

A formalisation of transcendence of e

Jujian Zhang

July 24, 2020

Abstract

The objective of this report is to present formalizations of some basic theorems from transcendental number theory with **Lean** and **mathlib** in the hope that it will serve as a motivation for mathematicians to be more curious about interactive theorem proving. The following theorems are formalized:

1. the set of algebraic numbers is countable, hence transcendental number exists:

```
1 theorem algebraic_set_countable : set.countable algebraic_set
2 theorem transcendental_number_exists :
3    $\exists x : \mathbb{R}, \text{transcendental } x$ 
```

2. all Liouville numbers are transcendental:

```
1 theorem liouville_numbers_transcendental :
2    $\forall x : \mathbb{R}, \text{liouville\_number } x \rightarrow \text{transcendental } x$ 
```

3. $\alpha := \sum_{i=0}^{\infty} \frac{1}{10^{i!}}$ is a Liouville number hence α is transcendental.

```
1 theorem liouville_alpha : liouville_number alpha
2 theorem transcendental_alpha : transcendental alpha :=
3   liouville_numbers_transcendental alpha liouville_alpha
```

4. e is transcendental:

```
1 theorem e_transcendental : transcendental e
```

Contents

1	Overview	2
1.1	Interactive theorem proving	2
1.2	History of transcendental numbers	3
2	Brief introduction to Lean	5
2.1	Simple type theory	5
2.1.1	Proposition as type	6
2.2	Lean and mathlib	7
2.2.1	prove a disjunction	11
2.2.2	prove an implication	11
2.2.3	prove an equivalence	12
2.2.4	prove a negation	12
2.2.5	prove a proposition with \forall	12
2.2.6	prove a proposition with \exists	12
2.3	An example	13
3	Formalisation using Lean	16
	Logistics of the formalisation	16
3.1	Countability argument	17
3.2	Liouville's theorem and Liouville number	23
	General theory about Liouville number	23
	Construction of a Liouville number	28
3.3	Hermite's proof of transcendence of e	32
	Bibliography	33

Chapter 1

Overview

1.1 Interactive theorem proving

Around 1920s, the German mathematician David Hilbert put forward the Hilbert programme to seek:

1. an axiomatic foundation of mathematics;
2. a proof of consistency of the said foundation;
3. Entscheidungsproblem: an algorithm to determine if any proposition is universally valid given a set of axioms.

The first two aims were later proved to be impossible by Gödel and the celebrated incompleteness theorems. Via the completeness of first order logic, the Entscheidungsproblem can also be interpreted as an algorithm for producing proofs using deduction rules. Even without a panacea approach for mathematics, computer still bears advantages against a carbon-based mathematician. Perhaps the most manifested advantage is the accuracy of a computer to execute its command and to recall its memories. Thus came the idea of **interactive theorem proving** — instead of hoping a computer algorithm to spit out some unfathomable proofs, assuming computers are given the ability to check correctness of proofs, human-comprehensible proofs can be verified by machines and thus guaranteed to be free of errors. With a collective effort, all theorems verified this way can be collected in an error-free library such that all mathematicians can utilise to prove further theorems which can then be added to the collection, ad infinitum [Boy+94]. Curry-Howard isomorphism provided the crucial relationship between mathematical proofs and computer programmes, more specifically relationship between propositions and types, to make such project feasible [KK11]. The idea will be explained in section 2 along with **Lean**.

The proof of “Kepler’s conjecture¹” will serve as an illustrative example of utility of interactive theorem proving. As early as 1998, Thomas Hales had

¹the most efficient way to pack spheres should be hexagonally

claimed a proof [Hal98; HUW14], however the proof is controversial in the sense that mathematician even with great effort could not guarantee its correctness. A collaborative project using **Isabelle**² and **HOL Light**³ verified the proof around 2014 and hence settled the controversy in 2017 [Hal+17]. There is also Georges Gonthier with his teams using **Coq**⁴ who formalised the four colour theorem and Feit-Thompson theorem where the latter is a step to the classification of simple groups [Gon08; Gon+13]. Using **Lean**⁵, Buzzard, Commelin, and Massot were able to formalise modern notion of perfectoid spaces [BCM20].

1.2 History of transcendental numbers

“Transcendence” as a mathematical jargon first appeared in a Leibniz’s 1682 paper where he proved that \sin is a transcendental function in the sense that for any natural number n there does not exist polynomials p_0, \dots, p_n such that

$$p_0(x) + p_1(x) \sin(x) + p_2(x) \sin(x)^2 + \dots + p_n(x) \sin(x)^n = 0$$

holds for all $x \in \mathbb{R}$ [Bou98]. The Swiss mathematician Johann Heinrich Lambert in his 1768 paper proved the irrationality of e and π where he also conjectured their transcendence [Lam04]. It is until 1844 that Joseph Liouville proved the existence of any transcendental numbers and until 1851 an explicit example of transcendental number is actually given by its decimal expansion:[Kem16]

$$\sum_{i=1}^{\infty} \frac{1}{10^{i!}} = 0.11000100000\dots.$$

However, this construction is still artificial in nature. The first example of a real number proven to be transcendental that is not constructed for the purpose of being transcendental was e . Charles Hermite proved the transcendence of e in 1873 with a method applicable with help of symmetric polynomial to transcendence of π in 1882 and later to be generalised to Lindemann-Weierstrass theorem in 1885 stating that if $\alpha_1, \dots, \alpha_n$ are distinct algebraic numbers then $e^{\alpha_1}, \dots, e^{\alpha_n}$ are linearly independent over the algebraic numbers [Bak90]. The transcendence of π was particularly celebrated because it immediately implied the impossibility of the ancient greek question of squaring the circle, i.e. it is not possible to construct a square, using compass and ruler only, with equal area to a circle. For this question is plainly equivalent to construct $\sqrt{\pi}$ which is not possible for otherwise π is algebraic. Georg Cantor in 1874 proved that algebraic numbers are countable hence not only did transcendental numbers exist, they exist in a ubiquitous manner – there is a bijection from the set of all transcendental numbers to \mathbb{R} [Can32; Can78].

²a theorem prover relies extensively on dependent type theory and Curry-Howard correspondence.

³ibid.

⁴ibid.

⁵ibid.

In 1900, Hilbert proposed twenty-three questions, the 7th of which is regarding transcendental numbers: Is a^b transcendental, for any algebraic number a that is not 0 or 1 and any irrational algebraic number b ? The answer is yes by Gelfond-Schneider theorem in 1934 [Gel34]. This has some immediate consequences such that

1. $2^{\sqrt{2}}$ and its square root $\sqrt{2^{\sqrt{2}}}$ are transcendental;
2. e^π is transcendental for $e^\pi = (e^{i\pi})^{-i} = (-1)^{-i}$;
3. $i^i = e^{-\frac{\pi}{2}}$ is transcendental etc.

In contrast, none of $\pi \pm e$, $\pi e, \frac{\pi}{e}, \pi^\pi, \pi^e$ etc are proven to be transcendental. It is also conjectured by Stephen Schanuel that given any n \mathbb{Q} -linearly independent $z_1, \dots, z_n \in \mathbb{C}$, then $\text{trdeg}(\mathbb{Q}(z_1, \dots, z_n, e^{z_1}, \dots, e^{z_n})/\mathbb{Q})$ is at least n [Lan66]. If this were proven, the algebraic independence of e and π would follow immediately by setting $z_1 = 1$ and $z_2 = \pi i$ with Euler's identity.

Chapter 2

Brief introduction to Lean

Lean is developed by Leonardo de Moura at Microsoft Research Redmond from 2013 using dependent type theory and calculus of inductive constraint [AMK15]. In this chapter, basic ideas of Curry-Howard isomorphism will be demonstrated by some basic examples of mathematical theorem expressed in Lean using dependent type theory.

