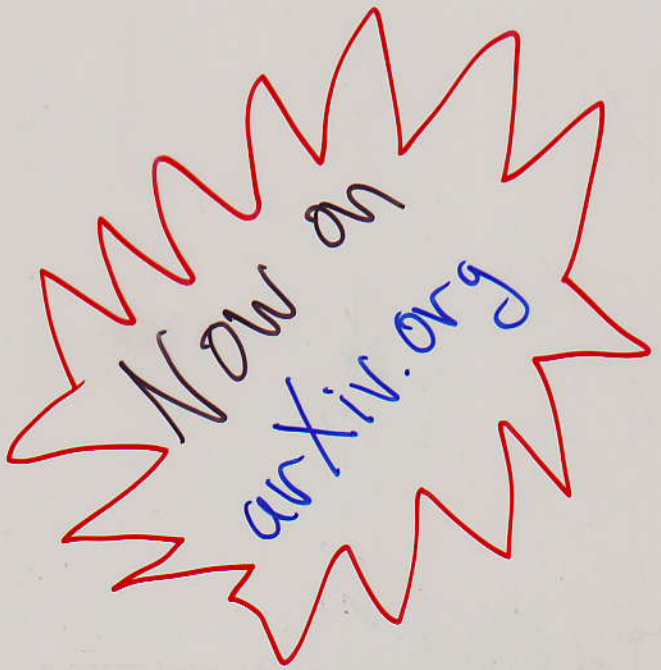
Deciding if a Quadratic Equation over a free group has a solution is NP-complete

O. Kharlampovich
A. M. Myasnikov
I. G. Lysenok
N. W. M. Touikan



Complexity:

P - problem $P(\text{input}) \to$ yes / no

e.g. P: "a+b=c?"
input: (3,7,11)
P(3,7,11) → No

$\mathcal{A}$ - algorithm $\mathcal{A}(\text{input}) \to$ yes / no after finitely many steps

---

Def: $\mathcal{A}$ - algorithm. The complexity $C_{\mathcal{A}}$ is a function

$C_{\mathcal{A}}(n) :=$ "max # of steps for $\mathcal{A}$ to terminate on an input of size n"

Def: A problem q is in $\mathcal{P}$ if $\exists$ algo $\mathcal{A}$; $A, B, N \in \mathbb{N}$ s.t.

$\mathcal{A}$ solves q & $C_{\mathcal{A}}(n) \leq A n^N + B$

Def: A problem q is in $\mathcal{NP}$ if $\exists$ a polynomial time algo $\mathcal{A}$ & a polynomial f such that for each input I, $q(I) \to$ yes iff there is a "certificate" $C(I)$ such that

$$\text{size}(C(I)) \leq f(\text{size}(I))$$

$$\& \ \mathcal{A}(C(I)) \longrightarrow \text{yes}$$

E.g. Is $p \in \mathbb{N}$ composite? certificate: (n, m) s.t. $n, m < p$
algo: $n \cdot m = p$?

Def: A problem q is $\mathcal{NP}$-hard if for any problem $p \in \mathcal{NP}$ there is a polynomial time reduction from p to q.

to show that a problem P is NP-hard it is enough to give a polynomial time reduction of an NP-hard problem q to P.

NP-hard is a lower bound on the complexity of a problem.

Def: NP + NP-hard = NP-complete

SYSTEMS OF EQUATIONS:

Hilbert's 10th problem: Is there an algorithm to decide solvability of a diophantine system of equations?

1970: Matiyasevich, Robinson, Davis, Putnam: No

Makanin 1982: There exists an algorithm that decides if an arbitrary system of equations over a free group has a solution.

2005: Diekert, Gutierrez, Haenah; based on Plandowski (2004) show that decidability of a sys of eqns over a free group is in PSPACE ← upper bound

2008: We give a lower bound...

F(A) - free gp on basis A.

