universität innsbruck

Recent Advances in Formalising Modern Number Theory

Manuel Eberl
Computational Logic Group

What is Isabelle/HOL?

What is Isabelle/HOL

(Almost) everything I will talk about today is formalised in the interactive theorem prover Isabelle/HOL.

What is Isabelle/HOL

(Almost) everything I will talk about today is formalised in the interactive theorem prover Isabelle/HOL.

- Enables you to write mathematical proofs in a formal language

What is Isabelle/HOL

(Almost) everything I will talk about today is formalised in the interactive theorem prover Isabelle/HOL.

- Enables you to write mathematical proofs in a formal language
- Checks proofs for correctness

What is Isabelle/HOL

(Almost) everything I will talk about today is formalised in the interactive theorem prover Isabelle/HOL.

- Enables you to write mathematical proofs in a formal language
- Checks proofs for correctness
- Every proof gets broken down to basic logical inferences in the kernel

What is Isabelle/HOL

(Almost) everything I will talk about today is formalised in the interactive theorem prover Isabelle/HOL.

- Enables you to write mathematical proofs in a formal language
- Checks proofs for correctness
- Every proof gets broken down to basic logical inferences in the kernel
- Large library of mathematics (especially analysis)

What is Isabelle/HOL

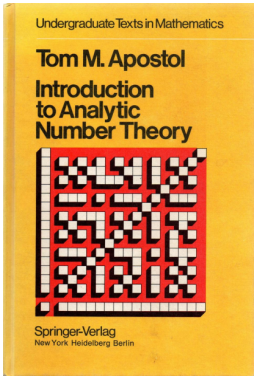
```
lemma
  fixes a b :: int
  assumes "even a ∨ even b"
  shows   "even (a * b)"
proof
  from assms show "even (a * b)"
  proof
    assume "even a"
    then obtain a' where a': "a = 2 * a'"

    have "even (2 * (a' * b))"

    also have "2 * (a' * b) = a * b"
      using a'
    finally show "even (a * b)"
  next
    assume "even b"
    (* ... *)
```

What is Isabelle/HOL

```
lemma
  fixes a b :: int
  assumes "even a ∨ even b"
  shows   "even (a * b)"
proof -
  from assms show "even (a * b)"
  proof
    assume "even a"
    then obtain a' where a': "a = 2 * a'"
      by (elim evenE)
    have "even (2 * (a' * b))"
      by simp
    also have "2 * (a' * b) = a * b"
      using a' by simp
    finally show "even (a * b)" .
  next
    assume "even b"
    (* ... *)
```

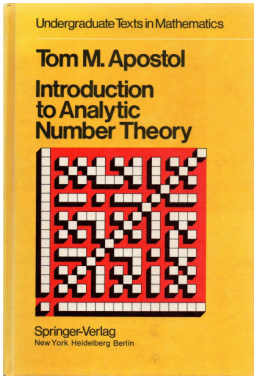
Current Work

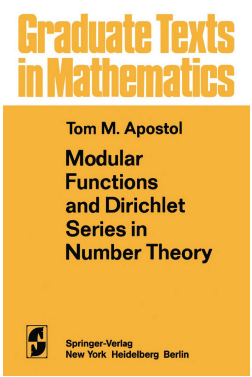
Part of my PhD: (almost) an entire undergraduate maths textbook formalised.



Current Work

Part of my PhD: (almost) an entire undergraduate maths textbook formalised.
Part of my Postdoc: formalise the graduate-level second book as well.





Complex Lattices

Complex Lattices

Given two complex numbers z_1, z_2 with $z_1/z_2 \notin \mathbb{Q}$, the *complex lattice* $\Lambda(z_1, z_2)$ is the set of all complex numbers of the form $az_1 + bz_2$ for $a, b \in \mathbb{Z}$.

Complex Lattices

Given two complex numbers z_1, z_2 with $z_1/z_2 \notin \mathbb{Q}$, the *complex lattice* $\Lambda(z_1, z_2)$ is the set of all complex numbers of the form $az_1 + bz_2$ for $a, b \in \mathbb{Z}$.

Definition (Elliptic functions)

A meromorphic function $f : \mathbb{C} \to \mathbb{C}$ is called elliptic w.r.t. a lattice Λ if it satisfies $f(z + \lambda) = f(z)$ for any $\lambda \in \Lambda$.

Complex Lattices

Given two complex numbers z_1, z_2 with $z_1/z_2 \notin \mathbb{Q}$, the *complex lattice* $\Lambda(z_1, z_2)$ is the set of all complex numbers of the form $az_1 + bz_2$ for $a, b \in \mathbb{Z}$.

Definition (Elliptic functions)

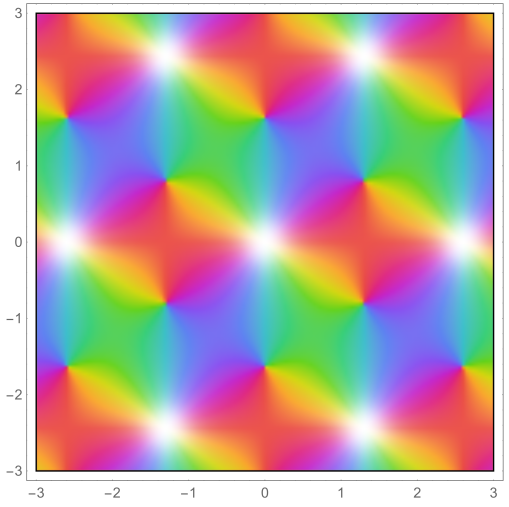
A meromorphic function $f : \mathbb{C} \to \mathbb{C}$ is called elliptic w.r.t. a lattice Λ if it satisfies $f(z + \lambda) = f(z)$ for any $\lambda \in \Lambda$.

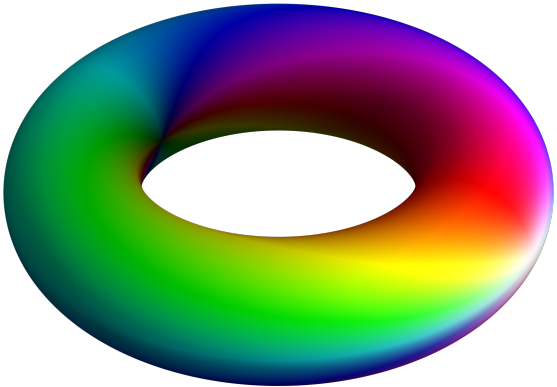
Fact

The Weierstraß function and its derivative are elliptic:

$$\wp(z) = \frac{1}{z^2} + \sum_{\lambda \in \Lambda \setminus \{0\}} \left(\frac{1}{(z-\lambda)^2} - \frac{1}{\lambda^2} \right) \qquad \wp'(z) = -2 \sum_{\lambda \in \Lambda} \frac{1}{(z-\lambda)^3}$$

All others can be expressed in terms of $\wp$ and $\wp'$.