2.1 Simple type theory

Unlike set theory where everything from natural numbers to modular forms is essentially a set. Type theory associate every expression with a **type**. In set theory, an element can belongs to different sets, for example 0 is simultaneously in $\mathbb{N} \subseteq \mathbb{Q} \subseteq \mathbb{R} \subseteq \mathbb{C}$. However an expression can only have one type. 0 without any context will have type \mathbb{N} and, to specify the zero with type \mathbb{R} we write $(0 : \mathbb{R})$. If a has type α , we write $a : \alpha$. By a universe of types we mean a collection of types. Types can be combined to form new types in the following way:

- let α and β be types then $\alpha \rightarrow \beta$ is the type of functions from α to β : the element of type $\alpha \rightarrow \beta$ is a function that for any element of α gives an element of β . For mathematician this loosely means that for any two classes α and β , there is a new class $\text{hom}(\alpha, \beta)$. Sometimes we are not bothered to give a function a name, we can use the λ notation: $(\lambda x : \alpha, \text{expression})$ has type $\alpha \rightarrow \dots$ depending on the content of expression. This can be thought of \mapsto . For example $(\lambda x : \mathbb{N}, x + 1) : \mathbb{N} \rightarrow \mathbb{N}$.
- let α and β be types then $\alpha \times \beta$ is the cartesian product of α and β : the element of type $\alpha \times \beta$ is an ordered tuple (a, b) where $a : \alpha$ and $b : \beta$.
- Let α be a type in universe \mathcal{U} and $\beta : \alpha \rightarrow \mathcal{U}$ be a family of type that for any $a : \alpha, \beta(a)$ is a type in \mathcal{U} . Then we can form the Π -type

$$\prod_{a:\alpha} \beta(a)$$

whose element is of the form $f : \prod_{a:\alpha} \beta(a)$ such that for any $x : \alpha$, $f(x) : \beta(x)$. Note that function type is actually an example of Π -type where β is a constant family of types. For this reason, we also call Π -types dependent functions. For example if $\text{Vec}(\mathbb{R}, n)$ is the type of \mathbb{R}^n , then

$$n \mapsto \underbrace{(1, \dots, 1)}_{n \text{ times}} : \prod_{m:\mathbb{N}} \text{Vec}(\mathbb{R}, m)$$

- We also have dependent cartesian product or Σ -type: Let α be a type in universe \mathcal{U} and $\beta : \alpha \rightarrow \mathcal{U}$ be a family of types in \mathcal{U} , then the Σ -type

$$\sum_{a:\alpha} \beta(a)$$

whose element is of the form $(x, y) : \sum_{a:\alpha} \beta(a)$ such that $x : \alpha$ and $y : \beta(x)$.

Similarly

$$\left(n, \underbrace{(1, \dots, 1)}_{n \text{ times}} \right) : \sum_{m:\mathbb{N}} \text{Vec}(\mathbb{R}, m)$$

2.1.1 Proposition as type

In type theory, a proposition p can be thought as a type whose elements is a proof of p .

Example 1. $1 + 1 = 2$ is a proposition. `rfl` is an element of type $1 + 1 = 2$ where `rfl` is the assertion that every term equals to itself.

Example 2. For two propositions p and q , the implication $p \implies q$ then can be interpreted as function $p \rightarrow q$. To say $\text{imp} : p \rightarrow q$ is to say for any $hp : p$ we have $\text{imp}(hp) : q$, or equivalently given any hp , a *proof* of proposition p , $\text{imp}(hp)$ is a proof of proposition q .

Example 3. If $p : \alpha \rightarrow \text{proposition}$ $\forall x : \alpha, p(x)$ can be interpreted as a Π -type $\prod_{x:\alpha} p(x)$. To prove $\forall x : \alpha, p(x)$, we need to find an element of type $\prod_{x:\alpha} p(x)$, equivalently for any $x : \alpha$, we need to find an element of type $p(x)$, equivalently for any $x : \alpha$, we need to find a proof of $p(x)$.

Similarly, $\exists x : \alpha, p(x)$ can be interpreted as a Σ -type $\sum_{x:\alpha} p(x)$. To prove $\exists x : \alpha, p(x)$ is to find an element x of type α and prove $p(x)$, equivalently to find an element $x : \alpha$ and an element of type $p(x)$ and this is precisely $(x, p(x)) : \sum_{a:\alpha} p(a)$.

Theorems are true propositions, using the interpretation above, theorems are inhabited types and to prove a theorem is to find an element of the required type.

2.2 Lean and mathlib

`mathlib` is *the* collection of mathematical definition, theorems, lemmas built on Lean. `mathlib` includes topics in algebra, topology, manifolds and combinatorics etc. In this section, we are going to explain briefly how to use Lean with `mathlib`.

In Lean, new definition can be introduced with the following syntax:

```
1 def name (arg1:type1) ... (argn:typen) : return_type := contents
2
3 def name' {arg1:type1} ... (argn:typen) : return_type := contents
```

`return_type` is optional when it can be inferred from `contents`. If an argument is surrounded by curly bracket instead of round bracket, then when the definition is invoked the said argument is implicit, i.e. `name' a2 ... an` where `ai:typei`. To explicitly mention the said argument, one needs to use `@name' a1 ... an` where `ai:typei`. One can use “if then else” to introduce a function whose value depends on the value of arguments:

```
1 def name args : return_type :=
2   if (h args)
3   then contents1
4   else contents2
5
6 def name args : return_type :=
7   ite (h args) contents1 contents2
```

New notations are introduced with the following syntax:

```
1 notation _`lhs`_ := _rhs_
```

so that Lean will treat every occurrence of `_`lhs`_` as `_rhs_` verbatim. For example `notation ℤ[X]` := polynomial ℤ` will replace the Lean type `polynomial ℤ` with a more family notation of $\mathbb{Z}[X]$.

For any type of α , we can introduce a subtype of α by:

```
1 def α' := {x : α // property_satisfied_by_x}
```

An element of type α' is of the form $\langle x, hx \rangle$ where $x : \alpha$ and hx is a proof that x satisfies the given property.

Theorems or lemmas are introduced with the following syntax:

```
1 theorem name (arg1:type1) ... (argn:typen) : content :=
2 begin
3   -- proof of the theorem
4 end
```

To write a proof understandable to Lean, one need to use *tactic mode*. In Lean, one can use

- proof by induction: if the goal is a proposition about natural number n , `induction n with n IH` is to prove the proposition by induction. This command will change the current goal to two goals. The first goal is to prove the proposition for $n = 0$ and the second goal is to prove the proposition $n + 1$ with the additional inductive hypothesis `IH`;

```

1 theorem awesome_theorem_about_natural_number (n : ℕ) :
  ↔ propositionn :=
2 begin
3   induction n with n IH,
4
5   a_proof_of_proposition0
6
7   -- (IH : propositionn) is now in context
8   a_proof_of_propositionn+1
9 end

```

- proof by contradiction: if the goal is to prove proposition H , `by_contra absurdum` will add `absurdum : ¬H` into the current context and turn the goal into proving `false`;

```

1 theorem awesome_theorem : awesome_proposition :=
2 begin
3   by_contra absurdum,
4
5   -- Now (absurdum : ¬ awesome_proposition) is in context and
6   ↔ the goal is to prove falsehood.
7   a_proof_of_falsehood
8 end

```

- proof in a forward manner i.e. introduce new theorem or convert known theorem in current context to approach the goal:
 - `have H := content` will introduce a new proposition whose proof is given by `content`.
 - `have H : some_proposition` will add one more goal of proving the proposition then introduce the proved proposition to the current context.
 - If `H` is in context then `replace H := content` will change `H` to (a proof of) the proposition that `content` is proving.
 - `replace H : some_proposition` will add one more goal of proving `some_proposition` and then replace `H` to the proposition proven.

- If H is in context, `simp at H` will simplify H to using small lemmas¹. `simp only [h1,...,hn]` is to simplify only using $h1 \dots hn$.
- `rw` is for term rewriting. If we have $h : lhs = rhs$ or $h : lhs \leftrightarrow rhs$ and another H in context, then `rw h at H` will replace every occurrence of lhs with rhs in H and `rw \leftarrow h` will replace every occurrence of rhs with lhs in H .
`rw [h1, h2, ..., hn] at H` is the same as `rw h1 at H, rw h2 at H, ..., rw hn at H`.
- Since `rw` and `simp` will change all occurrence, this sometimes would be inconvenient. If H is in context, `conv_lhs at H {tactics}` will confine the scope of `tactics` only to left hand side of H ; similarly `conv_rhs at H {tactics}` will confine the scope to right hand side of H .
- `generalise H : lhs = var_name` will set `var_name` to lhs and add (proof of) the proposition $H : lhs = var_name$ to the current context.
- If $H : \exists x : type, property_about_x$ is in the current context, `choose x hx using H` will introduce $x : type$ with the assumption `property_about_x` to the current context.
- If $H : p \wedge q$ is in the current context, then $H.1$ is (a proof of) p and $H.2$ is (a proof of) q .
- If $H : ite\ h1\ h2\ h3$ is in the current context, then `split_ifs at H` will turn the current goal into two goals, the first one is to prove the original goal with the additional assumption $h1$ and $h2$; the second one is to prove the original with goal with the additional assumption $\neg h1$ and $h3$.
- proof in a backward manner i.e. convert or replace the goal so that it is closer to what is known in context:
 - `unfold definition` is to unfold a definition to what is explicitly defined when the definition is introduced.
 - `simp`, `rw`, `conv_lhs {tactics}` and `conv_rhs {tactics}` is the same as above except now they change at goal.
 - Given (a proof of) proposition $H : h1 \rightarrow h2$, then `apply H` will change the goal of proving $h2$ to prove $h1$.
 - `suffices H : some_proposition` ask a proof of the current goal with additional H , then ask for a proof of H .
 - `norm_cast` is convert the type of numbers. For example the current goal is $(x : \mathbb{R}) < (y : \mathbb{R})$ where x and y are of type \mathbb{N} , then after

¹to be more precise, lemma with `@[simp]` tag, i.e. lemmas declared in the following syntax `@[simp] lemma lemma_name args : Prop`. These lemma are usually trivial in nature such as `nat.add_zero` which asserts that $\forall n : \mathbb{N}, n + 0 = n$.

