwanego „uogólnionym ciągiem pentagonalnym". Teraz możemy zdefiniować nasz początkowy ciąg (s) jako:

$$s_{<0} = 0$$
$$s_0 = 1$$
$$s_1 = 1$$
$$s_n = \sum_{n=1}^{\infty} s_{m_n} * (-1)^{\left\lfloor\frac{n-1}{2}\right\rfloor} = \sum_{n=1}^{\infty} s_{P\left((-1)^{n+1}\left\lceil\frac{n}{2}\right\rceil\right)} * (-1)^{\left\lfloor\frac{n-1}{2}\right\rfloor} =$$
$$= \sum_{n=1}^{\infty} s_{\frac{3\left((-1)^{n+1}\left\lceil\frac{n}{2}\right\rceil\right)^2 - \left((-1)^{n+1}\left\lceil\frac{n}{2}\right\rceil\right)}{2}} * (-1)^{\left\lfloor\frac{n-1}{2}\right\rfloor}$$

Jest to wzór skomplikowany, ale poprawny. Wiedząc teraz jak liczyć kolejne wyrazy w tym ciągu, spróbujmy znaleźć połączenie z liczbami pierwszymi. Otóż gdyby uznać, że s_0 nie musi koniecznie być równe 1, ale opierając się nadal na tej samej definicji, to powstaną nam nowe ciekawe ciągi. Jednym z nich jest ciąg, w którym wartość s_0 przyrównujemy do n. Oznacza to, że dla każdego kolejnego wyrazu w ciągu s_0 wzrasta o 1 (taki ciekawy ciąg, gdzie zerowy wyraz ciągu nie jest stały). Gdy będziemy tak kontynuować powstanie nam ciąg:

$$1, 2, 2, 3, 2, 4, 2, 4, 3, 4, 2, 6, 2, 4, 4, 5, 2, 6, 2, 6, 4, 4, 2, 8, 3, 4, 4, 6, \ldots$$

Może się wydawać, że ten ciąg jest nudny. Jednak po głębszym przyjrzeniu się można zauważyć, że jest to ciąg ilości dzielników n. Na przykład dla $n = 12$ wyraz ciągu przyjmuje wartość 6, ponieważ 12 ma 6 dzielników: $1, 2, 3, 4, 6, 12$. Oznacza to, że liczby pierwsze są oznaczone w tym ciągu cyfrą 2 (każda liczba pierwsza dzieli się tylko przez 1 i samą siebie). Znaleźliśmy więc sposób na szukanie liczb pierwszych! Jest to teoretycznie dobry sposób, gdyż nie sprawdzamy żadnych podzielności. Jednak w praktyce okazuje się być on niezbyt skuteczny, ponieważ aby obliczyć kolejny wyraz w ciągu trzeba znać wszystkie poprzednie. Kłopot nie jest tutaj w obliczaniu poprzedników, ale w przechowywaniu ich w pamięci. Na przykład chcąc obliczyć dzielniki $n = 1000000$ trzeba przechowywać wszystkie wcześniejsze wyrazy, których jest 999999.

11. Hipoteza Riemanna

Leonard Euler rozwiązał sformułowany sto lat wcześniej tak zwany Problem Bazylejski:

$$\sum_{n=1}^{\infty} \frac{1}{n^2} = \frac{1}{1^2} + \frac{1}{2^2} + \frac{1}{3^2} + \cdots = ?$$

Euler udowodnił, że powyższy ciąg „zbiega się" do $\frac{\pi^2}{6}$ (gdzie $\pi = 3{,}14..$). Dowód ten zainspirował innego słynnego matematyka Bernarda Riemanna, który uogólnił wzór z Problemu Bazylejskiego. Nową funkcję oznaczył grecką literą dzeta ζ.

$$\zeta(x) = \frac{1}{1^x} + \frac{1}{2^x} + \frac{1}{3^x} + \cdots$$

Dzisiaj w/w funkcja nosi nazwę funkcji dzeta-Riemanna. Jakie wartości przyjmuje ta funkcja? Euler udowodnił już, że $\zeta(2) = \frac{\pi^2}{6}$. Dla x = 3 otrzymujemy również szereg zbieżny $\zeta(3) = 1{,}202$ … Zobaczmy tabelę z kilkoma wynikami funkcji $\zeta(x)$:

x	$\zeta(x)$	Uwagi
2	1,644934067	$\frac{\pi^2}{6}$
3	1,202056903	
4	1,082323234	$\frac{\pi^4}{90}$
5	1,036927755	
10	1,000994575	$\frac{\pi^{10}}{9355}$
20	1,000000954	$\frac{174611\,\pi^{20}}{1531329465290625}$

Rysunek 43 Wartości funkcji dzeta-Riemana dla poszczególnych liczb naturalnych

Powyższa tabela pokazuje nam, że dla x ≥ 2 funkcja dzeta-Riemanna jest zbieżna. Stosowaliśmy tutaj funkcję $\zeta(x)$ dla $x \in \mathbb{N}$, ale przecież nic nie stoi na przeszkodzie, aby stosować, ją dla $x \in \mathbb{R}$. (Dlatego od początku oznaczałem zmienną przy funkcji dzeta jako x /rzeczywiste/, a nie n /naturalne/).

