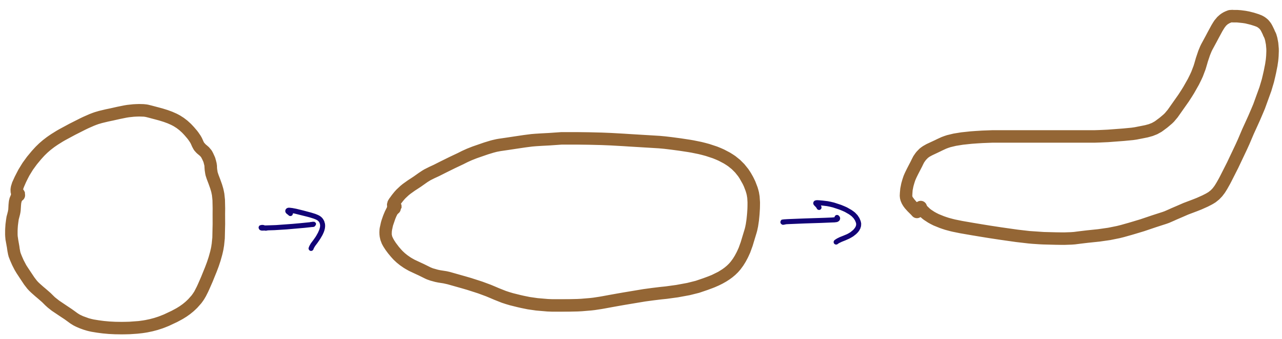
Knots and minimal surfaces

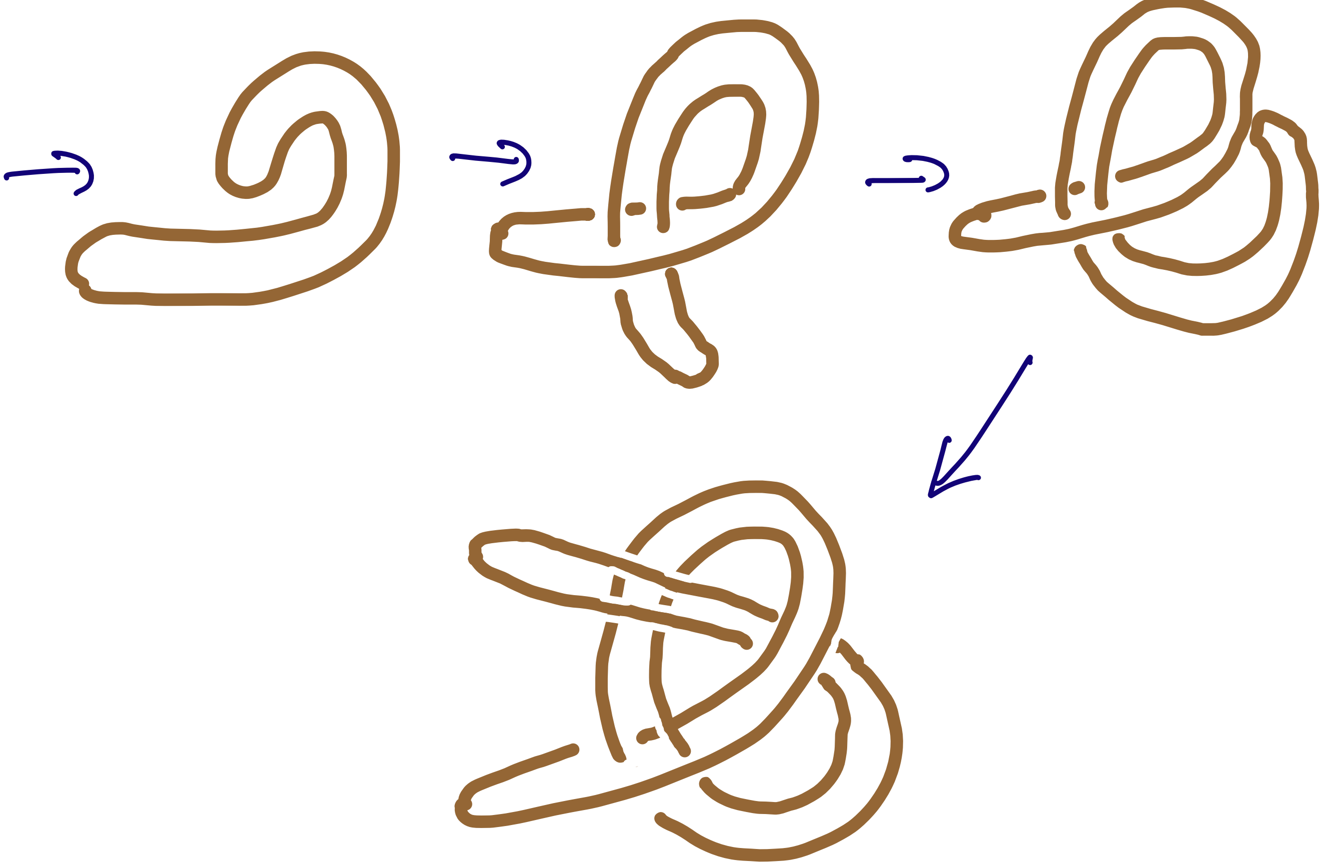
BSSM 2022

What is a knot?

Take a piece of string, twist it around itself and then seal the ends:

Two knots are isotopic if you can wiggle one without breaking the string until it becomes the other:



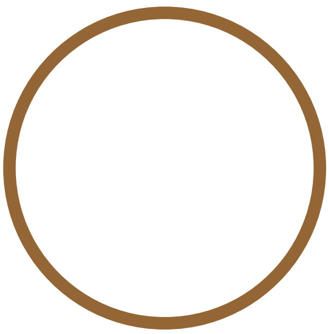


The string can stretch or shrink in length but never snap.

Knot theory is the study of knots up to isotopy.

It's a branch of TOPOLOGY

The simplest knot is called the unknot:



Here's some more:

TREFOIL

FIGURE 8.

Intuitively we can tell these knots are different, i.e. not isotopic.

But how can we prove it?

And what about this knot?

How can we tell knots apart?

Rigourous definitions

$$S^1 = \{ (x,y) \in \mathbb{R}^2 \mid x^2 + y^2 = 1 \}.$$
$$= \{ e^{i\theta} \in \mathbb{C} \mid \theta \in \mathbb{R} \}$$

A knot is the image of an embedding
$f : S^1 \to \mathbb{R}^3$

Embedding means:

(1) f is smooth. ie the map
$\theta \mapsto f(e^{i\theta})$ is smooth $\mathbb{R} \to \mathbb{R}^3$

(2) If $f(p) = f(q)$ then $p = q$

(3) $\frac{d}{d\theta} f(e^{i\theta})$ is never zero.

Suppose K_0, K_1 are two knots, the images of embeddings $f_0, f_1 : S^1 \longrightarrow \mathbb{R}^3$

K_0 and K_1 are isotopic if there is a map

$$F : [0,1] \times S^1 \longrightarrow \mathbb{R}^3$$

such that

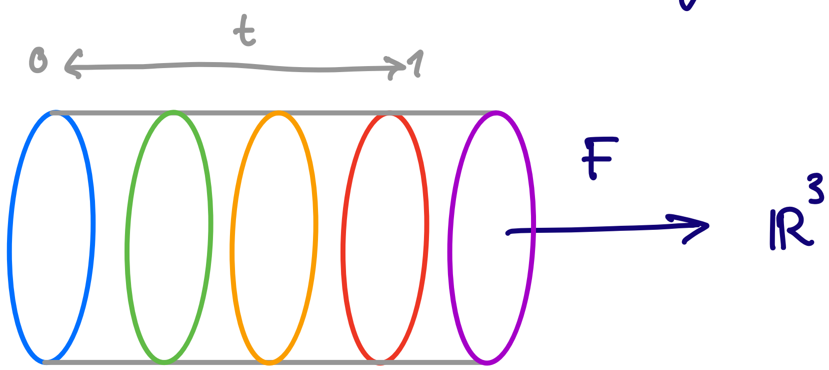
1) $F(0,p) = f_0(p)$ for all $p \in S^1$
and $F(1,p) = f_1(p)$ for all $p \in S^1$