`norm_cast` the goal will become $x < y$. This should be simpler because \mathbb{R} in `Lean` is equivalent classes of Cauchy sequence of \mathbb{Q} while natural number is much easier to work with.

`norm_num` is equivalent to `norm_cast`, `simp`.

- `ext` will convert the current goal with axioms of extensionality. For example if the goal is to prove equality of polynomial then after `ext` the goal would become to prove that every coefficient is equal; or if the goal is to prove equality of sets of type α $A = B$, then after `ext`, an arbitrary element `x` of type α will be introduced to context then the goal will become to prove $x \in A \iff x \in B$. `ext var_name` will force `Lean` to introduce new variable under the identifier `var_name`.
- If the goal is to prove `ite h1 h2 h3` (or `ite h1 h2 h3 = rhs`), then `split_ifs at H` will turn the current goal into two goals, the first one is to prove `h2` (`h3 = rhs` resp.) with additional assumption `h1`; the second one is to prove `h3` (`h3 = rhs` resp.) with additional assumption `¬h1`
- when the goal is easily provable, one can use the following to finish a goal:
 - `refl` (for reflexive) is used to prove proposition of the form `lhs = rhs` when `lhs` is **definitionally** equal to `rhs`. Definitional equality is more general than two string being literally identical but is less general than being (canonical) isomorphic. For example

$$\sum_{i=0}^{\infty} \frac{1}{2^i} = \sum_{j=0}^{\infty} \frac{1}{2^j}$$

is a definitional equality but

$$\mathbb{R}^n = \text{Func}(\{0, \dots, n-1\}, \mathbb{R})$$

is not a definitional equality (strictly speaking perhaps not an equality at all).

- `exact H` will prove current goal if the goal is definitionally equal to `H`.
- `ring` will try to prove the current goal using associativity and commutativity of addition and multiplication.
- `linarith` is used when proving inequality from context. `linarith` is semi-automated, so it can work with inequalities with symbols or variables but only to a degree. If `linarith` failed, one has to either provide `linarith` with more propositions or use other tactics to change goal into something more manageable for `linarith`.
`linarith [h1, ..., hn]` is equivalent to use `linarith` with additional (proofs of) propositions `h1 ... hn`.

- `tidy` is to ask `Lean` to try different tactics and finishes the goal if it possible.
- If there is multiple goals, one can use `{ }` to focus on the first one.
- If the entirety of proof is one line, one can replace `begin contents end` with `by contents`.

A proposition if not atomic is either a conjunction, a disjunction, an implication, an equivalence, a negation or a proposition with universal quantifier or existential quantifier.

prove a conjunction

If goal is to prove a conjunction of the form $h_1 \wedge h_2$, `split` is used. It will change the current goal to two goals of proving h_1 and h_2 respectively. Then the general pattern is

```
1 theorem how_to_prove_conjunction (h1 : Prop) (h2 : Prop) : h1 ∧ h2 :=
2 begin
3   split,
4
5   proof_of_h1
6
7   proof_of_h2
8 end
```

2.2.1 prove a disjunction

If the goal is to prove a disjunction of the form $h_1 \vee h_2$, one can use `left` to change the goal to prove h_1 or `right` to change the goal to prove h_2 . Let us assume h_1 is a true proposition :

```
1 theorem how_to_prove_disjunction (h1 : Prop) (h2 : Prop) : h1 ∨ h2 :=
2 begin
3   left,
4
5   proof_of_h1
6 end
```

2.2.2 prove an implication

If the goal is to prove an implication of the form $p \implies q$, one can use `intro` hp to add $hp:p$ a proof of p into the context and convert goal to prove q .

```
1 theorem how_to_prove_implication (p : Prop) (q : Prop) : p → q :=
2 begin
3   intro hp,
4
```

```

5 | proof_of_q
6 | end

```

If the goal is of the form $p_1 \rightarrow p_2 \rightarrow \dots p_n$, one can use `intros hp1 ...hpn` as an abbreviation of `intro hp1, intro hp2, ..., intro hpn`.

2.2.3 prove an equivalence

An equivalence of the form $p \iff q$ is by definition $p \implies q \wedge q \implies p$. Thus by `split` will change the goal to two goals, one to prove $p \implies q$, the other to prove $q \implies p$. Then use section 2.2.2.

2.2.4 prove a negation

A negation of the form $\neg p$ is by definition $p \implies \perp$. Thus `intro hp` will add `hp:p` to current context and convert the goal to prove a falsehood.

```

1 | theorem how_to_prove_negation (p : Prop) :  $\neg p$  :=
2 | begin
3 |   intro hp,
4 |
5 |   proof_of_falsehood
6 | end

```

2.2.5 prove a proposition with \forall

A proposition of the form $\forall a : \alpha, p(a)$ where α is a type and $p : \alpha \rightarrow \mathbf{Prop}$ can be proved also using `intro x0`. This will add an arbitrary $x_0 : \alpha$ to the current context and change the goal to prove $p(x_0)$.

```

1 | theorem how_to_proposition_with_universal_quantifier {α : Type} (p
2 |    $\hookrightarrow$  : α → Prop) :  $\forall a : \alpha, p\ a$  :=
3 | begin
4 |   intro x0,
5 |   a_proof_of_p(x0)
6 | end

```

If the goal is the form $\forall a_1 : \alpha_1, \forall a_2 : \alpha_2, \dots, \forall a_n : \alpha_n, p\ a_1\ a_2 \dots a_n$ can be proved using `intros a1 a2 ... an` as an abbreviation of `intro a1, intro a2, ..., intro an`.

2.2.6 prove a proposition with \exists

A proposition of the form $\exists a : \alpha, p(a)$ where α is a type and $p : \alpha \rightarrow \mathbf{Prop}$ can be proved by `use x0`. This will convert the goal to prove $p(x_0)$.

```

1 theorem how_to_proposition_with_universal_quantifier { $\alpha$  : Type} (p
   $\hookrightarrow$  :  $\alpha \rightarrow \mathbf{Prop}$ ) :  $\exists a : \alpha, p\ a :=$ 
2 begin
3   a_construction_of_x0
4
5   use x0,
6
7   a_proof_of_p(x0)
8 end

```

2.3 An example

To illustrate the above syntax and patterns, we present an example of defining mean and proving some basic properties thereof.

```

1 import data.real.basic
2 import tactic
3
4 noncomputable theory
5 open_locale classical
6
7 def mean (x y :  $\mathbb{R}$ ) :  $\mathbb{R}$  := (x + y) / 2
8
9 theorem min_le_mean :  $\forall x\ y : \mathbb{R}, \min x\ y \leq (\text{mean } x\ y) :=$ 
10 begin
11   intros x y,
12   have ineq1 :  $\min x\ y \leq x := \text{min\_le\_left } x\ y,$ 
13   have ineq2 :  $\min x\ y \leq y := \text{min\_le\_right } x\ y,$ 
14
15   unfold mean, rw le_div_iff, rw mul_two,
16   apply add_le_add,
17   exact ineq1, exact ineq2,
18
19   linarith,
20 end
21
22 theorem mean_le_max :  $\forall x\ y : \mathbb{R}, (\text{mean } x\ y) \leq \max x\ y :=$ 
23 begin
24   intros x y,
25   have ineq1 :  $x \leq \max x\ y := \text{le\_max\_left } x\ y,$ 
26   have ineq2 :  $y \leq \max x\ y := \text{le\_max\_right } x\ y,$ 
27
28   unfold mean, rw div_le_iff, rw mul_two,
29   apply add_le_add,
30   exact ineq1, exact ineq2,
31
32   linarith,
33 end
34
35 theorem a_number_in_between :
36    $\forall x\ y : \mathbb{R}, x \leq y \rightarrow \exists z : \mathbb{R}, x \leq z \wedge z \leq y :=$ 
37 begin
38   intros x y hxy,

```

```

39 | have ineq1 := min_le_mean x y,
40 | have ineq2 := mean_le_max x y,
41 | have min_eq_x := min_eq_left hxy,
42 | have max_eq_y := max_eq_right hxy,
43 | use mean x y,
44 | split,
45 |
46 | { conv_lhs {rw <-min_eq_x}, exact ineq1, },
47 | { conv_rhs {rw <-max_eq_y}, exact ineq2, },
48 | end

```

Line 1 will make basic properties of real available to use and line 2 will make all the tactics we discussed amongst other more advanced tactics available to use. We add line 4 so that `lean` would ignore the issue of computability and line 5 so that we can use proof by contradiction².

We define the mean value of two real numbers on line 7. Then `mean`³ has type $\mathbb{R} \rightarrow \mathbb{R} \rightarrow \mathbb{R}$, `mean 1` has type $\mathbb{R} \rightarrow \mathbb{R}$ and `mean 1 2` has type \mathbb{R} .