A quadratic equation E with variables $\{x_i, y_i, z_j\}$ and non-trivial coefficients $\{w_i, d\} \in F(A)$ is said to be in *standard form* if its coefficients are expressed as freely and cyclically reduced words in A^* and E has either the form:

$$\left(\prod_{i=1}^{g}[x_i, y_i]\right)\left(\prod_{j=1}^{m-1} z_j^{-1} w_j z_j\right) d = 1 \text{ or } \left(\prod_{i=1}^{g}[x_i, y_i]\right) d = 1 \tag{1}$$

where $[x, y] = x^{-1}y^{-1}xy$, in which case we say it is *orientable* or it has the form

$$\left(\prod_{i=1}^{g} x_i^2\right)\left(\prod_{j=1}^{m-1} z_j^{-1} w_j z_j\right) d = 1 \text{ or } \left(\prod_{i=1}^{g} x_i^2\right) d = 1 \tag{2}$$

in which case we say it is non-orientable. The *genus* of a quadratic equation is the number g in (1) and (2) and m is the number of coefficients. If $g = 0$ then we will define E to be orientable. If E is a quadratic equation we define its *reduced Euler characteristic*, $\overline{\chi}$ as follows:

$$\overline{\chi}(E) = \begin{cases} 2 - 2g \text{ if } E \text{ is orientable} \\ 2 - g \text{ if } E \text{ is not orientable} \end{cases}$$

We finally define the *length* of a quadratic equation E to be

$$\text{length}(E) = |w_1| + \ldots + |w_{n-1}| + d + 2(\text{number of variables})$$

It is a well known fact that an arbitrary quadratic equation over a free group can be brought to a standard form in time polynomial in its length.

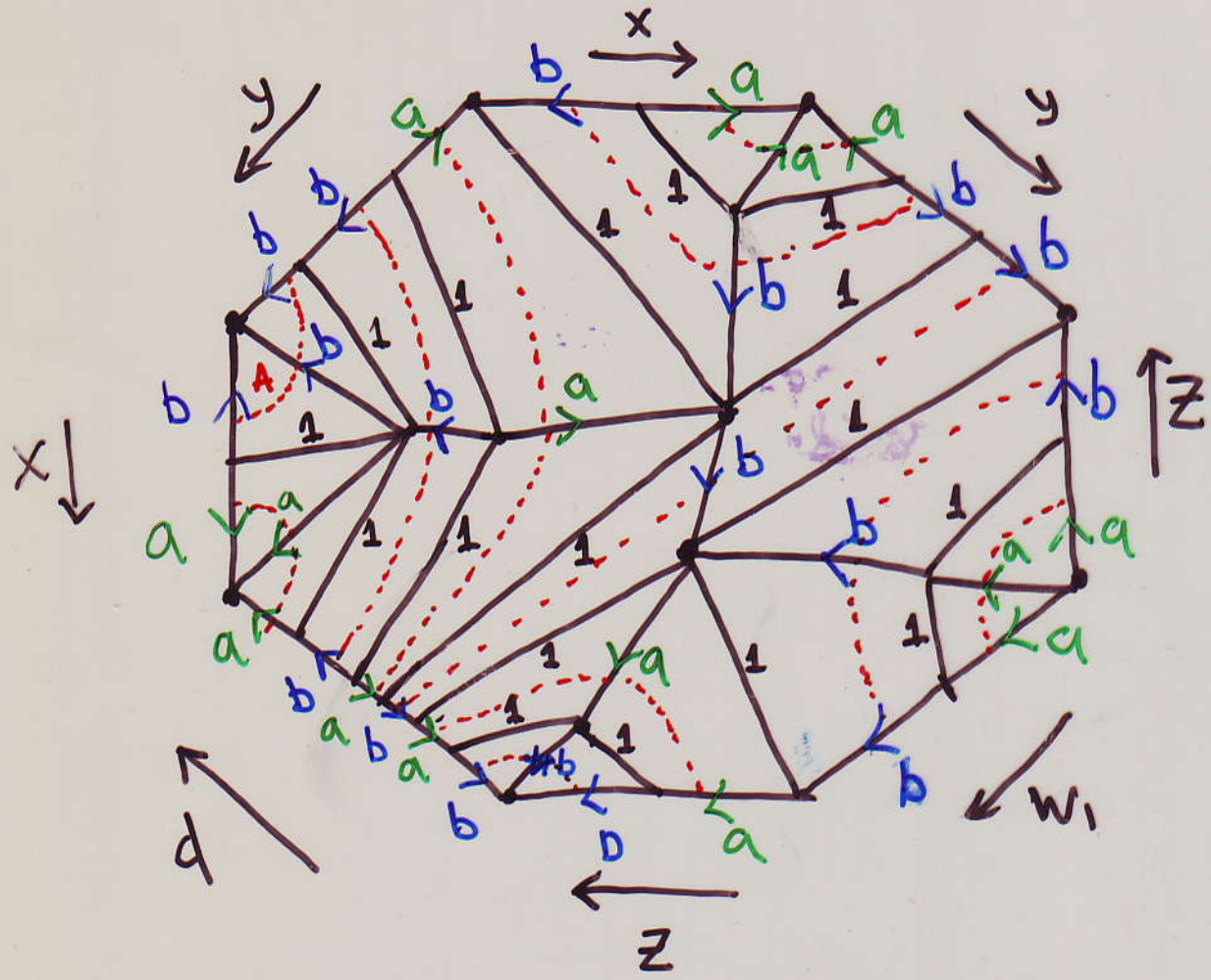
DIAGRAMS:

$x^{-1}y^{-1}xyz^{-1}\underbrace{ab}_{w_1}z\underbrace{b^{-1}a^{-1}b^{-1}a^{-1}ba}_{d}=1$

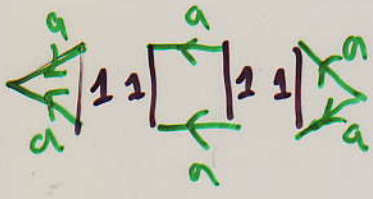
Variables: x, y, z ← coefs →

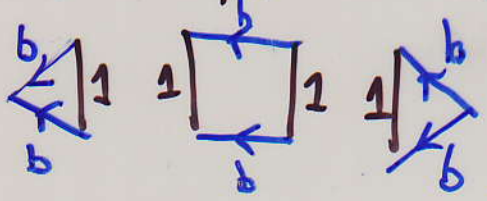
Solution: $x \mapsto \underbrace{b^{-1}a}_{B}$ $y \mapsto \underbrace{a^{-1}b^2}_{B}$ $z \mapsto \underbrace{ab}$

~~$a^{-1}bA\ b^{-1}\ b^{-1}a\ b^{-1}a\ a^{-1}bb\ b^{-1}a^{-1}ab\ ab\ b^{-1}a^{-1}b^2a^{-1}ba=1$~~



Solutions enable us to fill the polygon with cells

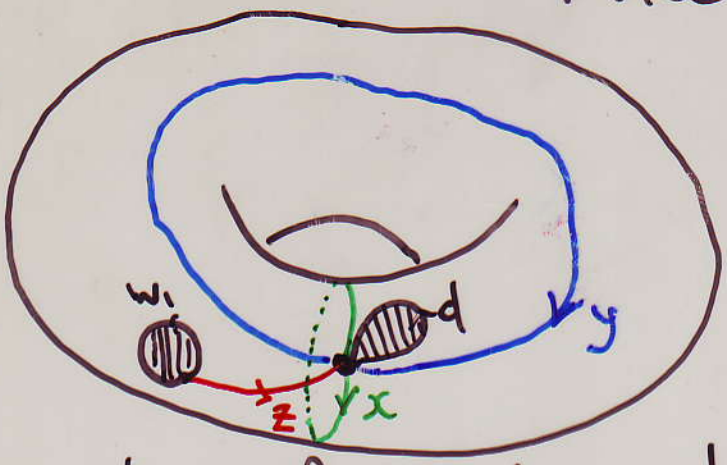




Olshanskii calls these 0-cells in his 1989 paper.

QUADRATIC $\leftrightarrow$ EACH VARIABLE OCCURS TWICE

So in our eg.

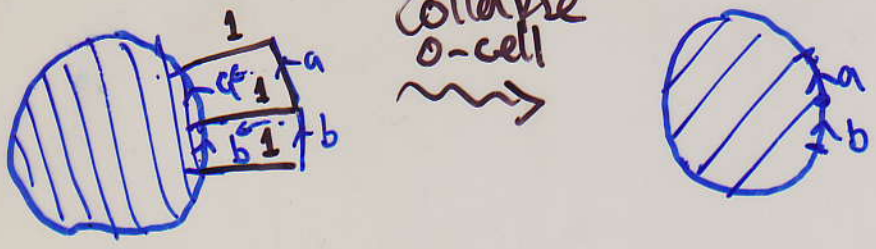


So we get a compact surface along whose boundary we can read the coefficients, moreover this surface is tiled by 0-cells (not shown). Conversely any such tiling by 0-cells gives a solution of the eqn. This tiling is called a diagram.

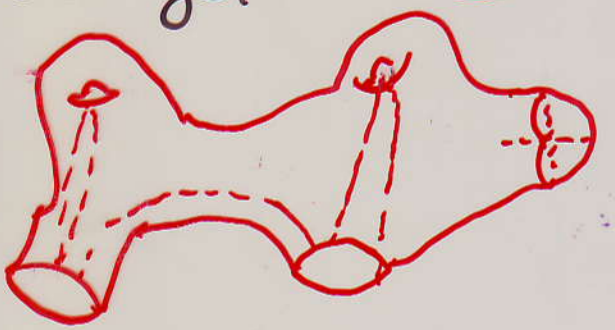
Def: A diagram Δ on a surface Σ is called simple if there are no embedded

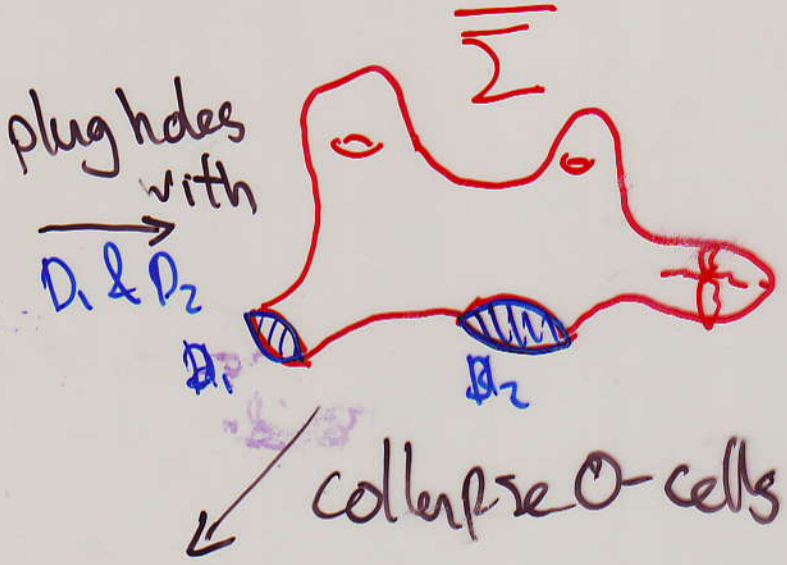
$$\chi(\Sigma_1) + \chi(\Sigma_2) = \chi(\Sigma_0) + 2$$

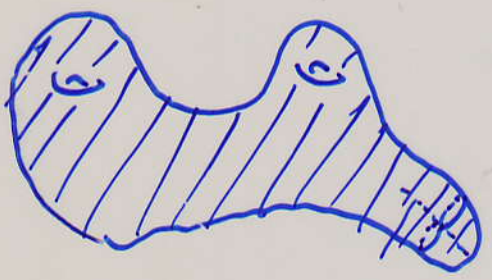
For a surface Σ with a simple diagram Δ we can plug $\partial\Sigma$ with discs & doing stuff like:



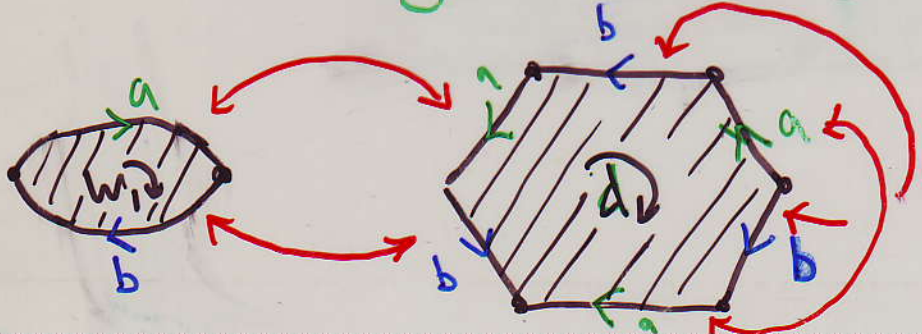
we get Σ







The surface $\overline{\Sigma}$ is realizable by attaching together discs with labeled boundaries along their boundaries respecting labels. Conversely realizing $\overline{\Sigma}$ in this way enables to get Σ & a diagram.



From our example

Theorem 0.1. *Let E be a quadratic equation over $F(A)$ in standard form. If $g = 0, m = 2$, or E is not orientable and $g = 1, m = 1$ then we set $N = 1$. Otherwise we set $N \leq 3(m - \overline{\chi}(E))$. E has a solution if and only if for some $n \leq N$;*

(i) there is a set $P = \{p_1, \ldots p_n\}$ of variables and a collection of m discs $D_1, \ldots, D_m$ such that,

(ii) the boundaries of these discs are circular 1-complexes with directed and labeled edges such that each edge has a label in P and each $p_j \in P$ occurs exactly twice in the union of boundaries;

(iii) if we glue the discs together by edges with the same label, respecting the edge orientations, then we will have a collection $\Sigma_0, \ldots, \Sigma_l$ of closed surfaces and the following inequalities: if E is orientable then each Σ_i is orientable and

$$\left(\sum_{i=0}^{l} \chi(\Sigma_i)\right) - 2l \geq \overline{\chi}(E)$$

if E is non-orientable either at least one Σ_i is non-orientable and

$$\left(\sum_{i=0}^{l} \chi(\Sigma_i)\right) - 2l \geq \overline{\chi}(E)$$

or, each Σ_i is orientable and

$$\left(\sum_{i=0}^{l} \chi(\Sigma_i)\right) - 2l \geq \overline{\chi}(E) + 2$$

and

(iv) there is a mapping $\overline{\psi} : P \to (A \cup A^{-1})^*$ such that upon substitution, the coefficients $w_1, \ldots, w_{m-1}$ and d can be read without cancellations around the boundaries of $D_1, \ldots, D_{m-1}$ and D_m, respectively; and finally that

(v) if E is orientable the discs $D_1, \ldots, D_m$ can be oriented so that w_i is read clockwise around ∂D_i and d is read clockwise around ∂D_m, moreover all these orientations must be compatible with the gluings.

Proved in "Homomorphism Diagrams of Surface Groups" Olshanskii 1989

Theorem 0.2. *There exists a polynomial time algorithm $\mathscr{A}$ such that a quadratic equation E over $F(A)$ in standard form has a solution if and only if there is a certificate c of size bounded by*

$$2(|w_1| + \ldots + |w_m - 1| + |d| + 3(2g + m)) \leq 8 * length(E)$$

such that $\mathscr{A}$ answers "yes" on the input (E, c).

The certificate will consist of the following:

1. A collection of variables $P = \{p_1, \ldots, p_n\}$ where $n \leq \max\{3(2g + m), 1\}$
2. A collection of substitutions $\overline{\psi} = \{p_i \mapsto a_i, i = 1 \ldots n\}$ where $a_i \in (A \cup A^{-1})^*$.
3. A collection of words in P^*

$$C = \begin{cases} C_1 = p_{11}^{\epsilon_{11}} \ldots p_{1l}^{\epsilon_{1j(l)}} \\ \ldots \\ C_m = p_{m1}^{\epsilon_{m1}} \ldots p_{mj(m)}^{\epsilon_{mj(m)}} \end{cases}$$

with $p_{ij} \in P, \epsilon_{ij} \in \{-1, 1\}$ and each $p_i \in P$ occurring exactly twice.

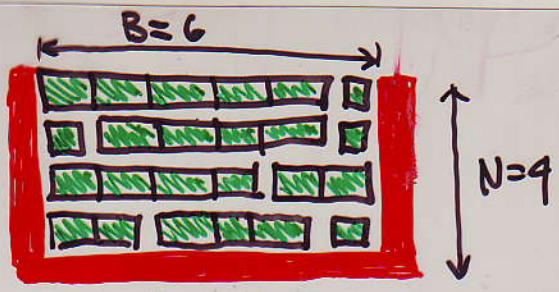
Got 

Step1: Check that upon the subst $p_i \xrightarrow{\overline{\psi}} F(A)$ we read w_i around ∂C_i: EASY

Step2: Verify the topology of the union of discs given by the p_i - gluing scheme.

Easy because Euler characteristic can be computed with local data. # of components & orientability also easy.

Bin Packing



Input: a k-tuple of integers $(r_1, \dots, r_k)$ & integers B, N s.t. $r_1 + \dots + r_k = NB$

Question: Is there a partition $\{1, \dots, k\} = S_1 \sqcup S_2 \sqcup \dots \sqcup S_N$ s.t. for $i = 1, \dots, N$ we have:

$$\sum_{j \in S_i} r_j = B$$

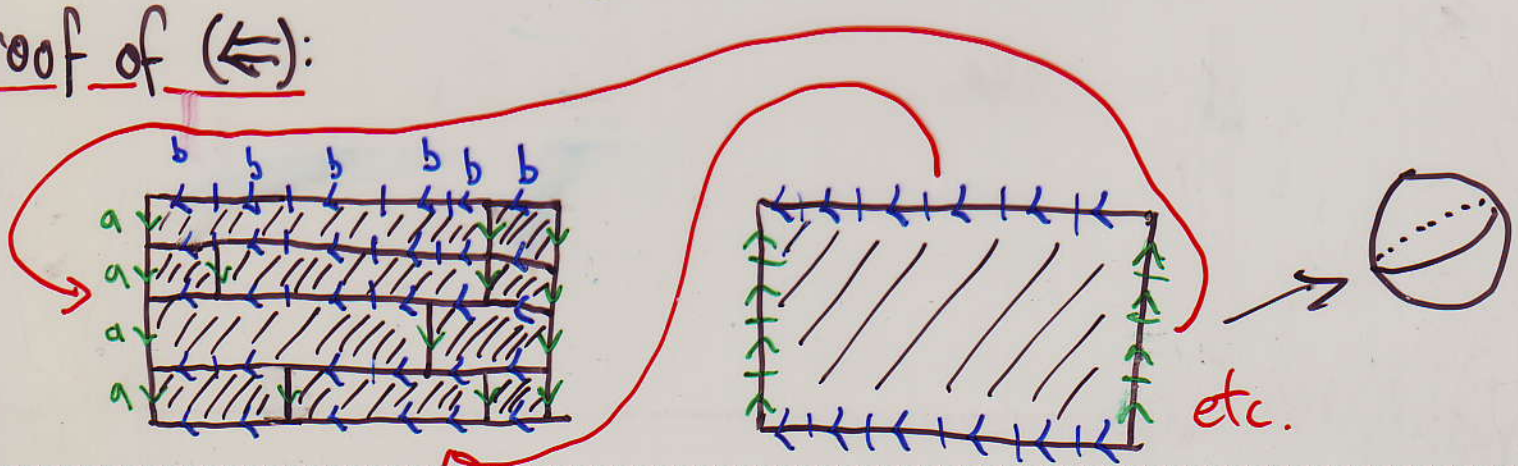
(strongly) NP-HARD

Thm: The equation:

$$\left(\prod_{j=1}^{k} z_j^{-1} [a, b^{r_j}] z_j \right) [a^N, b^B]^{-1} = 1$$

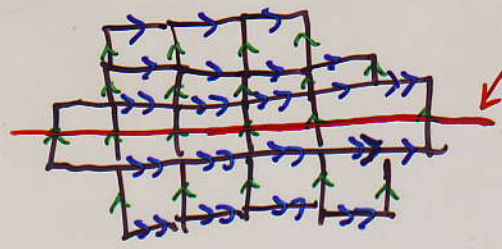
Has a solution iff the bin packing problem with inputs $(r_1, \dots r_k)$ & integers NB has a positive solution

proof of ($\Leftarrow$):

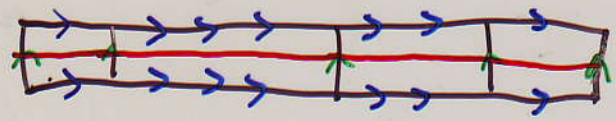


Consider a disc D in $\mathbb{E}^2$ tiled by $[a, b^n]$-discs

e.g. D

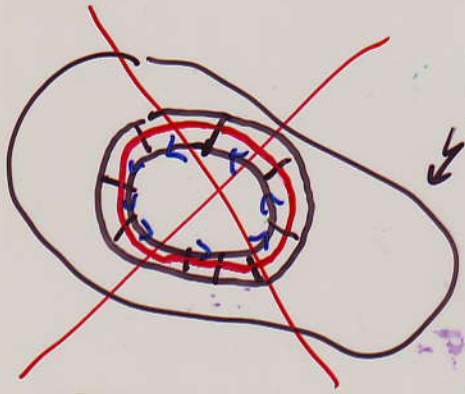


Def:

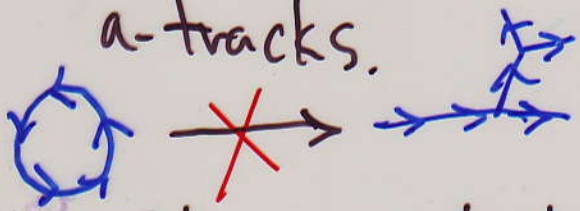


← Ribbon.

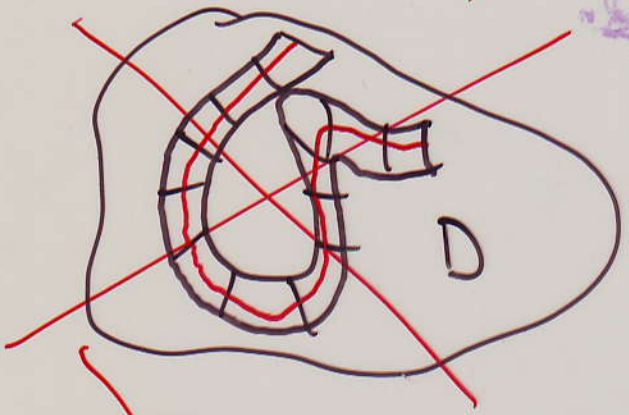
Lem 1:



No circular a-tracks.