x	$\zeta(x)$
$\frac{3}{2}$	2,612375349
$\sqrt{2}$	3,020737679
π	1,176241738

Rysunek 44 Wartości funkcji dzeta-Riemana dla poszczególnych liczb rzeczywistych

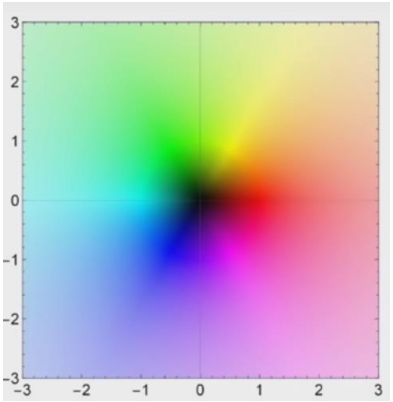
W powyższych tabelach zawsze x było większe od 1. Co się stanie jeżeli będziemy chcieli obliczyć wartość funkcji ζ dla liczb ujemnych? Dobrym przykładem jest $\zeta(-1)$:

$$\zeta(-1) = \frac{1}{1^{-1}} + \frac{1}{2^{-1}} + \frac{1}{3^{-1}} + \cdots = 1 + 2 + 3 + \cdots = \infty$$

Tak naprawdę to $\zeta(-1) = -\frac{1}{12}$, ale do takich obliczeń wykorzystano inną definicję. Okazuje się, że wykorzystując funkcję dzeta-Riemanna otrzymujemy szeregi zbieżne dla $x > 1$. Oczywiście mówimy tu o $x \in \mathbb{R}$. Riemann jednak spojrzał na problem bardziej ogólnie. Rozszerzył dziedzinę funkcji do liczb zespolonych $s \in \mathbb{C}$

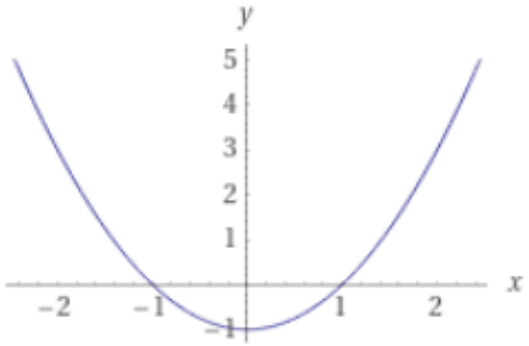
$$\zeta(s) = \frac{1}{1^s} + \frac{1}{2^s} + \frac{1}{3^s} + \cdots$$

Ale jak przedstawić „wykres" funkcji mającej za dziedzinę zbiór liczb zespolonych $\mathbb{C}$ i zbiór wartości również $\mathbb{C}$? Liczby zespolone przedstawiamy w dwóch wymiarach. Czyli dziedzina to dwa wymiary, przeciwdziedzina również. Więc taki „wykres" powinien mieć cztery wymiary. I tu przychodzą z pomocą komputery, a właściwie odpowiednie oprogramowanie, które zmienia nam wartości liczb zespolonych na kolory i ich jasności (dwa wymiary). Czyli tak: na płaszczyźnie mamy dziedzinę funkcji (oś X część rzeczywista, oś Y część urojona liczby zespolonej. Natomiast na tej samej płaszczyźnie zaznaczamy wartości funkcji różnymi barwami oraz ich kolorami. Poniżej zamieszczono próbkę takiego wykresu funkcji $f(s) = s$. Czyli właściwie definicja przypisania kolorów.



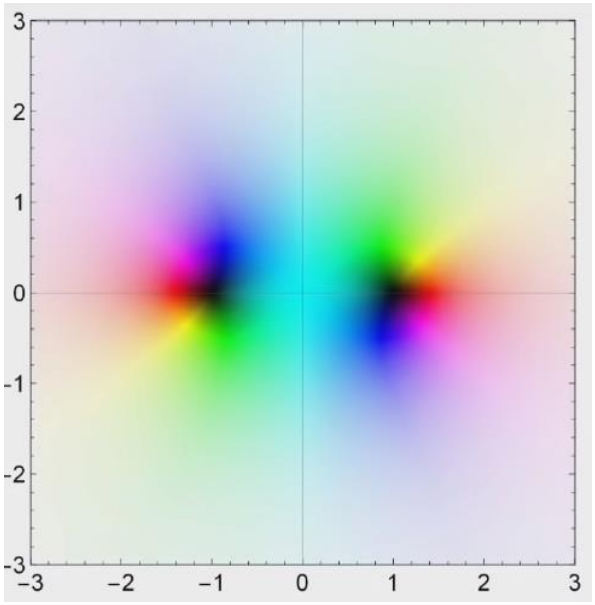
Rysunek 45 Kolorowany wykres czterowymiarowy funkcji f(s) = s

Jak narysować wykres funkcji np. $f(s) = s^2 - 1; s \in \mathbb{C}$? Wykres dla liczb rzeczywistych wyglądałby następująco:



Rysunek 46 Wykres (dwuwymiarowy) funkcji $f(x) = x^2 - 1$

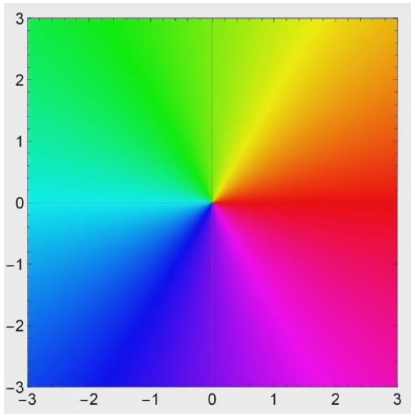
Jeżeli zastosujemy „kolorowanie” mające na celu przedstawienie wykresu czterowymiarowego, to taka parabola wyglądałaby następująco:



Rysunek 47 Wykres czterowymiarowy funkcji $f(s) = s^2 - 1$

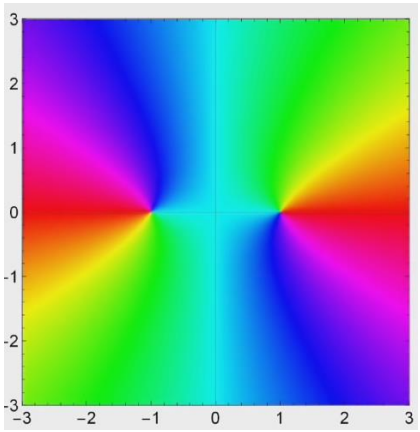
Zauważmy, że w obu przypadkach miejsca zerowe tej funkcji wynoszą $x_1 = -1 \wedge x_2 = 1$. Na rysunku 46 – na wykresie do którego jesteśmy przyzwyczajeni w szkole – miejsca zerowe widzimy od razu, natomiast na rysunku 47 musimy się dokładnie przyjrzeć. Miejsca zerowe są w punktach, które są czarne (im wartość wyniku jest mniejsza tym kolory są ciemniejsze). Jeżeli popatrzymy się na rysunek, to zobaczymy, że te „czarne punkty” mają współrzędne (-1,0) i (1,0).

Często aby jeszcze bardziej uprościć rysunek – tak aby stał się on bardziej czytelny rezygnujemy z „jasności” wykresu operując tylko kolorami. Tam gdzie obraz ciemnieje (wartości funkcji są stosunkowo niskie) obraz staje się nieczytelny.



Rysunek 48 Kolorowany wykres funkcji f(s) = s bez jasności kolorów

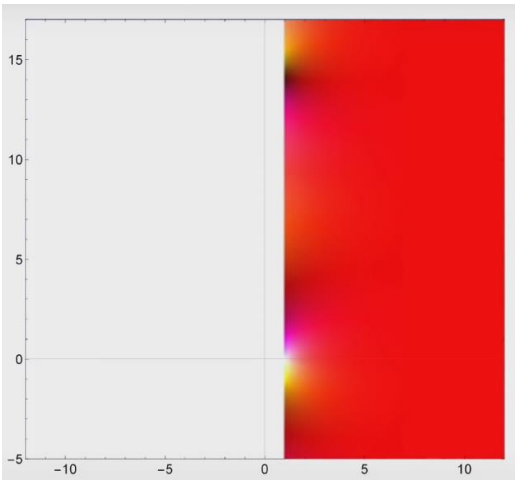
Wykres przedstawiony na rysunku 48 jest o wiele czytelniejszy. Miejsca zerowe na takim rysunku są w miejscu gdzie spotykają się kolory podstawowe. W tak uproszczony sposób wcześniejszą parabolę możemy przedstawić następująco:



Rysunek 49 Kolorowany wykres funkcji $f(s) = s^2 - 1$ *bez jasności kolorów*

Zauważmy, że miejsca zerowe funkcji $y = s^2 - 1$ są łatwiejsze do odczytania.

Jak zatem będzie wyglądał wykres $\zeta(s); s \in \mathbb{C}$



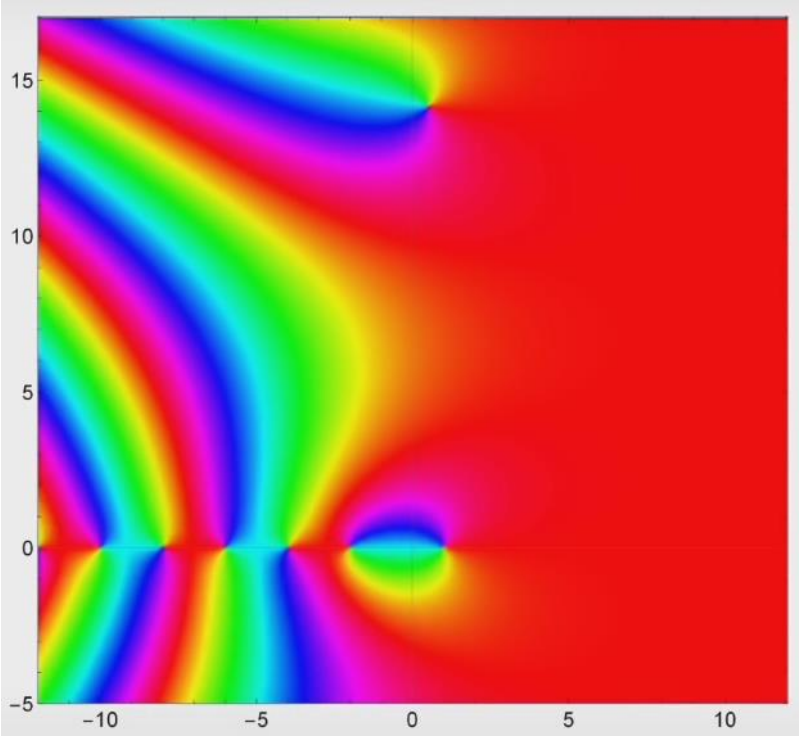
Rysunek 50 Wykres $\zeta(s); s \in \mathbb{C} \ (Re(s) > 1)$

Niestety dla liczb zespolonych podobnie jak w przypadku liczb rzeczywistych funkcja dzeta-Riemanna ma sens tylko dla $s > 1$. Riemann rozszerzył analitycznie dziedzinę funkcji ζ (tzw. funkcja holomorficzna) do całego zbioru liczb zespolonych. W swojej pracy pokazał jak rozszerzyć funkcję ζ tak, aby dla argumentów większych od 1 redukowała się do przedstawionego wcześniej wzoru Eulera. Niestety opis rozszerzenia funkcji dzeta jest zbyt skomplikowany, aby przedstawiać go w niniejszej pracy. Poniżej podamy tylko postać jaką przyjmuje ta funkcja rozszerzona na cały zbiór liczb zespolonych (za *wikipedia.org*):

$$\zeta(z) = \frac{1}{1-2^{1-z}} \cdot \sum_{n=0}^{\infty} \frac{1}{2^{n+1}} \sum_{k=0}^{n} (-1)^k \binom{n}{k} \cdot (k+1)^{-z}$$