3
2
1
0
−1
−2
−3
−3
−2
−1
0
1
2
3



Weierstraß $\wp$ and the Eisenstein Series

Fact

$\wp$ has the following series expansion at $z = 0$:

$$\wp(z) = \frac{1}{z^2} + \sum_{k=1}^{\infty} (k+1) G_{k+2} z^k \qquad \text{where } G_k = \sum_{\lambda \in \Lambda \setminus \{0\}} \frac{1}{\lambda^k} \text{ for } n \geq 3$$

Weierstraß $\wp$ and the Eisenstein Series

Fact

$\wp$ has the following series expansion at $z = 0$:

$$\wp(z) = \frac{1}{z^2} + \sum_{k=1}^{\infty} (k+1) G_{k+2} z^k \qquad \text{where } G_k = \sum_{\lambda \in \Lambda \setminus \{0\}} \frac{1}{\lambda^k} \text{ for } n \geq 3$$

G_k is called the Eisenstein series.

Weierstraß $\wp$ and the Eisenstein Series

Fact

$\wp$ has the following series expansion at $z = 0$:

$$\wp(z) = \frac{1}{z^2} + \sum_{k=1}^{\infty} (k+1) G_{k+2} z^k \qquad \text{where } G_k = \sum_{\lambda \in \Lambda \setminus \{0\}} \frac{1}{\lambda^k} \text{ for } n \geq 3$$

G_k is called the Eisenstein series. (note: $G_k = 0$ if k is odd)

Differential Equation of $\wp$

Theorem

$$\wp'(z)^2 = 4\wp(z)^3 - 60G_4\wp(z) - 140G_6$$

Differential Equation of $\wp$

Theorem

$$\wp'(z)^2 = 4\wp(z)^3 - 60G_4\wp(z) - 140G_6$$

Proof.

- Subtract both sides and compute the series expansion.

Differential Equation of $\wp$

Theorem

$$\wp'(z)^2 = 4\wp(z)^3 - 60G_4\wp(z) - 140G_6$$

Proof.

- Subtract both sides and compute the series expansion.
- The resulting function is an *entire* elliptic function (no more poles!)

Differential Equation of $\wp$

Theorem

$$\wp'(z)^2 = 4\wp(z)^3 - 60G_4\wp(z) - 140G_6$$

Proof.

- Subtract both sides and compute the series expansion.
- The resulting function is an *entire* elliptic function (no more poles!)
- *Liouville's Theorem:* A bounded entire function is constant.

Differential Equation of $\wp$

Theorem

$$\wp'(z)^2 = 4\wp(z)^3 - 60G_4\wp(z) - 140G_6$$

Proof.

- Subtract both sides and compute the series expansion.
- The resulting function is an *entire* elliptic function (no more poles!)
- *Liouville's Theorem:* A bounded entire function is constant.

□

Note: $y^2 = 4x^3 - ax - b$ is an elliptic curve.

Differential Equation of $\wp$

Theorem

$$\wp'(z)^2 = 4\wp(z)^3 - 60G_4\wp(z) - 140G_6$$

Proof.

- Subtract both sides and compute the series expansion.
- The resulting function is an *entire* elliptic function (no more poles!)
- *Liouville's Theorem:* A bounded entire function is constant.

□

Note: $y^2 = 4x^3 - ax - b$ is an elliptic curve.
$z \mapsto (\wp(z), \wp'(z))$ maps points on a torus $\mathbb{C}/\Lambda$ to points on a complex elliptic curve.

Example: Differential Equation of $\wp$

$$\wp'(z)^2 = 4\wp(z)^3 - 60G_4\wp(z) - 140G_6$$

```
theorem weierstrass_fun_ODE1':
  assumes "z ∉ Λ"
  shows   "℘' z ^ 2 = 4 * ℘ z ^ 3 - 60 * G₄ * ℘ z - 140 * G₆"
```

Example: Fourier Expansion of G_k

$$G_k(z) = 2(\zeta(k) + \frac{(2i\pi)^k}{(k-1)!} \sum_{n=0}^{\infty} \sigma_{k-1}(n) e^{2\pi i n z})$$

```
lemma Eisenstein_G_fourier_expansion:
  assumes "Im z > 0" and "k > 2" and "even k"
  shows "Eisenstein_G k z =
           2 * (zeta k + (2*i*pi)^k / fact (k-1) *
                  (∑n. divisor_sigma (k-1) n * exp (2*n*pi*i*z)))"
```

$\wp$ and the Eisenstein series

Theorem (Differential equation for $\wp$)

$$\wp'(z)^2 = 4\wp(z)^3 - 60G_4\wp(z) - 140G_6$$

$\wp$ and the Eisenstein series

Theorem (Differential equation for $\wp$)

$$\wp'(z)^2 = 4\wp(z)^3 - 60G_4\wp(z) - 140G_6$$

Expanding the series again and comparing coefficients, we get:

Theorem

$$(2n+3)(n-2)b_n = 3\sum_{k=1}^{n-2} b(k)b(n-k-1)$$

where $b_n = (2n+1)G_{2n+2}$ *for any* $n \geq 3$.

$\wp$ and the Eisenstein series

Theorem (Differential equation for $\wp$)

$$\wp'(z)^2 = 4\wp(z)^3 - 60G_4\wp(z) - 140G_6$$

Expanding the series again and comparing coefficients, we get:

Theorem

$$(2n+3)(n-2)b_n = 3\sum_{k=1}^{n-2} b(k)b(n-k-1)$$

where $b_n = (2n+1)G_{2n+2}$ *for any* $n \geq 3$.

This means: all the G_k really only depend on G_4 and G_6.

$\wp$ and the Eisenstein series

Theorem (Differential equation for $\wp$)

$$\wp'(z)^2 = 4\wp(z)^3 - 60G_4\wp(z) - 140G_6$$

Expanding the series again and comparing coefficients, we get:

Theorem

$$(2n+3)(n-2)b_n = 3\sum_{k=1}^{n-2} b(k)b(n-k-1)$$

where $b_n = (2n+1)G_{2n+2}$ *for any* $n \geq 3$.

This means: all the G_k really only depend on G_4 and G_6.

$\wp$ and the Eisenstein series

Theorem (Differential equation for $\wp$)

$$\wp'(z)^2 = 4\wp(z)^3 - 60G_4\wp(z) - 140G_6$$

Expanding the series again and comparing coefficients, we get:

Theorem

$$(2n+3)(n-2)b_n = 3\sum_{k=1}^{n-2} b(k)b(n-k-1)$$

where $b_n = (2n+1)G_{2n+2}$ *for any* $n \geq 3$.

This means: all the G_k really only depend on G_4 and G_6.
G_4 and G_6 are the important ones.

