The Congruent Number Theorem

Gymnasiearbete i matematik vid Vetenskapens Hus

Sjoerd Wijnand de Vries, Elin Ottergren

Hösten 2024

VETENSKAPENS HUS

Innehåll

Träff 1 - 11/9 Introduktion

Vi enades om att ses på onsdagar kl 16:30. Det här dokumentet kan användas av alla som vill lära sig att skriva med LaTeX för att hjälpas åt med minnesanteckningar från träffarna. Förslagsvis kan anteckningsansvar fördelas i någon form av turordning.

Kongruenta tal

Definition 0.0.1. *Inom talteori kallas ett postivit heltal för ett kongruent tal om det är arean av en rätvinklig triangel med rationella sidlängder.*

Mer generellt kan vi utöka defintionen till att också gälla alla rationella tal som uppfyller denna egenskap. Listan över kongruenta heltal börjar med:

5, 6, 7, 13, 14, 15, 20, 21, 22, 23, 24,
28, 29, 30, 31, 34, 37, 38, 39, 41, 45, 46,
47, 52, 53, 54, 55, 56, 60, 61, 62, 63, 65,
69, 70, 71, 77, 78, 79, 80, 84, 85, 86, 87,
88, 92, 93, 94, 95, 96, 101, 102, 103, 109,
110, 111, 112, 116, 117, 118, 119, 120...

Fyll på med era egna anteckningar om det ni vet om kongruenta tal...

Att bestämma vilka de kongruenta talen är är det som kallas för The Congruent Number Problem"eller ibland CNP som förkortning. Matematisk teori som kan bli aktuell för att angripa detta problem är:

- mängdlära
- gruppteori
- geometri
- modulär aritmetik
- elliptiska kurvor

Problemet har en lång historia och beksrevs redan för omkring 2000 år sedan inom arabisk kultur och vi kunde se att det sättet att se problemet är ekvivalent med vår beskrivning här ovan med hjälp av trianglar. Vi kom även in på hur den mer moderna matematiken som beskriver elliptiska kurvor kan vara ytterligare ett sätt att se på detta problem. Fyll på avsnittet med det ni har antecknat under träffen...

Ett exempel på hur bevis och satser kan skrivas

Vi ska nu visa att:

Sats 0.1. $\sqrt{2}$ *är inte ett rationellt tal*

Bevis. Antag att $\sqrt{2}$ är rationellt (för att komma till en motsägelse). Då kan vi skriva $\sqrt{2}$ som

$$\sqrt{2} = \frac{p}{q},\ p, q \in \mathbb{Z}$$

Där p och q saknar gemensamma delare (vi kan förkorta bråket så långt det går). Det ger oss att:

$$2q^2 = p^2$$

och således måste p^2 vara delbart med 2. Det innbär att p måste vara ett jämt tal (eftersom kvadraten av ett udda tal alltid också blir udda). Därför kan p skrivas som $p = 2k$ för något heltal k. Vi sätter in detta i vårt uttryck och får:

$$2q^2 = (2k)^2 = 4k^2$$

Nu kan vi dela båda sidor med 2 och få att

$$q^2 = 2k^2$$

Därför är även q ett jämnt tal (med samma argument som för p). Vi har nu visat att både p och q är delbara med 2 men vi antog att de saknade gemensamma delare och vi har därför en motsägelse. □

Aritmetikens fundamentalsats

säger att varje heltal större än 1 har en unik faktorisering eller primtalsfaktorisering, säger att varje heltal större än 1 kan entydigt delas upp som en produkt av primtal, bortsett från faktorerna ordning. Med andra ord, oavsett hur du faktorisera ett heltal större än 1, kommer du alltid att få samma uppsättning primtal, även om de kan vara ordnade i olika följd.

Definition 0.1.1. *Ett heltal n kallas prima om dess enda delare är ± 1 och $\pm n$.*

Sats 0.2 (Unik primtalsfaktorisering)**.** *Varje heltal är en entydig (unik upp till tecken) produkt av primtal.*

Fortsätt detta avsnitt genom att bevisa att $\sqrt{2}$ inte är ett rationellt tal genom att använda unik primtalsuppdelning...

Ekvivalenta formuleringar

Det finns en del egenskaper hos tal som man kan undersöka som har visat sig vara ekvivalenta med att vara ett kongruent tal.

Sats 0.3 (Araberna ($\sim$900 e.Kr.))**.** *Araberna undrade vilka heltal n som har egenskapen att det finns ett rationellt tal x så att $x^2 - n$ och $x^2 + n$ båda är rationella tal. Detta är ekvivalent med att säga att n är ett kongruent tal.*

Bevis. Låt n vara ett kongruent tal. $\implies \exists a, b \in \mathbb{Q} : \frac{ab}{2} = n \ \ \& \ \ \sqrt{a^2 + b^2} \in \mathbb{Q} \implies t = \frac{\sqrt{a^2 + b^2}}{2} \in \mathbb{Q}$. Vi bevisar att detta leder till att n har egenskapen arberna udnersökte genom att bevisa att detta leder till att det finns rationella s och u så att $u^2 - t^2 = n, |s^2 - t^2| = n$. Låt $u = \frac{a+b}{2}$ och $s = \frac{a-b}{2}$. Då får vi:

$$u^2 - t^2 = \frac{a^2 + 2ab + b^2}{4} - \frac{a^2 + b^2}{4} = \frac{ab}{2} = n$$

och

$$|s^2 - t^2| = |\frac{a^2 - 2ab + b^2}{4} - \frac{a^2 + b^2}{4}| = |\frac{-ab}{2}| = n$$

vilket bevisar implikation åt ena hållet. (TODO: bevisa implikation åt andra hållet) □

Sats 0.4 (Elliptiska kurvor)**.** *Att n är ett kongruent tal är ekvivalent med att säga att den Elliptiska kurvan $E_n : y^2 = x^3 - n^2x$ har positiv rank.*

Liten ordlista från träff 1

- Integer - heltal
- Rational number - rationella tal (bråktal)
- Prime number - primtal
- $\mathbb{N}$ symbol för de naturliga talen
- $\mathbb{Z}$ symbol för heltalen
- $\mathbb{Q}$ symbol för de rationella talen
- $\mathbb{C}$ symbol för de rationella talen
- $\in$ visar att något tillhör en mängd
- $P \implies Q$ visar logiskt implikation (om P så Q)
- s.t. så att (such that)

Fyll på listan med ord, begrepp eller symboler som ni undrar över!

Träff 2 - 18/9 Mängdlära 1

Frågor om förra träffen

Sats 0.5 (Problem 1)**.** $\sqrt{n} \in \mathbb{Q} \iff n = m^2$ *där* $m \in \mathbb{Z}$

Bevis. Om $n = m^2$ är ju $\sqrt{n} = \sqrt{m^2} = m$ vilket är rationellt då m är ett heltal.

Antag nu att $\sqrt{n}$ är rationellt alltså finns det heltals p och q ($q \neq 0$) så att $\sqrt{n} = \frac{p}{q}$.

$$\sqrt{n} = \frac{p}{q} \implies nq^2 = p^2$$

Låt oss nu primtalsfaktorisera p och q på så sätt att

$$p = w_1^{e_1} w_2^{e_2} ... w_k^{e_k} \implies p^2 = w_1^{2e_1} w_2^{2e_2} ... w_k^{2e_k}$$

$$q = W_1^{E_1} W_2^{E_2} ... W_k^{E_m} \implies q^2 = W_1^{2E_1} W_2^{2E_2} ... W_k^{2E_m}$$

där alla w_i och W_i är distinkta primtal och alla e_i och E_i är heltal över 0. Vi vet från sats 0.2 att för varje W_i måste de finnas ett j så att $W_i = w_j$ annars skulle ha en till primtalsfakorisering som innehåller W_i. Vi kan därför nummerera våra primtal så att $W_i = w_i$ för alla i det finns ett W_i, vi vet då även $m < k$. Detta ger:

$$q^2 = w_1^{2E_1} w_2^{2E_2} ... w_m^{2E_m}$$

alltså får vi:

$$\frac{p^2}{q^2} = \frac{w_1^{2e_1} w_2^{2e_2} ... w_k^{2e_k}}{w_1^{2E_1} w_2^{2E_2} ... w_k^{2E_m}} = w_1^{2v_1} w_2^{2v_2} ... w_k^{2v_k} = n \text{ där } v_i \in \mathbb{N} \implies n = (w_1^{v_1} w_2^{v_2} ... w_k^{v_k})^2$$

$$\implies n = m^2$$

□

Mängder

Definition 0.5.1. *En mängd är en kollektion av element där ordning eller upprepade element inte spelar någon roll. Ett element kan vara vad som helst.*

Mängder skrivs med dina element innan för { och } till exempel så är {1,3,2} mängden som innehåller elementen 1,2 och 3. Den tomma mängden {} skrivs $\emptyset$.

Notation

Det finns en del olika sätt att notera mängder. Man kan som vi skrev tidigare skriva alla element mellan { och }. Man även skriva till exempel $\{x \in \mathbb{Z} \mid x > 0\}$ vilket betyder mängden av tal x i $\mathbb{Z}$ som är större än 0. Detta skulle alltså vara mängden {1,2,3,...}.

För att beskriva ett intervall av reala tal finns det notation för det också. Vill man till exempel beskriva intervallet av alla tal från och med 0 till och med 1 kan man skriva [0,1].

Operationer

- Medlemskap ($\in$)
  $x \in S \iff x$ finns i mängden S
- Innehållande ($\subset$ eller $\subseteq$)
  $S \subseteq T \iff$ alla s i S finns i T
  $S \subset T \iff S \subseteq T$ och $S \neq T$
- Kardinalitet ($\#S$ eller $|S|$)
  $\#S = |S| =$ antalet element i S
- Power set
  $\mathcal{P}(S) = \{T \mid T \subseteq S\} =$ mängden av alla delmängder av S

- Produkt ($\times$)
  $S \times T = \{(s,t) \mid s \in S, t \in T\}$
  där (s,t) är en lista där ordning spelar roll dvs: (s,t) $\neq$ (t,s).
  $S^n = \underbrace{S \times S \times ... \times S}_{n \text{ gånger}}$
- Union ($\cup$)
  $S \cup T = \{e \mid e \in S \text{ eller } e \in T\}$
- Snitt ($\cap$)
  $S \cap T = \{e \mid e \in S \text{ och } e \in T\}$

Funktioner

En funktion från mängden S till mängden T, $f : S \to T$, är en regel som tilldelar varje element i S till ett och endast ett element i T.

Definition 0.5.2. *En funktion, $f : S \to T$, är delmängd $\Gamma_f \subseteq S \times T$ så att det existerar ett och endast ett t för all s så att $(s,t) \in \Gamma_f$.*

Om vi till exempel har en funktion $f : \{*\} \to T$ är det samma sak som att välja ett element från mängden T. På samma sätt är en funktion $f : \{1,2\} \to T$ ekvivalent med att välja två element, kan vara samma, från T och numerera dem 1-2.

Definition 0.5.3. *Förbilden f^{-1} av en funktion $f : S \to T$ är en funktion $f^{-1} : T \to \mathcal{P}(S)$. Förbilden av f av ett t är alla s så att $f(s) = t$ dvs $f^{-1}(t) = \{s \in S \mid f(s) = t\}$.*

Definition 0.5.4. *En funktion $f : S \to T$ är injektiv om $\#f^{-1}(s) \leq 1$ för alla $s \in S$.*

Definition 0.5.5. *En funktion $f : S \to T$ är surjektiv om $\#f^{-1}(s) \geq 1$ för alla $s \in S$.*

Definition 0.5.6. *En funktion är bijektiv om den är både injektiv och surjektiv.*

Sats 0.6. *Det finns alltid en bijektion från en mängd S till samma mängd S.*

Bevis. Mängden $\Gamma_f = \{(s,s) \in S^2\}$ är en bijektiv funktion då $f^{-1}(s) = \{s\} \implies \#f^{-1}(s) = 1$ för alla s. □

Sats 0.7. *f är injektiv är ekvivalent med att säga att f har egenskapen $f(s) = f(s') \iff s = s'$.*

Bevis. Låt $f : S \to T$ vara en injektiv funktion. Vi vet per definition att $\#f^{-1}(t) \leq 1$ för alla $t \in T$. Då vet vi att för alla $s \in S$ att $s \in f^{-1}(f(s))$ så alltså har vi $f^{-1}(f(s)) = \{s\}$. Så om $f(s) = f(s') \iff s = s'$.

Åt andra hållet så antar vi att vi har en funktion $f : S \to T$ så att $f(s) = f(s') \iff s = s'$. Då vet vi att $f^{-1}(f(s)) = \{s\}$. Om $t \in T$ inte kan skrivas $t = f(s)$ så får vi $f^{-1}(s) = \emptyset$ alltså $\#f^{-1}(t) = 0$, och om $t \in T$ kan skrivas $t = f(s)$ så får vi $f^{-1}(t) = \{s' \in S \mid f(s') = t\} = \{s' \in S \mid f(s') = f(s)\} = \{s\}$ alltså $\#f^{-1}(t) = 1$. Så f är injektiv. □

Sats 0.8. *Let $f : S \to T$ be a function between finite sets. If f is bijective, then $\#S = \#T$.*

The converse also holds: if S and T are finite sets such that $\#S = \#T$, then there is a bijection $S \to T$. However, if S and T are both infinite, there may not be a bijection between them. We will not go into this further in this course, but it might be something to explore in your gymnasiearbete if you are interested in the mathematical meaning of infinity, and different kinds of infinities.

Träff 3 - 25/9 Mängdlära 2 & Gruppteori 1

Composition of functions

Given two functions $f : S \to T$ and $g : T \to U$, we can *compose* them to get a new function $S \to U$. The new function is denoted by $g \circ f$, and is defined by the rule

$$(g \circ f)(s) = g(f(s)), \qquad s \in S.$$

We call $g \circ f$ the *composition* of f and g. The notation $g \circ f$ can be pronounced "g composed with $f$ör "g after f".