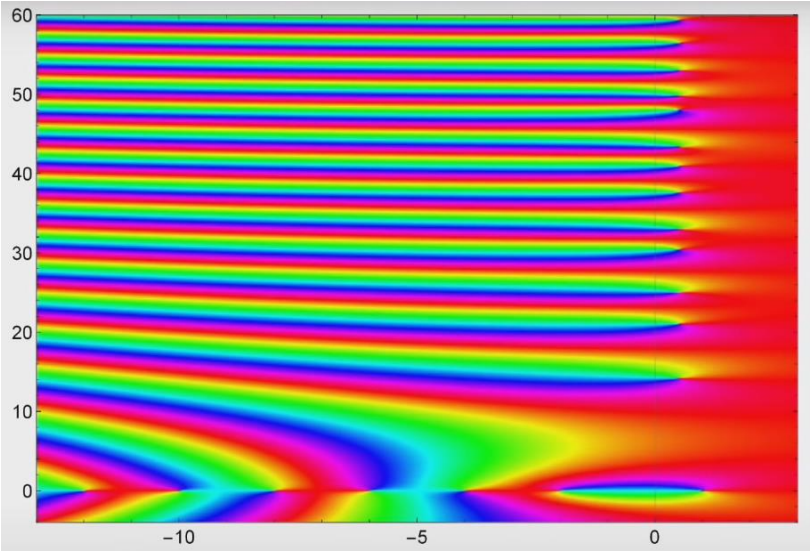
Piękne, prawda? Riemannowi tak udało się rozszerzyć dziedzinę, że tylko w punkcie (1,0) (jedynka rzeczywista i zero urojone) funkcja nie ma wartości (widać to przy pierwszym ułamku powyższej sumy, gdzie dla z = (1,0) mianownik przyjmuje wartość zero). Co więcej dla wszystkich argumentów (-2n, 0) gdzie $n \in \mathbb{N}$, czyli wszystkich liczb mających za część rzeczywistą parzyste liczby ujemne i zero dla części urojonej, funkcja „zeruje się”. Wszystkie miejsca zerowe takiej postaci nazywamy „trywialnymi” miejscami zerowymi.

Zobaczmy teraz jak wyglądałby wykres funkcji ζ.



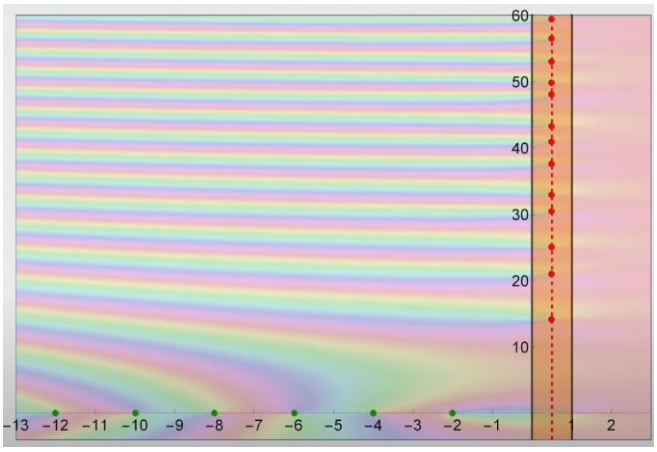
Rysunek 51 Wykres $\zeta(s); s \in \mathbb{C}$

Lub jeżeli oddalimy się trochę od wykresu i go przeskalujemy:



Rysunek 52 Przeskalowany wykres $\zeta(s); s \in \mathbb{C}$

Widzimy, że oprócz miejsc zerowych na osi X (trywialnych) dość dużo „zer” jest także na prawo od osi Y. Przypomnimy, że na tego typu wykresach miejsca zerowe są tam, gdzie spotykają się wszystkie kolory podstawowe. Matematycy określili dla funkcji dzeta-Riemanna pas pomiędzy 0 i 1 na osi rzeczywistej (X) mianem paska krytycznego, a linię biegnącą środkiem tego pasa jako linię krytyczną. Zaznaczmy teraz kilka pierwszych miejsc zerowych opisywanej funkcji.



Rysunek 53 Wykres $\zeta(s)$ z zaznaczoną prostą krytyczną

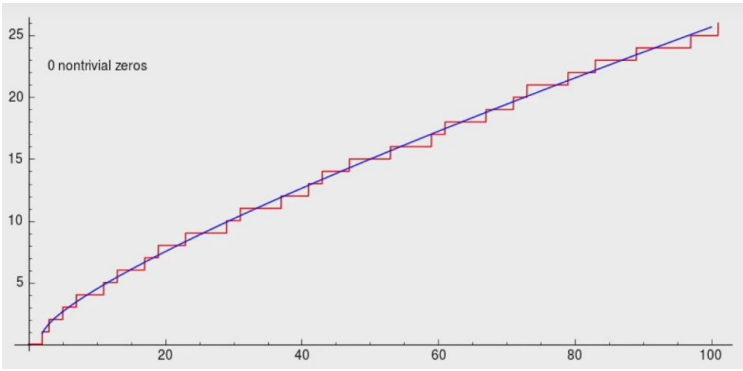
Sformułujmy słynną hipotezę Riemanna

Twierdzenie 11.1 Hipoteza Riemana

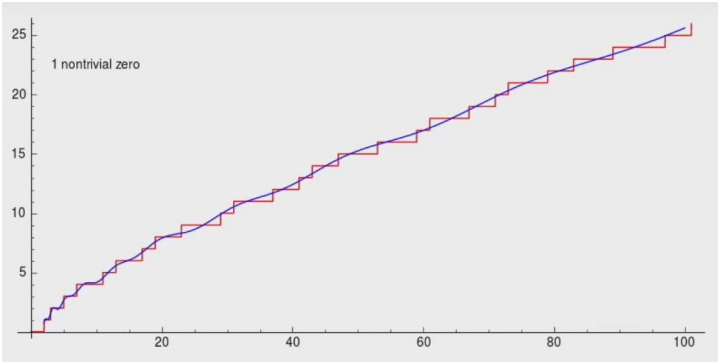
Wszystkie nietrywialne „zera” funkcji ζ leżą na prostej krytycznej.