Example:

```
theorem eisenstein_series_as_polynomials:
  defines "b ≡ λn. (2*n + 1) * (G (2*n + 2))"
  assumes "n ≥ 3"
  shows   "(2 * n + 3) * (n - 2) * b n = 3 * (∑i=1..n-2. b i * b (n - i - 1))"
```

G_k as a function of Λ

Observations:

- Rotation/scaling in $\mathbb{C}$ corresponds to multiplication with a constant $c \in \mathbb{C}$

G_k as a function of Λ

Observations:

- Rotation/scaling in $\mathbb{C}$ corresponds to multiplication with a constant $c \in \mathbb{C}$
- G_k is homogenous of degree $-k$; that is: $G_k(c\Lambda) = c^{-k} G_k(\Lambda)$

G_k as a function of Λ

Observations:

- Rotation/scaling in $\mathbb{C}$ corresponds to multiplication with a constant $c \in \mathbb{C}$
- G_k is homogenous of degree $-k$; that is: $G_k(c\Lambda) = c^{-k} G_k(\Lambda)$
- So if two lattices are the same modulo rotation/scaling, their G_k are not the same, but at least related.

G_k as a function of Λ

Observations:

- Rotation/scaling in $\mathbb{C}$ corresponds to multiplication with a constant $c \in \mathbb{C}$
- G_k is homogenous of degree $-k$; that is: $G_k(c\Lambda) = c^{-k} G_k(\Lambda)$
- So if two lattices are the same modulo rotation/scaling, their G_k are not the same, but at least related.
- It can be shown that $\Delta = (60G_4)^3 - (420G_6)^2 \neq 0$.
  This is called the *modular discriminant*.

G_k as a function of Λ

Observations:

- Rotation/scaling in $\mathbb{C}$ corresponds to multiplication with a constant $c \in \mathbb{C}$
- G_k is homogenous of degree $-k$; that is: $G_k(c\Lambda) = c^{-k} G_k(\Lambda)$
- So if two lattices are the same modulo rotation/scaling, their G_k are not the same, but at least related.
- It can be shown that $\Delta = (60G_4)^3 - (420G_6)^2 \neq 0$.
  This is called the *modular discriminant*.
- Δ is homogenous of degree -12: $\Delta(c\Lambda) = c^{-12}\Delta(\Lambda)$

G_k as a function of Λ

Observations:

- Rotation/scaling in $\mathbb{C}$ corresponds to multiplication with a constant $c \in \mathbb{C}$
- G_k is homogenous of degree $-k$; that is: $G_k(c\Lambda) = c^{-k} G_k(\Lambda)$
- So if two lattices are the same modulo rotation/scaling, their G_k are not the same, but at least related.
- It can be shown that $\Delta = (60G_4)^3 - (420G_6)^2 \neq 0$.
  This is called the *modular discriminant*.
- Δ is homogenous of degree -12: $\Delta(c\Lambda) = c^{-12}\Delta(\Lambda)$

Fact

The Klein J invariant $J = (60G_4)^3/\Delta$ *satisfies* $J(c\Lambda) = \Lambda$.

G_k as a function of Λ

Observations:

- Rotation/scaling in $\mathbb{C}$ corresponds to multiplication with a constant $c \in \mathbb{C}$
- G_k is homogenous of degree $-k$; that is: $G_k(c\Lambda) = c^{-k} G_k(\Lambda)$
- So if two lattices are the same modulo rotation/scaling, their G_k are not the same, but at least related.
- It can be shown that $\Delta = (60G_4)^3 - (420G_6)^2 \neq 0$.
  This is called the *modular discriminant*.
- Δ is homogenous of degree -12: $\Delta(c\Lambda) = c^{-12}\Delta(\Lambda)$

Fact

The Klein J invariant $J = (60G_4)^3/\Delta$ satisfies $J(c\Lambda) = \Lambda$.
It can be shown that $J(\Lambda_1) = J(\Lambda_2)$ iff Λ_1 and Λ_2 are homothetic.

Modular Forms

Functions like G_k, Δ, J that are smooth maps $\Lambda(\mathbb{C}) \to \mathbb{C}$ and satisfy an equation of the form $f(c\Lambda) = c^k f(\Lambda)$ for some weight $k \in \mathbb{Z}$ are called *modular forms*.

Modular Forms

Functions like G_k, Δ, J that are smooth maps $\Lambda(\mathbb{C}) \to \mathbb{C}$ and satisfy an equation of the form $f(c\Lambda) = c^k f(\Lambda)$ for some weight $k \in \mathbb{Z}$ are called *modular forms*.

Since we don't care about scaling/rotation, we can wlog look only at lattices of the form $\Lambda(1, \tau)$ with $\mathrm{Im}(\tau) > 0$.

Modular Forms

Functions like G_k, Δ, J that are smooth maps $\Lambda(\mathbb{C}) \to \mathbb{C}$ and satisfy an equation of the form $f(c\Lambda) = c^k f(\Lambda)$ for some weight $k \in \mathbb{Z}$ are called *modular forms*.

Since we don't care about scaling/rotation, we can wlog look only at lattices of the form $\Lambda(1, \tau)$ with $\mathrm{Im}(\tau) > 0$.

Fact: τ, τ' define the same lattice iff $\tau' = \frac{a\tau+b}{c\tau+d}$ for some $a, b, c, d \in \mathbb{Z}$ with $ad - bc = 1$.

Modular Forms

Functions like G_k, Δ, J that are smooth maps $\Lambda(\mathbb{C}) \to \mathbb{C}$ and satisfy an equation of the form $f(c\Lambda) = c^k f(\Lambda)$ for some weight $k \in \mathbb{Z}$ are called *modular forms*.

Since we don't care about scaling/rotation, we can wlog look only at lattices of the form $\Lambda(1, \tau)$ with $\mathrm{Im}(\tau) > 0$.

Fact: τ, τ' define the same lattice iff $\tau' = \frac{a\tau+b}{c\tau+d}$ for some $a, b, c, d \in \mathbb{Z}$ with $ad - bc = 1$.

Thus we can view f as a map $\mathbb{H} \to \mathbb{C}$ that

Modular Forms

Functions like G_k, Δ, J that are smooth maps $\Lambda(\mathbb{C}) \to \mathbb{C}$ and satisfy an equation of the form $f(c\Lambda) = c^k f(\Lambda)$ for some weight $k \in \mathbb{Z}$ are called *modular forms*.

Since we don't care about scaling/rotation, we can wlog look only at lattices of the form $\Lambda(1, \tau)$ with $\mathrm{Im}(\tau) > 0$.

Fact: τ, τ' define the same lattice iff $\tau' = \frac{a\tau+b}{c\tau+d}$ for some $a, b, c, d \in \mathbb{Z}$ with $ad - bc = 1$.