We can rephrase this as follows. For sets S and T, define a new set

$$\mathrm{Hom}(S, T) = \{\text{functions } S \to T\}.$$

Then we get, for any three sets S, T and U, a function

$$\circ : \mathrm{Hom}(S, T) \times \mathrm{Hom}(T, U) \longrightarrow \mathrm{Hom}(S, U)$$

sending $(f, g) \mapsto g \circ f$.

(Exercise: if $\#S = n$ and $\#T = m$ for $n, m \in \mathbb{N}$, what is the cardinality of $\mathrm{Hom}(S, T)$?)

Equivalence relations

The next thing we will formalise is the notion of identifying elements of a set.

Definition 0.8.1. *Let S be a set. A subset $R \subseteq S \times S$ is called a* relation *on S. A relation is called*

1. reflexive *if for all $s \in S$, we have $(s, s) \in R$;*

2. symmetric *if for all* $s, s' \in S$, $(s, s') \in R$ *implies* $(s', s) \in R$;

3. transitive *if for all* $s, s', s'' \in S$, $(s, s') \in R$ *and* $(s', s'') \in R$ *implies* $(s, s'') \in R$.

A relation which is reflexive, symmetric and transitive is called an equivalence relation *on* S.

Informally, we usually write $s \sim s'$ to mean $(s, s') \in R$, and think about this as s being related to s'. A general relation on S can behave in counterintuitive ways, but an equivalence relation behaves "nicely": when $\sim$ is an equivalence relation, we can truly think of related elements as being the same. We will formalise this later by the notion of a quotient by an equivalence relation.

Exempel 0.1. *1. Let S be any set. Then the identity function, seen as the subset $\{(s, s) \mid s \in S\}$ of $S \times S$, is an equivalence relation. For this relation, an element in S is related only to itself, so it is rather boring.*

2. *Let S be any set. Then $S \times S$ is an equivalence relation. That is, $s \sim s'$ for any two elements $s, s' \in S$, so all elements are related.*

3. *Let $S = \{f : \mathbb{R} \to \mathbb{R}\}$, and define $f \sim g$ if and only if $f(0) = g(0)$. This is an equivalence relation. To prove this, we just have to check the three properties. Reflexivity: for any function f, we have $f(0) = f(0)$, so $f \sim f$. Symmetry: if $f \sim g$, then $f(0) = g(0)$, so $g(0) = f(0)$, so $g \sim f$. Transitivity: if $f \sim g$ and $g \sim h$, then $f(0) = g(0) = h(0)$, so $f \sim h$.*

4. *Let $S = \{X \subseteq \mathbb{Z} \mid \#X = \infty\}$, and $X \sim Y \iff \#(X \cap Y) = \infty$. Then $\sim$ is reflexive and symmetric, but not transitive, so $\sim$ is not an equivalence relation.*

5. *Let $S = \mathbb{Z} \times (\mathbb{Z} \setminus \{0\})$, and define $\sim$ via $(a, b) \sim (c, d) \iff ad = bc$. This is an equivalence relation.*

We now come to the main application of equivalence relations: they allow us to divide out" by the relation. More precisely, if S is a set and $\sim$ is an equivalence relation on S, then we can construct a new set $S/\sim$ whose elements are the elements of S, except if two elements are related in S, they become the same element in $S/\sim$.

Definition 0.8.2. *Let $\sim$ be an equivalence relation on S. For $s \in S$, define the* equivalence class *of s to be*

$$[s] := \{s' \in S \mid s' \sim s\}.$$

Define the quotient of S by $\sim$ *to be the set*

$$S/\sim := \{[s] \mid s \in S\}.$$

Sats 0.9. *The map* $\pi : S \to S/\sim$ *defined by* $\pi(s) = [s]$ *is surjective, and* $\pi^{-1}([s]) = [s]$ *for all* $[s] \in S/\sim$.

Gruppteori

Groups are fundamental objects in mathematics. They form the basis of many other constructions, and they are also used in physics, chemistry, and biology. To motivate the definition of groups, we need to understand what is meant by "endowing a set with additional structure". A set is, by definition, nothing more than a collection of elements. But sometimes we have an idea of what a set looks like: for instance, $\mathbb{R}^2$ can be visualised as the xy-plane. In our minds, the elements $(0,0)$ and $(0,1)$ are closer together than the elements $(0,0)$ and $(5,5)$, but set-theoretically there is no interplay between any of these elements. What we can do is endow the set $\mathbb{R}^2$ with a certain "distance function" $d : \mathbb{R}^2 \times \mathbb{R}^2 \to \mathbb{R}_{\geq 0}$, where $d(p_1, p_2)$ gives the distance between p_1 and p_2. (Can you write down an explicit formula for this distance function in terms of the coordinates of p_1 and p_2?)

This distance function formalizes our intuition about $\mathbb{R}^2$ as a space, rather than a set, and when one wants to study $\mathbb{R}^2$ as a space, it is convenient to study the pair $(\mathbb{R}^2, d)$ – i.e. to always see $\mathbb{R}^2$ not just as a set, but as a set endowed with the additional structure given by the function d.

This is perhaps still a bit vague, but the idea is that sets themselves don't contain much information. Here is another example, which is more in the spirit of group theory. Consider the set of integers $\mathbb{Z}$. This is a set of numbers, and when we think about numbers, we never just think about them as being arbitrary elements. Rather, there are relations between different numbers, for example $1 + 1 = 2$. However, when we view $\mathbb{Z}$ as a set, mathematically there is no addition rule. If we want to have the relation $1 + 1 = 2$ between the elements 1 and 2 in $\mathbb{Z}$, we need to endow $\mathbb{Z}$ with the additional structure of an addition rule.

So how do we do this? Well, addition is just a function $\mathbb{Z} \times \mathbb{Z} \to \mathbb{Z}$: it takes as input two integers and gives one integer as an output. But it's not just any function -

it satisfies some special properties. For example, $a + b = b + a$ for all $a, b \in \mathbb{Z}$. Another property is that $a + 0 = a$ for all $a \in \mathbb{Z}$. In fact, the pair $(\mathbb{Z}, +)$ is an example of an *abelian group*. There are so many interesting mathematical objects which come equipped with a certain "addition rule" or "multiplication rule" that mathematicians decided to create a name for them, and this is what a group is.

Definition 1.0.1. *Let S be a set. A* binary operation *on S is a function $\cdot : S \times S \to S$. We write $\cdot((s, s'))$ as $s \cdot s'$ or simply ss'.*

Certainly, any kind of multiplication or addition rule will be a binary operation, but there are many binary operations which do not behave nicely. Let's see some examples.

Exempel 1.1.

1. *Let $S = \mathbb{N}$ and $\cdot = +$, the usual addition of natural numbers. This is a binary operation: for $(n, m) \in \mathbb{N} \times \mathbb{N}$, we have $n + m \in \mathbb{N}$.*
2. *The previous example also works with $\mathbb{N}$ replaced by $\mathbb{Z}, \mathbb{Q}, \mathbb{R}$ or $\mathbb{C}$. On all these sets, multiplication defines a second binary operation.*
3. *Let S be any set, and let $\pi_1 : S \times S \to S$ be the function $\pi_1(s, s') = s$.*
4. *Let $S = Hom(X, X)$ for some set X, i.e. $S = \{f : X \to X\}$. Then composition defines a binary operation $S \times S \to S$.*
5. *Let S be any set. Then union and intersection define binary operations on the power set $\mathcal{P}(S)$.*

Definition 1.0.2. *Let $\cdot : S \times S \to S$ be a binary operation.*

1. *We say $\cdot$ is associative if for all $a, b, c \in S$, we have $a \cdot (b \cdot c) = (a \cdot b) \cdot c$.*
2. *We say $\cdot$ is commutative if for all $a, b \in S$, we have $a \cdot b = b \cdot a$.*

Definition 1.0.3. *A* group *is a pair $(G, \cdot)$, where G is a set and $\cdot$ is a binary operation on G, such that the following conditions are satisfied:*

(G1) The binary operation is associative;

(G2) There exists an element $e \in G$ such that $e \cdot g = g = g \cdot e$ for all $g \in G$;

(G3) For every $g \in G$, there exists some $g^{-1} \in G$ such that $g \cdot g^{-1} = e = g^{-1} \cdot g$.

If the binary operation is also commutative, we say G is an abelian group.

We call $e \in G$ the *identity element* of G. If $g \in G$ is any element, we call g^{-1} the *inverse* of g. Note that $e^{-1} = e$.

The cardinality or size of a group G is usually called the *order* of G, which is still denoted by $|G|$ or $\#G$.

As an exercise, you should check which binary operations from Example 1.1 endow S with a group structure. For instance, $(\mathbb{N}, +)$ is not a group: there are no inverses. On the other hand, $(\mathbb{Z}, +)$ is a group. (What is the identity element?) None of the sets $\mathbb{N}, \mathbb{Z}, \mathbb{Q}, \mathbb{R}$, and $\mathbb{C}$ are groups under multiplication, but one can easily fix it in the last three examples by removing zero (the only element which would not have an inverse).

Träff 4 - 2/10 Gruppteori 2

Examples

Let's have a look at some important classes of groups.

I.1 Cyclic groups

Let $n \in \mathbb{Z}_{\geq 1}$. Define the *cyclic group of order* n to be the set

$$C_n = \{\zeta \in \mathbb{C} \mid \zeta^n = 1\}.$$

These are the n-th roots of unity in $\mathbb{C}$. They all lie on the unit circle and form a group under multiplication.

More explicitly, we can describe the set C_n by fixing a primitive n-th root of unity ξ (this means that $\xi^n = 1$, but $\xi^m \neq 1$ for $1 \leq m < n$), for instance $\xi = e^{2\pi i/n}$, and defining

$$C_n = \{\xi^m \mid 1 \leq m \leq n\}.$$

Check for yourself that C_n is a group under multiplication with identity element 1. (You may assume that multiplication of complex numbers is associative.)

I.2 The integers modulo n

Let $n \geq 1$. Define an equivalence relation on $\mathbb{Z}$ as follows. For $x, y \in \mathbb{Z}$, let

$$x \sim y \iff n \mid x - y.$$

Here $n \mid x - y$ means that n divides $x - y$; that is, there exists an integer k such that $kn = x - y$. (Check that this is indeed an equivalence relation.) Then $\mathbb{Z}/n\mathbb{Z} := \mathbb{Z}/\sim$ are the *integers modulo* n. They form a group under addition, as we will see in a minute.

As a set, we have

$$\mathbb{Z}/n\mathbb{Z} = \{[0], [1], \ldots, [n-1]\},$$

where e.g. $[0]$ denotes the equivalence class of 0. To see this, we can use division with remainder: any $m \in \mathbb{Z}$ can be written in a unique way as

$$m = kn + r, \qquad k \in \mathbb{Z}, \qquad 0 \leq r < n - 1.$$

Therefore $m - r = kn$, so by definition $m \sim r$, which means $[m] = [r]$. What this tells us is that there are at most n elements in the set $\mathbb{Z}/n\mathbb{Z}$, and it is not hard to see that the listed elements $[0], [1], \ldots, [n-1]$ are all distinct. Explicitly, we have

$$[r] = \{m \in \mathbb{Z} \mid m = r + kn \text{ for some } k \in \mathbb{Z}\}.$$

To be even more concrete, let's take the example $n = 3$. Then

$$\begin{aligned}
[0] &= \{\ldots, -6, -3, 0, 3, 6, 9, \ldots\}, \\
[1] &= \{\ldots, -5, -2, 1, 4, 7, 10, \ldots\}, \\
[2] &= \{\ldots, -4, -1, 2, 5, 8, 11, \ldots\}.
\end{aligned}$$

We see that every integer lies in at least one of these sets and the sets are disjoint, so $\mathbb{Z}/3\mathbb{Z} = \{[0], [1], [2]\}$.

We now define a binary operation on $\mathbb{Z}/n\mathbb{Z}$ as follows: for $x, y \in \mathbb{Z}$, we set

$$[x] + [y] := [x + y].$$

This defines a group law essentially because $(\mathbb{Z}, +)$ is a group. For instance, we can prove associativity as follows:

$$[x]+([y]+[z]) = [x]+[y+z] = [x+(y+z)] = [(x+y)+z] = [x+y]+[z] = ([x]+[y])+[z].$$

The identity element is $[0]$, and the inverse of $[r]$ is $[-r] = [n - r]$.

Notationally, it is often easier to omit the brackets and instead write equations as $\ldots \equiv \ldots \pmod{n}$ to denote the fact that we work in $\mathbb{Z}/n\mathbb{Z}$ rather than $\mathbb{Z}$. For example, we can write

$$2 + 2 \equiv 1 \pmod{3}$$

to mean that $[2] + [2] = [1]$ in $\mathbb{Z}/3\mathbb{Z}$. (This is because $[2] + [2] = [4] = [1]$, since $4 = 1 + 1 \cdot 3$.)

I.3 The Klein four group

Consider the set $K_4 = \{e, a, b, c\}$ with the following binary operation:

$$\begin{gathered}
e \cdot x = x \quad \forall x \in K_4, \\
a \cdot b = c, \qquad a \cdot c = b, \qquad b \cdot c = a, \\
a^2 = e, \qquad b^2 = e, \qquad c^2 = e.
\end{gathered}$$

and we moreover impose that $\cdot$ is commutative, which then determines all the other multiplications.

You can prove that $(K_4, \cdot)$ is associative by hand, but it's not a very inspiring exercise. Instead, we will see later that this is a group. (Of course, the identity is e, and every element it its own inverse.)