We can introduce and prove theorems about `mean` that the mean value of two numbers is greater than or equal to the minimum of the two numbers but less than the maximum of the two numbers. This is from line 9 to line 33 where

- `min_le_left` is a proof of the proposition $\forall(x y : \alpha), \min(x, y) \leq x$ where α is an implicit argument with a linear order. In this case, `Lean` infers from context that α is \mathbb{R} . Thus `min_le_left x y` is a proof of `min x y ≤ x`.
- `min_le_right` is a proof of the proposition $\forall(x y : \alpha), \min(x, y) \leq y$ In this case, `min_le_right x y` is a proof of `min x y ≤ y`.
- Similarly, `le_max_left` is a proof of the proposition $\forall(x y : \alpha), x \leq \max(x, y)$ where α is an implicit argument with a linear order. In this case, `le_max_left` is a proof of `x ≤ max x y`.
- Similarly, `le_max_right` is a proof of the proposition $\forall(x y : \alpha), y \leq \max(x, y)$ where α is an implicit argument with a linear order. In this case, `le_max_right` is a proof of `y ≤ max x y`.
- `le_div_iff` is a proof that $0 < c \rightarrow (a \leq \frac{b}{c} \iff a \times c \leq b)$ where a, b, c are elements of a type with a linear ordered field structure. So by `rw le_div_iff`, the goal would change from `min x y ≤ (x + y) / 2` to `min x y * 2 ≤ x + y`. Since `le_div_iff` requires the assumption that $0 < c$, a new goal to prove that $0 < 2$ is created after the original goal. This goal is proved by the final `linarith`.

²`Lean` by default use constructivism where $\neg\neg p \implies p$ is not an axiom of deduction. Thus the law of excluded middle is not by default a tautology.

³`mean` is not a function $\mathbb{R}^2 \rightarrow \mathbb{R}$ but a function $\mathbb{R} \rightarrow \text{Func}(\mathbb{R}, \mathbb{R})$. This is called currying.

- `div_le_iff` is proof that $0 < b \implies (\frac{a}{b} \leq c \iff a \leq c \times b)$ where a, b, c are elements of a type with a linear ordered field structure. So by `rw div_le_iff` the goal would change from $(x + y) / 2 \leq \max x y$ to $x + y \leq \max x y * 2$. Since `div_le_iff` requires the assumption that $0 < b$, a new goal to prove $0 < 2$ is created after the original goal. This goal is proved by the final `linarith`.
- `mul_two` proves the lemma that $\forall n : \alpha, n \times 2 = n + n$ where α is a semiring. Thus `rw mul_two` would change the goal of proving $\min x y * 2 \leq x + y$ ($x + y \leq \max x y * 2$ resp.) to $\min x y + \min x y \leq x + y$ ($x + y \leq \max x y + \max x y$ resp.).
- `add_le_add` proves the lemma that $a \leq b \rightarrow c \leq d \rightarrow a + c \leq b + d$ where a, b, c and d are elements of an ordered additive commutative monoid. Since the goal now is to prove $\min x y + \min x y \leq x + y$, by apply `add_le_add`, goal will be replaced by two goals of proving $\min x y \leq x$ and $\min x y \leq y$. These are *exactly* `ineq1` and `ineq2`.

Chapter 3

Formalisation using Lean

Logistics of the formalisation

There are five main files in the formalisation where

1. `small_things.lean` formalised results about the trivial embedding of $\mathbb{Z}[X] \subset \mathbb{R}[X]$ and manipulation of inequality in real numbers common to all three parts;
2. `algebraic_over_Z.lean` formalised countability of algebraic numbers. In this file we made an extensive use of Schröder-Berstein theorem.
3. `liouville.lean` formalised Liouville's theorem and a construction of a Liouville's number;
4. `e_trans_helpers2.lean` formalised some results about differentiation and integration. Especially the formalisations of

$$\frac{d^n}{dx^n} uv = \sum_{i=0}^n \binom{n}{i} \frac{d^i u}{dx^i} \frac{d^{n-i} v}{dx^{n-i}}$$

where u and v are differentiable function from \mathbb{R} to \mathbb{R} and

$$\int_0^t e^{t-x} f(x) dx = e^t \sum_{i=0}^m f^{(i)}(0) - \sum_{i=0}^m f^{(i)}(t)$$

where $f(X) \in \mathbb{Z}[X]$;

5. `e_transcendental.lean` formalised transcendence of e by assuming the algebraicity of e which resulted in two contradictory bounds using the results from `e_trans_helpers2.lean`.

3.1 Countability argument

The main caveat in this part is internal specification of `mathlib`. A real number x is in `Lean` is algebraic over \mathbb{Z} if and only if there exists a nonzero polynomial $p(X) \in \mathbb{Z}[X]$ such that p is in the kernel of the unique \mathbb{Z} -algebra homomorphism $\mathbb{Z}[X] \rightarrow \mathbb{R}$ given by $X \mapsto x$.

```
1  ∃ (p : ℤ[X]), p ≠ 0 ∧ ↑(polynomial.aeval ℤ ℝ x) p = 0
```

Here the \mathbb{Z} -algebra homomorphism is `polynomial.aeval ℤ ℝ x`. `↑` is to convert the homomorphism to a function applicable to `p`. The reason that a conversion is necessary is because algebra homomorphism contains more information than a function, it is a structure containing the map and other fields containing (proofs of) properties of algebra homomorphism. However in polynomial library of `mathlib`, the definition of root is as following :

```
1  def is_root (p : polynomial R) (a : R) : Prop := p.eval a = 0
```

Thus the first part of this formalisation is to unify the two evaluation methods – denote i to be the trivial embedding $\mathbb{Z}[X] \subseteq \mathbb{R}[X]$ and ι_x to be the unique \mathbb{Z} -algebra homomorphism $\iota_x : \mathbb{Z}[X] \rightarrow \mathbb{R}$ given by $X \mapsto x$ then for all polynomial $p(X) \in \mathbb{Z}[X]$, then $\forall x \in \mathbb{R}, (ip)(x) = \iota_x p$:

```
1  -- the trivial embedding ℤ[X] ⊆ ℝ[X]
2  def poly_int_to_poly_real (p : ℤ[X]) : polynomial ℝ :=
   ↪ polynomial.map Zembℝ p
3
4  def poly_int_to_poly_real_wd (p : ℤ[X]) :=
5    ∀ x : ℝ, polynomial.aeval ℤ ℝ x p = (poly_int_to_poly_real
   ↪ p).eval x
6
7  theorem poly_int_to_poly_real_well_defined
8    (x : ℝ) (p : ℤ[X]) : poly_int_to_poly_real_wd p :=
9    begin
10     proof_omitted
11   end
```

Source Code 3.1: unifying two ways of evaluation

For any $p \in \mathbb{Z}[X]$, we can define the set of roots to be $\{x \in \mathbb{R} \mid (ip)(x) = 0\}$ or $\{x \in \mathbb{R} \mid \iota_x p = 0\}$ where the former is builtin as `↑(poly_int_to_poly_real p).roots`¹ and the latter is defined as line 1 in source code 3.2. By line 7 in source code 3.1, two sets must be equal, then two sets are have finite cardinality:

```
1  def roots_real (p : ℤ[X]) : set ℝ :=
2    {x | polynomial.aeval ℤ ℝ x p}
```

¹ `_.roots` in fact has type `finset ℝ`. The type `finset` is a `set` with a proof of finite cardinality. Here `↑` is used to convert a `finset` to `set` by discarding the proof of finite cardinality.

```

3
4 theorem roots_real_eq_roots (p :  $\mathbb{Z}[X]$ ) (hp : p ≠ 0) :
5   roots_real p = ↑(poly_int_to_poly_real p).roots :=
6 begin
7   proof_omitted
8 end
9
10 theorem roots_finite (p :  $\mathbb{Z}[X]$ ) (hp : p ≠ 0) :
11   set.finite (roots_real p) :=
12 begin
13   proof_omitted
14 end

```

Source Code 3.2: two ways of defining roots

We defined the set of all algebraic numbers over \mathbb{Z} to be

```

1 def algebraic_set : set  $\mathbb{R}$  := {x | is_algebraic  $\mathbb{Z}$  x}

```

To investigate the countability of `algebraic_set`, we compare it with

$$\bigcup_{n \in \mathbb{N}} \bigcup_{\substack{p \in \mathbb{Z}[X] \\ p \neq 0 \\ \deg p < n+1}} \{x \in \mathbb{R} \mid \iota_x p = 0\}. \quad (3.1)$$

To this end, we introduce some types of interest:

```

1 notation `int_n` n := fin n →  $\mathbb{Z}$ 
2 notation `nat_n` n := fin n →  $\mathbb{N}$ 
3 notation `poly_n'` n := {p :  $\mathbb{Z}[X]$  // p ≠ 0 ∧ p.nat_degree < n}
4 notation `int_n'` n := {f : fin n →  $\mathbb{Z}$  // f ≠ 0}
5 notation `int'` := {r :  $\mathbb{Z}$  // r ≠ 0}

```

where $\langle m, hm \rangle$ is an element of `fin n` if and only if m is a natural number and hm is a proof of $m < n$. Then `fin n` is the type of only n elements. Thus

- `int_n n` is \mathbb{Z}^n ;
- `int_n' n` is $\mathbb{Z}^n - \{(0, \dots, 0)\}$;
- `int'` is $\mathbb{Z} - \{0\}$;
- `nat_n n` is \mathbb{N}^n ;
- `poly_n' n` is the type of non-zero integer polynomials with degree less than n .