Thus we can view f as a map $\mathbb{H} \to \mathbb{C}$ that

- satisfies $f(\frac{az+b}{cz+d}) = (cz+d)^k f(z)$ for $a, b, c, d \in \mathbb{Z}$ with $ad - bc = 1$

Modular Forms

Functions like G_k, Δ, J that are smooth maps $\Lambda(\mathbb{C}) \to \mathbb{C}$ and satisfy an equation of the form $f(c\Lambda) = c^k f(\Lambda)$ for some weight $k \in \mathbb{Z}$ are called *modular forms*.

Since we don't care about scaling/rotation, we can wlog look only at lattices of the form $\Lambda(1, \tau)$ with $\mathrm{Im}(\tau) > 0$.

Fact: τ, τ' define the same lattice iff $\tau' = \frac{a\tau+b}{c\tau+d}$ for some $a, b, c, d \in \mathbb{Z}$ with $ad - bc = 1$.

Thus we can view f as a map $\mathbb{H} \to \mathbb{C}$ that

- satisfies $f(\frac{az+b}{cz+d}) = (cz+d)^k f(z)$ for $a, b, c, d \in \mathbb{Z}$ with $ad - bc = 1$
- is meromorphic ("nice" everywhere except for perhaps some poles)

Modular Forms

Functions like G_k, Δ, J that are smooth maps $\Lambda(\mathbb{C}) \to \mathbb{C}$ and satisfy an equation of the form $f(c\Lambda) = c^k f(\Lambda)$ for some weight $k \in \mathbb{Z}$ are called *modular forms*.

Since we don't care about scaling/rotation, we can wlog look only at lattices of the form $\Lambda(1, \tau)$ with $\mathrm{Im}(\tau) > 0$.

Fact: τ, τ' define the same lattice iff $\tau' = \frac{a\tau+b}{c\tau+d}$ for some $a, b, c, d \in \mathbb{Z}$ with $ad - bc = 1$.

Thus we can view f as a map $\mathbb{H} \to \mathbb{C}$ that

- satisfies $f(\frac{az+b}{cz+d}) = (cz+d)^k f(z)$ for $a, b, c, d \in \mathbb{Z}$ with $ad - bc = 1$
- is meromorphic ("nice" everywhere except for perhaps some poles)
- is meromorphic at the cusp $z = i\infty$

Modular Forms

Functions like G_k, Δ, J that are smooth maps $\Lambda(\mathbb{C}) \to \mathbb{C}$ and satisfy an equation of the form $f(c\Lambda) = c^k f(\Lambda)$ for some weight $k \in \mathbb{Z}$ are called *modular forms*.

Since we don't care about scaling/rotation, we can wlog look only at lattices of the form $\Lambda(1, \tau)$ with $\mathrm{Im}(\tau) > 0$.

Fact: τ, τ' define the same lattice iff $\tau' = \frac{a\tau+b}{c\tau+d}$ for some $a, b, c, d \in \mathbb{Z}$ with $ad - bc = 1$.

Thus we can view f as a map $\mathbb{H} \to \mathbb{C}$ that

- satisfies $f(\frac{az+b}{cz+d}) = (cz+d)^k f(z)$ for $a, b, c, d \in \mathbb{Z}$ with $ad - bc = 1$
- is meromorphic ("nice" everywhere except for perhaps some poles)
- is meromorphic at the cusp $z = i\infty$

Bottom line: invariant under modular transformations $z \mapsto z + 1$ and $z \mapsto -\frac{1}{z}$ and doesn't do anything "crazy" in $\mathbb{H} \cup \{i\infty\}$.

Modular Transformations and Fundamental Domain

The transformations $z \mapsto \frac{az+b}{cz+d}$ for $a, b, c, d \in \mathbb{Z}$ with $ad - bc = 1$ with function composition form a group – the so-called modular group.

Modular Transformations and Fundamental Domain

The transformations $z \mapsto \frac{az+b}{cz+d}$ for $a, b, c, d \in \mathbb{Z}$ with $ad - bc = 1$ with function composition form a group – the so-called modular group.

Any points in $\mathbb{H}$ that get mapped to each other by some modular transformation can be considered equivalent.

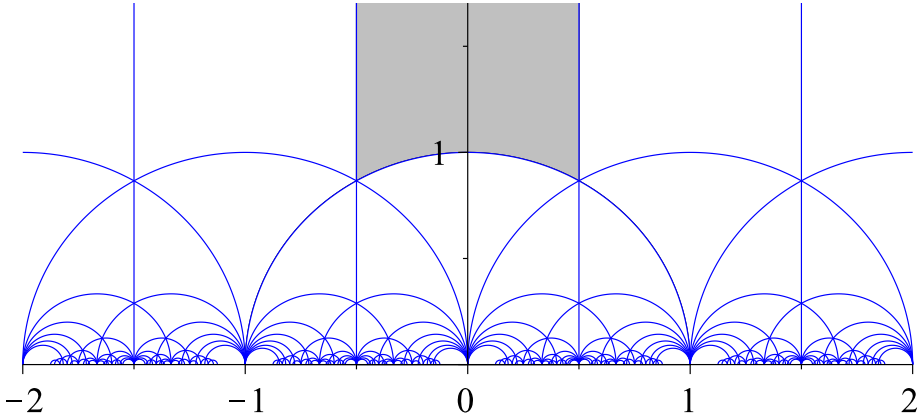
Modular Transformations and Fundamental Domain

The transformations $z \mapsto \frac{az+b}{cz+d}$ for $a, b, c, d \in \mathbb{Z}$ with $ad - bc = 1$ with function composition form a group – the so-called modular group.

Any points in $\mathbb{H}$ that get mapped to each other by some modular transformation can be considered equivalent.

Each point in $\mathbb{H}$ has a canonical representative in the fundamental domain $|z| \geq 1$, $|\mathrm{Re}(z)| \leq \frac{1}{2}$.

The Fundamental Domain



Source: https://commons.wikimedia.org/wiki/File:ModularGroup-FundamentalDomain.svg

Valence Formula for Modular Forms

Theorem

If f is a modular form of weight k, the number of zeros minus the number of poles of f in the fundamental domain is equal to $\frac{k}{12}$.

Valence Formula for Modular Forms

Theorem

If f is a modular form of weight k, the number of zeros minus the number of poles of f in the fundamental domain is equal to $\frac{k}{12}$.

Caveat:

- Multiplicity has to be taken into account – a double zero counts twice!

Valence Formula for Modular Forms

Theorem

If f is a modular form of weight k, the number of zeros minus the number of poles of f in the fundamental domain is equal to $\frac{k}{12}$.

Caveat:

- Multiplicity has to be taken into account – a double zero counts twice!
- Points on the border of the fundamental region must only be counted once

Valence Formula for Modular Forms

Theorem

If f is a modular form of weight k, the number of zeros minus the number of poles of f in the fundamental domain is equal to $\frac{k}{12}$.