I.4 Dihedral groups

For every $n \geq 3$, there is a group D_{2n} called the *dihedral group of order* $2n$. It is defined as the group of symmetries of a regular n-gon. For example, if $n = 3$, the group D_6 consists of the symmetries of an equilateral triangle. Here a *symmetry* is defined to be a function from the n-gon to itself which sends vertices to vertices and which preserves distances between points. Since functions can be composed, so can symmetries, and this composition law defines a group structure on D_{2n}.

The dihedral group is always generated by a rotation σ over $360/n$ degrees and a reflection τ through an axis of symmetry; in other words, all other symmetries are obtained by composing these symmetries in some order. Then $\sigma^n = \mathrm{id}$ and $\tau^2 = \mathrm{id}$, but there are also other elements such as $\sigma\tau$, $\tau\sigma^2\tau\sigma^{-1}$, etc. Play around with this a bit and see if you can prove that D_{2n} always has order $2n$.

I.5 Symmetric groups

Let X be a finite set. Define the *symmetric group* $(\mathrm{Sym}(X), \circ)$ whose elements are bijections $f : X \to X$ and whose binary operation is given by composition. This is clearly associative: the functions $f \circ (g \circ h)$ and $(f \circ g) \circ h$ both send $x \in X$ to $f(g(h(x)))$, so they define the same function. Moreover, the identity function is a bijection, and every bijection has an inverse, so this really is a group.

If $X = \{1, 2, \ldots, n\}$ is the set of the first n positive integers, then $\mathrm{Sym}(X)$ is usually denoted by S_n and called the *symmetric group on n letters*. It has order $n!$ (exercise).

Elements of S_n are conveniently written as *products of cycles*: e.g. the cycle $(12) \in S_2$ is the bijection sending $1 \mapsto 2$ and $2 \mapsto 1$, and the product of cycles $(152)(34)$ in S_5 is the bijection sending $1 \mapsto 5 \mapsto 2 \mapsto 1$ and $3 \mapsto 4 \mapsto 3$.

I.6 Braid groups

For any $n \geq 1$, let B_n denote the *braid group on n strands*, defined as follows. The elements of B_n are configurations of strands connecting n points $x_1, \ldots, x_n$ to n

points $y_1, \ldots, y_n$, in some order. The strands are allowed to go over and under each other, but they can't loop backwards: if $x_1, \ldots, x_n$ are on the left and $y_1, \ldots, y_n$ are on the right, then every strand has to move from left to right. Two configurations of strands are called equivalent if they can be continuously deformed into one another, and in this case they define the same element in the braid group.

The composition law is given by connecting strands: if $x_1, \ldots, x_n$ are connected to $y_1, \ldots, y_n$ and $y_1, \ldots, y_n$ are connected to $z_1, \ldots, z_n$, then we get an induced configuration of strands between $x_1, \ldots, x_n$ and $z_1, \ldots, z_n$. Pictures help: see for example https://en.wikipedia.org/wiki/Braid_group.

Some examples: if $n = 1$, up to deformation there is only the straight braid between two points x_1 and y_1, so $B_1 = \{\mathrm{id}\}$, the trivial group. If $n \geq 2$, the braid group is infinite: for instance, denote by t the braid configuration where the top strand goes over the lower strand, connecting x_1 to y_2 and x_2 to y_1. Then t^{-1} is the braid configuration where the top strand goes under the lower strand, as the configuration $t \circ t^{-1}$ can be deformed into two straight strands. But t^2 can not be continuously deformed into the identity, as the strands have become "braided". In general, t^n is not the identity for any $n \geq 1$, so B_2 is infinite. This also implies that B_n is infinite for any $n > 2$, since we can reconfigure the first two strands in infinitely many ways while keeping the other $n - 2$ strands fixed.

Products

If $(G, \cdot_G)$ and $(H, \cdot_H)$ are groups then $G \times H$ is a group with a binary operation $\cdot_G \times \cdot_H$. That is to say, the binary operation $(G \times H) \times (G \times H) \to (G \times H)$ is given by $((g_1, h_1,)(g_2, h_2)) \mapsto (g_1 \cdot_G g_2, h_1 \cdot_H h_2)$.

The identity element for the new group is (e_G, e_H) and the inverse is $(g, h)^{-1} = (g^{-1}, h^{-1})$.

With this method we can create a lot of new groups to study, for example $B_n \times D_{2m}$ for any $n \geq 1$ and $m \geq 3$.

Sats 1.1. *Let $g, h_1, h_2 \in G$. Suppose $gh_1 = gh_2$. Then $h_1 = h_2$.*

Bevis. We assume $gh_1 = gh_2$ and multiply both sides with g^{-1} on the left. This gives

$$g^{-1} \cdot (gh_1) = g^{-1} \cdot (gh_2) \iff (g^{-1}g)h_1 = (g^{-1}g)h_1 \iff eh_1 = eh_2 \iff h_1 = h_2.$$

□

Note that the proof used all three properties of groups: associativity, identity and inverses. If any one of these is not satisfied, we can in general not cancel out g in the equation $gh_1 = gh_2$. Note also that it is important that g is on the left: in the equation $gh_1 = h_2 g$, we can *not* cancel out g unless the group is abelian.

Subgroups

Definition 1.1.1. *Let G be a group and $H \subseteq G$. Then H is a* subgroup *of G if:*

- *1) For every $h, h_1 \in H$ $h \cdot h_1 \in H$*
- *2) $e \in H$*
- *3) $h \in H$ and $h^{-1} \in H$*

Definition 1.1.2. *Let $g \in G$. Then define* the subgroup generated by g *to be*

$$\langle g \rangle = \{g^n \mid n \in \mathbb{Z}\} \subseteq G.$$

If $\langle g \rangle = G$, we say g is a generator *of G.*

Exercise: show that $\langle g \rangle$ is always a subgroup. It certainly need not be infinite, even though it looks that way: for instance, if $e \in G$ is the identity, then $\langle e \rangle = \{e\}$ is the trivial subgroup.

Definition 1.1.3. *Let $g \in G$. The* order *of g is min $\{n \in \mathbb{N} | g^n = e\}$ or ∞ if no such n exists. The order of g is denoted ord(g).*

The following proposition shows that the term "orderdoes not clash with the order of a group: indeed, the order (as defined above) of g is equal to the order (= cardinality) of the subgroup generated by g.

Proposition 1.1.1. *For $g \in G$ we have ord$(g) = |\langle g \rangle|$.*

Bevis. We prove this only when both sides are finite; you can prove as an exercise that if one side is infinite, then so is the other.

Let $\langle g \rangle$ be finite, say $|\langle g \rangle| = n$. We will show that $n = ord(g)$. This means that we need to show that $g^n = e$ and $g^m \neq e$ for any $m \in \mathbb{N}$, $m < n$

$|\langle g \rangle| = n \iff |\{g^m \mid m \in \mathbb{Z}\}| = n$

Suppose that $m \in \mathbb{N}$ satisfies $g^m = e$. Then:

for any $k \in \mathbb{Z}$, write $k = k' \cdot m + r$, $0 \leq r < m$

So $g^k = g^{k'm+r} = g^{k'm} + g^r = (g^m)^{k'} \cdot g^r = e^{k'} \cdot g^r = e \cdot g^r = g^r$

So $\langle g \rangle \subseteq \{g^0 = e, g^1, ..., g^{m-1}\}$

So $|\langle g \rangle| \leq m$

So $g^m = e \implies m \geq n$

Since $g^{\text{ord}(g)} = e$ by definition, this shows that $n = |\langle g \rangle| \leq ord(g)$.

To show equality, we need to show that the $\text{ord}(g)$ elements $g^0, \ldots, g^{\text{ord}(g)-1}$ are distinct.

So we need to show that $g^i \neq g^j$ for any $i \neq j, 0 \leq i, j < \text{ord}(g)$.

If $g^i = g^j, i < j$ then

$e = g^{j-i} \implies j - i \geq \text{ord}(g)$ or $j - i = 0$

The first case can't happen if i and j are both less than $\text{ord}(g)$, so we get $j - i = 0$, so $j = i$, which contradicts the assumption that $i < j$. Therefore $g^i = g^j$ is impossible and so $ord(g) = |\langle g \rangle|$. $\square$

Träff 5 - 16/10 Gruppteori 3

Let's see Proposition 1.1.1 in action.

Exempel 1.2.
1. $G = C_4 = \{1, i, -1, -i\}$. *In this case, we have*

$$\begin{aligned} ord(1) &= 1 && \text{since } 1^1 = 1; \\ ord(i) &= 4 && \text{since } i^1 = i \neq 1; i^2 = -1 \neq 1; i^3 = -i \neq 1; i^4 = 1; \\ ord(-1) &= 2 && \text{since } (-1)^1 = -1 \neq 1; (-1)^2 = 1; \\ ord(-i) &= 4 && \text{since } (-i)^1 = -i \neq 1; (-i)^2 = -1 \neq 1; (-i)^3 = i \neq 1; (-i)^4 = 1. \end{aligned}$$

Similarly, we have $\langle 1 \rangle = \{1\}$, $\langle i \rangle = C_4$, $\langle -1 \rangle = \{1, -1\} = C_2$, $\langle -i \rangle = C_4$. *In particular,* C_4 *has two generators, namely* i *and* $-i$.

2. $G = \mathbb{Z}/4\mathbb{Z} = \{0, 1, 2, 3\}$, *where we write* n *to mean the equivalence class* $[n]$, *for notational ease. Note that the group operation is now addition, so for* $g \in \mathbb{Z}/4\mathbb{Z}$ *and* $m \in \mathbb{Z}$, *we have* $g^m = g + g + \ldots + g = m \cdot g$. *The orders of the elements of* G *look as follows:*

$$\begin{aligned} ord(0) &= 1 && \text{since } 1 \cdot 0 = 0; \\ ord(1) &= 4 && \text{since } 1 \cdot 1 = 1 \neq 0; 2 \cdot 1 = 2 \neq 0; 3 \cdot 1 = 3 \neq 0; 4 \cdot 1 = 0; \\ ord(2) &= 2 && \text{since } 1 \cdot 2 = 2 \neq 0; 2 \cdot 2 = 0; \\ ord(3) &= 4 && \text{since } 1 \cdot 3 = 3 \neq 0; 2 \cdot 3 = 2 \neq 0; 3 \cdot 3 = 1 \neq 0; 4 \cdot 3 = 0. \end{aligned}$$

You can check by hand that 1 *and* 3 *are generators of* $\mathbb{Z}/4\mathbb{Z}$, *while* 0 *and* 2 *are not.*

3. $G = \mathbb{Z}$. *In this case,* $ord(0) = 1$ *and* $ord(n) = \infty$ *for any* $n \neq 0$. *Indeed,* $mn \neq 0$ *for any* $m \geq 1$ *if* $n \neq 0$. *The subgroup generated by* n *is*

$$\langle n \rangle = n\mathbb{Z} = \{m \in \mathbb{Z} \mid n \text{ divides } m\}.$$

In particular, 1 *and* -1 *are the only generators of* $\mathbb{Z}$.

4. $G = S_3 = \{e, (12), (13), (23), (123), (132)\}$. *The order of any 2-cycle* (ab) *is 2, and the order of any 3-cycle* (abc) *is 3. In particular,* G *has no generators (since a generator would have to have order* $\#G = 6$*).*

Looking at the above examples, you may note that the order of an element always divides the order of the group. This is always true. It follows from the following result.

Sats 1.2 (Lagrange's Theorem)**.** *Let G be a finite group and let H be a subgroup of G. Then $\#H$ divides $\#G$.*

The proof of Lagrange's theorem uses certain equivalence relations which turn out to behave especially nicely when dealing with finite groups. We will not cover the proof in the course, but let's simply note that the observation about the orders of elements follows directly:

Följdsats 1.2.1. *Let G be a finite group and let $g \in G$. Then $ord(g)$ divides $\#G$.*

Bevis. The subgroup $\langle g \rangle$ of G has order $\text{ord}(g)$ by Proposition 1.1.1, so we can apply Lagrange's theorem with $H = \langle g \rangle$. □

Homomorphisms

It has been said that in mathematics, it is more important how an object relates to other objects than what the object looks like itself. We can study how objects relate to each other through the functions between the objects. However, we don't just want any functions, but only those functions which respect the additional structure we have put on our sets. In the case of groups, we want the functions to take into account the group structure. This leads us to the notion of a homomorphism.

Definition 1.2.1. *Let $(G, \cdot_G)$ and $(H, \cdot_H)$ be groups. A* group homomorphism *from G to H is a function $\varphi : G \to H$ such that for all $g, g' \in G$, we have*

$$\varphi(g \cdot_G g') = \varphi(g) \cdot_H \varphi(g').$$

Usually we will not write the subscripts G and H in $\cdot_G$ and $\cdot_H$, since it should be clear from the context in which group the multiplication occurs, but sometimes the subscripts can serve as a helpful reminder.

The following objects will be helpful in the study of homomorphisms.

Definition 1.2.2. *Let $\varphi : G \to H$ be a homomorphism. The* image *of φ is the subset*

$$Im(\varphi) := \{\varphi(g) \mid g \in G\} \subseteq H.$$

The kernel *of φ is the subset*

$$Ker(\varphi) := \{g \in G \mid \varphi(g) = e\}.$$

Exempel 1.3.
1. Let G be any group. Then there is a unique function $\varphi : G \to \{e\}$ from G to the trivial group, and this is a group homomorphism: for any $g, g' \in G$, we have

$$\varphi(g \cdot g') = e = e \cdot e = \varphi(g) \cdot \varphi(g').$$

2. Let G be any group. Then the set of functions $f : \{e\} \to G$ is in natural bijection with G: for every element $g \in G$, there is a function $\{e\} \to G$ sending $e \mapsto g$. Out of these functions, there is only one homomorphism, namely the function $e \mapsto e_G$, where e_G denotes the identity element of G. (We will see in Lemma 1.2.1 that any homomorphism must preserve the identity element, and it is straightforward to verify that the function $e \mapsto e_G$ is indeed a homomorphism.)
3. Let G and H be any groups. Then the function $\varphi : G \to H$ sending $g \mapsto e$ for every $g \in G$ is a group homomorphism: indeed,

$$\varphi(g \cdot g') = e = e \cdot e = \varphi(g) \cdot \varphi(g') \text{ for all } g, g' \in G.$$

Note that φ is equal to the composition $G \to \{e\} \to H$ of the homomorphisms from examples 1 and 2.