Then $\mathbb{Z} \simeq \mathbb{Z} - \{0\}$ by the bijective function $s : \mathbb{Z} \rightarrow \mathbb{Z} - \{0\}$:

$$n \mapsto \begin{cases} m & \text{if } m < 0 \\ m + 1 & \text{if } m \geq 0 \end{cases}$$

```

1 def strange_fun :  $\mathbb{Z} \rightarrow \text{int}' :=$ 
2    $\lambda m, \text{if } h : m < 0$ 
3     then  $\langle m, \text{by linarith} \rangle$ 
4     else  $\langle m + 1, \text{by linarith} \rangle$ 
5
6 theorem strange_fun_inj :
7   function.injective strange_fun :=
8 begin
9   proof_omitted
10 end
11
12 theorem strange_fun_sur :
13   function.surjective strange_fun :=
14 begin
15   proof_omitted
16 end
17
18 theorem int_equiv_int' :  $\mathbb{Z} \simeq \text{int}' :=$ 
19 begin
20   apply equiv.of_bijective strange_fun,
21   split,
22   exact strange_fun_inj,
23   exact strange_fun_sur,
24 end

```

Source Code 3.3: $\mathbb{Z} \simeq \mathbb{Z} - \{0\}$

Then we prove that for all non-zero $n : \mathbb{N}$, non-zero integer polynomials of degree less than n bijectively correspond to $\mathbb{Z}^n - \{(0, \dots, 0)\}$ via the function: $p \mapsto \mathbf{z}$ where the i -th coordinate of \mathbf{z} is the i -th coefficient of p .

```

1 def identify (n : nat) : (poly_n' n)  $\rightarrow$  (int_n' n) :=
2    $\lambda p, \langle \lambda m, p.1.\text{coeff } m.1, \text{a\_proof\_z\_is\_not\_zero} \rangle$ 
3
4 theorem sur_identify_n (n : nat) (hn : n  $\neq$  0) :
5   function.surjective (identify n) :=
6 begin
7   proof_omitted
8 end
9
10 theorem inj_identify_n (n : nat) (hn : n  $\neq$  0) :
11   function.injective (identify n) :=
12 begin
13   proof_omitted
14 end
15
16 theorem poly_n'_equiv_int_n' (n : nat) :
17   (poly_n' n.succ)  $\simeq$  (int_n' n.succ) :=
18 begin
19   apply equiv.of_bijective (identify n.succ),
20   split,
21   exact inj_identify_n n.succ (nat.succ_ne_zero n),
22   exact sur_identify_n n.succ (nat.succ_ne_zero n),
23 end

```

Source Code 3.4: non-zero integer polynomial with degree less than n has the same cardinality as $\mathbb{Z}^n - \{(0, \dots, 0)\}$, here $\mathbf{n.succ}$ means $n + 1$.

Then we define two injective functions $F : \mathbb{Z}^{n+1} \rightarrow \mathbb{Z}^{n+1} - \{(0, \dots, 0)\}$ and $G : \mathbb{Z}^{n+1} - \{(0, \dots, 0)\} \rightarrow \mathbb{Z}^{n+1}$ by:

$$F(m_0, \dots, m_n) = (s(m_0), \dots, s(m_n))$$

$$G(m_0, \dots, m_n) = (m_0, \dots, m_n)$$

where $s : \mathbb{Z} \rightarrow \mathbb{Z} - \{0\}$ is defined previously. By Schröder-Berstein theorem, there is then a bijection $\mathbb{Z}^{n+1} \rightarrow \mathbb{Z}^{n+1} - \{(0, \dots, 0)\}$ and thus $\mathbb{Z}^{n+1} \simeq \mathbb{Z}^{n+1} - \{(0, \dots, 0)\}$:

```

1 def F (n : nat) : (int_n n.succ) → (int_n' n.succ) :=
2   λ f, ⟨λ m, (strange_fun (f m)).1,
3     a_proof_of_(s(m₀),...,s(mₙ))_non-zero⟩
4 theorem F_inj (n : nat) : function.injective (F n) :=
5 begin
6   proof_omitted
7 end
8
9 def G (n : nat) : (int_n' n.succ) → (int_n n.succ) :=
10  λ f m, (f.1 m)
11 theorem G_inj (n : nat) : function.injective (G n) :=
12 begin
13   proof_omitted
14 end
15
16 theorem int_n_equiv_int_n' (n : nat) :
17   (int_n n.succ) ≃ int_n' n.succ :=
18 begin
19   choose B HB using function.embedding.schroeder_bernstein (F_inj
20     ↪ n) (G_inj n),
21   apply equiv.of_bijective B HB,
end

```

Source Code 3.5: $\mathbb{Z}^{n+1} \simeq \mathbb{Z}^{n+1} - \{(0, \dots, 0)\}$

For any natural number $n \geq 1$, we then construct two injective function $f_n : \mathbb{Z}^{n+2} \rightarrow \mathbb{Z}^{n+1} \times \mathbb{Z}$ and $g_n : \mathbb{Z}^{n+1} \times \mathbb{Z} \rightarrow \mathbb{Z}^{n+2}$:

$$f_n((m_0, \dots, m_{n+1})) = ((m_0, \dots, m_n), m_{n+1})$$

$$g_n(((m_0, \dots, m_n), m_{n+1})) = (m_0, \dots, m_{n+1})$$

Then by Schröder-Berstein theorem $\mathbb{Z}^{n+2} \simeq \mathbb{Z}^{n+1} \times \mathbb{Z}$ for all $n \geq 1$.

```

1 def fn (n : nat) :
2   (int_n n.succ.succ) → (int_n n.succ) × ℤ := λ r,
3   ⟨λ m, r (⟨m.1, nat.lt_trans m.2 (nat.lt_succ_self n.succ))),
4     r (⟨n.succ, nat.lt_succ_self n.succ⟩)⟩
5 theorem fn_inj (n : ℕ) : function.injective (fn n) :=

```

```

6  begin
7    proof_omitted
8  end
9
10 def gn (n : nat) : (int_n n.succ) × ℤ → (int_n n.succ.succ) := λ r
11   ↪ m,
12 begin
13   by_cases (m.1 = n.succ),
14   exact r.2,
15   exact r.1 ((m.1, lt_of_le_of_ne (fin.le_last m) h)),
16 end
17 theorem gn_inj (n : nat) : function.injective (gn n) :=
18 begin
19   proof_omitted
20 end
21
22 theorem aux_int_n (n : nat) :
23   (int_n n.succ.succ) ≃ (int_n n.succ) × ℤ :=
24 begin
25   choose B HB using function.embedding.schroeder_bernstein (fn_inj n)
26   ↪ (gn_inj n),
27   apply equiv.of_bijective B HB,
28 end

```

Source Code 3.6: $\mathbb{Z}^{n+2} \simeq \mathbb{Z}^{n+1} \times \mathbb{Z}$ for all $n \geq 1$

Now we are finally in the position of using formula 3.1 to prove the countability of all algebraic numbers. We first define the set of real roots of non-zero integer polynomial of degree less than n to be:

```

1  def algebraic_set'_n (n : ℕ) : set ℝ :=
2    ⋃ p : (poly_n' n.succ), roots_real p.1

```

Hence by taking union over all natural numbers we can obtain an equivalent definition of all algebraic number over \mathbb{Z} :

```

1  def algebraic_set' : set real :=
2    ⋃ n : ℕ, algebraic_set'_n n.succ
3
4  theorem algebraic_set'_eq_algebraic_set :
5    algebraic_set' = algebraic_set :=
6  begin
7    proof_omitted
8  end

```

We prove by induction that for any $n \in \mathbb{N}$, \mathbb{Z}^{n+1} is denumerable (i.e. countably infinite) where the base case is $\mathbb{Z}^1 \simeq \mathbb{Z}$ and the inductive step is to prove \mathbb{Z}^{n+2} is denumerable using denumerability of \mathbb{Z}^{n+1} . Since non-zero integer polynomials of degree less than $n + 1$ bijectively corresponds to \mathbb{Z}^{n+1} , we have non-zero integer polynomials of degree less than $n + 1$ is denumerable hence is countable. Then the result of taking union over the countable set \mathbb{N} , the result