Dlaczego hipoteza Riemanna jest tak ważna? I dlaczego zajmujemy się nią w pracy dotyczącej liczb pierwszych? Przecież do tej pory przy opisie funkcji dzeta-Riemanna nie padło ani jedno słowo dotyczące liczb pierwszych. Riemann definiując funkcję dzeta chciał jeszcze dokładniej przybliżyć funkcję $\pi(n)$ – ilość liczb pierwszych nie większych od n. Okazuje się, że stosując pewne poprawki związane z miejscami zerowymi można bardzo dokładnie określić ilość liczb pierwszych w pewnych zakresach.

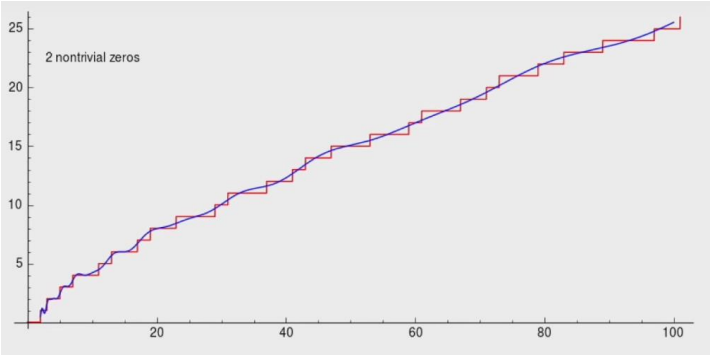
Jeżeli pokażemy wykresy funkcji $\pi(x)$ oraz jej przybliżenia wzorem Riemanna dla poszczególnych przybliżeń nietrywialnymi miejscami zerowymi to okaże się, że wzór Riemanna bardzo dokładnie określa funkcję $\pi(x)$.



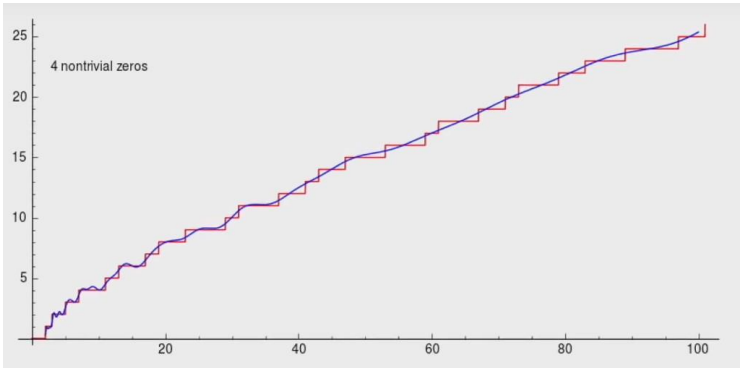
Rysunek 54 Wykres funkcji $\pi(x)$ bez poprawek



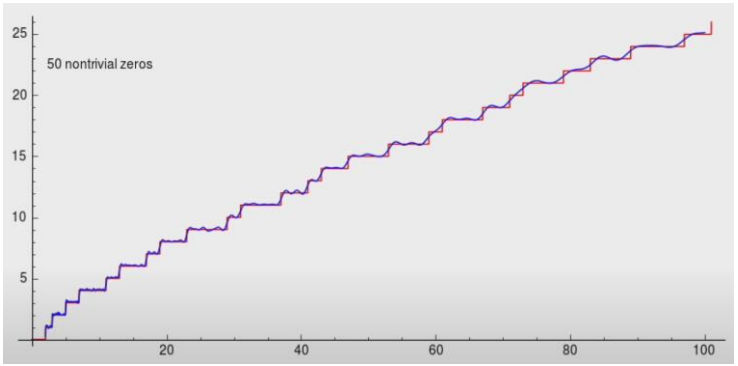
Rysunek 55 Wykres funkcji $\pi(x)$ z 1 poprawką



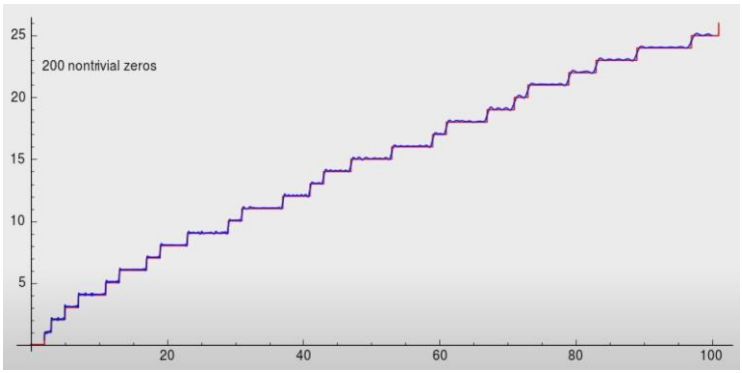
Rysunek 56 Wykres funkcji $\pi(x)$ z 2 poprawkami



Rysunek 57 Wykres funkcji $\pi(x)$ z 4 poprawkami



Rysunek 58 Wykres funkcji z π(x) z 50 poprawkami



Rysunek 59 Wykres funkcji $\pi(x)$ z 200 poprawkami

Widzimy, że wykres wzoru Riemanna i funkcji $\pi(x)$ po zastosowaniu 200 poprawek z nietrywialnych miejsc zerowych funkcji dzeta-Riemanna jest już właściwie identyczny z wykresem funkcji $\pi(x)$.