The above examples are in some sense trivial. We will see more interesting examples of homomorphisms as the course progresses.

Hjälpsats 1.2.1. *Let $\varphi : G \to H$ be a homomorphism. Then*

1. *$\varphi(e_G) = e_H$, where e_G denotes the identity in G and e_H the identity in H;*
2. *For all $g \in G$, we have $\varphi(g^{-1}) = \varphi(g)^{-1}$.*

Bevis. 1. We will use the fact that φ is a homomorphism and that multiplication with the identity element is the identity map (i.e. $e_G \cdot_G g = g$ for all $g \in G$). We have

$$\varphi(e_G) = \varphi(e_G \cdot_G e_G) = \varphi(e_G) \cdot_H \varphi(e_G).$$

Multiplying both sides of the equation with $(\varphi(e_G))^{-1}$ (which we know exists because H is a group, and groups have inverses), we get

$$e_H = \varphi(e_G),$$

as we wanted.

2. On the one hand,
$$\varphi(g \cdot_G g^{-1}) = \varphi(g) \cdot_H \varphi(g^{-1}),$$
and on the other hand
$$\varphi(g \cdot_G g^{-1}) = \varphi(e_G) = e_H,$$
using part 1. Hence
$$\varphi(g) \cdot_H \varphi(g^{-1}) = e_H,$$
and multiplying this equation on the left by $\varphi(g)^{-1}$ gives
$$\varphi(g^{-1}) = \varphi(g)^{-1}.$$

□

We can now prove that images and kernels of homomorphisms are not just subsets, but subgroups.

Proposition 1.2.1. *Let $\varphi : G \to H$ be a homomorphism. Then $Im(\varphi)$ is a subgroup of H and $Ker(\varphi)$ is a subgroup of G.*

Bevis. We prove that $\mathrm{Ker}(\varphi)$ is a subgroup and leave $\mathrm{Im}(\varphi)$ as an exercise. We need to show that $\mathrm{Ker}(\varphi)$ contains the identity, is closed under multiplication, and is closed under inverses. We do this using Lemma 1.2.1.

Since $\varphi(e_G) = e_H$, we have $e_G \in \mathrm{Ker}(\varphi)$.

If $\varphi(g) = e_H$ and $\varphi(g') = e_H$, then $\varphi(gg') = \varphi(g)\varphi(g') = e_H e_H = e_H$, so $gg' \in \mathrm{Ker}(\varphi)$.

If $\varphi(g) = e_H$, then $\varphi(g^{-1}) = \varphi(g)^{-1} = e_H^{-1} = e_H$, so $g^{-1} \in \mathrm{Ker}(\varphi)$. □

Kernels are useful tools in the study of homomorphisms, as shown by the following

Proposition 1.2.2. *Let $\varphi : G \to H$ be a group homomorphism. Then φ is injective if and only if $Ker(\varphi) = \{e\}$.*

Bevis. If φ is injective, then $\#\mathrm{Ker}(\varphi) = \#\varphi^{-1}(e_H) \leq 1$. But we know $e_G \in \mathrm{Ker}(\varphi)$, so $\mathrm{Ker}(\varphi) = \{e_G\}$.

Conversely, suppose $\mathrm{Ker}(\varphi) = \{e_G\}$. We need to show that if $\varphi(g) = \varphi(g')$, then $g = g'$. So suppose $\varphi(g) = \varphi(g')$. Then
$$\varphi(g^{-1}g') = \varphi(g)^{-1}\varphi(g') = \varphi(g')^{-1}\varphi(g') = e_H,$$
so $g^{-1}g' \in \mathrm{Ker}(\varphi)$. But $\mathrm{Ker}(\varphi) = e_G$ by assumption, so $g^{-1}g' = e_G$ and multiplying this equation by g on the left gives $g' = g$, as required. □

In some sense, one can think of kernels as measuring "how injective" a homomorphism is: if the kernel is large, it is far away from being injective, in the sense that many elements of G get mapped to the same element in H.

We now come to a very important notion in mathematics.

Definition 1.2.3. *A bijective group homomorphism $\varphi : G \to H$ is called an* isomorphism. *If there exists an isomorphism $G \to H$, we say G and H are* isomorphic. *This is denoted by $G \cong H$ or $\varphi : G \xrightarrow{\sim} H$ or $G \xrightarrow{\sim} H$.*

If $\varphi : G \to H$ is an isomorphism, we know there is an inverse function $\varphi^{-1} : H \to G$. One can show that φ^{-1} is a group homomorphism as well, so both φ and φ^{-1} are isomorphisms. Therefore one could also write $H \xrightarrow{\sim} G$ to mean that H and G are isomorphic. The notation $G \cong H$ is usually preferred over $G \xrightarrow{\sim} H$ unless there is an explicit isomorphism $G \to H$.

The notion of isomorphism is similar to the notion of bijection for sets. If two groups are isomorphic, it means that the underlying sets are bijective, *and* that the group structures are the same (there are groups whose underlying sets are in bijection, but which are not isomorphic). So group-theoretically, the properties of isomorphic groups are the same.

When encountered with a new group, mathematicians want to find out which familiar group it is isomorphic to. In the same spirit, they want to find out which possible groups exists, up to isomorphism. This question has essentially been answered for finite groups over the past century or so, but the classification is very complicated and the proof of the classification is over 10.000 pages long.

Let's see some examples of isomorphic and non-isomorphic groups.

Exempel 1.4.
1. Recall the Klein four group $\{e, a, b, c\}$. This is not isomorphic to the cyclic group C_4. Indeed, if this were the case, there would have to be an element of order 4 in the Klein four group (why?), but there is no such element.
2. The Klein four group is isomorphic to $\mathbb{Z}/2\mathbb{Z} \times \mathbb{Z}/2\mathbb{Z}$. An isomorphism is given by

$$\begin{aligned} \{e, a, b, c\} &\longrightarrow \mathbb{Z}/2\mathbb{Z} \times \mathbb{Z}/2\mathbb{Z} \\ e \longmapsto (0,0), \quad a \longmapsto (1,0), \quad b \longmapsto (0,1), \quad c \longmapsto (1,1). \end{aligned}$$

3. For any $n \geq 1$, we have $\mathbb{Z}/n\mathbb{Z} \cong C_n$. Indeed, the function

$$\begin{aligned} \varphi : \mathbb{Z}/n\mathbb{Z} &\longrightarrow C_n, \\ [m] &\longmapsto e^{2\pi i m/n} \end{aligned}$$

is a well-defined isomorphism. To see this, we first need to check that if $[m] = [m']$ *for integers* $m, m' \in \mathbb{Z}$*, then* $\varphi(m) = \varphi(m')$*. This is true because* $[m] = [m'] \iff n \mid m - m'$*, so that (for some* $k \in \mathbb{Z}$*)*

$$e^{2\pi i m/n} = e^{2\pi i(m'+kn)/n} = e^{2\pi i k} e^{2\pi i m'/n} = e^{2\pi i m'/n}.$$

Next, φ *is a homomorphism because*

$$\varphi([m] + [m']) = e^{2\pi i(m+m')/n} = e^{2\pi i m/n} \cdot e^{2\pi i m'/n} = \varphi([m]) \cdot \varphi([m']).$$

Finally, we know both C_n *and* $\mathbb{Z}/n\mathbb{Z}$ *have* n *elements. This implies that if* φ *is injective, it is also surjective and hence an isomorphism. To check that it is injective, we compute*

$$Ker(\varphi) = \{[m] \mid e^{2\pi i m/n} = 1\} = \{[m] \mid n \text{ divides } m\} = \{[0]\},$$

which suffices by Proposition 1.2.2.

The fact that $C_n \cong \mathbb{Z}/n\mathbb{Z}$ explains why we got the same answers for the orders of elements in C_4 and $\mathbb{Z}/4\mathbb{Z}$.

Elliptic curves

We have already seen that elliptic curves come up in the congruent number problem: in particular, we saw that n is a congruent number if and only if the elliptic curve $y^2 = x^3 - n^2x$ has rational points with $y \neq 0$. Another main reason why elliptic curves are interesting is because they give rise to interesting groups. These groups are given by equipping the set of points on the elliptic curve with a certain addition rule. We just have to add one point "at infinity" for this to work. Let's be more precise about what this all means.

Definition 2.0.1. *Let* $K \in \{\mathbb{Q}, \mathbb{R}, \mathbb{C}\}$*. An* elliptic curve over K *is an equation of the form*

$$E : y^2 = x^3 + ax + b,$$

such that $a, b \in K$ *and* $\Delta := 4a^3 + 27b^2 \neq 0$.

The requirement that $\Delta \neq 0$ is saying precisely that E is *smooth*, i.e. that the graph is differentiable everywhere.

Definition 2.0.2. *Let E be an elliptic curve over K. The K-*points *of the elliptic curve are defined as the set*

$$E(K) := \{(x, y) \in K^2 \mid y^2 = x^3 + ax + b\} \cup \{\infty\}.$$

If $K = \mathbb{Q}$, we also call $E(\mathbb{Q})$ the set of rational points.

In the above definition, ∞ can be understood as a formal symbol for an extra element that we are adding to the curve. However, it has a geometric interpretation as a point which lies at the end of the tails of the elliptic curve, infinitely far up (or down) the y-axis. Any vertical line intersects the point ∞. We will return to this construction when we talk about projective geometry.

Sats 2.1. *Let E be an elliptic curve over K. Then $E(K)$ is an abelian group.*

Bevis. We define the group law as follows. Suppose P and Q are points in $E(K)$. Let Z be the third point of intersection of the line through P and Q with E (if $P = Q$, the line through P and Q is the tangent line; if one of P, Q equals ∞, the line is vertical). Draw a vertical line through the point Z; it intersects E in another point. This is the point $P + Q$.

The above construction defines a binary operation on $E(K)$, but we still need to show that it defines an abelian group structure on $E(\mathbb{Q})$. Check the following:

1. The identity is $\infty \in E(K)$.
2. If $P \in E(K) \setminus \{\infty\}$, then $-P$ is the point of intersection of E with the vertical line through P.
3. For all P and Q in $E(K)$, we have $P + Q = Q + P$.
4. Forget about associativity.

Associativity is hard to prove for this group structure, but I promise it works. □

If $K = \mathbb{R}$ or $\mathbb{C}$, the set $E(K)$ is uncountable, which means it has the same cardinality of $\mathbb{R}$. In this case, the points are "continuous". If $K = \mathbb{Q}$, this is no longer the case: the points on E with rational coordinates are spaced apart. In fact, it is

not clear whether there are any rational points at all; this depends on the specific elliptic curve we are dealing with.

The behaviour of the set of rational points is characterised by the *rank* of the elliptic curve, which we now define. It is based on the following celebrated result of Mordell from 1922.

Sats 2.2 (Mordell's theorem)**.** *Let E be an elliptic curve over $\mathbb{Q}$. Then there exists an integer $r \geq 0$, called the* rank *of E, such that*

$$E(\mathbb{Q}) \cong \mathbb{Z}^r \times F,$$

where F is a finite abelian group.

The content of the theorem is that one can find such an integer r, which implies that the rational points on E are in some sense "discrete" or "sparse". In contrast, consider the group $(\mathbb{Q}, +)$: there is no way to write $\mathbb{Q} \cong \mathbb{Z}^r$ for some $r \geq 0$, as this would imply that all elements in $\mathbb{Q}$ have denominator bounded by some integer. We also know that elements of $\mathbb{Q}$ can be arbitrarily close together: just consider $1/n$ and $1/(n+1)$ and let $n \to \infty$. Mordell's theorem implies that rational points on elliptic curves can not be arbitrarily close together.

Another characterisation of the rank is the following.

Definition 2.2.1. *Let E be an elliptic curve over $\mathbb{Q}$. The rank of E is the maximal integer $n \geq 0$ such that there exist n distinct elements $P_1, P_2, \ldots, P_n \in E(\mathbb{Q})$ such that*

1. *$ord(P_i) = \infty$ for all $i = 1, \ldots, n$;*
2. *If we have an equality*
$$P_i = \sum_{j=1}^{n} m_j P_j$$
with all $m_j \in \mathbb{Z}$, then we must have $m_i = 1$ and $m_j = 0$ for all $j \neq i$.

Note that the rank of E is zero if and only if $E(\mathbb{Q})$ has no elements of infinite order. By Mordell's theorem, this is equivalent to saying that $E(\mathbb{Q})$ is a finite group. Again, this is not obvious: there exist infinite abelian groups without elements of infinit order.

Mathematicians have been trying to understand ranks of elliptic curves for many decades. Perhaps the most famous open problem on ranks of elliptic curves is

the Birch and Swinnerton-Dyer conjecture. It says something about the difficulty of the congruent number problem that n is a congruent number if and only if $E : y^2 = x^3 - n^2x$ has positive rank.

The largest known rank of any elliptic curve over $\mathbb{Q}$ is "at least 28", and it is not known if there are elliptic curves with arbitrarily high rank, or if there is some uniform bound. Moreover, it is believed that 50% of all elliptic curves have rank 0 and 50% of all elliptic curves have rank 1.