$\bigcup_{n \in \mathbb{N}} \bigcup_{p \in \mathbb{Z}[X], p \neq 0, \deg p < n+1} \{x \in \mathbb{R} \mid \iota_x p = 0\}$ is still countable. Then finally the set of all algebraic numbers over \mathbb{Z} is countable. Since \mathbb{R} is uncountable, transcendental number must exist:

```

1  theorem int_1_equiv_int : (int_n 1)  $\simeq$   $\mathbb{Z}$  :=
2  begin
3    proof_omitted
4  end
5
6  theorem int_n_denumerable {n : nat} :
7    denumerable (int_n n.succ) :=
8  begin
9    proof_omitted
10 end
11
12 theorem poly_n'_denumerable (n : nat) :
13   denumerable (poly_n' n.succ) :=
14 begin
15   proof_omitted
16 end
17
18 theorem algebraic_set'_n_countable (n : nat) :
19   set.countable (algebraic_set'_n n) :=
20 begin
21   proof_omitted
22 end
23
24 theorem algebraic_set'_countable :
25   set.countable algebraic_set' :=
26   set.countable_Union
27     ( $\lambda$  n, algebraic_set'_n_countable n.succ)
28
29 theorem algebraic_set_countable :
30   set.countable algebraic_set :=
31 begin
32   rw  $\leftarrow$  algebraic_set'.eq_algebraic_set,
33   exact algebraic_set'_countable
34 end
35
36 theorem transcendental_number_exists :
37    $\exists$  x :  $\mathbb{R}$ , transcendental x :=
38 begin
39   proof_omitted
40 end

```

Source Code 3.7: algebraic numbers are countable, hence transcendental numbers exists

3.2 Liouville's theorem and Liouville number

General theory about Liouville number

A Liouville number is a real number that is “almost rational”, i.e. for any $n \in \mathbb{N}$ there is a rational number² $\frac{a}{b} \in \mathbb{Q}$ such that $b > 1$ and $0 < |x - \frac{a}{b}| < \frac{1}{b^n}$.

```

1 def liouville_number (x : ℝ) :=
2   ∀ n : ℕ, ∃ a b : ℤ,
3     b > 1 ∧
4     0 < abs(x - a / b) ∧ abs(x - a / b) < 1/b^n

```

Source Code 3.8: Definition of Liouville number

We first prove a lemma about irrational root of an integer polynomial:

Lemma 3.2.1. if f is an integer polynomial with degree $m > 1$ and α is an irrational root for $i(f)$ where $i : \mathbb{Z}[X] \rightarrow \mathbb{R}[X]$ is the trivial embedding, then there is a positive real number A such that for every rational number $\frac{a}{b}$,

$$\left| \alpha - \frac{a}{b} \right| > \frac{A}{b^m}:$$

```

1 lemma about_irrational_root (α : ℝ)
2   (hα : irrational α) (f : ℤ[X])
3   (f_deg : f.nat_degree > 1)
4   (α_root : f_eval_on_ℝ f α = 0) :
5   ∃ A : ℝ, A > 0 ∧ ∀ a b : ℤ, b > 0 → abs(α - a/b) >
    ↪ (A/b^(f.nat_degree)) :=

```

Proof. We will abuse the notation to denote f as $i(f) \in \mathbb{R}[X]$

```

6 begin
7   have f_nonzero : f ≠ 0,
8     proof_omitted
9   generalize hfℝ: f.map Zembℝ = f_ℝ,
10  have hfℝ_nonzero : f_ℝ ≠ 0,
11    proof_omitted
12  generalize hDf: f_ℝ.derivative = Df_ℝ,

```

Since $\text{abs} \circ Df : \mathbb{R} \rightarrow \mathbb{R}$ given by $x \mapsto \left| \frac{d}{dt} \right|_{t=x} f(t)$ is a continuous function and $[\alpha - 1, \alpha + 1]$ is a non-empty compact subset of \mathbb{R} , $\text{abs} \circ Df$ attains a maximum on $[\alpha - 1, \alpha + 1]$ denote it by M .

```

13 have H := is_compact.exists_forall_ge
14   a_proof_of_[α - 1, α + 1]_compact

```

²Without lose of generality, we are always assuming the denominator is a strictly positive natural number.

```

15         a_proof_of_ $[\alpha - 1, \alpha + 1]$ _not_empty
16         a_proof_of_abs  $\circ$  Df_continuous,
17
18     choose x_max hx_max using H,
19     generalize M_def: abs (Df_ℝ.eval x_max) = M,
20     have hM := hx_max.2, rw M_def at hM,
21     have M_non_zero : M ≠ 0,
22     proof_omitted
23     have M_pos : M > 0,
24     proof_omitted

```

Let us consider the smallest element B of the set $\{1, \frac{1}{M}\} \cup \{|\alpha - x| \mid f(x) = 0 \wedge x \neq \alpha\}$, then $B > 0$.

```

25     generalize roots_def : f_ℝ.roots = f_roots,
26     generalize roots'_def : f_roots.erase α = f_roots',
27     generalize roots_distance_to_α : f_roots'.image (λ x, abs (α -
28         ↪ x)) = distances,
29     generalize hdistances' : insert (1/M) (insert (1:ℝ) distances) =
30         ↪ distances',
31     have hnon_empty: distances'.nonempty,
32     proof_omitted
33     generalize hB : finset.min' distances' hnon_empty = B,
34     have allpos : ∀ x : ℝ, x ∈ distances' → x > 0,
35     proof_omitted
36     have B_pos : B > 0,
37     proof_omitted

```

Let $A = \frac{B}{2}$ then $A > B > 0$. We claim that A satisfies the lemma, i.e. $A > 0$ and for every rational number $\frac{a}{b}$, $|\alpha - a/b| > \frac{A}{b^m}$ where m is the degree of f .

```

36     generalize hA : B / 2 = A,
37     use A, split,
38     a_proof_of_A > 0

```

We proceed by assuming that there exists a rational number $\frac{a}{b}$ such that $\left|\alpha - \frac{a}{b}\right| \leq \frac{A}{b^m}$ for a contradiction. Since $b \geq 1$, we have $\left|\alpha - \frac{a}{b}\right| \leq A < B$. Then $\frac{a}{b}$ is not root of f because otherwise $B \leq \left|\alpha - \frac{a}{b}\right|$.

```

39     by_contra absurd,
40     simp only [gt_iff_lt, classical.not_forall, not_lt,
41         ↪ classical.not_imp] at absurd,
42     choose a ha using absurd,
43     choose b hb using ha,
44     have hb2 : b ^ f.nat_degree ≥ 1,
45     proof_omitted

```

```

45 have hb21 : abs (α - a / b) ≤ A,
46   proof_omitted
47 have hb22 : abs (α - a/b) < B,
48   proof_omitted
49 have hab0 : (a/b:ℝ) ∈ set.Icc (α-1) (α+1),
50   proof_omitted
51 have hab1 : (a/b:ℝ) ≠ α,
52   proof_omitted
53 have hab2 : (a/b:ℝ) ∉ f_roots,
54   proof_omitted

```

Since $\alpha \neq \frac{a}{b}$, we can assume without loss of generality that $\frac{a}{b} < \alpha$. Since $\text{eval}_f : \mathbb{R} \rightarrow \mathbb{R}$ given by $x \mapsto f(x)$ is differentiable, we can use mean value theorem to find $x_0 \in (\frac{a}{b}, \alpha)$ such that

$$\begin{aligned}
Df(x_0) &= \frac{\text{eval}_f(\alpha) - \text{eval}_f(\frac{a}{b})}{\alpha - \frac{a}{b}} \quad [\text{Mean value theorem}] \\
&= -\frac{\text{eval}_f(\frac{a}{b})}{\alpha - \frac{a}{b}} \quad [\alpha \text{ is a root of } i(f)]
\end{aligned}$$

```

55 have hab3 := ne_iff_lt_or_gt.1 hab1,
56 cases hab3,
57 have H :=
58   exists_deriv_eq_slope (λ x, f_ℝ.eval x) hab3 _ _,
59   choose x0 hx0 using H,
60   have hx0r := hx0.2,
61   rw [polynomial.deriv, hDf, ←hfℝ] at hx0r,
62   rw [f_eval_on_ℝ] at α_root, rw [α_root, hfℝ] at hx0r, simp only
   ↪ [zero_sub] at hx0r,

```

Then $|Df(x_0)| > 0$ hence $\left| \alpha - \frac{a}{b} \right| = \left| \frac{\text{eval}_f(\frac{a}{b})}{Df(x_0)} \right|$ is non-zero. Since M is the maximum of $\text{abs} \circ Df$ on $[\alpha - 1, \alpha + 1]$. We have $|Df(x_0)| \leq M$ and thus $\left| \alpha - \frac{a}{b} \right| \geq \frac{|\text{eval}_f(\frac{a}{b})|}{M}$. If we write $f(X)$ as $\sum_{j=0}^m \lambda_j X^j$ then

$$\left| \text{eval}_f \left(\frac{a}{b} \right) \right| = \left| \sum_{j=0}^m \lambda_j \frac{a^j}{b^j} \right| = \frac{1}{b^m} \left| \sum_{j=0}^m \lambda_j a^j b^{m-j} \right| \geq \frac{1}{b^m}$$