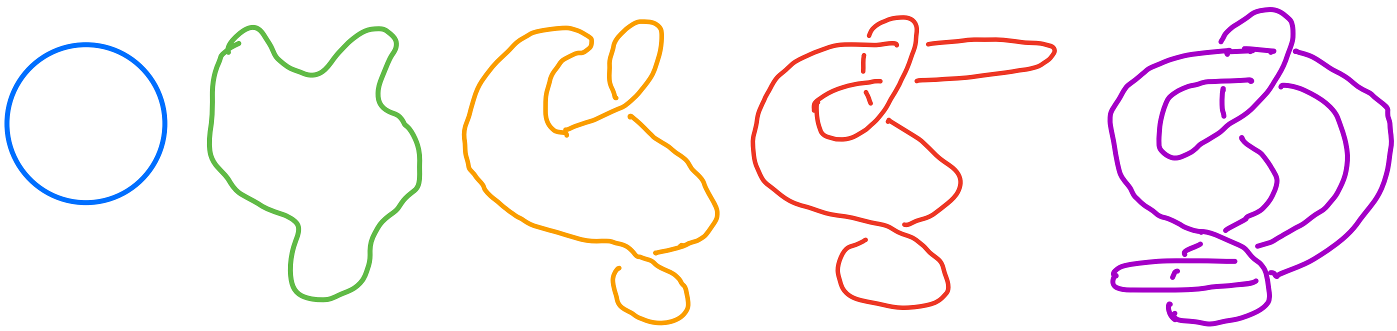
2) $(t, \theta) \mapsto F(t, e^{i\theta})$ is a smooth map $[0,1] \times \mathbb{R} \to \mathbb{R}^3$

3) for each t the map $f_t : S^1 \to \mathbb{R}^3$ given by

$$f_t(p) = F(t,p)$$

is an embedding.





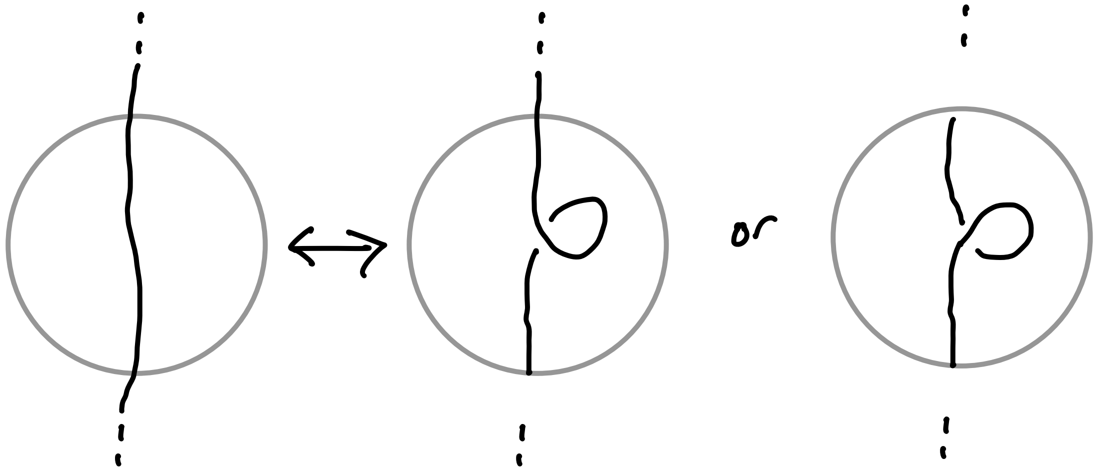
The drawings we have been using of knots, as closed paths on the plane with crossings, are called projections.

Clearly, one knot has many projections

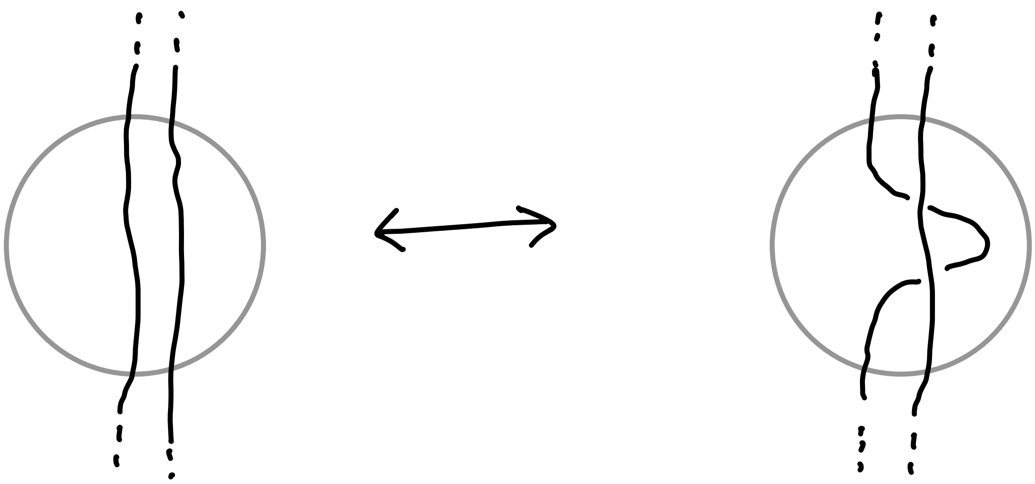
Reidermeister moves

Given a projection, here are three simple moves we can make which don't change the knot (i.e. up to isotopy).

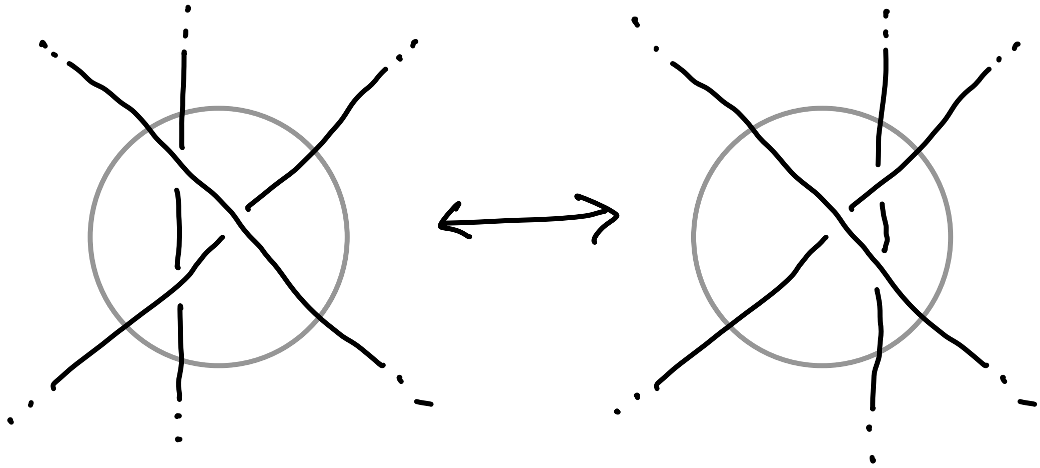
R_1



R_2



R_3



R_1, R_2, R_3 are the three **Reidemeister moves**

Reidermeister's Theorem

Two projections represent isotopic knots if and only if they are joined by a series of Reidermeister moves

In practice it's very hard work to prove knots are isotopic this way, and nearly impossible to show knots are distinct.

Instead we can use this as follows.

Suppose we have a certain property of knot projections which is left unchanged by the Reidermeister moves. Then we can use it to tell knots apart!