Träff 6 - 23/10 Aritmetisk Geometri 1

Torsion points on elliptic curves

Last time we discussed Mordell's theorem, which states that for an elliptic curve $E/\mathbb{Q}$, we have

$$E(\mathbb{Q}) \cong \mathbb{Z}^r \times F,$$

where $r = \mathrm{rk}(E) \geq 0$ is the rank of E and F is a finite abelian group. The group F can naturally be identified with a subgroup of $E(\mathbb{Q})$ as follows: if $0 := (0, 0, \ldots, 0) \in \mathbb{Z}^r$ denotes the identity element in $\mathbb{Z}^r$, we have an injective group homomorphism

$$\varphi : F \xrightarrow{\sim} \{0\} \times F \subseteq \mathbb{Z}^r \times F \xrightarrow{\sim} E(\mathbb{Q}).$$

Note that the inclusion $\subseteq$ can be seen as a function, so this composition makes sense.

The image of φ is equal to the subset of rational points of E with finite order. These points are called torsion points.

Definition 2.2.2. *Let $P \in E(\mathbb{Q})$ for some elliptic curve $E/\mathbb{Q}$. Let $n \geq 1$. We say P is an* n-torsion point *if $nP = \infty$. The set of n-torsion points forms a subgroup of $E(\mathbb{Q})$ which is denoted by $E[n](\mathbb{Q})$. The union of all n-torsion points is called the* torsion subgroup *of E, which is denoted by $E(\mathbb{Q})_{tors}$.*

By definition, $E(\mathbb{Q})_{tors}$ consists of the rational points with finite order. What we have said above is equivalent to saying that

$$F \cong E(\mathbb{Q})_{tors}.$$

In particular, the torsion subgroup is a finite group (which is not a priori obvious). Mordell's theorem can thus also be stated as

$$E(\mathbb{Q}) \cong \mathbb{Z}^{\mathrm{rk}(E)} \times E(\mathbb{Q})_{tors}.$$

I.1 2-torsion points

Let's have a closer look at $E[2](\mathbb{Q})$. This finite subgroup consists of the points $P \in E(\mathbb{Q})$ such that $2P = \infty$, or equivalently, $P = -P$. But we know how to describe $-P$: this is the point P reflected in the x-axis. Thus, if $\infty \neq P = (x, y)$, we have

$$P = -P \iff (x, y) = (x, -y) \iff y = -y \iff y = 0.$$

Thus, $E[2](\mathbb{Q})$ consists of ∞ and the $\mathbb{Q}$-rational points $(x, 0)$ where x satisfies $x^3 + ax + b = 0$.

This is a cubic equation in one variable, so it has at most three distinct roots. Therefore, $\#E[2](\mathbb{Q}) \leq 4$ for any elliptic curve E.

(More generally, one can prove that $\#E[n](\mathbb{C}) = n^2$, so $\#E[n](\mathbb{Q}) \leq n^2$.)

Proposition 2.2.1. *Let $n \geq 1$. Denote by E_n the elliptic curve defined by $y^2 = x^3 - n^2x$. Then we have*

$$E_n[2](\mathbb{Q}) \cong \mathbb{Z}/2\mathbb{Z} \times \mathbb{Z}/2\mathbb{Z}.$$

Bevis. We have seen that $E_n(\mathbb{Q})$ consists of the points ∞ and the points $(x, 0)$ where $x^3 - n^2x = 0$. The latter equation can be written as

$$x(x + n)(x - n) = 0,$$

so it has three rational solutions $x = 0$, $x = -n$ and $x = n$. Hence

$$E[2](\mathbb{Q}) = \{\infty, (0, 0), (0, n), (0, -n)\}.$$

This is a group of order 4 in which every non-identity element has order 2, so it is isomorphic to the Klein four group. □

In the next lectures, we will work our way up to the following result.

Sats 2.3. *Let $n \geq 1$. Then*

$$E_n(\mathbb{Q})_{tors} = E_n[2](\mathbb{Q}).$$

In other words, the elliptic curves of the form $E_n : y^2 = x^3 - n^2x$ have no torsion points besides the four 2-torsion points. (This is a special property of the elliptic curves E_n; there are many elliptic curves $E/\mathbb{Q}$ whose torsion subgroups have a different structure.) Theorem 2.3 is important because it allows us to prove the claimed characterisation of congruent numbers (Theorem 0.4):

Sats 2.4. *Let $n \geq 1$. Then n is a congruent number if and only if $rk(E_n) \geq 1$.*

Bevis. On the problem sheets, you have shown that there is a bijection between the sets

$$S_1 := \left\{ (a, b, c) \in \mathbb{Q}^3_{>0} \mid a^2 + b^2 = c^2 \text{ and } \frac{ab}{2} = n \right\}$$

and

$$S_2 := \{(x, y) \in \mathbb{Q}^2_{>0} \mid y^2 = x^3 - n^2 x\} \subset E_n(\mathbb{Q}).$$

In particular, n is a congruent number if and only if $S_1 \neq \emptyset$ if and only if $S_2 \neq \emptyset$.

Suppose first that n is congruent. Then $S_1 \neq \emptyset$, so the bijection gives us a point $P \in E_n(\mathbb{Q})$ with $y \neq 0$. Thus $P \notin E_n[2](\mathbb{Q})$, so by Theorem 2.3, P is not a torsion point; in other words, P has infinite order. Hence $\mathbb{Z}^{\mathrm{rk}(E_n)} \neq \{0\}$, i.e., $\mathrm{rk}(E_n) \geq 1$.

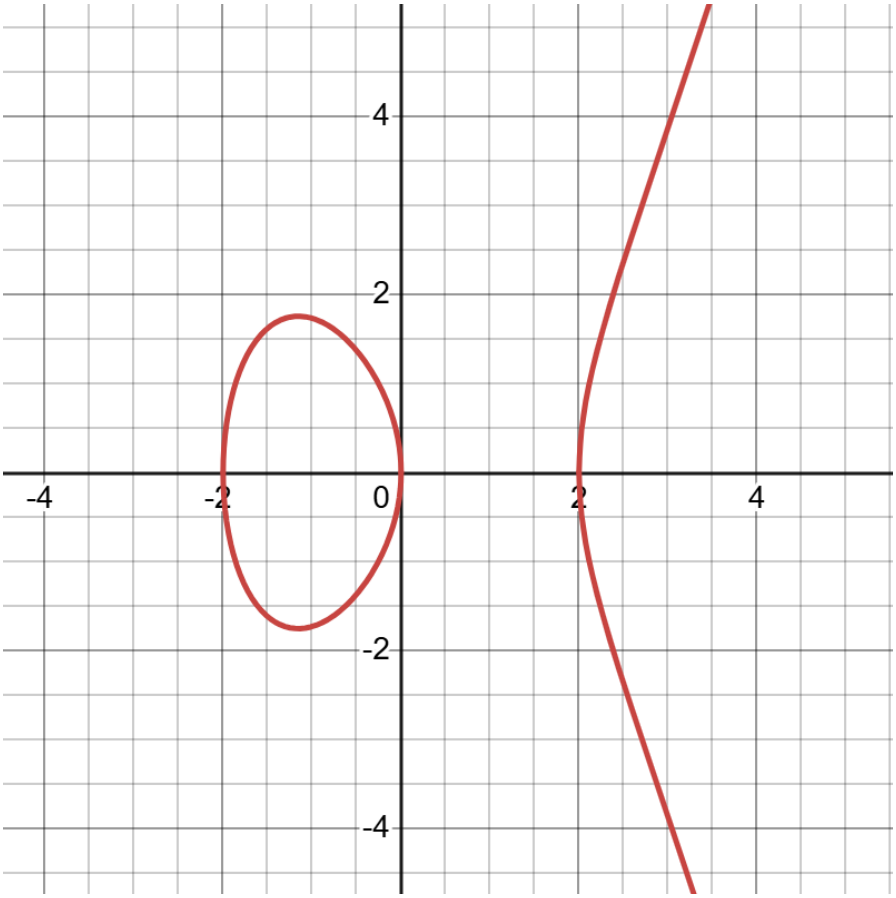
Conversely, suppose that $\mathrm{rk}(E_n) \geq 1$. Then there is a point $P \in E(\mathbb{Q})$ which is not a torsion point; in particular, $P = (x, y)$ with $y \neq 0$. By replacing P with $-P$ if necessary, we may assume $y > 0$. If also $x > 0$, then $S_2 \neq \emptyset$ and we are done, so suppose $x \leq 0$. Since $y \neq 0$, we in fact have $-n < x < 0$.

Thus, P lies on the closed loop which makes up part of the graph of E_n (see Figure I.1). The fact that the round part is convex means that the tangent line at P does not intersect the closed loop in another point besides P. Hence, the point of intersection of the tangent line at P with E_n, which is the point $-2P = (x_2, y_2)$, must satisfy $x_2 > 0$. If $y_2 > 0$ then $-2P \in S_2$; if $y_2 < 0$ then $2P \in S_2$. In either case, $S_2 \neq \emptyset$ so n is a congruent number. □

The proof of Theorem 2.3 will require us to consider points on elliptic curves modulo p. Before going into this, let's stay in the world of rational points a little longer.

Euclidean geometry and Pythagorean triples

In classical geometry, we work inside n-dimensional Euclidean space, which is defined to be the set $\mathbb{R}^n$ for some positive integer n. For $n = 1$, this is just the real line. For $n = 2$, it's the xy-plane. For $n = 3$, it's three-dimensional space, and after that it gets harder to visualise, but mathematically not much changes: we can consider points $(x, y, z, w) \in \mathbb{R}^4$, for example.



Figur 1: The real points of the elliptic curve E_2.

We actually see $\mathbb{R}^n$ not just a set, but as a space in which we can talk about distance. This distance is defined in terms of a *metric* $d : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$, or a distance function, defined as follows:

$$d((x_1, \ldots, x_n), (y_1, \ldots, y_n)) := \sqrt{(x_1 - y_1)^2 + \ldots + (x_n - y_n)^2}.$$

For $n = 1, 2$, and 3, this really gives our intuitive notion of distance between points, by Pythagoras's theorem. We won't usually write this metric on $\mathbb{R}^n$ explicitly, but when we talk about distance in this space, we are actually referring to this function.

Now let's focus on the case $n = 2$, so we are inside the xy-plane. Here we can already do a lot of interesting mathematics. For example, we can see functions $\mathbb{R} \to \mathbb{R}$ geometrically by plotting its graph and studying the function that way. But we can also draw shapes which don't come from functions, such as triangles and circles. An explicit example is the unit circle:

$$S^1 := \{(x, y) \in \mathbb{R}^2 \mid x^2 + y^2 = 1\}.$$[1]

Geometry can be useful because humans naturally have geometric intuition. We can often look at shapes and draw conclusions from them more easily than from long lines of equations. However, it's still important to be able to translate your geometric intuition into rigorous mathematics, to make sure that you aren't making any mistakes: maths can be counter-intuitive sometimes!

Let's now look at a nice application of geometry to solve a problem of number-theoretic nature, namely: how many Pythagorean triples are there, and can we find all of them?

Definition 2.4.1. *A Pythagorean triple is a tuple (a, b, c) of positive integers such that $a^2 + b^2 = c^2$. We say the triple is* primitive *if a, b and c have no common factor.*

We will restrict ourselves to finding the primitive triples, since any Pythagorean triple is a multiple of a primitive one (in the sense that it is of the form (na, nb, nc) for some $n \geq 1$ and (a, b, c) a primitive Pythagorean triple).

Here's how one can solve this problem. We want to find all the (primitive) solutions $(a, b, c) \in \mathbb{Z}^3_{>0}$ of the equation $X^2 + Y^2 = Z^2$. Dividing the equation by Z^2 (which we can do if $Z \neq 0$), this gives the equivalent equation

$$\left(\frac{X}{Z}\right)^2 + \left(\frac{Y}{Z}\right)^2 = 1.$$

But if we make the substitution $x := \frac{X}{Z}$ and $y := \frac{Y}{Z}$, this just says $x^2+y^2 = 1$. Thus, we have reduced the number of variables in the equation by one. But this comes at a cost: where before we were looking for *integer* solutions to $X^2 + Y^2 = Z^2$, we are now looking for *rational* solutions to $x^2 + y^2 = 1$. In other words, finding all Pythagorean triples is equivalent to answering the following question:

Can we find all points (x, y) on the unit circle with rational coordinates?

There are four points on the unit circle which obviously have rational coordinates, namely $\{(\pm 1, 0), (0, \pm 1)\}$. Miraculously, if we pick one of these, we can find all the other rational points by drawing lines with rational slope through the chosen point and intersecting them with the circle.

[1] This looks a lot like the graph of a function, but it isn't: there is no function $f : \mathbb{R} \to \mathbb{R}$ such that $S^1 = \Gamma_f$. Instead, one can view S^1 as the vanishing locus of the function $f(x, y) = x^2+y^2-1$, which is a function $f : \mathbb{R}^2 \to \mathbb{R}$.

Sats 2.5. *Let $P = (-1, 0)$. Then there is a bijection*

$$\left\{ \begin{array}{c} \textit{lines through } P \\ \textit{with rational slope} \end{array} \right\} \overset{1:1}{\longleftrightarrow} \left\{ \begin{array}{c} (x, y) \in \mathbb{Q}^2 \setminus \{P\} \\ \textit{such that } x^2 + y^2 = 1 \end{array} \right\}$$

Bevis. The bijection works as follows: for any line l through P with rational slope, l intersects the circle in exactly one other point. We claim that this point has rational coefficients, and moreover that any point on the circle with rational coefficients can be obtained in this way.