Hence we have $\left| \alpha - \frac{a}{b} \right| \geq \frac{1}{Mb^m} > \frac{A}{b^m}$. But we assumed $\left| \alpha - \frac{a}{b} \right| < \frac{A}{b^m}$ to start with, this is the desired contradiction.

```

63 have Df_x0_nonzero : Df_ℝ.eval x0 ≠ 0,
64   proof_omitted
65 have H2 : abs(α - a/b) = abs((f_ℝ.eval (a/b:ℝ)) / (Df_ℝ.eval
   ↪ x0)),

```

```

66     proof_omitted
67
68     have ineq' : polynomial.eval (a/b:ℝ) (polynomial.map ℤembℝ f) ≠
        ↪ 0,
69     proof_omitted
70     have ineq : abs (α - a/b) ≥ 1/(M*b^(f.nat_degree)),
        proof_omitted
71     have ineq2 : 1/(M*b^(f.nat_degree)) > A / (b^f.nat_degree),
        proof_omitted
72     have ineq3 : abs (α - a / b) > A / b ^ f.nat_degree,
        proof_omitted
73     have ineq4 : abs (α - a / b) > abs (α - a / b),
        proof_omitted
74     linarith,
75
76     We omit the proof of differentiability of evalf, continuity of abs ∘ Df and the case when
77     ↪  $\frac{a}{b} > \alpha$ 
78     rest_omitted
79
80 end
81
82

```

□

We then prove irrationality of Liouville number.

Lemma 3.2.2. Every Liouville number is irrational

```

1 lemma liouville_numbers_irrational: ∀ (x : ℝ), (liouville_number x)
  ↪ → irrational x :=

```

Proof. Let x be an arbitrary Liouville number and suppose for a contradiction that $x = \frac{a}{b}$, write $n = b + 1$ then $2^{n-1} > b$.

```

2 begin
3   intros x liouville_x a b hb rid,
4   replace rid : x = ↑a / ↑b, linarith,
5   generalize hn : b.nat_abs + 1 = n,
6   have b_ineq : 2 ^ (n-1) > b,
7   proof_omitted

```

Since $x = \frac{a}{b}$ is a Liouville number we can find a rational number $\frac{p}{q}$ such that $q > 1$ and $0 < \left| \frac{a}{b} - \frac{p}{q} \right| < \frac{1}{q^n}$ or equivalently $0 < \frac{|aq - bp|}{bq} < \frac{1}{q^n}$. If $aq - bp = 0$, then $0 < 0$ is the desired contradiction.

```

8   choose p hp using liouville_x n,
9   choose q hq using hp, rw rid at hq,
10  have q_pos : q > 0 := by linarith,
11  rw [div_sub_div at hq, abs_div at hq],
12
13  by_cases (abs (a*q-b*p:ℝ) = 0),

```

```

14 intermediate_step_omitted
15 linarith,

```

If $aq - bp \neq 0$ then $\frac{1}{bq} \leq \frac{|aq - bp|}{bq}$. But we also have $b < 2^{n-1}$ and $2^{n-1}q \leq q^n$ because $q \geq 2$. Hence $bq < q^n$, then $\frac{|aq - bp|}{bq} > \frac{1}{q^n}$. This is the desired contradiction.

```

16 have ineq4 : 1 / (b * q : ℝ) ≤ (abs(a * q - b * p : ℝ)) / (b * q),
17   proof_omitted
18 have b_ineq'' : (b * q : ℝ) < (2 : ℝ)^(n-1) * (q : ℝ),
19   proof_omitted
20 have q_ineq3 : 2 ^ (n - 1) * q ≤ q ^ n,
21   proof_omitted
22 have b_ineq2 : b * q < q ^ n, linarith,
23 have rid'' :
24   abs (a * q - b * p : ℝ) / (b * q : ℝ) > 1 / q ^ n,
25   proof_omitted,
26
27 have hq22 := hq2.2,
28 linarith,
29
30 We manipulated inequalities involving division and multiplication hence we need to prove
31   ↪ several things to be positive.
32 proofs_omitted
end

```

□

With the above lemmas, we are ready to prove the transcendence of Liouville numbers.

Theorem 3.2.1. *Every Liouville number is transcendental*

```

1 theorem liouville_numbers_transcendental : ∀ x : ℝ,
  ↪ liouville_number x → transcendental x :=

```

Proof. Let x be an arbitrary Liouville number then x is irrational. Assume for a contradiction that x is algebraic, let f be the non-zero integer polynomial admitting x as root as a \mathbb{R} -polynomial. Then since x is irrational, f has degree at least 2.

```

2 begin
3   intros x liouville_x,
4   have irrational_x : irrational x := liouville_numbers_irrational
5     ↪ x liouville_x,
6   intros rid, rw is_algebraic at rid,
7   choose f hf using rid,
8   have f_deg : f.nat_degree > 1,
9     proof_omitted

```

By using lemma 3.2.1 we can find a real number $A > 0$ such that for any rational number $\frac{p}{q}$, $\left|x - \frac{p}{q}\right| > \frac{A}{q^n}$ where n is the degree of f .

```

9   have about_root : f_eval_on_ℝ f x = 0,
10   proof_omitted
11   choose A hA using about_irrational_root x irrational_x f f_deg
    ↪ about_root,
12   have A_pos := hA.1,

```

Since \mathbb{R} is an Archimedean field, we can find an $r \in \mathbb{N}$ such that $\frac{1}{A} \leq 2^r$. Then consider $m := r + n$. Since x is a Liouville number, there is a rational number $\frac{a}{b}$ such that $b > 1$ and $0 < \left|x - \frac{a}{b}\right| < \frac{1}{b^m} = \frac{1}{b^r b^n}$.

```

13  have exists_r := pow_big_enough A A_pos,
14  choose r hr using exists_r,
15  have hr' : 1/(2^r) ≤ A,
16  proof_omitted
17  generalize hm : r + f.nat_degree = m,
18  replace liouville_x := liouville_x m,
19  choose a ha using liouville_x,
20  choose b hb using ha,
21
22  have ineq : abs (x-a/b:ℝ) < 1/((b:ℝ)^r)*(1/(b:ℝ)^f.nat_degree),
23  proof_omitted

```

Since $b \geq 2$, we have $\frac{1}{b^r} \leq \frac{1}{2^r} \leq A$. Thus $\left|x - \frac{a}{b}\right| < \frac{1}{b^r b^n} \leq \frac{A}{b^n}$. This contradicts lemma 3.2.1 stating that $\left|x - \frac{a}{b}\right| > \frac{A}{q^n}$.

```

24  have ineq3 : 1/(b:ℝ)^r ≤ A,
25  proof_omitted,
26  have ineq4 : 1/(b:ℝ)^r * (1/(b:ℝ)^f.nat_degree) ≤ (A /
    ↪ (b:ℝ)^f.nat_degree),
27  proof_omitted
28  have ineq5 : abs (x - a/b:ℝ) < A/(b:ℝ)^f.nat_degree, linarith,
29  have rid := hA.2 a b _, linarith, linarith,
30  end

```

□

Construction of a Liouville number

Knowing that all Liouville numbers are transcendental, we now focus on constructing a Liouville number

$$\alpha = \sum_{j=0}^{\infty} \frac{1}{10^{j!}}$$

hence obtain a concrete example of transcendental number α .

Lemma 3.2.3. α converges.

Proof. Since for any $n \in \mathbb{N}$ we have $\frac{1}{10^n}$ is none-negative and $\frac{1}{10^n} \leq \frac{1}{10^{n!}}$, we can use comparison test against $\sum_{j=0}^{\infty} \frac{1}{10^j}$ to deduce the convergence of α .

```

1 def ten_pow_n_fact_inverse (n : ℕ) : ℝ :=
2   (1/10)^n.fact
3 def ten_pow_n_inverse (n : ℕ) : ℝ :=
4   (1/10)^n
5
6 lemma summable_ten_pow_n_fact_inverse : summable
7   ↪ ten_pow_n_fact_inverse :=
8   begin
9     exact @summable_of_nonneg_of_le _
10      ten_pow_n_inverse
11      ten_pow_n_fact_inverse
12      a_proof_of_1_10^n ≥ 0
13      a_proof_of_1_10^n ≤ 1/10^n!
14      a_proof_of_∑_{j=0}^∞ 1/10^j converges,
15
16 end
17
18 def α := ∑' n, ten_pow_n_fact_inverse n

```

†: In Lean, \sum' is to indicate infinite sum while \sum is for finite sum. \square

Lemma 3.2.4. For every $k \in \mathbb{N}$, there exists some $p_k \in \mathbb{N}$ such that

$$\sum_{j=0}^k \frac{1}{j^{k!}} = \frac{p_k}{10^{k!}}$$

```

1 notation `α_k` k := ∑ ii in finset.range(k+1),
2   ↪ ten_pow_n_fact_inverse ii
3
4 notation `α_k_rest` k :=
5   ∑ ii, ten_pow_n_fact_inverse (ii+(k+1))
6
7 theorem α_k_rat (k:ℕ) :
8   ∃ (p:ℕ), α_k k = (p:ℝ)/((10:ℝ)^k.fact) :=