Tricolourability

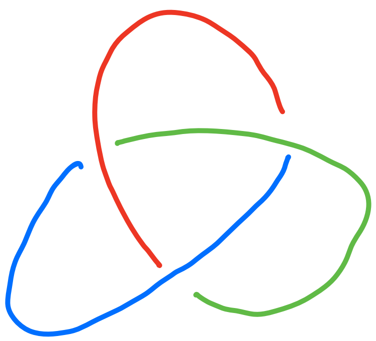
Pick three colours RED GREEN BLUE

We use them to colour strands of a knot projection

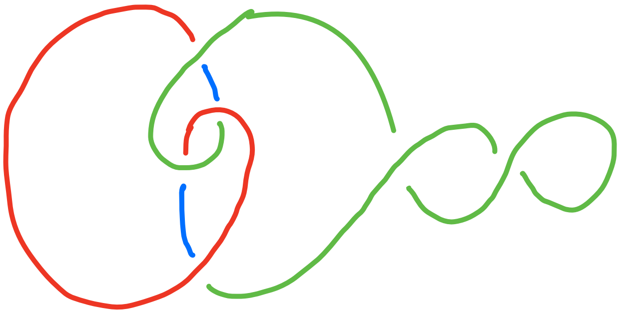
A strand is an unbroken arc running between two undercrossings

A tricolouring is a colouring with three colours for which at each crossing either all three colours appear, or only one colour appears.

If the projection has a tricolouring then it is called tricolourable.



The 'standard' projection of the trefoil is tricolourable



Another less standard projection of the trefoil which is also tricolourable.

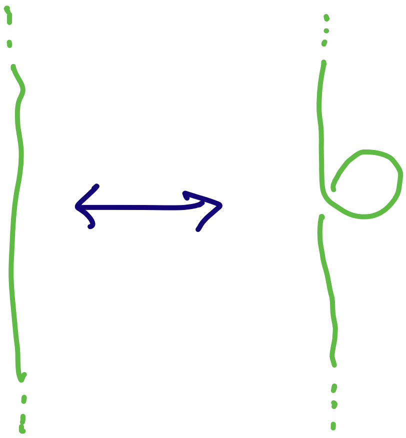
Theorem

If one projection of a knot is tricolourable then they all are.

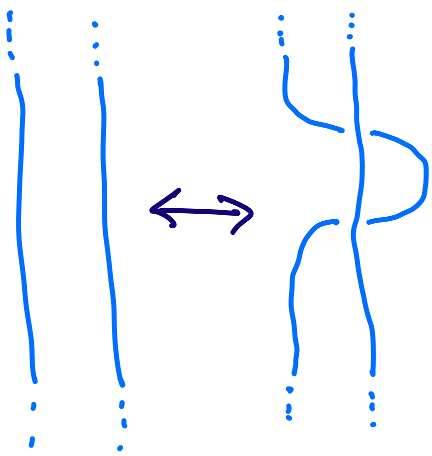
Sketch of proof.

We must check that tricolourability is preserved by the Reidemeister moves.

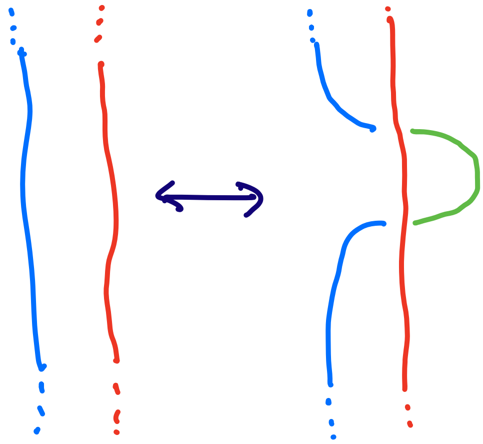
R_1:



R_2 two cases



or



R_3 several cases, left as an exercise for the reader!

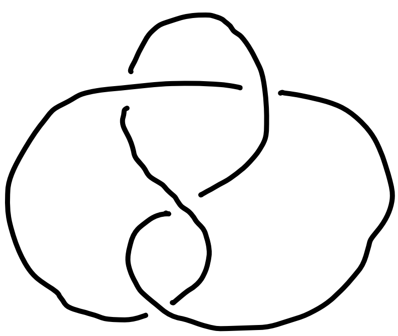
A knot is called tricolourable if all its projections are

The trefoil is tricolourable

The unknot is not tricolourable

So they are not isotopic!

The figure 8 knot:



Not tricolourable!

So, figure 8 and trefoil are different

But how can we prove that the figure 8 is not the unknot?

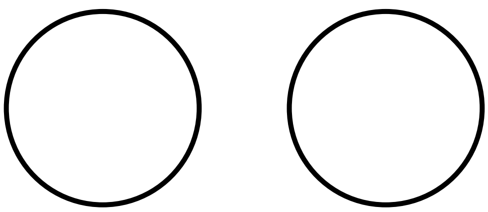
The Alexander polynomial

Discovered in 1928 by Alexander

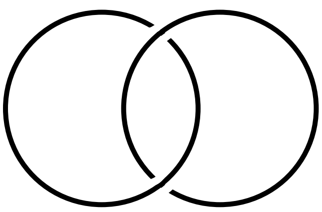
Revitalised in 1969 by Conway.

We need to use links

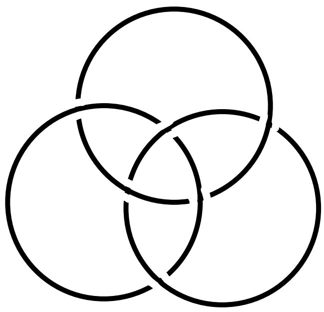
A link is a disjoint union of knots:



the 2-component unlink

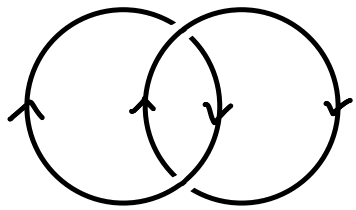


the Hopf link.



The Borromean rings (from 1300s crest of the Borromeo family)

We also need to orient our links



An oriented Hopf link

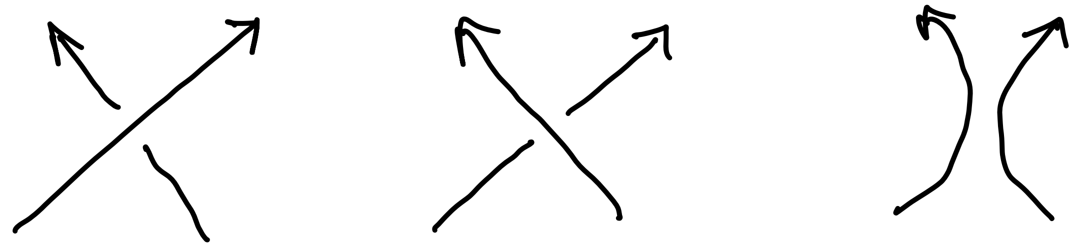
(If a link has c components it has 2^c orientations).

Theorem

The following two rules uniquely determine a polynomial $A_L(z)$ for each isotopy class of link L.

(1) If U is the unknot, then $A_U(z) = 1$.

(2) Let L_+, L_- and L_0 be three links which are related by changing a single crossing as indicated:



L_+ L_- L_0

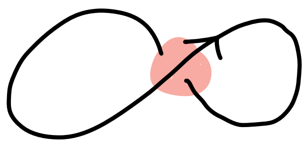
Then $A_{L_+} - A_{L_-} = z\, A_{L_0}$

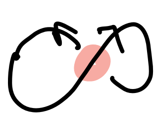
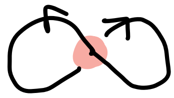
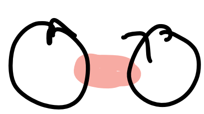
"SKEIN RELATION"

Examples

1) The 2-component unlink, $U \sqcup U$

We start with a diagram of the unknot:



$A\big($  $\big) - A\big($  $\big) = z\, A\big($  $\big)$

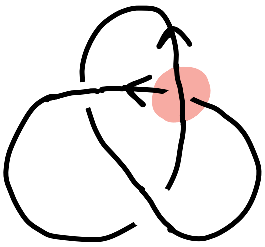
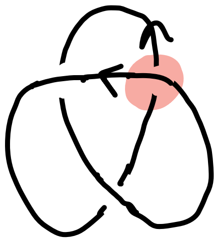
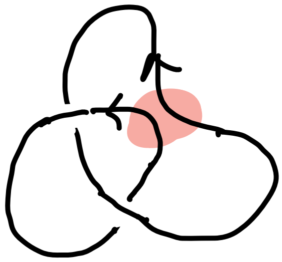
L_+	L_-	L_0
unknot	unknot	2-comp. unlink

$$A_U - A_U = z\, A_{U \sqcup U}$$

i.e. $A_{U\sqcup U} = 0$.