Możemy powiedzieć, że Riemann podał dokładny wzór na rozmieszczenie liczb pierwszych, co do tej pory było świętym Graalem dla matematyków. Ale jest jeden problem. Hipoteza Riemanna nie jest do tej pory udowodniona. Tak więc przeniósł on problem z jednego poziomu na drugi. Teraz to udowodnienie Hipotezy Riemanna jest świętym Graalem dla matematyków.

12. Kryptografia

Można sobie zadać pytanie dlaczego liczby pierwsze, a w szczególności Hipoteza Riemanna, są takie ważne. Przecież tego typu rozważania są pasjonujące, ale pewnie tylko dla pewnej grupy ludzi, których można nazwać entuzjastami (może nawet fanatykami) Matematyki (przez duże M). Ale dla reszty? Otóż okazuje się, że liczby pierwsze mają zastosowanie w naszym codziennym życiu. Do czego są potrzebne? Do szyfrowania! A szyfrowanie przydaje się do codziennych naszych kontaktów. Na przykład przy logowaniu się na facebook'a, czy do banku. Nie mówiąc już o zastosowaniach militarnych.

Jedna z metod szyfrowania tzw. RSA (od nazwisk Rivest, Shamir, Adleman) oparta jest na liczbach pierwszych i tak zwanej arytmetyce modulo. Jak wygląda takie szyfrowanie? Sam algorytm jest dość prosty. Poniżej przedstawimy podstawowy opis tej metody. Niektóre odmiany tej metody różnią się od siebie, natomiast tutaj chcemy przedstawić tylko wykorzystanie liczb pierwszych, a nie sam sposób szyfrowania.

Weźmy dwie liczby pierwsze p i q. Niech n będzie iloczynem tych dwóch liczb $n = p \cdot q$, a $\varphi(n) = (p-1)(q-1)$ Weźmy także liczbę e taką, że $e \cdot k = 1 \, mod \, \varphi(n)$. Liczby n oraz e stanowią tak zwany klucz publiczny, a liczba k jest kluczem prywatnym. Klucz publiczny dostępny jest dla każdego, kto chce na przykład przesłać zaszyfrowane dane do banku. Jeżeli znamy klucz publiczny i mamy jakąś informację do przekazania (dla uproszczenia niech informacja będzie jakąś liczbą oznaczoną literą J) to przesyłamy ją według algorytmu:

Algorytm szyfrowania RSA za pomocą klucza publicznego n oraz e jednostki informacji (liczby) oznaczonej J

$$C = J^e \, (mod \, n)$$

I już. To już koniec algorytmu. Jednostka informacji (liczba) C jest już zaszyfrowana. Znajomość klucza publicznego nie pozwala nam na odszyfrowanie wiadomości. Do tego potrzebny jest nam klucz prywatny. Poniżej przedstawiam algorytm deszyfrujący:

Algorytm deszyfrowania RSA za pomocą klucza prywatnego k oraz klucza publicznego n oraz e jednostki informacji zaszyfrowanej (liczby) oznaczonej C

Niech $k = \frac{(p-1)\cdot(q-1)+1}{e}$

Wtedy $J = C^k \, (mod \, n)$

Jednostka informacji (liczba) J jest już odszyfrowana. Także bardzo prosty algorytm, który wykorzystamy w następującym przykładzie:

Weźmy literę **M**, którą chcemy zaszyfrować. Liczby pierwsze, które będą nam generować zarówno klucz publiczny, jak i prywatny, wybierzemy małe, aby można było prowadzić obliczenia.

Niech $p = 11, q = 17$. Wtedy $n = p \cdot q = 187, \varphi(n) = (p-1)(q-1) = 160$. Wybieramy takie e aby $e \cdot k = 1 \; mod \; \varphi(n)$ dla pewnego k. Np. niech $e = 7$. Teraz musimy przekształcić daną do zaszyfrowania na liczbę. Możemy do tego wykorzystać na przykład tablicę ASCII (patrz rozdział 2). W tabeli tej dużej literze M odpowiada liczba 77. Otrzymujemy:

$$C = J^e \; (mod \; n) = 77^7 \; (mod \; 187)$$

Do dalszych obliczeń wykorzystujemy zależność $a^{n+m} = \; a^n \cdot a^m$ oraz wzór z arytmetyki modulo: $(a \cdot b) mod \; n = \big((\, a \; mod \; n) \cdot (b \; mod \;\; n)\big) mod \; n$

$$\begin{aligned} C = 77^7 \; (mod \; 187) &= (77^3 \; mod \; 187) \cdot (77^4 \; mod \; 187) mod \; 187 \\ &= (456533 \; mod \; 187) \cdot (35153041 \; mod \; 187) \; mod \; 187 \\ &= (66 \cdot 33) mod \; 187 = 2178 \; mod \; 187 = \mathbf{121} \end{aligned}$$

A więc litera M, której w tabeli ASCII odpowiada liczba 77 została zaszyfrowana jako liczba 121.