Let $y = a(x + 1)$ be the line through P with slope $a \in \mathbb{Q}$. Then we calculate the point of intersection as follows:

$$\begin{aligned} x^2 + y^2 = 1 \text{ and } y = a(x+1) &\implies x^2 + a^2(x+1)^2 = 1 \\ &\iff (x+1)^2 - 2x - 1 + a^2(x+1)^2 = 1 \\ &\iff (x+1)^2(1+a^2) - 2(x+1) = 0 \\ &\iff (x+1)((x+1)(1+a^2) - 2) = 0 \\ &\iff x = -1 \text{ or } x = -1 + \frac{2}{a^2+1} = \frac{1-a^2}{1+a^2}. \end{aligned}$$

The solution $x = -1$ corresponds to the point P, whereas $x = -1 + 2/(1 + a^2)$ gives

$$y = a\left(-1 + \frac{2}{1+a^2} + 1\right) = \frac{2a}{1+a^2}$$

This shows that if $a \in \mathbb{Q}$, then indeed x and y are also rational numbers!
To show that the map is surjective, suppose that (u, v) is a rational point on the unit circle. Then the line through P and (u, v) is $y = \frac{v}{u+1}(x + 1)$, which has rational slope. Hence (u, v) is obtained by intersecting a line through P with the unit circle. □

Let's use this result to explicitly get a formula for the Pythagorean triples.

Sats 2.6. *Any primitive Pythagorean triple is of the form*

$$\left(q^2 - p^2, 2pq, q^2 + p^2\right)$$

for some positive integers $0 < p < q$.

Bevis. For any rational number $a \in \mathbb{Q}$, we can define a line $l_a : y = a(x+1)$ through P with rational slope a. By Theorem 2.5, such a line gives a rational point (x, y) on the unit circle, with

$$(x, y) = \left(\frac{1-a^2}{1+a^2}, \frac{2a}{1+a^2} \right).$$

This corresponds to a Pythagorean triple as follows: we had reduced $X^2+Y^2 = Z^2$ to $x^2+y^2 = 1$ by dividing the equation by Z^2. This procedure kills common factors between X, Y and Z, so going in the other direction might not give us all Pythagorean triples anymore, but it will at least give the primitive ones (and some more - can you give a criterion which says which non-primitive triples occur?). So we need to recover X, Y and Z from x and y. Note that we are not interested in points with $x = 0$ or $y = 0$, since this will give $X = 0$ or $Y = 0$.

Write the rational number a as $a = p/q$, where p and q have no common factors. Then we have

$$x = \frac{1-\frac{p^2}{q^2}}{1+\frac{p^2}{q^2}} = \frac{q^2-p^2}{q^2+p^2}; \qquad y = \frac{2\frac{p}{q}}{1+\frac{p^2}{q^2}} = \frac{2pq}{q^2+p^2}.$$

Now these are fractions of integers, so we get $X = q^2-p^2, Y = 2pq, Z = q^2+p^2$. If we want X, Y and Z to all be positive, we need $0 < a < 1$, i.e. p and q have the same sign and $p < q$. Thus, it suffices to take $p, q \in \mathbb{Z}$ with $0 < p < q$. □

The well-known triple $(3, 4, 5)$ is obtained for $(p, q) = (1, 2)$, and the triple $(5, 12, 13)$ is obtained for $(p, q) = (2, 3)$. But we can now also easily generate big Pythagorean triples. For example, $p = 1000, q = 1717$ gives the (primitive) triple

$$(1948089, 3434000, 3948089),$$

corresponding to the fact that

$$3795050751921 + 11792356000000 = 15587406751921.$$

One can generalise the above situation by considering the equation $x^2+y^2 = a$ for a general real number a. This has no real solutions if $a < 0$, one solution if $a = 0$, and infinitely many solutions if $a = 1$ (these are the Pythagorean triples). But what about other values of a? This is something you could explore in a project. I will make the following statement, without further arguments why this should be true (or even what all the words mean):

For degree 2 equations in two variables over $\mathbb{Q}$, there are either no solutions or infinitely many solutions in $\mathbb{Q}^2$, unless the equation is singular, in which case one could get a finite, non-zero number of solutions. If the equation is non-singular, then all solutions can be obtained from a single solution by drawing lines with rational slope and intersecting them with the curve.

An example of a circle without rational points is the circle of radius $\sqrt{3}$, given by $x^2 + y^2 = 3$. This is not at all obvious at first glance! To prove this fact, we have to dive into the world of modular arithmetic.

Modular arithmetic

In the past, we have defined the finite set $\mathbb{Z}/n\mathbb{Z}$ as the set of equivalence classes of integers under the relation $a \sim b \iff n \mid a - b$. Subsequently we defined a binary operation (addition) on $\mathbb{Z}/n\mathbb{Z}$ and saw that this binary operation defined a group structure such that $\mathbb{Z}/n\mathbb{Z}$ is isomorphic to the cyclic group of order n. Our next aim is to define another binary operation (multiplication) on $\mathbb{Z}/n\mathbb{Z}$ and study what happens when we do this.

The definition of multiplication is completely analogous to the addition: we simply multiply in $\mathbb{Z}$. In other words, define

$$\begin{aligned} \cdot : \mathbb{Z}/n\mathbb{Z} \times \mathbb{Z}/n\mathbb{Z} &\longrightarrow \mathbb{Z}/n\mathbb{Z} \\ ([a], [b]) &\longmapsto [ab]. \end{aligned}$$

This is well-defined, meaning that if $a \sim a'$ and $b \sim b'$ then $[ab] = [a'b']$ (can you prove this?).

Multiplication is associative and has a unit element, namely [1]. However, inverses do not exist; for instance, for any $a \in \mathbb{Z}$, $[0] \cdot [a] = [0] \neq [1]$ whenever $n > 1$; hence [0] does not have an inverse. But this is not the only problem: if $n = 6$, we have $[2] \cdot [3] = [6] = [0]$, which means [2] and [3] cannot have inverses either. However, note that $[5]^2 = [25] = [1]$, so [5] is its own inverse (note also that $5 \equiv -1 \pmod 6$).

In conclusion, multiplication does *not* define a group law on $\mathbb{Z}/n\mathbb{Z}$. However, if we restrict to those elements which have multiplicative inverses, we do get a group.

In what follows, we will omit the brackets $[a]$ and simply write a to denote elements in $\mathbb{Z}/n\mathbb{Z}$. It should be clear from the context where a given element lives.

Definition 3.0.1. *Let $n \geq 1$. The* group of units modulo n *is the group*

$$(\mathbb{Z}/n\mathbb{Z})^\times := \{a \in \mathbb{Z}/n\mathbb{Z} \mid \exists b \in \mathbb{Z}/n\mathbb{Z} \text{ such that } ab = 1\}$$

whose binary operation is given by multiplication.

The above definition gives a group because inverses now exist by definition. However, it does not tell you what $(\mathbb{Z}/n\mathbb{Z})^\times$ looks like explicitly, so in that sense the definition is unsatisfying. The following proposition remedies this.

Proposition 3.0.1. *Let $n \geq 1$. Then*

$$(\mathbb{Z}/n\mathbb{Z})^\times = \{a \in \mathbb{Z}/n\mathbb{Z} \mid gcd(a, n) = 1\}.$$

Here $\gcd(a, n)$ denotes the *greatest common divisor* of a and n. The condition $\gcd(a, n) = 1$ means precisely that a and n have no non-trivial common divisors. In this case we also say that a and n are *coprime*.

The proof of Proposition 3.0.1 relies on Bézout's Lemma, which we state here without proof.

Hjälpsats 3.0.1 (Bézout)**.** *Let a and b be integers with $g := gcd(a, b)$. Then there exist integers p and q such that*

$$ap + bq = g.$$

It is a good exercise to try to prove Proposition 3.0.1 using Bézout's Lemma. If you are not familiar with greatest common divisors, try to compute some examples first.

Följdsats 3.0.1. *Let p be a prime number. Then*

$$(\mathbb{Z}/p\mathbb{Z})^\times = \mathbb{Z}/p\mathbb{Z} \setminus \{0\}.$$

In other words, every non-zero element has a multiplicative inverse.

Bevis. By Proposition 3.0.1, an element $a \in \mathbb{Z}/p\mathbb{Z}$ is a unit in $\mathbb{Z}/p\mathbb{Z}$ if and only if $\gcd(a, p) = 1$. But since p is prime, we either have $\gcd(a, p) = 1$ or $\gcd(a, p) = p$, and the latter occurs if and only if $p \mid a$, i.e. $a = 0$ in $\mathbb{Z}/p\mathbb{Z}$. □

Exempel 3.1.

1. *The units in* $\mathbb{Z}/4\mathbb{Z}$ *are* 1 *and* 3*. The element* 2 *is not a unit because* $gcd(2,4) = 2$*. We also see that* $2 \cdot 2 = 4 = 0 \pmod 4$*, which means* 2 *can't be a unit. On the other hand, the fact that* 1 *and* 3 *are units should not be a surprise: 1 is always a unit, and* $3 \equiv -1 \pmod 4$*, and* -1 *is also always a unit. One can also see directly that* $3^2 = 9 \equiv 1 \pmod 4$ *(since* $9 = 1 + 2 \cdot 4$*).*

2. *The units in* $\mathbb{Z}/7\mathbb{Z}$ *are* $1, 2, 3, 4, 5$ *and* 6*, since 7 is a prime number. We can verify this explicitly: 1 and 6 are units because* $6 = -1$*. For* 2*, we have*

$$2 \cdot 4 = 8 \equiv 1 \pmod 7,$$

so $2^{-1} = 4$*. For* 3*, we have*

$$3 \cdot 5 = 15 \equiv 1 \pmod 7,$$

so $3^{-1} = 5$*. The above equations also imply that* $4^{-1} = 2$ *and* $5^{-1} = 3$*, so we have verified that all these elements are units.*

3. *Let* $n = 10$*. Since the prime factorisation of 10 is* $10 = 2 \cdot 5$*, the elements in* $(\mathbb{Z}/10\mathbb{Z})^\times$ *are the elements which are not divisible by 2 or 5. Thus,*

$$(\mathbb{Z}/10\mathbb{Z})^\times = \{1, 3, 7, 9\}.$$

When p is prime, every non-zero element is a unit. Hence the group of units has order $p-1$. By Lagrange's theorem, the order of any element of the multiplicative group divides $p-1$, and so every element in $(\mathbb{Z}/p\mathbb{Z})^\times$ satisfies the equation

$$X^{p-1} = 1.$$

This looks a lot like the cyclic group of order $p-1$: these are the complex solutions to $X^{p-1} = 1$. It turns out that $(\mathbb{Z}/p\mathbb{Z})^\times$ is in fact a cyclic group. (What I said above is not a proof, but it can be turned into a proof if one has a bit more background.)

Sats 3.1. *Let* p *be a prime number. Then* $(\mathbb{Z}/p\mathbb{Z})^\times \cong C_{p-1}$*. In particular,* $(\mathbb{Z}/p\mathbb{Z})^\times$ *has a generator.*

Interestingly enough, it is hard to explicitly construct a generator given a prime number p: no formula for a generator is known. We will use Theorem 3.1 quite a lot going forward.

We can now prove the claim that there are no rational points on the circle with radius $\sqrt{3}$.

Proposition 3.1.1. *There are no rational solutions (x, y) to the equation*

$$x^2 + y^2 = 3.$$

Bevis. Suppose the contrary, namely that there are such rational numbers x and y. Then we can write $x = a/c$ and $y = b/c$ for some integers a, b, c without common factor. This gives

$$a^2 + b^2 = 3c^2.$$

Reducing this equation modulo 3 gives $a^2 + b^2 \equiv 0 \pmod 3$.

If we consider the squares in $\mathbb{Z}/3\mathbb{Z}$, we see that $0^2 = 0$, $1^2 = 1$, and $2^2 = 4 \equiv 1 \pmod 3$; hence only 0 and 1 are squares modulo 3. Therefore, to have a solution to $a^2 + b^2 \equiv 0 \pmod 3$, we must have $a^2 \equiv b^2 \equiv 0$, and thus $a \equiv b \equiv 0$. So $3 \mid a$ and $3 \mid b$.

But if this is the case, then $3^2 \mid a^2 + b^2 = 3c^2$, so $3 \mid c^2$, so $3 \mid c$. This contradicts the assumption that a, b and c have no common factor. □

The next result answers the following question: when does the equation $X^2 = -1$ have a solution in $\mathbb{Z}/p\mathbb{Z}$?

Sats 3.2. *Let $p \geq 3$ be an odd prime. Then -1 is a square modulo p if and only if $p \equiv 1 \pmod 4$.*

Bevis. Let $\zeta \in (\mathbb{Z}/p\mathbb{Z})^\times$ be a generator of the group of units modulo p. Thus, every non-zero element in $\mathbb{Z}/p\mathbb{Z}$ can be written as ζ^n for a unique $n \in \{1, \ldots, p-1\}$.

Suppose -1 is a square. Then $-1 = \zeta^n$ for some $1 < n < p-1$, and there exists some $\beta \in (\mathbb{Z}/p\mathbb{Z})^\times$ such that $\beta^2 = -1$. But $\beta = \zeta^m$ for some m, so $\beta^2 = -1$ gives

$$n \equiv 2m \pmod{p-1}.$$

Hence $n = 2m + k(p-1)$ for some $k \in \mathbb{Z}$. Since $p-1$ is even, n is even. But $(-1)^2 = 1 = \zeta^{p-1} = \zeta^{2n}$, so we have $p - 1 = 2n$ with n even, so $p - 1 = 4n'$ for $n' = n/2$. Hence $p = 1 + 4n'$, so $p \equiv 1 \pmod 4$.

Conversely, if $p \equiv 1 \pmod 4$, then $p - 1$ is divisible by 4. Then

$$(\zeta^{\frac{p-1}{4}})^2 = \zeta^{\frac{p-1}{2}} = -1,$$

where the last equality follows because $\zeta^{(p-1)/2} \neq 1$ and the equation $X^2 = 1$ has precisely two solutions (namely 1 and -1) in $\mathbb{Z}/p\mathbb{Z}$, so we must have $\zeta^{(p-1)/2} = -1$. Hence $\zeta^{(p-1)/4}$ is a square root of -1. □

Remark 3.2.1. *Note that every odd number is either congruent to 1 or 3 modulo 4. If one were to count all the primes which are 1 mod 4 and the primes which are 3 mod 4, their numbers approach a ratio of* $1:1$ *as the number of primes tends to infinity. However, primes congruent to 3 mod 4 seem to appear more often. This is an example of what's called* Chebyshev bias.