So Alexander polynomial of 2-comp unlink vanishes!

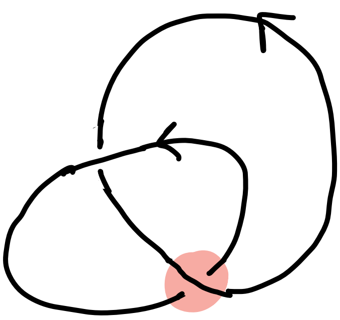
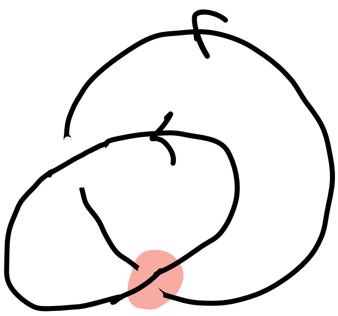
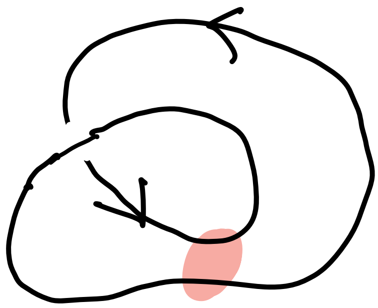
2) The trefoil

L_+ $\qquad$ L_- $\qquad$ L_0

Trefoil T $\qquad$ unknot U $\qquad$ Hopf link H

So $\qquad A_T = 1 + z A_H$

L_+ $\qquad$ L_- $\qquad$ L_0

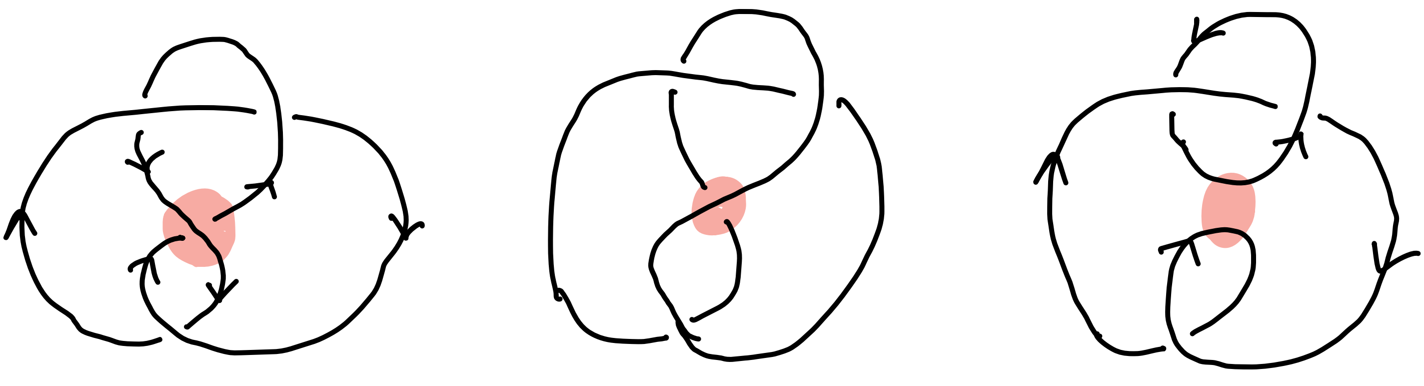
H $\qquad$ 2 comp unlink $\qquad$ U

So $A_H = z A_U = z$ and

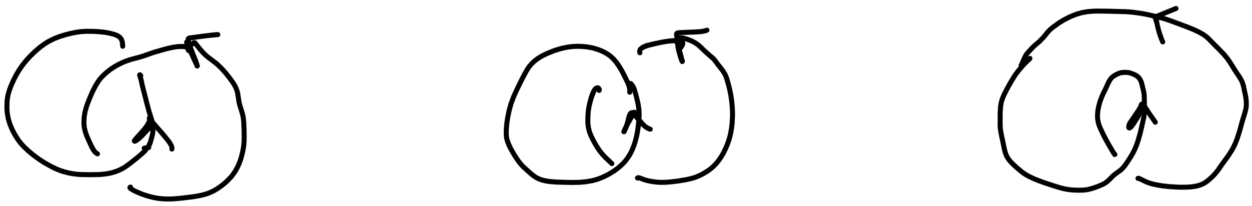
$$A_T = 1 + z A_H = 1 + z^2$$

This is another proof that the trefoil is not isotopic to the unknot.

3) The figure 8



L_+ $\qquad$ U unknot $\qquad$ $\hat{H}$ Hopf but NOT the same orientation as before



$\hat{H} = L_-$ $\qquad$ L_+ $\qquad$ $U = L_0$

So $A_{\hat{H}} = A_{L_+} - z A_{L_0}$

$= -z$

$A_{\text{fig } 8} = A_U + z A_{\hat{H}}$

$= 1 - z^2$

So we have shown that the figure 8 is not isotopic to the unknot!

The HOMFLYPT polynomial

Named after the discoverers in 1980s:

Hoste, Ocneau, Millett,
Freyd, Lickorish, Yetter,
Przytycki, Traczyk.

Theorem The two rules below uniquely determine a 2-variable polynomial associated to each isotopy class of oriented link:

1) $P_{U}(a,z) = 1.$

2) $a P_{L_+} - a^{-1} P_{L_-} = z P_{L_0}$

P_L is called the HOMFLYPT polynomial of L.

(It's a Laurent polynomial)

Notice that if you set $a = 1$ you recover the Alexander polynomial:

$$P_L(1,z) = A_L(z)$$

Exercise: compute the HOMFLYPT for the trefoil and the figure 8.

There are other link polynomials: the Jones polynomial, the Kauffman polynomial, ... and you (and I!) should

Learn about all of them !

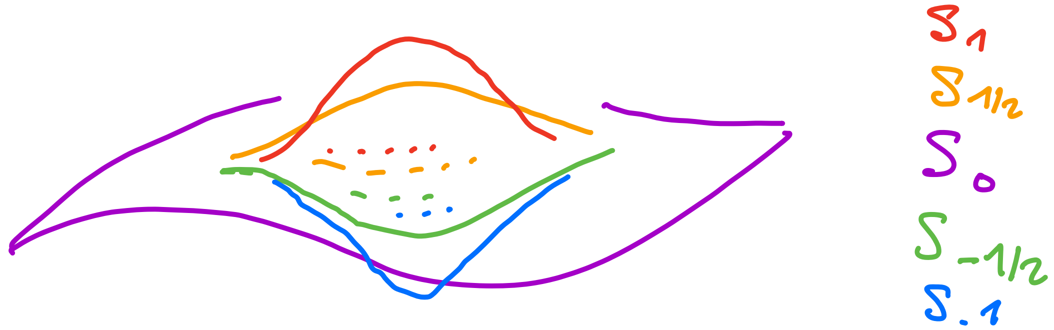
What is a minimal surface?

Let $S \subseteq \mathbb{R}^3$ be a smooth surface. S is called minimal if for every compactly supported perturbation S_t of S, we have

$$\frac{d}{dt} \mathrm{Area}(S_t) \Big|_{t=0} = 0$$

Here a "compactly supported perturbation S_t" means S_t is a path of surfaces for $t \in (-\varepsilon, \varepsilon)$ satisfying:

- $S_0 = S$,
- There is a compact set $C \subseteq \mathbb{R}^3$ st for all t, $S_t \setminus C = S \setminus C$



Let $\mathcal{S} = \{$ all surfaces in $\mathbb{R}^3 \}$

Area defines a function $\mathcal{S} \xrightarrow{\text{Area}} [0, \infty)$

A minimal surface is a critical point of this function.

Warning: surfaces can have infinite area (eg the plane $\mathbb{R}^2 \subseteq \mathbb{R}^3$) so the above story has to be carefully interpreted.

We can define "minimal" by only computing the area of $S_t \cap C$ where $S_t \setminus C = S \setminus C$, and C is compact.

Minimal surfaces are an important meeting point of geometry and analysis.

Suppose $\underset{\substack{\cap \text{ open} \\ \mathbb{R}^2}}{f : U} \to \mathbb{R}$ is smooth

$S = \{ (x, y, f(x,y)) \mid (x,y) \in U \}$

The graph of f.