Jak możemy odszyfrować zakodowaną wiadomość? Przy pomocy klucza prywatnego k. Obliczamy:

$$k = \frac{(p-1)(q-1)+1}{e} = \frac{10 \cdot 16 + 1}{7} = 23$$

Liczba k to klucz prywatny potrzebny do deszyfracji.

$$J \; = \; C^k mod \; n = 121^{23} \; mod \; 187 = 77$$

Oczywiście kodując informację nie szyfrujemy każdej litery osobno. Wtedy można byłoby złamać szyfr stosując inne metody łamania. Na przykład wykorzystując częstotliwość występowania poszczególnych liter w danym alfabecie. Szyfrując np. wyraz Matematyka dzielimy go na grupy liter i te dopiero szyfrujemy. W praktyce, aby utrudnić złamanie szyfru stosuje się grupy zawierające nawet kilkadziesiąt znaków.

Na czym polega genialność tego rozwiązania? Otóż mamy dwa osobne klucze, jeden do szyfrowania, drugi do deszyfracji. Klucz publiczny, którym szyfrujemy, nie może być wykorzystany do deszyfracji. Dlaczego? Przecież jeżeli znamy n oraz e możemy próbować rozłożyć n na czynniki pierwsze i odwrócić proces. Nic prostszego. Oczywiście, jeżeli n jest niewielkie możemy próbować, ale jeżeli liczby pierwsze p, q weźmiemy znacznie większe, takie na przykład 150 cyfrowe, to liczba n będzie 300 cyfrowa, a to już większy kłopot. W tej chwili do szyfrowania używa się kluczy, które w zapisie dwójkowym mają 2048 bitów, czyli dziesiętnie jest to liczba rzędu nawet 2^{2048}. Dziesiętnie są to liczby o 617 cyfrach, a znalezienie czynników pierwszych takich liczb w realnym czasie jest właściwie niemożliwe.

Metody faktoryzacji są coraz bardziej zaawansowane. Złamano już klucz 512 bitowy. Kilka lat temu pewnej grupie akademickiej udało się złamać klucz 768 bitowy. Zajęło im to około 2 lata i wykorzystywali w tym celu wszystkie komputery, które były w ich zasobach. Zaznaczyli przy

tym, że wydłużenie tego klucza chociażby o kilka bitów wydłużyłoby ten proces do nawet kilku tysięcy lat.

Co ma z tym wszystkim wspólnego Hipoteza Riemanna? Otóż, gdyby udało się ją udowodnić, to jak wspomniałem w poprzednim rozdziale poznalibyśmy rozmieszczenie liczb pierwszych. A to pozwoliłoby udoskonalić i znacznie przyspieszyć proces faktoryzacji. Czyli moglibyśmy łamać szyfrowanie RSA. A to byłoby już bardzo niebezpieczne. Właściwie wszystkie metody szyfrowania we współczesnym świecie wykorzystują metodę RSA. Logowanie do komputerów, do mediów społecznościowych, do banków, instytucji rządowych, a nawet do zasobów wojskowych stałoby się niebezpieczne. Dlatego liczby pierwsze, a w szczególności Hipoteza Riemanna są tak ważne nie tylko dla matematyków.

13. Zakończenie

Na tym kończę swoją pracę na temat liczb pierwszych. Mimo, że poruszyłem tylko kilka tematów z nimi związanych, dzieło to rozrosło się do niemałych rozmiarów. Zająłem się właściwie tylko tymi problemami, które zainteresowały mnie w większym stopniu niż pozostałe. Jest to bardzo trudny temat (chyba najbardziej skomplikowany z jakim do tej pory się spotkałem), mimo to bardzo fascynujący. Liczby pierwsze występują właściwie w każdym dziale matematyki. Spotykamy się z nimi nawet przy wydawałoby się zupełnie odrębnych tematach – jak chociażby wyliczanie liczby π (rozdział 7) czy e (nieujęte w tej pracy). Nic w tym dziwnego, wszak liczby pierwsze stanowią jakby „budulec" wszystkich pozostały liczb - porównuje się je do atomów w fizyce. Współpraca fizyków i matematyków doprowadziła jeszcze do jednego odkrycia. Otóż rozkład protonów w jądrze atomów ciężkich jest związany z nietrywialnymi miejscami zerowymi funkcji dzeta-Riemanna, a więc z rozkładem liczb pierwszych.

Jest to koniec mojej pracy, ale na pewno nie koniec mojej przygody z liczbami pierwszymi. Pozostało jeszcze dużo tematów do poznania i do odkrycia …