```

Proof. We prove by induction on k . For $k = 0$, the zeroth partial sum is $\frac{1}{10^{0!}} = \frac{1}{10}$. Thus we can pick $p_0 = 1$.

```

1 begin
2   induction k with k IH,
3
4   simp only [ten_pow_n_fact_inverse, pow_one, finset.sum_singleton,
5     ↪ finset.range_one, nat.fact_zero],
6   use 1, norm_cast,

```

Assuming that $\sum_{j=0}^k \frac{1}{10^j!} = \frac{p_k}{10^{k!}}$, let $m := 10^{(k+1)!-k!}$, then we can set $p_{k+1} := p_k m + 1$ then

$$\sum_{j=0}^{k+1} \frac{1}{10^j!} = \frac{p_k}{10^{k!}} + \frac{1}{10^{(k+1)!}} = \frac{p_k m + 1}{10^{(k+1)!}} = \frac{p_{k+1}}{10^{(k+1)!}}$$

```

6   choose pk hk using IH,
7   rw α_k at hk ⊢,
8   generalize hm : 10^((k+1).fact - k.fact) = m,
9   generalize hp : pk * m + 1 = p,
10  use p,
11  proof_omitted
12 end

```

† : At line 1 and 4 above, we use `ii` as indexing variable is to avoid clashes. \square

Theorem 3.2.2. α is a Liouville number

```

1 theorem liouville_α : liouville_number α :=

```

Proof. We need to prove that for an arbitrary $n \in \mathbb{N}$, there exists a rational number $\frac{p(n)}{q(n)}$ such that $p(n) > 1$ and $0 < \left| \alpha - \frac{p(n)}{q(n)} \right| < \frac{1}{q(n)^n}$. By lemma 3.2.4 We know that for some $p \in \mathbb{N}$,

$$\alpha = \sum_{j=0}^n \frac{1}{10^j!} + \sum_{j=0}^{\infty} \frac{1}{10^{(j+n+1)!}} = \frac{p}{10^{n!}} + \sum_{j=0}^{\infty} \frac{1}{10^{(j+n+1)!}}.$$

We take $p(n)$ to be p and $q(n)$ to be $10^{n!}$. Then $10^{n!} > 1$, thus it suffices to prove $0 < \left| \sum_{j=0}^{\infty} \frac{1}{10^{(j+n+1)!}} \right| < \left(\frac{1}{10^{n!}} \right)^n$

```

2 begin
3   intro n,
4   have lemma1 := α_k_rat n,
5   have lemma2 : (α_k_rest n) = α - α_k n,
6   proof_omitted

```



```

7 | choose p hp using lemma1,
8 | use p, use 10^(n.fact),
9 | suffices : 0 < abs (α_k_rest n) ∧
10 |   abs (α_k_rest n) < 1/(10^n.fact)^n,
11 |   split,
12 |   a_proof_of_10^n! > 1,
13 |   tidy,
14 |   split,

```

Since each summand is strictly positive, $\left| \sum_{j=0}^{\infty} \frac{1}{10^{(j+n+1)!}} \right| = \sum_{j=0}^{\infty} \frac{1}{10^{(j+n+1)!}} > 0$. Then we prove $\left| \sum_{j=0}^{\infty} \frac{1}{10^{(j+n+1)!}} \right| < \left(\frac{1}{10^n} \right)^n$, or equivalently $\sum_{j=0}^{\infty} \frac{1}{10^{(j+n+1)!}} < \left(\frac{1}{10^n} \right)^n$ instead. Because for all $j \in \mathbb{N}$, $10^j \times 10^{(n+1)!} \leq 10^{(j+(n+1))!}$, we have

$$\begin{aligned}
\sum_{j=0}^{\infty} \frac{1}{10^{(j+(n+1))!}} &\leq \sum_{j=0}^{\infty} \left(\frac{1}{10^j} \frac{1}{10^{(n+1)!}} \right) = \frac{1}{10^{(n+1)!}} \sum_{j=0}^{\infty} \frac{1}{10^j} \\
&= \frac{10}{9} \frac{1}{10^{(n+1)!}} < \frac{2}{10^{(n+1)!}} < \left(\frac{1}{10^n} \right)^n
\end{aligned}$$

```

15 | rw [α_k_rest, abs_of_pos (α_k_rest_pos n)],
16 |
17 | have ineq2 :
18 |   (∑' (j : ℕ), ten_pow_n_fact_inverse (j+(n+1))) ≤
19 |   (∑' (i : ℕ), (1/10 : ℝ)^i * (1/10 : ℝ)^(n+1).fact),
20 |   proof_omitted
21 | have ineq3 :
22 |   (∑' (i : ℕ), (1/10 : ℝ)^i * (1/10 : ℝ)^(n.fact*n.succ)) ≤
23 |   (2/10^n.succ.fact : ℝ),
24 |   proof_omitted
25 | have ineq4 : (2 / 10 ^ (n.fact*n.succ) : ℝ) <
26 |   (1/((10 : ℝ)^n.fact)^n),
27 |   proof_omitted,
28 |   tidy,
29 | end
30 |

```

□

Transcendence of α follows immediately from theorem 3.2.1 and theorem 3.2.2.

Corollary 3.2.1. α is a transcendental number.

t **theorem** transcendental_α : transcendental α :=
 ↪ liouville_numbers_transcendental α liouville_α

3.3 Hermite's proof of transcendence of e

Bibliography

- [AMK15] Jeremy Avigad, Leonardo de Moura, and Soonho Kong. *Theorem proving in Lean*. 2015.
- [AMR16] Jeremy Avigad, Leonardo de Moura de Moura, and Jared Roesch. *Programming in Lean*. 2016.
- [Bak90] Alan Baker. *Transcendental number theory*. Cambridge university press, 1990.
- [BCM20] Kevin Buzzard, Johan Commelin, and Patrick Massot. “Formalising perfectoid spaces”. In: *Proceedings of the 9th ACM SIGPLAN International Conference on Certified Programs and Proofs*. 2020, pp. 299–312.
- [Bou98] Nicolas Bourbaki. *Elements of the History of Mathematics*. Springer Science & Business Media, 1998.
- [Boy+94] Robert Boyer et al. “The QED manifesto”. In: *Automated Deduction—CADE 12* (1994), pp. 238–251.
- [Can32] G Cantor. *Über eine Eigenschaft des Inbegriffes aller reellen algebraischen Zahlen*. *Crelles J.* 77 (1874), 258-262; also in *Gesammelte A bhandlungen*. 1932.
- [Can78] Georg Cantor. “Ein beitrage zur mannigfaltigkeitslehre”. In: *Journal für die reine und angewandte Mathematik (Crelles Journal)* 1878.84 (1878), pp. 242–258.
- [Gel34] Aleksandr Gelfond. “Sur le septieme probleme de Hilbert”. In: *Известия Российской академии наук. Серия математическая* 4 (1934), pp. 623–634.
- [Gon+13] Georges Gonthier et al. “A machine-checked proof of the odd order theorem”. In: *International Conference on Interactive Theorem Proving*. Springer. 2013, pp. 163–179.
- [Gon08] Georges Gonthier. “Formal proof—the four-color theorem”. In: *Notices of the AMS* 55.11 (2008), pp. 1382–1393.
- [Hal+17] Thomas Hales et al. “A formal proof of the Kepler conjecture”. In: *Forum of mathematics, Pi*. Vol. 5. Cambridge University Press. 2017.

- [Hal98] Thomas C Hales. “The kepler conjecture”. In: *arXiv preprint math.MG/9811078* (1998).
- [HUW14] John Harrison, Josef Urban, and Freek Wiedijk. “History of Interactive Theorem Proving.” In: *Computational Logic*. Vol. 9. 2014, pp. 135–214.
- [Kem16] Aubrey J. Kempner. “On Transcendental Numbers”. In: *Transactions of the American Mathematical Society* 17.4 (1916), pp. 476–482. ISSN: 00029947. URL: <http://www.jstor.org/stable/1988833>.
- [KK11] Juliette Kennedy and Roman Kossak. *Set Theory, Arithmetic, and Foundations of Mathematics: Theorems, Philosophies*. Vol. 36. Cambridge University Press, 2011.
- [Lam04] M Lambert. “Mémoire sur quelques propriétés remarquables des quantités transcendentes circulaires et logarithmiques”. In: *Pi: A Source Book*. Springer, 2004, pp. 129–140.
- [Lan66] Serge Lang. *Introduction to transcendental numbers*. Addison-Wesley Pub. Co., 1966.
- [Zha20] Jujian Zhang. *e is a transcendental number*. 2020. URL: https://jjaassoonn.github.io/e_transcendence_doc.html.
- [Zor00] Vladimir A Zorich. *Mathematical Analysis I, volume 1. Universitext*. 2000.