S is minimal if and only if f solves a difficult partial differential equation:

$$\operatorname{div}\left(\frac{\nabla f}{\sqrt{1+|\nabla f|^2}}\right) = 0$$

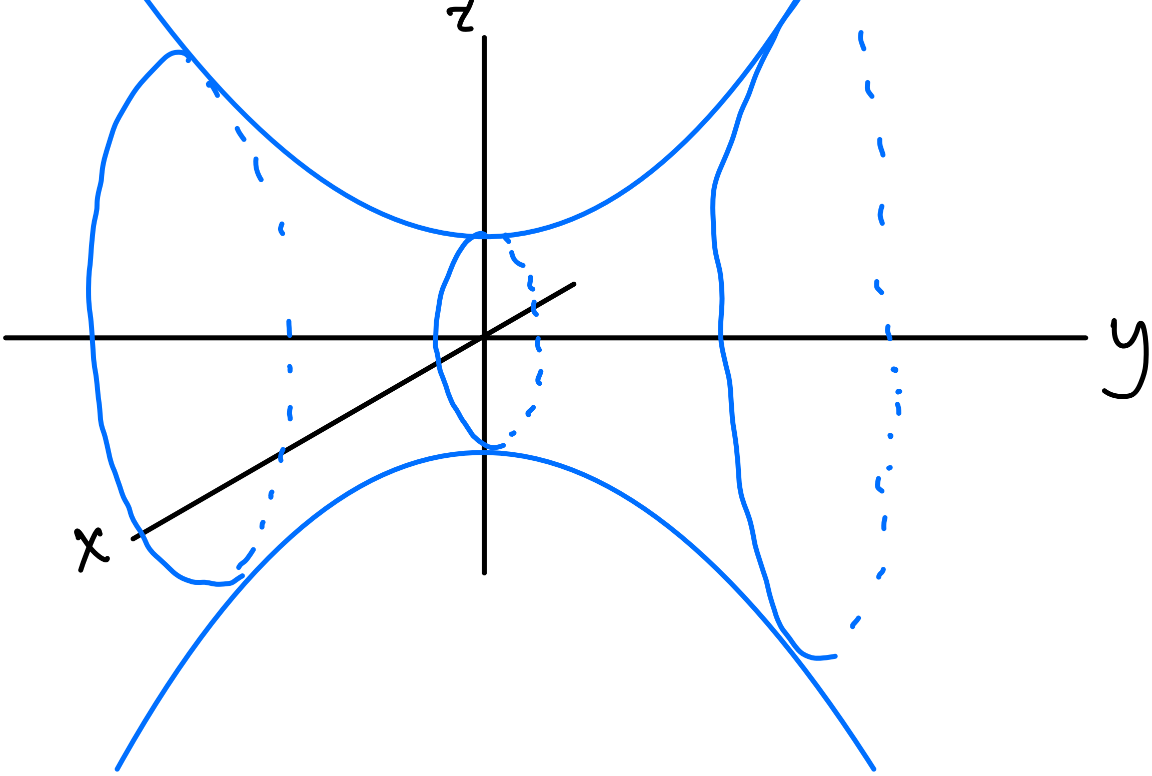
This equation is non-linear, second-order.

To find solutions you need geometry AND analysis.

One "easy" solution: take a curve in the (y,z) plane and rotate to get a surface of revolution S

Asking for S to be minimal gives an ODE for your curve, with an easy solution: $z = \cosh(y)$.

The surface of revolution is called the catenoid.



Minimal surfaces in $\mathbb{R}^3$ is a subject with a long history and is still right at the forefront of research today.

We need something a little more exotic however.

What is hyperbolic space?

Use coordinates $x_1, \dots, x_{n-1}, y$ on the half-space $\mathbb{H}^n = \{(x,y) \in \mathbb{R}^n \mid y > 0\}$.

If we have a curve $\gamma : [a,b] \to \mathbb{H}^n$ we can define the Euclidean length by

$$L_{Euc}(\gamma) = \int_a^b |\gamma'(t)| \, dt$$

We define the hyperbolic length by:

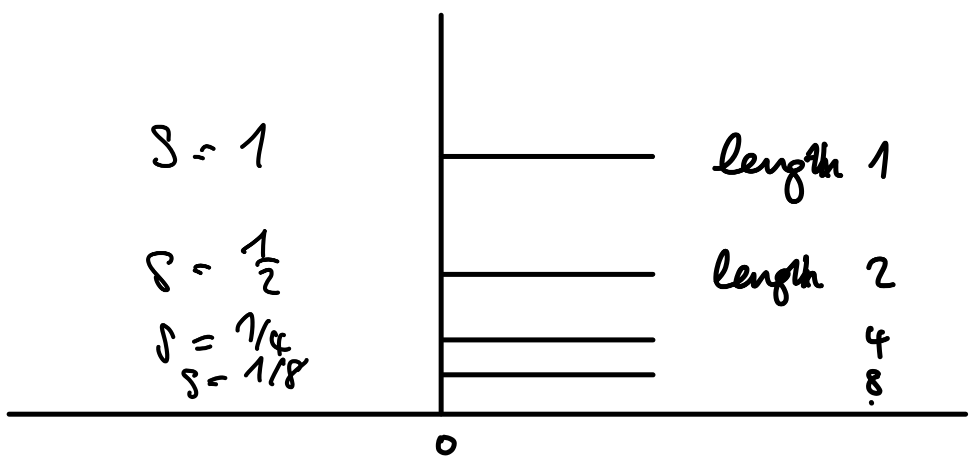
$$L_{hyp}(\gamma) = \int_a^b \frac{1}{y(t)} |\gamma'(t)| \, dt$$

where $y(t)$ is the y-coord of $\gamma(t)$.

This has the effect of magnifying the lengths of curves where y is small.

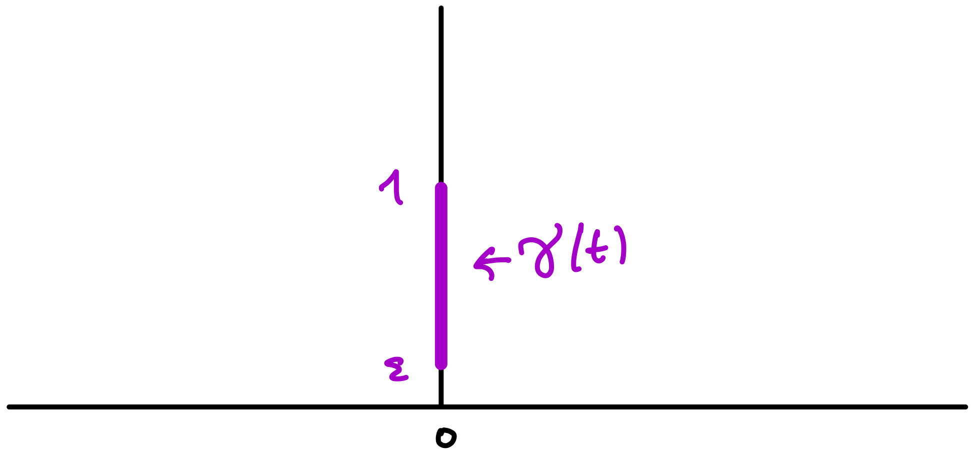
Eg $\gamma(t) = (t, 0, \dots, 0, s)$ for $t \in [0,1]$ and $s \in (0,\infty)$

$$L_{hyp}(\gamma) = \frac{1}{s}$$



Another example: $\gamma(t) = (0, \dots, 0, t)$ for $t \in [\varepsilon, 1]$

$$L_{hyp}(\gamma) = \int_\varepsilon^1 \frac{1}{t} \, dt = \log \frac{1}{\varepsilon}$$



As $\varepsilon \to 0$, $\log \frac{1}{\varepsilon} \longrightarrow \infty$

So even though the curve has length $1-\varepsilon$ Euclidean terms, in hyperbolic terms it has length tending to ∞ as $\varepsilon \to 0$.

Moral: the boundary $y=0$ is infinitely far away

Moral: space explodes as you get closer and closer to the boundary.