14. Spis Rysunków

Rysunek 1 Gdzie na osi liczbowej przedstawić symbol i? 8
Rysunek 2 Przedstawienie liczb zespolonych na układzie współrzędnych 8
Rysunek 3 Graficzna interpretacja arytmetyki modularnej 10
Rysunek 4 Tablica ASCII 11
Rysunek 5 Tabela z ilością liczb pierwszych w kolejnych dziesiątkach liczb naturalnych 15
Rysunek 6 Ilość liczb pierwszych w przedziale 1-100 w podziale na dziesiątki 15
Rysunek 7 ilość liczb pierwszych w przedziale 1-1000 w podziale na setki 15
Rysunek 8 Ilość liczb pierwszych (narastająco) w kilku pierwszych dziesiątkach 16
Rysunek 9 Sito Eratostenesa 17
Rysunek 10 Rozkład liczby 29 (pokolorowany) 28
Rysunek 11 Rozkład liczby 29 (niepokolorowany) 28
Rysunek 12 Pojedyncze graficzne rozwiązanie 29
Rysunek 13 Punkty kratowe 32
Rysunek 14 Punkty kratowe określające pole koła 33
Rysunek 15 Okręgi współśrodkowe, o różnicy promienia 0,1 33
Rysunek 16 Okręgi współśrodkowe o promienia będącymi pierwiastkami liczb naturalnych 34
Rysunek 17 Rozkład 5525 na czynniki pierwsze Gaussa 35
Rysunek 18 Rozkład 975 na czynniki pierwsze Gaussa 35
Rysunek 19 Wykres ilości liczb pierwszych w przedziale od 1 do 10 41
Rysunek 20 Wykres ilości liczb pierwszych w przedziale od 1 do 50 41
Rysunek 21Wykres ilości liczb pierwszych w kolejnych przedziałach 42
Rysunek 22 Wykresy ilości liczb pierwszych oraz funkcji $\boldsymbol{n2\ i\ n}$ 42
Rysunek 23 Wartości funkcji $\boldsymbol{\pi n, \pi nn}$ 43
Rysunek 24 Wartości funkcji $\boldsymbol{\pi n, \pi nn, n\pi(n)}$ 44
Rysunek 25 Wartości funkcji $\boldsymbol{\pi n, nln\ (n)}$ oraz błąd bezwzględny 44
Rysunek 26 Wartości funkcji $\boldsymbol{\pi n, nln\ (n)}$ błąd bezwzględny oraz błąd względny i dokładność 45
Rysunek 27 Wykresy funkcji $\boldsymbol{\pi n, nln\ (n)}$ w przedziale 1-100 45
Rysunek 28 Wykres funkcji $\boldsymbol{\pi n, nln\ (n)}$ w przedziale 1-10 000 000 46
Rysunek 29 Liczby wpisane w spiral e kwadratową 47
Rysunek 30 Spirala Ulama 47
Rysunek 31 Spirala Ulama 200x200 48
Rysunek 32 Losowy rozkład liczb w zakresie 200x200 48
Rysunek 33 Sposób przełożenia osi liczbowej na spiralę 49
Rysunek 34 Graficzna reprezentacja wyliczeń funkcji kwadratowej dla liczb 5, 18, 39, 68, 105 50
Rysunek 35 Graficzna reprezentacja wyliczeń funkcji kwadratowej dla liczb 7, 21, 43, 73, 111 51
Rysunek 36 Funkcja kwadratowa dla liczb 1, 9, 25 , 49, 81, 121 52
Rysunek 37 Funkcja kwadratowa dla liczb 4, 16, 36, 64, 100 52
Rysunek 38 Funkcja kwadratowa dla liczb 2, 11, 28, 53, 86 52
Rysunek 39 Funkcja kwadratowa dla liczb 3, 13, 31, 57, 91 52
Rysunek 40 Nie wszystkie wartości początkowe funkcji leżą na prostej 52
Rysunek 41 Liczby pierwsze generowane przez wielomian n^2+n+41 w zakresie 0-99 53
Rysunek 42 Liczby pierwsze generowane przez wielomian n^2+n+41 w zakresie 0-999 53
Rysunek 43 Wartości funkcji dzeta-Riemana dla poszczególnych liczb naturalnych 59
Rysunek 44 Wartości funkcji dzeta-Riemana dla poszczególnych liczb rzeczywistych 60
Rysunek 45 Kolorowany wykres czterowymiarowy funkcji f(s) = s 60
Rysunek 46 Wykres (dwuwymiarowy) funkcji $f(x)=x^2-1$ 61
Rysunek 47 Wykres czterowymiarowy funkcji $f(s)=s^2-1$ 61
Rysunek 48 Kolorowany wykres funkcji f(s) = s bez jasności kolorów 62
Rysunek 49 Kolorowany wykres funkcji $fs= s^2-1$ bez jasności kolorów 62
Rysunek 50 Wykres $\boldsymbol{\zeta s; s \in \mathbb{C}\ (Res > 1)}$ 62
Rysunek 51 Wykres $\boldsymbol{\zeta s; s \in \mathbb{C}}$ 63
Rysunek 52 Przeskalowany wykres $\boldsymbol{\zeta s; s \in \mathbb{C}}$ 64
Rysunek 53 Wykres $\boldsymbol{\zeta s}$ z zaznaczoną prostą krytyczną 64
Rysunek 54 Wykres funkcji $\boldsymbol{\pi(x)}$ bez poprawek 65
Rysunek 55 Wykres funkcji $\boldsymbol{\pi(x)}$ z 1 poprawką 65
Rysunek 56 Wykres funkcji $\boldsymbol{\pi(x)}$ z 2 poprawkami 65
Rysunek 57 Wykres funkcji $\boldsymbol{\pi(x)}$ z 4 poprawkami 66
Rysunek 58 Wykres funkcji z π(x) z 50 poprawkami 66
Rysunek 59 Wykres funkcji $\boldsymbol{\pi(x)}$ z 200 poprawkami 66

15. Bibliografia

1. https://www.3blue1brown.com/
2. https://www.youtube.com/c/Mathologer/
3. Ian Stewart *Wielkie Problemy Matematyczne* wyd. Prószyński i Spółka
4. https://www.mersenne.org/
5. Wykład Tomasza Milera z cyklu #Naukanażywo
6. https://mathworld.wolfram.com/
7. oeis.org
8. https://wikipedia.org
9. K. Kłaczkow, M. Kurczab, E. Świda, *Analiza matematyczna dla licealistów*, Oficyna Edukacyjna 2002
10. Tom Davis, *Pick's Theorem*, 2003. (http://www.geometer.org/mathcircles/pick.pdf)

16. Spis treści

1. Wstęp 5
2. Używane pojęcia 8
3. Trochę historii 12
4. Zacznijmy od początku 14
5. Rodzaje liczb pierwszych 23
6. Twierdzenie Fermata o Dwóch Kwadratach 27
7. $Liczba\ \pi$ a liczby pierwsze 32
8. Gęstość liczb pierwszych 41
9. Spirala Ulama 47
10. Wzór Pentagonalny Eulera 55
11. Hipoteza Riemanna 59
12. Kryptografia 67
13. Zakończenie 70
14. Spis Rysunków 71
15. Bibliografia 72
16. Spis treści 73