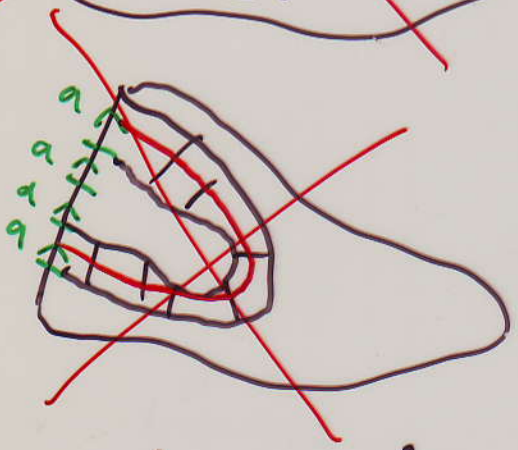


Lem 2:



ribbons embed

Lem 3:



Proof: glue mirror image contradict Lem 1.

Lem 4: Cannot tile S^2 with coherently oriented $[a, b^n]$-discs.

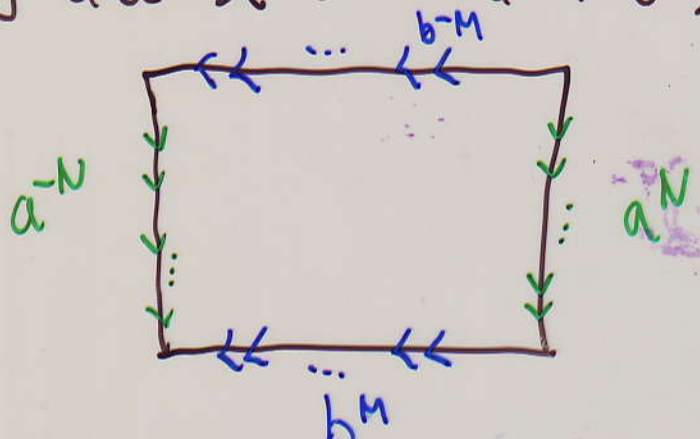
Proof: All a-tracks are circles. Remove a disc, embed into $\mathbb{E}^2$ contradict lem 1. //

proof of ($\Rightarrow$)

spse $\left(\prod_{j=1}^{k} z_j^{-1}[a, b^{r_i}]z_j\right)[a^N, b^B]^{-1} = 1$ has a solution.

Then we have spheres $S_1, \ldots S_m$ tiled by $[a, b^{r_i}]$-discs & 1 $[a^N, b^B]$-disc. If $m \geq 2$, then one sphere must be covered by $[a, b^{r_i}]$-discs contradicting Lemma.

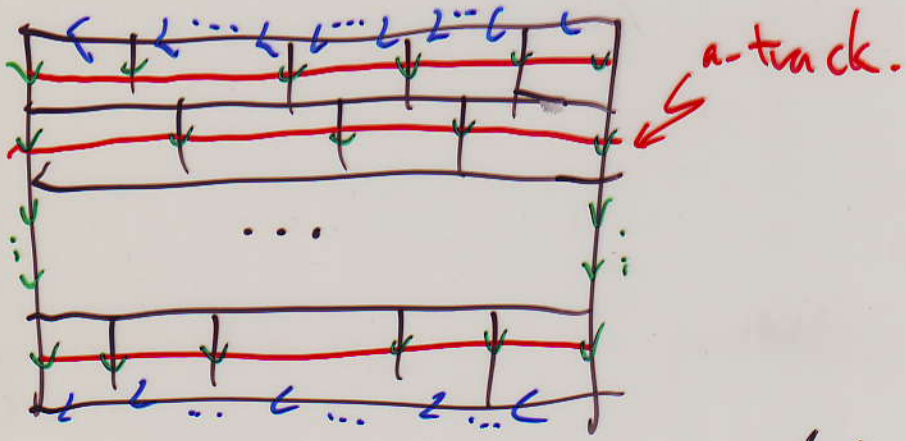
Let S_1 be the unique sphere, remove the $[a^N, b^B]$-disc & embed into $\mathbb{E}^2$. We get a disc



tiled by

- discs.

By Lemmas 1,2,3 the only possibility is



//