Caveat:

- Multiplicity has to be taken into account – a double zero counts twice!
- Points on the border of the fundamental region must only be counted once
- Any zero/pole at $z = i\infty$ also has to be counted

Valence Formula for Modular Forms

Theorem

If f is a modular form of weight k, the number of zeros minus the number of poles of f in the fundamental domain is equal to $\frac{k}{12}$.

Caveat:

- Multiplicity has to be taken into account – a double zero counts twice!
- Points on the border of the fundamental region must only be counted once
- Any zero/pole at $z = i\infty$ also has to be counted
- Any zero/pole at $z = i$ and $z = \rho$ has to be weighted with $\frac{1}{2}$ resp. $\frac{1}{3}$

Valence Formula for Modular Forms

Theorem

If f is a modular form of weight k, the number of zeros minus the number of poles of f in the fundamental domain is equal to $\frac{k}{12}$.

Caveat:

- Multiplicity has to be taken into account – a double zero counts twice!
- Points on the border of the fundamental region must only be counted once
- Any zero/pole at $z = i\infty$ also has to be counted
- Any zero/pole at $z = i$ and $z = \rho$ has to be weighted with $\frac{1}{2}$ resp. $\frac{1}{3}$

Consequence: entire modular forms are linear combinations of generators $G_{k-12r}\Delta^r$

Example: Valence Formula

```
theorem valence_formula:
   "(∑z∈𝒟Γ'∩(poles∪zeros)-{i,ϱ}. zorder f z) +
       zorder f ϱ / 3 + zorder f i / 2 + zorder_at_cusp = k / 12"
```

Consequences for the Klein J Invariant

For J, the valence formula also has some interesting consequences:

Consequences for the Klein J Invariant

For J, the valence formula also has some interesting consequences:

- We already know that J has no poles in $\mathbb{H}$.

Consequences for the Klein J Invariant

For J, the valence formula also has some interesting consequences:

- We already know that J has no poles in $\mathbb{H}$.
- Fourier expansion at $z = i\infty$ shows that it has a simple pole at $z = i\infty$.

Consequences for the Klein J Invariant

For J, the valence formula also has some interesting consequences:

- We already know that J has no poles in $\mathbb{H}$.
- Fourier expansion at $z = i\infty$ shows that it has a simple pole at $z = i\infty$.
- Thus, J must have precisely one zero inside the fundamental domain.

Consequences for the Klein J Invariant

For J, the valence formula also has some interesting consequences:

- We already know that J has no poles in $\mathbb{H}$.
- Fourier expansion at $z = i\infty$ shows that it has a simple pole at $z = i\infty$.
- Thus, J must have precisely one zero inside the fundamental domain.
- A simple computation shows that $J(\rho) = 0$.

Consequences for the Klein J Invariant

For J, the valence formula also has some interesting consequences:

- We already know that J has no poles in $\mathbb{H}$.
- Fourier expansion at $z = i\infty$ shows that it has a simple pole at $z = i\infty$.
- Thus, J must have precisely one zero inside the fundamental domain.
- A simple computation shows that $J(\rho) = 0$.
- Since this zero's multiplicity only counts half, it must be a *double zero*.

Consequences for the Klein J Invariant

For J, the valence formula also has some interesting consequences:

- We already know that J has no poles in $\mathbb{H}$.
- Fourier expansion at $z = i\infty$ shows that it has a simple pole at $z = i\infty$.
- Thus, J must have precisely one zero inside the fundamental domain.
- A simple computation shows that $J(\rho) = 0$.
- Since this zero's multiplicity only counts half, it must be a *double zero*.
- Applying the same reasoning to $z \mapsto J(z) + c$, we obtain that J takes on every value $c \in \mathbb{C}$ *exactly once* in the fundamental domain.

Consequences for the Klein J Invariant

For J, the valence formula also has some interesting consequences:

- We already know that J has no poles in $\mathbb{H}$.
- Fourier expansion at $z = i\infty$ shows that it has a simple pole at $z = i\infty$.
- Thus, J must have precisely one zero inside the fundamental domain.
- A simple computation shows that $J(\rho) = 0$.
- Since this zero's multiplicity only counts half, it must be a *double zero*.
- Applying the same reasoning to $z \mapsto J(z) + c$, we obtain that J takes on every value $c \in \mathbb{C}$ *exactly once* in the fundamental domain.
- J is a smooth bijection between the fundamental domain and the whole complex plane

Consequences for the Klein J Invariant

For J, the valence formula also has some interesting consequences:

- We already know that J has no poles in $\mathbb{H}$.
- Fourier expansion at $z = i\infty$ shows that it has a simple pole at $z = i\infty$.
- Thus, J must have precisely one zero inside the fundamental domain.
- A simple computation shows that $J(\rho) = 0$.
- Since this zero's multiplicity only counts half, it must be a *double zero*.
- Applying the same reasoning to $z \mapsto J(z) + c$, we obtain that J takes on every value $c \in \mathbb{C}$ *exactly once* in the fundamental domain.
- J is a smooth bijection between the fundamental domain and the whole complex plane
- It can even be shown that J is locally injective everywhere except $z = \rho$ and $z = i$ (thus $J'(z) = 0 \leftrightarrow z \in \{\rho, i\}$)

Some Examples

Any modular function is a rational function of Klein's *J* invariant:

```
theorem (in modular_function') in_terms_of_Klein_j:
  obtains p q :: "complex poly"
  where "q ≠ 0" "⋀z. Im z > 0 ⟹ z ∉ poles ⟹
                        f z = poly p (Klein_j z) / poly q (Klein_j z)"
```

Little Picard

Theorem

A non-constant differentiable function $f : \mathbb{C} \to \mathbb{C}$ attains every complex value at least once, with at most one exception.

Little Picard

Theorem

A non-constant differentiable function $f : \mathbb{C} \to \mathbb{C}$ attains every complex value at least once, with at most one exception.

Proof.

Assume wlog that $0, 1 \notin f(\mathbb{C})$. Then there exists a well-defined smooth function $h : \mathbb{C} \to \mathbb{H}, h(z) = J^{-1}(f(z))$.

Little Picard

Theorem

A non-constant differentiable function $f : \mathbb{C} \to \mathbb{C}$ attains every complex value at least once, with at most one exception.

Proof.

Assume wlog that $0, 1 \notin f(\mathbb{C})$. Then there exists a well-defined smooth function $h : \mathbb{C} \to \mathbb{H}, h(z) = J^{-1}(f(z))$.

But because $z \in \mathbb{H}$ we have $\mathrm{Im}(z) < 0$, so $|e^{ih(z)}| = e^{-\mathrm{Im}(h(z))} < 1$. But then $e^{ih(z)}$ is bounded and entire

Little Picard

Theorem

A non-constant differentiable function $f : \mathbb{C} \to \mathbb{C}$ attains every complex value at least once, with at most one exception.

Proof.