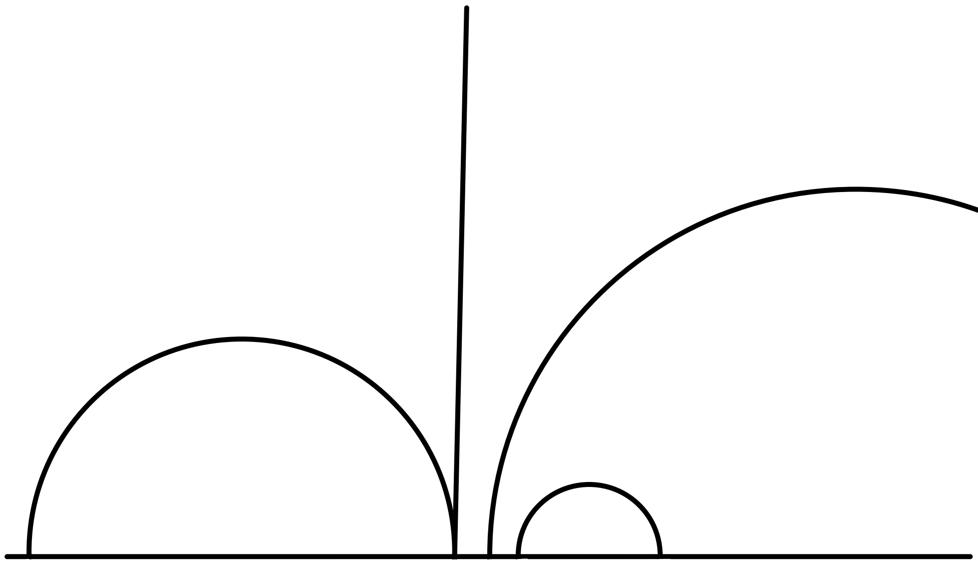
Geodesics

In Euclidean space, straight lines minimize the distance between points.

In hyperbolic geometry this role is played by the geodesics

A geodesic is a semi-circle in $\mathbb{H}^n$ whose centre lies on the boundary $y=0$

We include vertical rays as "semi-circles of infinite radius".



This is the start of non-Euclidean geometry and Riemannian geometry

These a branches of differential geometry which are right at the forefront of current research, both in mathematics and theoretical physics.

We will be interested in minimal surfaces in $\mathbb{H}^4$.

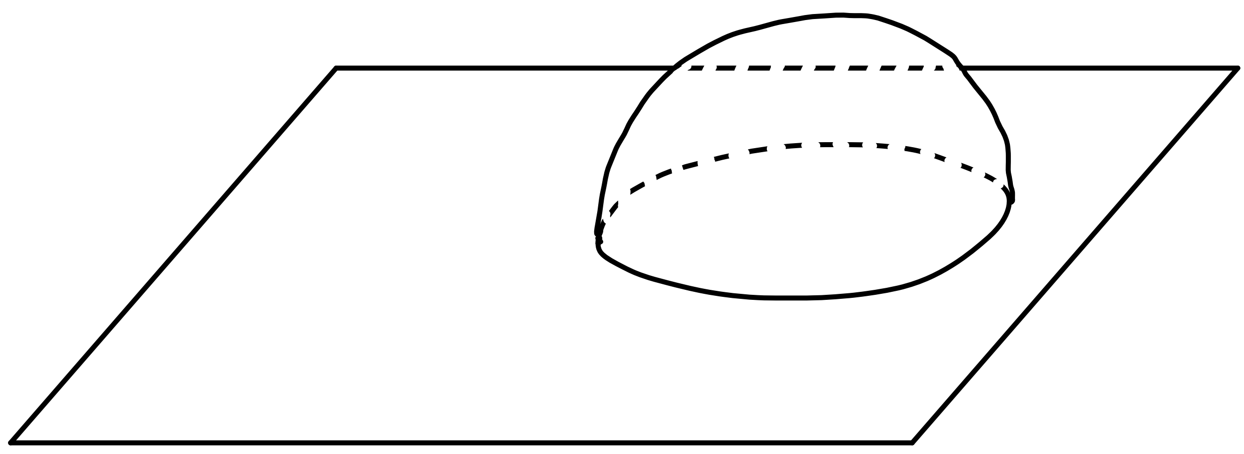
When computing the area of a surface in $\mathbb{R}^4$

we use a double integral. In hyperbolic geometry we weight the integral by $\frac{1}{y^2}$

Then the same definition makes sense: S is minimal if for every compactly supported perturbation, S_t,

$$\frac{d}{dt} \mathrm{Area}_{hyp}(S_t)\Big|_{t=0} = 0.$$

Examples: take a hemi-sphere in $\mathbb{H}^n$ centred on the boundary $y=0$. This gives a minimal surface:



A minimal surface in $\mathbb{H}^4$ runs out to infinity where it meets $\mathbb{R}^3 = \{y=0\}$ at right angles in a knot or a link.

The main conjecture

1) Given a link $L \subseteq \mathbb{R}^3$ you can count the minimal surfaces in $\mathbb{H}^4$ which meet the boundary in L.

2) This count is an INVARIANT. I.e. it doesn't change as you carry out an isotopy of L

3) These invariants can be put together to give a known link polynomial.

Finding minimal surfaces is very hard — you have to solve a non-linear PDE.

Computing knot polynomials from diagrams is relatively easy, we've already done a few!

So an easy calculation in knot theory would imply the existence of minimal surfaces!

This conjecture is in the style of a large body of modern mathematics (20th & 21st century) connecting topology, geometry and analysis.

Ideas go back to Stephen Smale in 60s and then a true revolution started by Simon Donaldson in 80s

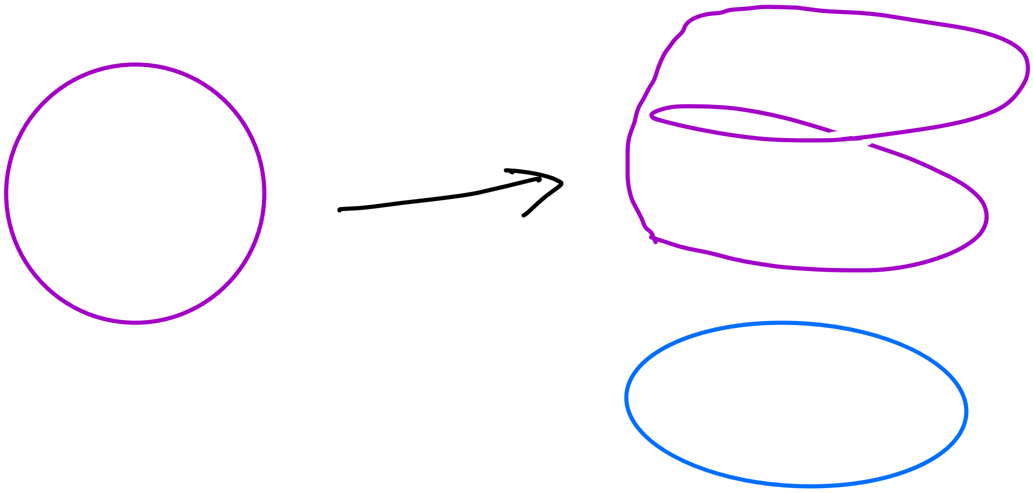
Moral: counting solutions to geometric PDEs can tell you a lot of things about topology and vice-versa!

Overall idea: take a technique from topology, apply it to the space of solutions to a geometric PDE, the conclusion will hopefully be something significant about the space on which the PDE is defined

The technique I want to use is the degree of a continuous map $f: S^1 \to S^1$.