In a very similar way, we can prove the following statement.

Proposition 3.2.1. *Let* p *be a prime such that* $p \equiv 3 \pmod 4$ *and fix* $a \in (\mathbb{Z}/p\mathbb{Z})^\times$. *Then the equation*

$$X^2 \equiv a \pmod p$$

has precisely 0 or 2 solutions in $\mathbb{Z}/p\mathbb{Z}$. *Moreover, if it has 0 (resp. 2) solutions, then the equation*

$$X^2 \equiv -a \pmod p$$

has 2 (resp. 0) solutions.

Bevis. Note that p is odd, which implies that $b \neq -b \pmod p$ for any $b \neq 0$. Therefore, if $X = b$ is a solution to the equation, then so is $X = -b$. Since a quadratic equation has at most two solutions, this shows that the equation has precisely 0 or 2 solutions.

To prove the second part, we again fix a generator ζ of the multiplicative group. The point is now that

$$-1 = \zeta^{\frac{p-1}{2}}$$

(because squaring this gives $\zeta^{p-1} = 1$), and the condition $p \equiv 3 \pmod 4$ means precisely that $(p-1)/2$ is odd. If we write $a = \zeta^n$ for some $1 \leq n \leq p-1$, we have a solution to $X^2 \equiv a \pmod p$ if and only if n is even, namely $X = \zeta^{n/2}$. So if $X^2 = a$ has no solution, then n is odd, but in that case

$$-a = (-1) \cdot a = \zeta^{\frac{p-1}{2}} \zeta^n = \zeta^{n+\frac{p-1}{2}},$$

and $n + (p-1)/2$ is even; therefore $X^2 = -a$ has a solution (and therefore, as mentioned at the start of the proof, it has two solutions). □

Träff 7 - 06/11 Elliptic curves modulo p

Our goal is still to prove Theorem 2.3. The strategy will be to relate the torsion points on the elliptic curve $E_n/\mathbb{Q}$ to points on "reductions modulo $p$öf this curve. Let's formalize what we mean by this.

Definition 3.2.1. *Fix a prime number $p > 3$. An* elliptic curve over $\mathbb{Z}/p\mathbb{Z}$ *is an equation $E : y^2 = x^3 + ax + b$ with $a, b \in \mathbb{Z}/p\mathbb{Z}$ such that $\Delta = 4a^3 + 27b^2 \not\equiv 0 \pmod{p}$. The* points *of E are given by*

$$E(\mathbb{Z}/p\mathbb{Z}) = \{(x, y) \in (\mathbb{Z}/p\mathbb{Z})^2 \mid y^2 = x^3 + ax + b\} \cup \{\infty\}.$$

The definition looks exactly like the one we had before, except now the number system is replaced by $\mathbb{Z}/p\mathbb{Z}$. The reason why this is a good idea is because $\mathbb{Z}/p\mathbb{Z}$ behaves a lot like $\mathbb{Q}$, $\mathbb{R}$ and $\mathbb{C}$ in the sense that we have an addition, a multiplication, and every non-zero element has a multiplicative inverse.

Note that an elliptic curve over $\mathbb{Z}/p\mathbb{Z}$ always has at most $p^2 + 1$ points, which are finitely many.

The set $E(\mathbb{Z}/p\mathbb{Z})$ can be turned into a group in exactly the same way as before: it is still true that every line (with coefficients in $\mathbb{Z}/p\mathbb{Z}$) intersects the elliptic curve in precisely three points, so we essentially define the group law as we did earlier.

Now suppose that $E : y^2 = x^3+ax+b$ is an elliptic curve over $\mathbb{Q}$ and suppose further that $a, b \in \mathbb{Z}$. Then we can turn E into an elliptic curve over $\mathbb{Z}/p\mathbb{Z}$ by reducing the equation modulo p, provided that p does not divide the discriminant Δ. We get an associated reduction map on points, which is actually a homomorphism (by the way we constructed the group law).

Definition 3.2.2. *Let E be as above and suppose $p > 3$ is a prime number such that $p \nmid \Delta$. Define the* reduction map $\pi : E(\mathbb{Q}) \longrightarrow E(\mathbb{Z}/p\mathbb{Z})$ *as follows. Firstly, we define $\pi(\infty) = \infty$.*

Secondly, if $(x, y) \in E(\mathbb{Q}) \setminus \{\infty\}$, write $x = A/C$ and $y = B/C$ such that A, B and C have no common divisor. If $p \nmid C$, then C has an inverse modulo p, and we define

$$\pi((x, y)) = ([A][C]^{-1}, [B][C]^{-1}),$$

where the brackets denote equivalence classes modulo p.

Lastly, if (x, y) is as above but $p \mid C$, then define $\pi((x, y)) = \infty$.

Note that the definition of the reduction map makes sense: we showed that for any prime p, an element $x \in \mathbb{Z}/p\mathbb{Z}$ has a multiplicative inverse if and only if $x \neq 0$. In particular, $[C]^{-1}$ is a well-defined element in $\mathbb{Z}/p\mathbb{Z}$ whenever $p \nmid C$. It seems ad-hoc to define $\pi((x, y)) = \infty$ if $p \mid [C]$, but in fact this is a very natural choice if one puts the notion of "the point at infinity" on firmer footing through the theory of projective geometry.

We will now prove two lemmas which show that the reductions of elliptic curves can be extremely useful.

Hjälpsats 3.2.1. *Suppose $E : y^2 = x^3 + ax + b$ is an elliptic curve with $a, b \in \mathbb{Z}$. Then the restriction of the reduction map to the torsion subgroup defines a homomorphism*

$$r : E(\mathbb{Q})_{tors} \longrightarrow E(\mathbb{Z}/p\mathbb{Z})$$

for any $p \nmid \Delta$. For all but finitely many such p, the map r is injective.

Remark 3.2.2. *It follows from the Nagell-Lutz theorem (which we will not discuss) that in fact r is injective for* all *primes p.*

Bevis. By Mordell's theorem, $E(\mathbb{Q})_{tors}$ is finite, say of order $m + 1$. Then we can write

$$E(\mathbb{Q})_{tors} = \{\infty, P_1, P_2, \ldots, P_m\}$$

with $P_i = (x_i, y_i)$ rational points on E. As before, we can define integers A_i, B_i and C_i such that they have no common divisor and such that $x_i = A_i/C_i$, $y_i = B_i/C_i$.

Let now $N := \max\{C_i \mid 1 \leq i \leq m\}$. Then for all $p > N$, we have that $C_i \neq 0 \pmod{p}$ for any i. In particular, by definition of the reduction map, we have $r(P_i) \neq \infty$ for any i. But this means that $\ker(r) = \{\infty\}$, so by Proposition 1.2.2, r is injective. Since there are only finitely many primes less than N, this completes the proof. □

Recall that the elliptic curves $E_n/\mathbb{Q}$ are given by the equations $y^2 = x^3 - n^2x$.

Hjälpsats 3.2.2. *Fix an integer $n \geq 1$. Suppose $p > 3$ is a prime such that $p \equiv 3 \pmod{4}$ and $p \nmid 2n$. Then we have*

$$\#E_n(\mathbb{Z}/p\mathbb{Z}) = p + 1.$$

Bevis. The discriminant of E_n is $\Delta = -4n^6$, so a prime p satisfies $p \nmid \Delta$ if and only if $p \nmid 2n$. (Since we assume $p > 3$, we may equivalently ask that $p \nmid n$.) Therefore, E_n defines an elliptic curve over $\mathbb{Z}/p\mathbb{Z}$.

$E_n(\mathbb{Z}/p\mathbb{Z})$ clearly contains the 4 points $\infty, (0,0), (n,0)$, and $(-n,0)$, which are all distinct: indeed, $n \equiv 0 \pmod{p}$ would imply $p \mid n$, and $n \equiv -n \pmod{p}$ would imply $p \mid 2n$; by assumption, neither of these are the case.

Let now $b \in \mathbb{Z}/p\mathbb{Z} \setminus \{0, n, -n\}$. Put $a = b^3 - n^2 b$. Then we have

$$(-b)^3 - n^2(-b) = -b^3 + n^2 b = -a.$$

By Proposition 3.2.1, we conclude that there are precisely 2 points in $E(\mathbb{Z}/p\mathbb{Z})$ whose x-coordinate lies in the set $\{b, -b\}$. Since b was arbitrary and there were $p-3$ choices for b, we get

$$\#E_n(\mathbb{Z}/p\mathbb{Z}) = 4 + 2 \cdot \frac{p-3}{2} = p + 1.$$

□

We are finally almost ready to prove that E_n only has 4 torsion points. The last ingredient is the following celebrated theorem by Dirichlet.

Sats 3.3 (Dirichlet's theorem on primes in arithmetic progressions)**.** *Let a and n be positive integers which have no common divisor. Then the set*

$$\{a, a+n, a+2n, a+3n, \ldots\}$$

contains infinitely many primes.

Dirichlet's theorem is quite remarkable: it is usually not easy to construct sets with many primes. Note that the condition on a and n is really necessary: for instance, if $a = n = 2$, the described set is the set of positive even numbers, which only contains the prime 2.

As an application, we see that there are infinitely many primes congruent to 1 modulo 4 and infinitely many primes congruent to 3 modulo 4 (we get this by setting $n = 4$ and $a = 1$, resp. $a = 3$ in Dirichlet's theorem). This is something we claimed before.

We will not prove Dirichlet's theorem, as the methods used to prove it are not very similar to what we've been doing so far. However, we will apply it in the next proof.

Proof of Theorem 2.3. Suppose for contradiction that $\#E_n(\mathbb{Q})_{tors} = M > 4$. Since $E_n[2](\mathbb{Q})$ is a subgroup of order 4, we have $4 \mid M$, so either $8 \mid M$ or m has an

odd divisor m. In particular, we may assume that $E_n(\mathbb{Q})$ contains a subgroup of order 8 or order m for some odd m.

We proved that the reduction map

$$r : E_n(\mathbb{Q})_{tors} \to E_n(\mathbb{Z}/p\mathbb{Z})$$

is injective for all primes p which are large enough. Since the image of a homomorphism is a subgroup, Lagrange's theorem then tells us that

$$8 \mid \#E_n(\mathbb{Z}/p\mathbb{Z}) \qquad \text{or} \qquad m \mid \#E_n(\mathbb{Z}/p\mathbb{Z}),\ m \text{ odd}.$$

In particular, if p is a large enough prime with $p \equiv 3 \pmod 4$, we get

$$8 \mid p+1 \qquad \text{or} \qquad m \mid p+1,\ m \text{ odd}.$$

We show that in either case, we get a contradiction.

In the first case, since 8 is coprime to 3, Dirichlet's theorem tells us that there are infinitely many primes of the form $8k+3$. In particular, there is a prime $p = 8k+3$ which is large enough such that the reduction map $E_n(\mathbb{Q})_{tors} \to E_n(\mathbb{Z}/p\mathbb{Z})$ is injective. Since $p \equiv 3 \pmod 4$, this gives $8 \mid p+1$, i.e. $p+1 \equiv 0 \pmod 8$, i.e. $p \equiv 7 \pmod 8$. But by assumption, $p = 8k+3 \equiv 3 \pmod 8$. These can't both be true, so we have a contradiction.

In the second case, suppose m is odd and suppose further that $3 \nmid m$. Then $4m$ is coprime to 3, so by Dirichlet's theorem, there are infinitely many primes of the form $4mk+3$. Arguing as above, we have for some $p = 4mk+3$ large enough that

$$p \equiv 3 \pmod m \qquad \text{and} \qquad p \equiv -1 \pmod m.$$

But these can't both be true: otherwise $3 \equiv -1 \pmod m$, so $4 \equiv 0 \pmod m$, so $m \mid 4$, so $m \in \{1, 2, 4\}$; but we assumed that m was odd.

Finally, suppose that m is odd and $3 \mid m$. By Dirichlet's theorem, there are infinitely many primes of the form $12k+7$. These primes are again all congruent to 3 mod 4. Therefore there is a large enough prime p such that

$$p \equiv -1 \pmod m \qquad \text{and} \qquad p \equiv 12k+7 \pmod m.$$

But this gives $12k \equiv -8 \pmod m$. This is impossible: since m is odd, we have $\gcd(-8, m) = 1$, so -8 is a unit modulo 8. On the other hand, since $3 \mid m$, we have $3 \mid \gcd(12k, m)$, and so $12k$ is not a unit modulo m. This contradiction finishes the proof. $\square$

Träff 8 - 13/11 L-functions

We have now proved Theorem 0.4. In this last lecture, we will see that this is not just abstract nonsense, but that the theory of elliptic curves really gives new avenues of attack on the congruent number problem. Let's start with a first application.

Sats 3.4 (Fermat)**.** *The positive integer 1 is not a congruent number.*

Bevis. It suffices to show that the elliptic curve $E_1 : y^2 = x^3 - x$ has rank zero. Equivalently, the equation $y^2 = x^3 - x$ should have no solutions $(x, y) \in \mathbb{Q}^2$ with $xy \neq 0$.

Suppose for contradiction that such a solution does exist. Since both x and y are non-zero, we can write $y = tx$ for some $t = m/n \in \mathbb{Q}$, where $m, n \in \mathbb{Z}$ are coprime integers.

This gives

$$t^2x^2 = x^3 - x \iff x(x^2 - t^2x - 1) = 0.$$

Since $x \neq 0$, this gives $x^2 - t^2x - 1 = 0$. Viewing x as a variable, this equation is assumed to have a rational solution, so it factors into two linear factors over $\mathbb{Q}$. This is equivalent to saying that the discriminant is a square. In this case, the discriminant is given by $t^4 + 4$, so there exists a rational number w_0 such that

$$t^4 + 4 = w_0^2 \iff (m/n)^4 + 4 = w_0^2.$$

Multiplying by n^4 and simplifying gives

$$m^4 + 4n^4 = n^4w_0^2 = (n^2w_0)^2 = w^2,$$

where we define $w := n^2w_0$. Since the left-hand side is an integer, also w is an integer.

The above equation is assumed to have a solution $(m, n, w) \in \mathbb{Z}^3$ such that $mnw \neq 0$. (To check the last assertion, note that $t = m/n$ so $n \neq 0$; if $m = 0$ then $t = 0$ so $y = tx = 0$, but we assumed $y \neq 0$; and finally if $w = 0$ then $w_0 = 0$ so $t^4 = 4$, but this is a contradiction since t was assumed to be rational.)

Therefore, it suffices to show that the equation

$$x^4 + 4y^4 = z^2$$

has no integer solutions (x, y, z) with $xyz \neq 0$. We will prove this by Fermat's technique of *descent*. This means that we assume a solution (x, y, z) exists (assume for simplicity that all of x, y, z are positive), and from this we will construct a new integer solution (p, q, r) where $0 < r < z$. Since this solution (p, q, r) would then by the same argument give rise to a solution (p', q', r') with $0 < r' < r$, and so on, this gives a contradiction, since there are only finitely many integers between 0 and z.

First suppose $\gcd(x, y) = g > 1$. Then $g^4 \mid x^4 + 4y^4 = z^2$, so $g^2 \mid z$, and clearly $(x/g, y/g, z/g^2)$ is a new solution with $0 < z/g^2 < z$. So we may assume that $\gcd(x, y) = 1$.

If $2 \mid x$, then $2 \mid z$, and we get

$$16(x/2)^2 + 4y^2 = 4(z/2)^2 \iff y^4 + 4(x/2)^4 = (z/2)^2,$$

so in this case $(y, x/2, z/2)$ is the solution we were looking for. So we may also assume that x is odd.

In this case, $(x^2, 2y^2, z)$ is a primitive Pythagorean triple! Indeed, the equation can be written as

$$(x^2)^2 + (2y^2)^2 = z^2,$$

and since $\gcd(x, y) = 1$ and x is odd, there is no common divisor between x^2, $2y^2$ and z. But we have classified Pythagorean triples. In particular, we know that there exist integers a and b with $0 < a < b$ such that

$$\begin{cases} x^2 = b^2 - a^2, \\ 2y^2 = 2ab, \\ z = a^2 + b^2. \end{cases}$$

The fact that the triple is primitive means that $\gcd(a, b) = 1$ (this was on an exercise sheet). Thus, $y^2 = ab$ and so both a and b are squares, say $a = c^2$ and $b = d^2$. Plugging this into the expression for x gives

$$x^2 = d^4 - c^4 \iff x^2 + c^4 = d^4,$$

so we again get a primitive Pythagorean triple (x, c^2, d^2). Thus there exist integers $0 < e < f$ such that

$$\begin{cases} x = f^2 - e^2, \\ c^2 = 2ef, \\ d^2 = e^2 + f^2. \end{cases}$$

Again $\gcd(e, f) = 1$, and combined with the equation $c^2 = 2ef$ this implies that one of e, f is a square and the other one is twice a square. Hence we can write

$$\{e, f\} = \{p^2, 2q^2\}$$

for some integers p, q. The equation for d^2 then reads

$$p^4 + 4q^4 = d^2,$$

so finally we have found a new solution (p, q, d) to the original equation. Since $0 < d \leq d^2 = b < z$, this solution satisfies the descent condition, so we are done. □

A downside of the above proof is that it does not easily generalize to (dis)prove that other integers n are congruent numbers. This reflects the fact that we don't have a good way of computing the rank of an elliptic curve just by looking at the equation of the curve.

In the rest of the lecture, I'll try to give an overview of a more streamlined approach to the congruent number problem using elliptic curves.

The Riemann zeta function

The Riemann zeta function is usually defined via

$$\zeta(s) = \sum_{n=1}^{\infty} n^{-s} = \sum_{n=1}^{\infty} \frac{1}{n^s} = 1 + \frac{1}{2^s} + \frac{1}{3^s} + \frac{1}{4^s} + \dots$$

But when does this expression really make sense? In other words, for which values of $s \in \mathbb{C}$ do we get a sensible answer (meaning another complex number) out of this expression? After all, we would like the zeta function to be a function $\zeta : \mathbb{C} \to \mathbb{C}$.

In general, there are criteria which determine when an infinite sum evaluates to something finite, which you would learn in the first year of a bachelor's degree in maths. Let's just note the following. If $s = 1$, the series is

$$\sum_{n=1}^{\infty} \frac{1}{n} = 1 + \frac{1}{2} + \frac{1}{3} + \dots,$$

which famously *diverges*, i.e. becomes arbitrarily large. This can be seen by grouping terms together; e.g. $1/3 + 1/4 > 1/2$, and $1/5 + 1/6 + 1/7 + 1/8 > 1/2$, and

so on, so that the sum is at least "infinity times 1/2". Thus, the above expression for $\zeta(s)$ does *not* make sense at $s = 1$.

If $s = 2$, we get the sum

$$\sum_{n=1}^{\infty} \frac{1}{n^2} = 1 + \frac{1}{4} + \frac{1}{9} + \dots,$$

and now the fractions get smaller much faster. Evaluating the above sum was an open problem for a while, called the *Basel problem*. Euler showed in 1735 that this sum is equal to $\pi^2/6$. Thus,

$$\zeta(2) = \pi^2/6.$$

Note also that if $s = a + bi$ is a complex number, then the absolute value $|n^{-s}|$ equals n^{-a}, i.e., the absolute value only depends on the real part of s. This is because for any real number $x > 0$, we have $x^i = e^{i \log(x)}$ and so x^i lies on the unit circle, which means it has norm 1; it follows that

$$|n^{-s}| = |n^{-a-bi}| = |n^{-a}| \cdot |n^{-bi}| = |n^{-a}| \cdot |n^i|^{-b} = |n^{-a}|.$$

Using general theorems about convergence of infinite sums, one can show that in fact the expression for $\zeta(s)$ yields a well-defined complex number whenever $s \in \mathbb{C}$ with $\mathrm{Re}(s) > 1$.

However, it turns out that the definition can be extended to the whole complex plane in a canonical way; this is called *analytic continuation*. To see how the zeta function can be extended to the strip $0 < \mathrm{Re}(s) < 1$, note that we have

$$2^{-s}\zeta(s) = \sum_{n=1}^{\infty} \frac{1}{(2n)^s},$$

so we get each second term in the original zeta function; hence we can get an alternating sum by taking

$$(1 - 2 \cdot 2^{-s})\zeta(s) = \sum_{n=1}^{\infty} (-1)^{n+1} n^{-s}.$$

Alternating sums have better convergence properties: for instance,

$$1 - \frac{1}{2} + \frac{1}{3} - \frac{1}{4} + \dots$$

is clearly bounded above by 1. In fact the above sum is equal to $\ln(2)$. In any case, for any $s \neq 1$ with $\mathrm{Re}(s) > 0$, one can now set

$$\zeta(s) := \frac{1}{1 - 2^{1-s}} \sum_{n=1}^{\infty} (-1)^{n+1} n^{-s},$$

which defines the zeta function on a slightly larger domain. However, we still have $\zeta(1) = \infty$; the zeta function has a *pole* at $s = 1$. However, it is well-defined everywhere else. In this sense, the zeta function is a bit like the function $y = 1/x$.

I.1 Euler product

There is also an expression for $\zeta(s)$ as an infinite product instead of an infinite sum, namely

$$\zeta(s) = \prod_{p \text{ prime}} \frac{1}{1 - p^{-s}}, \qquad \mathrm{Re}(s) > 1.$$

The fact that this agrees with the old definition comes down to unique prime factorization and the fact that

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

where x is treated as a variable. (Check that multiplying the right-hand side by $1 - x$ gives 1.) Since every positive integer n is a unique product of primes, we get n^{-s} as some combination of primes $p_i^{-n_i s}$, where $n_i \geq 0$ are integers; this allows one to show that the two expressions are equal.

The L-function of an elliptic curve will be defined similarly as a product over primes.

I.2 The Riemann hypothesis

The Riemann zeta function is perhaps best known for the question about its zeroes. Indeed, one can ask: for which $s \in \mathbb{C}$ do we have $\zeta(s) = 0$? (Now ζ denotes the analytic continuation of the previous expressions.)

It turns out that $\zeta(-2n) = 0$ for any integer $n \geq 1$, and every other zero satisfies $0 < \mathrm{Re}(s) < 1$. The Riemann hypothesis is the following conjecture:

Riemann Hypothesis. If $s \in \mathbb{C}$ is a zero of the Riemann zeta function, then either $s = -2n$ for some $n \geq 1$ or $\mathrm{Re}(s) = 1/2$.

Due in part to the Euler product, the Riemann hypothesis is related to the distribution of prime numbers on the number line. The Riemann hypothesis is a wide-open problem. Solving it would earn you a million dollars, as it is one of the seven Millennium Prize Problems.

Elliptic curve L-functions

Given an elliptic curve $E/\mathbb{Q}$, one can construct an analogue of the Riemann zeta function for E, called an L-function. This works as follows. Let Δ denote the discriminant of E. Then for any prime $p \nmid \Delta$, we get an elliptic curve E_p over $\mathbb{Z}/p\mathbb{Z}$ by the reduction mod $p$procedure we saw before. Counting the (finite) number of points of this curve, we can define

$$a_p(E) := p + 1 - \#E_p(\mathbb{Z}/p\mathbb{Z}).$$

This number $a_p(E)$ is very important. It is called the *Frobenius trace* of E_p. It satisfies something called the Hasse bound:

Sats 3.5 (Hasse)**.** *For any prime p, we have $|a_p(E)| \leq 2\sqrt{p}$.*

Exempel 3.2. *If $p = 41$, then $2\sqrt{p} \approx 12.8$, so any elliptic curve E over $\mathbb{Z}/41\mathbb{Z}$ satisfies*

$$30 \leq \#E(\mathbb{Z}/41\mathbb{Z}) \leq 54.$$

Similarly to ζ, we define the L-function as a product over primes: for a complex variable s, define

$$L(E, s) = \prod_{p \nmid \Delta} \frac{1}{1 - a_p(E)p^{-s} + p^{1-2s}} \prod_{p|\Delta} \epsilon(p),$$

where $\epsilon(p)$ lies in the set $\{1 - p^{-s}, 1 + p^{-s}, 1\}$, depending on what kind of curve E becomes modulo p.

Using Hasse's theorem, one can show that $L(E, s)$ converges for $\mathrm{Re}(s) > 3/2$. Moreover, $L(E, s)$ has an analytic continuation to all of $\mathbb{C}$ (this is a deep result proved only in recent decades). The following conjecture is often called the BSD conjecture, after the names of the mathematicians who came up with it.

Conjecture 3.5.1 (Birch–Swinnerton-Dyer Conjecture)**.** *Let E be any elliptic curve over $\mathbb{Q}$. Then*

$$rk(E) = ord_{s=1} L(E, s),$$

where the right-hand side denotes the integer n such that $L(E, 1) = 0$, $L'(E, 1) = 0$, ..., $L^{(n-1)}(E, 1) = 0$, and $L^{(n)}(E, 1) \neq 0$.

In analogy with the Riemann hypothesis, what can we say about the zeroes of $L(E, s)$? One can show that there are trivial zeroes for $s = 0, -1, -2, \ldots$. The BSD conjecture suggests that $s = 1$ is a zero if and only if $\mathrm{rk}(E) > 0$. In general, experts expect the following version of the Riemann hypothesis to hold:

Conjecture 3.5.2 (Generalized Riemann Hypothesis)**.** *Suppose* $L(E,s) = 0$*. Then either* $s \in \{0, -1, -2, \ldots\}$ *or* $Re(s) = 1$*.*

The BSD conjecture says that ranks of elliptic curves are closely related to L-functions. In particular, if the BSD conjecture is true, one can precisely predict when a given number n is congruent.

Sats 3.6 (Tunnell, 1983)**.** *Let* $n \in \mathbb{N}$ *be a square-free integer. If* n *is an odd congruent number, then*

$$\#\{(x,y,z) \in \mathbb{Z}^3 \mid 2x^2+y^2+32z^2 = n\} = \frac{1}{2}\#\{(x,y,z) \in \mathbb{Z}^3 \mid 2x^2+y^2+8z^2 = n\}.$$

If n *is an even congruent number, then*

$$\#\{(x,y,z) \in \mathbb{Z}^3 \mid 4x^2+y^2+32z^2 = \frac{n}{2}\} = \frac{1}{2}\#\{(x,y,z) \in \mathbb{Z}^3 \mid 4x^2+y^2+8z^2 = \frac{n}{2}\}.$$

Moreover, if the BSD conjecture is true, then the converse implications also hold; i.e., if the equalities hold then n *is a congruent number.*

Tunnell's theorem may look complicated at first glance, but it is actually very easy to compute the cardinalities of the sets on both sides because all the variables are squared and hence non-negative. For example, here is a second proof that 1 is not a congruent number: we have

$$2x^2 + y^2 + 32z^2 = 1 \iff (x,y,z) = (0, \pm 1, 0)$$

and

$$2x^2 + y^2 + 8z^2 = 1 \iff (x,y,z) = (0, \pm 1, 0).$$

Since $2 \neq \frac{1}{2} \cdot 2$, 1 cannot be a congruent number, and we can conlude this even without knowing the BSD conjecture.