Assume wlog that $0, 1 \notin f(\mathbb{C})$. Then there exists a well-defined smooth function $h : \mathbb{C} \to \mathbb{H}, h(z) = J^{-1}(f(z))$.

But because $z \in \mathbb{H}$ we have $\mathrm{Im}(z) < 0$, so $|e^{ih(z)}| = e^{-\mathrm{Im}(h(z))} < 1$. But then $e^{ih(z)}$ is bounded and entire and thus constant by Liouville's theorem,

Little Picard

Theorem

A non-constant differentiable function $f : \mathbb{C} \to \mathbb{C}$ attains every complex value at least once, with at most one exception.

Proof.

Assume wlog that $0, 1 \notin f(\mathbb{C})$. Then there exists a well-defined smooth function $h : \mathbb{C} \to \mathbb{H}, h(z) = J^{-1}(f(z))$.

But because $z \in \mathbb{H}$ we have $\mathrm{Im}(z) < 0$, so $|e^{ih(z)}| = e^{-\mathrm{Im}(h(z))} < 1$. But then $e^{ih(z)}$ is bounded and entire and thus constant by Liouville's theorem, so $h(z)$ and thus also $f(z)$ must also be constant.

Little Picard

Theorem

A non-constant differentiable function $f : \mathbb{C} \to \mathbb{C}$ attains every complex value at least once, with at most one exception.

Proof.

Assume wlog that $0, 1 \notin f(\mathbb{C})$. Then there exists a well-defined smooth function $h : \mathbb{C} \to \mathbb{H}, h(z) = J^{-1}(f(z))$.

But because $z \in \mathbb{H}$ we have $\mathrm{Im}(z) < 0$, so $|e^{ih(z)}| = e^{-\mathrm{Im}(h(z))} < 1$. But then $e^{ih(z)}$ is bounded and entire and thus constant by Liouville's theorem, so $h(z)$ and thus also $f(z)$ must also be constant. □

Little Picard

Theorem

A non-constant differentiable function $f : \mathbb{C} \to \mathbb{C}$ attains every complex value at least once, with at most one exception.

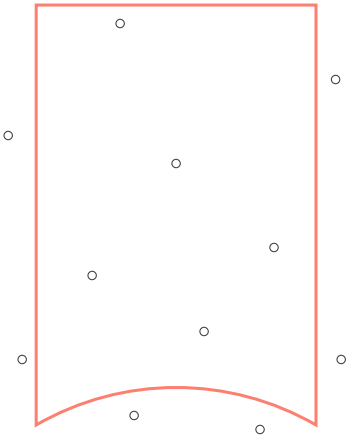
Proof.

Assume wlog that $0, 1 \notin f(\mathbb{C})$. Then there exists a well-defined smooth function $h : \mathbb{C} \to \mathbb{H}, h(z) = J^{-1}(f(z))$.

But because $z \in \mathbb{H}$ we have $\mathrm{Im}(z) < 0$, so $|e^{ih(z)}| = e^{-\mathrm{Im}(h(z))} < 1$. But then $e^{ih(z)}$ is bounded and entire and thus constant by Liouville's theorem, so $h(z)$ and thus also $f(z)$ must also be constant. □

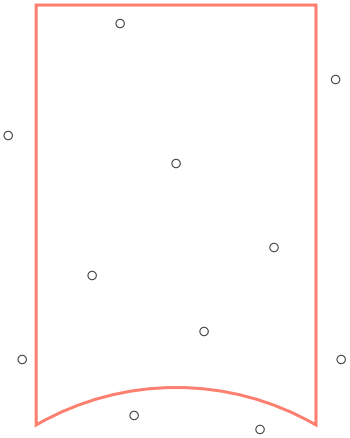
The proof in the book is not much longer than this sketch here, but full of subtleties that make formalisation very tricky.

Proving the Valence Formula



We want to count the zeros/poles inside the fundamental region.

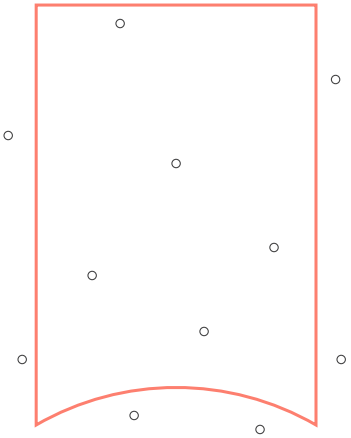
Proving the Valence Formula



We want to count the zeros/poles inside the fundamental region.

Argument Principle: $\#\text{zeros} - \#\text{poles} = \int_\gamma \frac{f'(z)}{f(z)}\,dz.$

Proving the Valence Formula

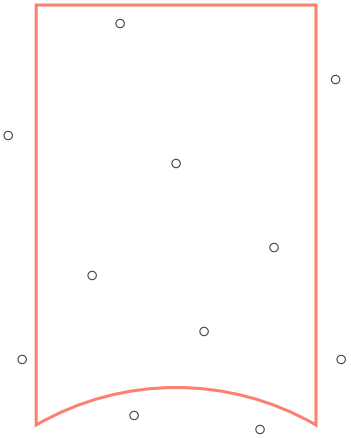


We want to count the zeros/poles inside the fundamental region.

Argument Principle: $\#\text{zeros} - \#\text{poles} = \int_\gamma \frac{f'(z)}{f(z)}\,dz.$

- The two vertical lines cancel because one is mapped to the other by $z \mapsto z+1$

Proving the Valence Formula

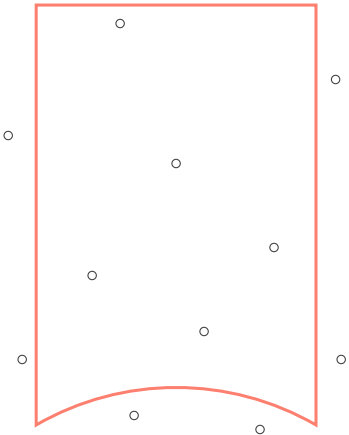


We want to count the zeros/poles inside the fundamental region.

Argument Principle: $\#\text{zeros} - \#\text{poles} = \int_\gamma \frac{f'(z)}{f(z)}\,dz.$

- The two vertical lines cancel because one is mapped to the other by $z \mapsto z+1$
- The two halves of the arc cancel up to a factor $\frac{k}{12}$ because one is mapped to the other by $z \mapsto -\frac{1}{z}$

Proving the Valence Formula

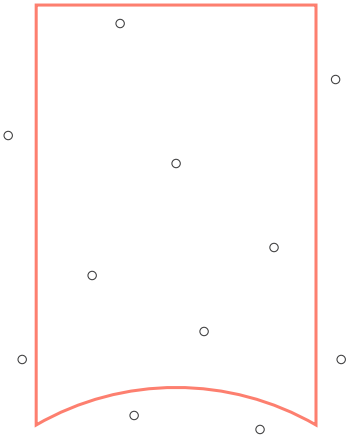


We want to count the zeros/poles inside the fundamental region.

Argument Principle: $\#\text{zeros} - \#\text{poles} = \int_\gamma \frac{f'(z)}{f(z)}\,dz.$

- The two vertical lines cancel because one is mapped to the other by $z \mapsto z+1$
- The two halves of the arc cancel up to a factor $\frac{k}{12}$ because one is mapped to the other by $z \mapsto -\frac{1}{z}$
- The contribution of the horizontal line vanishes

Proving the Valence Formula



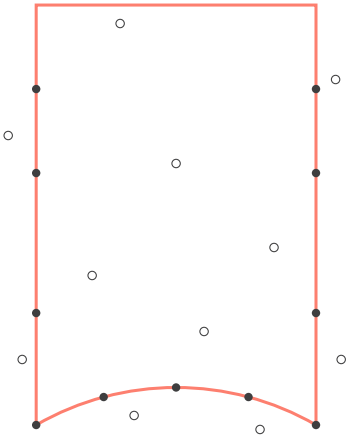
We want to count the zeros/poles inside the fundamental region.

Argument Principle: $\#\text{zeros} - \#\text{poles} = \int_\gamma \frac{f'(z)}{f(z)}\,dz.$

- The two vertical lines cancel because one is mapped to the other by $z \mapsto z+1$
- The two halves of the arc cancel up to a factor $\frac{k}{12}$ because one is mapped to the other by $z \mapsto -\frac{1}{z}$
- The contribution of the horizontal line vanishes

Problem: This integral is not defined if there are zeros/poles directly on γ!

Proving the Valence Formula



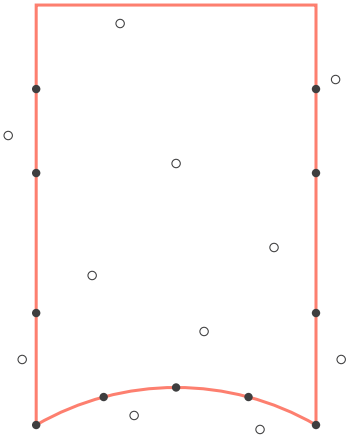
We want to count the zeros/poles inside the fundamental region.

Argument Principle: $\#\text{zeros} - \#\text{poles} = \int_\gamma \frac{f'(z)}{f(z)}\,dz.$

- The two vertical lines cancel because one is mapped to the other by $z \mapsto z+1$
- The two halves of the arc cancel up to a factor $\frac{k}{12}$ because one is mapped to the other by $z \mapsto -\frac{1}{z}$
- The contribution of the horizontal line vanishes

Problem: This integral is not defined if there are zeros/poles directly on γ!

Proving the Valence Formula



We want to count the zeros/poles inside the fundamental region.

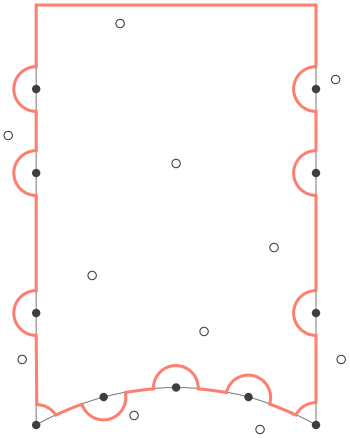
Argument Principle: $\#\text{zeros} - \#\text{poles} = \int_\gamma \frac{f'(z)}{f(z)}\,dz.$

- The two vertical lines cancel because one is mapped to the other by $z \mapsto z+1$
- The two halves of the arc cancel up to a factor $\frac{k}{12}$ because one is mapped to the other by $z \mapsto -\frac{1}{z}$
- The contribution of the horizontal line vanishes

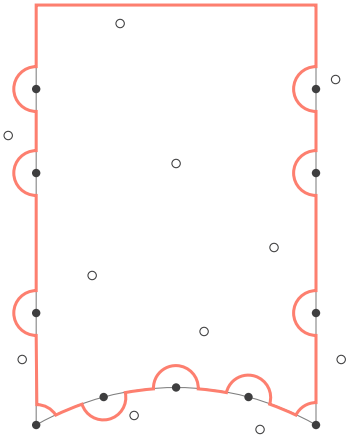
Problem: This integral is not defined if there are zeros/poles directly on γ!

Solution: Deform γ with small arcs of radius $\varepsilon \to 0$.

Proving the Valence Formula

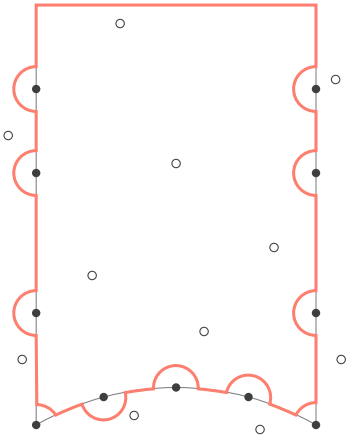


Proving the Valence Formula



Unfortunately, this countour is now very complicated.

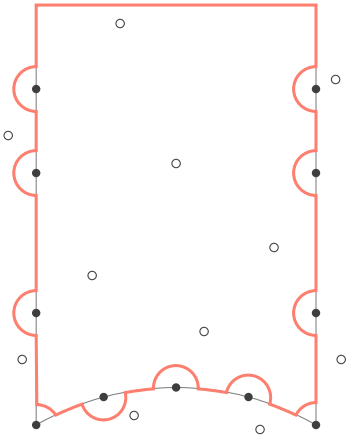
Proving the Valence Formula



Unfortunately, this countour is now very complicated.

To apply the Argument Principle, we need to prove lots of “obvious” facts about it, such as:

Proving the Valence Formula

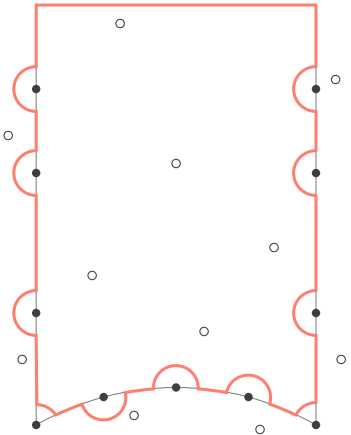


Unfortunately, this countour is now very complicated.

To apply the Argument Principle, we need to prove lots of “obvious” facts about it, such as:

- All the points that are supposed to be inside are inside

Proving the Valence Formula

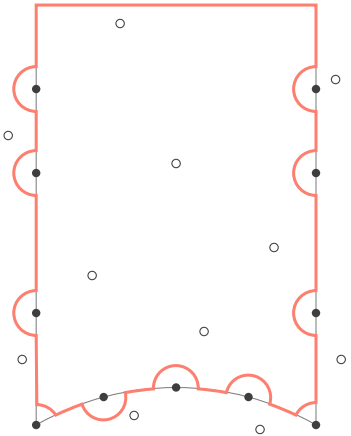


Unfortunately, this countour is now very complicated.

To apply the Argument Principle, we need to prove lots of “obvious” facts about it, such as:

- All the points that are supposed to be inside are inside
- The curve winds around the points inside of it once

Proving the Valence Formula

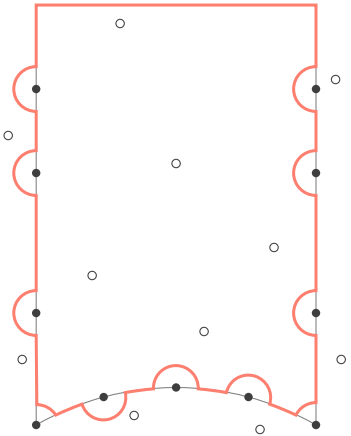


Unfortunately, this countour is now very complicated.

To apply the Argument Principle, we need to prove lots of “obvious” facts about it, such as:

- All the points that are supposed to be inside are inside
- The curve winds around the points inside of it once
- For small enough ε, there are no more bad points on the curve

Proving the Valence Formula



Unfortunately, this countour is now very complicated.

To apply the Argument Principle, we need to prove lots of “obvious” facts about it, such as:

- All the points that are supposed to be inside are inside
- The curve winds around the points inside of it once
- For small enough ε, there are no more bad points on the curve

This is already tedious for “easy” curves and completely unfeasible for something like this.

The Wiggle Framework

Solution: a relation $\gamma \underset{L/R}{\approx} \gamma_\varepsilon$ where:

The Wiggle Framework

Solution: a relation $\gamma \underset{L/R}{\approx} \gamma_\varepsilon$ where:

- γ is the original curve.

The Wiggle Framework

Solution: a relation $\gamma \underset{L/R}{\approx} \gamma_\varepsilon$ where:

- γ is the original curve.
- γ_ε is a family of deformed curves

The Wiggle Framework

Solution: a relation $\gamma \underset{L/R}{\approx} \gamma_\varepsilon$ where:

- γ is the original curve.
- γ_ε is a family of deformed curves
- L and R are points to avoid.

The Wiggle Framework

Solution: a relation $\gamma \underset{L/R}{\approx} \gamma_\varepsilon$ where:

- γ is the original curve.
- γ_ε is a family of deformed curves
- L and R are points to avoid.
- γ_ε avoids all points in L by "swerving left"  within some ε-neighbourhood

The Wiggle Framework

Solution: a relation $\gamma \underset{L/R}{\approx} \gamma_\varepsilon$ where:

- γ is the original curve.
- γ_ε is a family of deformed curves
- L and R are points to avoid.
- γ_ε avoids all points in L by "swerving left" within some ε-neighbourhood
- and analogously for R 

The Wiggle Framework

Solution: a relation $\gamma \underset{L/R}{\approx} \gamma_\varepsilon$ where:

- γ is the original curve.
- γ_ε is a family of deformed curves
- L and R are points to avoid.
- γ_ε avoids all points in L by "swerving left"  within some ε-neighbourhood
- and analogously for R

If we know that $\gamma \underset{L/R}{\approx} \gamma_\varepsilon$, we know everything we need to know to rescue our argument from before.

The Wiggle Framework

Solution: a relation $\gamma \underset{L/R}{\approx} \gamma_\varepsilon$ where:

- γ is the original curve.
- γ_ε is a family of deformed curves
- L and R are points to avoid.
- γ_ε avoids all points in L by "swerving left"  within some ε-neighbourhood
- and analogously for R

If we know that $\gamma \underset{L/R}{\approx} \gamma_\varepsilon$, we know everything we need to know to rescue our argument from before.

But how to show $\gamma \underset{L/R}{\approx} \gamma_\varepsilon$?

Compositionality

$$\frac{}{r \underset{\emptyset/\emptyset}{\approx} r}\ \textbf{REFL}$$

Compositionality

$$\frac{}{\gamma \underset{\emptyset/\emptyset}{\approx} \gamma}\ \textbf{REFL} \qquad \frac{\gamma \underset{L/R}{\approx} \gamma_\varepsilon}{\gamma^{\mathsf{rev}} \underset{R/L}{\approx} \gamma_\varepsilon^{\mathsf{rev}}}\ \textbf{REVERSE}$$

Compositionality

$$\frac{}{\gamma \underset{\emptyset/\emptyset}{\approx} \gamma}\ \textsc{refl} \qquad \frac{\gamma \underset{L/R}{\approx} \gamma_\varepsilon}{\gamma^{\mathsf{rev}} \underset{R/L}{\approx} \gamma_\varepsilon^{\mathsf{rev}}}\ \textsc{reverse}$$

$$\frac{\mathsf{end}(\gamma) = \mathsf{start}(\eta) \qquad L_1 \cap L_2 = \emptyset \qquad R_1 \cap R_2 = \emptyset \qquad \gamma \underset{L_1/R_1}{\approx} \gamma_\varepsilon \qquad \eta \underset{L_2/R_2}{\approx} \eta_\varepsilon}{\gamma \multimap \eta \underset{L_1 \cup L_2 / R_1 \cup R_2}{\approx} \gamma_\varepsilon \multimap \eta_\varepsilon}\ \textsc{join}$$

Basic Avoidance Patterns

Straight line: 

Basic Avoidance Patterns

Straight line: 

Circular arc: 

Basic Avoidance Patterns

Straight line: 

Circular arc: 

Straight line corner: 

Basic Avoidance Patterns

Straight line:

Circular arc:

Straight line corner:

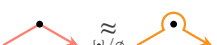
Line/arc corner:

Basic Avoidance Patterns

Straight line:

Circular arc:

Straight line corner:



Line/arc corner:

These rules have only trivial side conditions and are thus very easy to apply.

Conclusion

Of course, everything is more complicated than what I presented here.

Conclusion

Of course, everything is more complicated than what I presented here.

But:

- We can do very advanced graduate-level maths in a theorem prover

Conclusion

Of course, everything is more complicated than what I presented here.

But:

- We can do very advanced graduate-level maths in a theorem prover
- It doesn't require a huge team (most of this I did alone in $\approx$ 4 months' time)

Conclusion

Of course, everything is more complicated than what I presented here.

But:

- We can do very advanced graduate-level maths in a theorem prover
- It doesn't require a huge team (most of this I did alone in $\approx$ 4 months' time)
- Formalisation unearthes many problems and subtleties present in paper proofs.

Conclusion

Of course, everything is more complicated than what I presented here.

But:

- We can do very advanced graduate-level maths in a theorem prover
- It doesn't require a huge team (most of this I did alone in $\approx$ 4 months' time)
- Formalisation unearthes many problems and subtleties present in paper proofs.
- But it also clarifies many aspects.