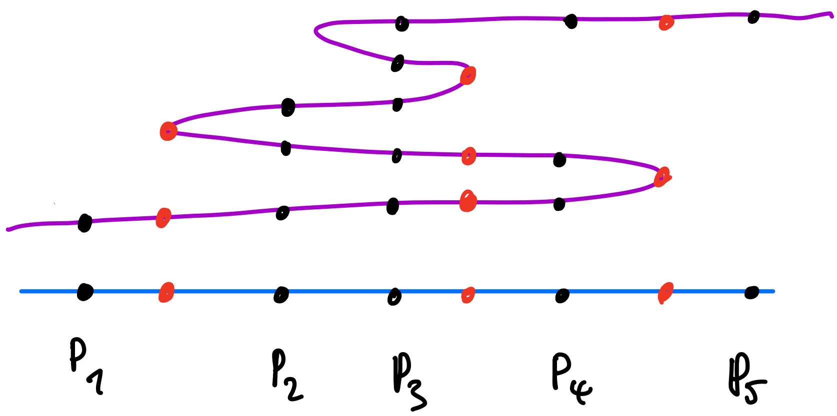
The degree counts the number of times

the image of f winds around the circle.

$f:$ 

$$\deg(f) = 2$$

We can compute the degree by counting the number of points in a preimage of a single point:

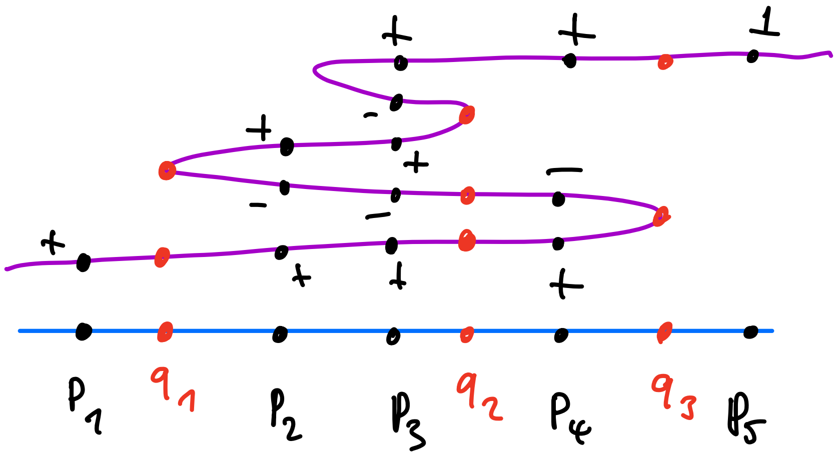


$f^{-1}(p_1)$ = single point

$f^{-1}(p_3)$ = three points

...

But $f^{-1}(p_3)$ has points where f goes in DIFFERENT directions.



The count WITH SIGNS taking into account the direction of f is the same for all of $f^{-1}(p_1)$, $f^{-1}(p_2), \ldots, f^{-1}(p_5)$!

However, there is a problem with the red points q_1, q_2, etc.

E.g. $f^{-1}(q_1)$ = two points and one of them doesn't have a sign

The problem is that df vanishes at this point ie its a critical point of f

Moral degree of f is signed count of points in $f^{-1}(p)$ as long as there are no critical points there.

For "generic" p this is always the case.

Important consequence: if p_1, p_2 are generic then the number of points in $f^{-1}(p_1)$ and in $f^{-1}(p_2)$ (counted with sign) IS THE SAME

Back to minimal surfaces in $\mathbb{H}^4$ filling knots or links.

Let $\mathcal{M}$ be space of all minimal surfaces in $\mathbb{H}^4$

Let $\mathcal{L}$ be space of all links in $\mathbb{R}^3$

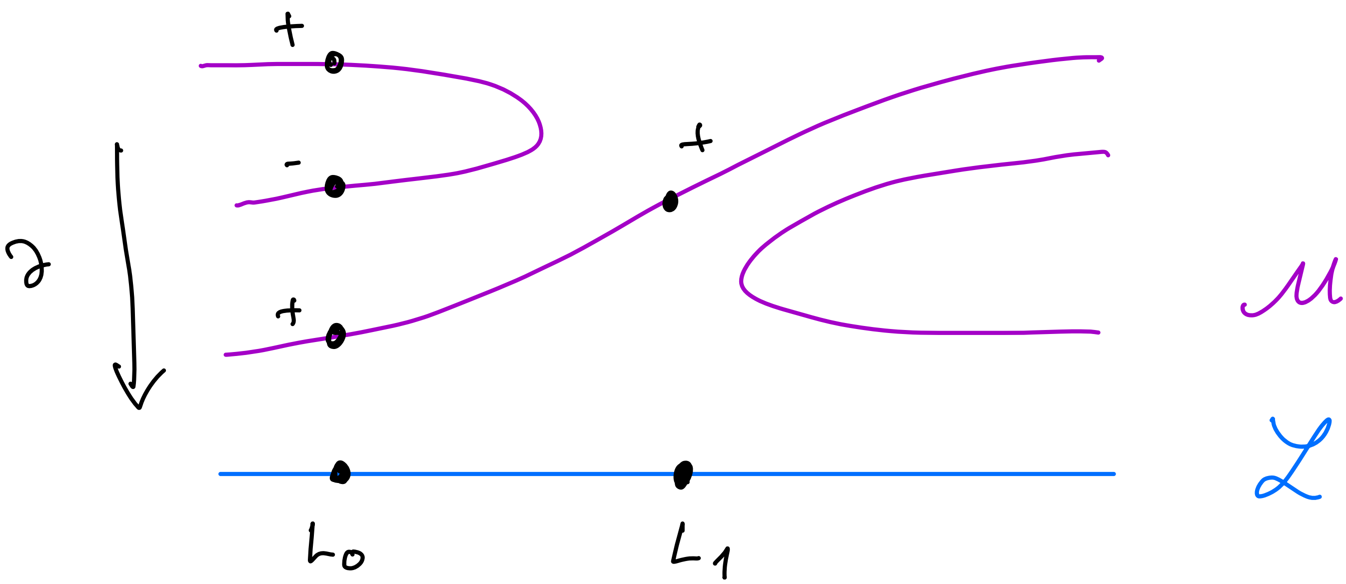
Given a minimal surface S, its boundary ∂S is a link ie we have a map

$$\partial : \mathcal{M} \longrightarrow \mathcal{L}$$

$$S \longmapsto \partial S$$

To count the minimal surfaces which fill a link $L \in \mathcal{L}$ we want to count number of points in $\partial^{-1}(L)$

$\boxed{\text{We want to define the degree of } \partial\text{!}}$



generic links in same isotopy class

In this picture $\deg(\partial) = 1$

What has to be done to make this work?

1) Want to ~~talk~~ about critical points of $\bar\partial : \mathcal{M} \to \mathcal{L}$

So we need $\mathcal{M}$ and $\mathcal{L}$ to have structure where we can talk about smooth maps

Theorem $\mathcal{M}$ and $\mathcal{L}$ are Banach manifolds (infinite dimensional!)

Also need $\bar\partial$ to be a "nice" smooth map

Theorem $\bar\partial$ is Fredholm of degree 0

Need to be able to define signs of points in $\bar\partial^{-1}(p)$

Theorem This can be done...

This is still not enough though.

Suppose L is generic. How do we know $\bar\partial^{-1}(L)$ is finite?

Suppose L_0, L_1 are generic and isotopic $\mathcal{L}$. How do we know $\partial^{-1}(L_0)$ and $\partial^{-1}(L_1)$ have same number of points

Want to prove properness!

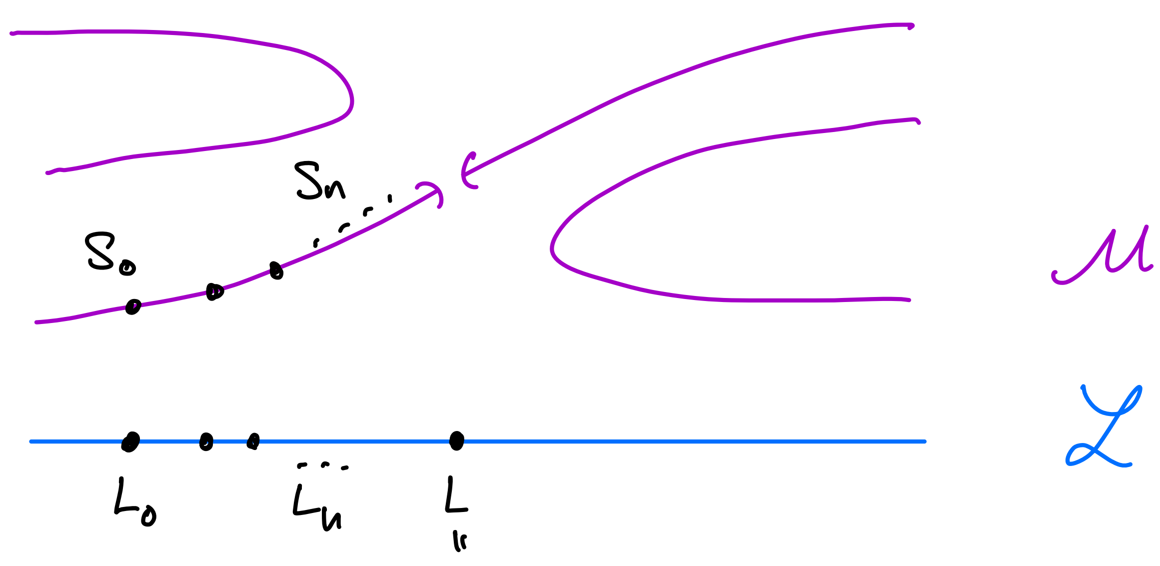
$S_n \in \mathcal{M}$ an infinite sequence of minimal surfaces

$L_n = \partial S_n$ their boundaries

Suppose $L_n \to L$

Want to prove (a subsequence of) the S_n converge to a minimal surface S with $\partial S = L$

Eg want to avoid this:

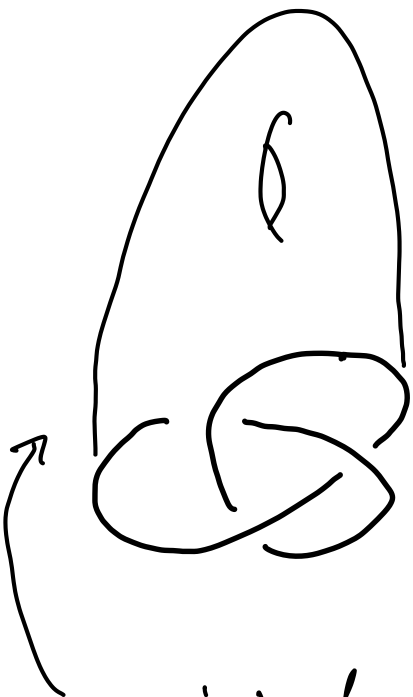


$$\lim L_n$$

$$\lim S_n \text{ doesn't exist !}$$

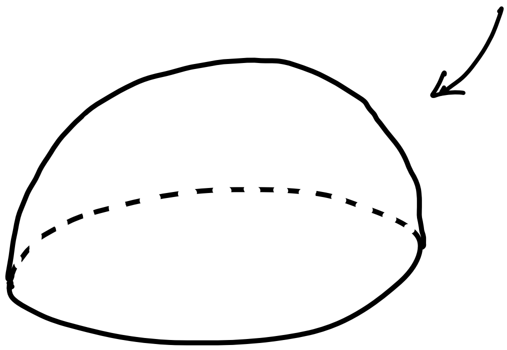
Until now we haven't talked about the topology of our minimal surfaces

Now it becomes important!

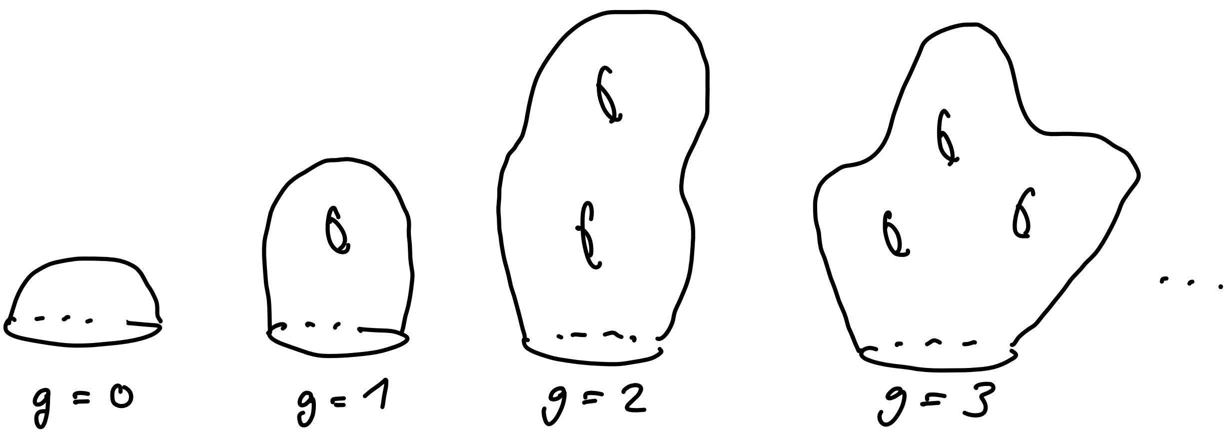


minimal surface of genus 1 filling a trefoil

minimal disk filling an unknot



"Genus" is "number of holes"



Theorem Let $\mathcal{M}_0$ be the space of minimal disks (genus 0).
The boundary map $\partial : \mathcal{M}_0 \to \mathcal{L}$ is **proper**

This means that we have a well defined knot invariant

$N(K)$ = # minimal disks in $\mathbb{H}^4$ with boundary K (counted w/. sign and K generic in isotopy class).

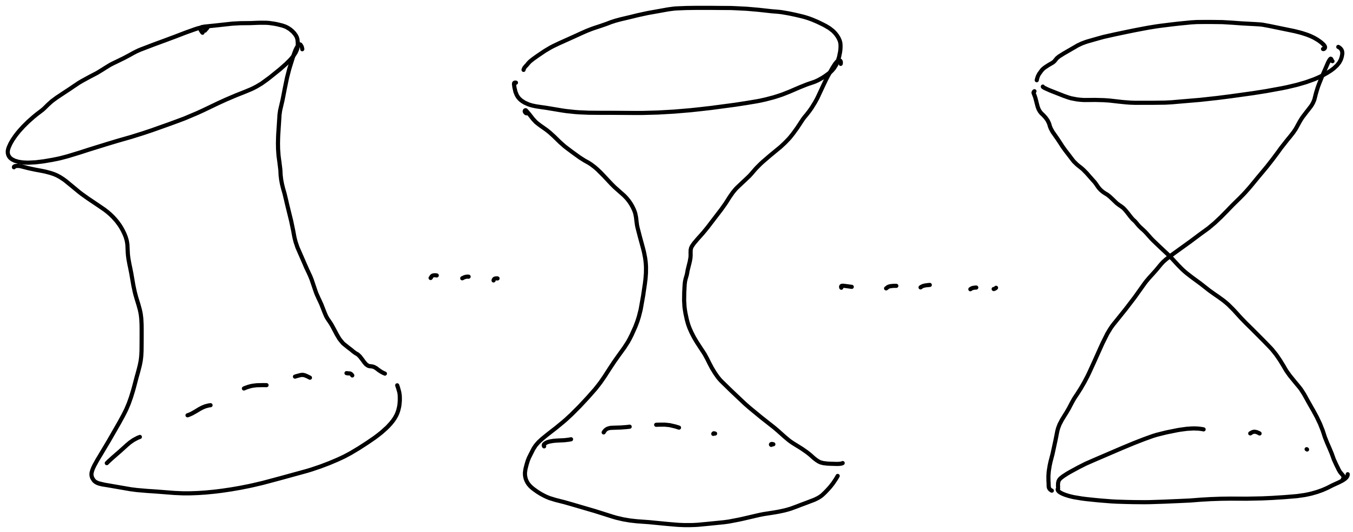
Big question: Is this a known knot invariant eg a coefficient in a knot polynomial?

And for higher genus or more boundary components?

Problem ∂ is not proper!

Beautiful example of this due to Tien Nguyen.

There is a path of minimal annuli which fill Hopf links:



S_0, $\partial S_0 = H_0$ $\qquad$ S_n, $\partial S_n = H_n$ $\qquad$ $H_n \to H$

$S_n \to S$

The limiting minimal surface is SINGULAR

It has a different topology.

If we want to count minimal annuli we need to deal with this phenomenon.

And a skein relation?...

Need to understand what happens to the minimal surfaces when the boundary develops a crossing: