THE SET OF SLICE DOODLES IS NP-COMPLETE

Mikhail Khovanov

Department of Mathematics, Yale University, New Haven, CT 06520, USA

ABSTRACT. A doodle is a collection of closed triple point free piecewise-linear curves on a surface M. A doodle is called slice if there is a handlebody with boundary Mand a triple point free proper piecewise-linear mapping of a union of discs to the handlebody such that the boundary of discs maps (one-to-one except for a finite number of points) to the union of curves representing doodle. We prove that the set of slice doodles is NP-complete.

§0. Introduction

The boundary of a four-dimensional ball is a three-sphere. A knot in the threesphere is called *slice* if it bounds a disc embedded in the four-dimensional ball. The problem whether a knot is slice is very complicated and no solution is known at this time.

It turns out that many problems in knot theory admit a "toy" framework where, instead of considering isotopy classes of embeddings of circles in S^3 (links), we study triple point free curves on the 2-sphere up to homotopies that do not create triple points. Such curves are called *doodles*. In the piecewise-linear case, the local moves of doodles are depicted on figure 1.

Doodles have analogues of the braid group and fundamental group of knot complement (see [4]). In this paper we define slice doodles and prove that the problem whether a doodle is slice is NP-complete.

We were motivated by the work of Carter [1] where an algorithm for computing the genus of a triple point free curve on a surface is given (slice \cong genus is 0). The algorithm runs in exponential time.

 $\S1$ is based on the ideas of [1].

In §1 we transform the problem whether a doodle is slice to a certain combinatorial problem, COMBINATORIAL SLICE DOODLE. In §2 we reduce 3-SATISFIABILITY to COMBINATORIAL SLICE DOODLE, proving, therefore, NP-completeness of deciding whether a doodle is slice.

See [2],[5] for the definition of NP-completeness.

Doodles were originally defined by Fenn and Taylor in [3]. Our definition differs from theirs in that the curves are not required to be simple.

§1. Combinatorics of slice doodles.

Throughout the paper we work in the piecewise-linear category. The term *surface* will always mean closed oriented 2-manifold.

Definition 1.1. A double is a finite collection of triple point free closed curves on a surface M.

Underlying curves of a doodle constitute a graph on M with nodes of the graph being double points of the doodle and edges – pieces of curves between double points. All nodes have valency four and the graph might have loops and multiple edges. This graph is called a diagram of doodle.

Two doodles are called *equivalent* if there is a homotopy on M of underlying systems of curves with no triple points at any time during homotopy.

In terms of diagrams, there are two elementary moves depicted on figure 1 that preserve doodle. Two diagrams represent the same doodle iff they can be connected by a sequence of elementary moves.

Let $(C_1, ..., C_k)$ be the curves that constitute doodle.

Definition 1.2. A doodle $(C_1, ..., C_k)$ is called slice if for some diagram Δ there is a handlebody $H, \partial H = M$ and a triple point free proper map

$$\psi:\cup_{i=1}^k\mathbf{D}_i^2\to H$$

with $\psi(\partial \mathbf{D}_i) = C_i \subset M$ and the map

$$\psi(\partial \mathbf{D}_i) \to M, \quad 1 \le i \le k$$

is one-to-one except for the double points of selfintersections of C_i .

It is easy to see that if this is true (i.e. such ψ exists) for some diagram of doodle Δ , it is true for any diagram of Δ . A diagram is called *slice* if it is a diagram of a slice doodle.

Definition 1.3. SLICE DOODLE is a YES/NO problem that on input: surface M and a 4-valent graph Δ on M answers YES iff Δ is a diagram of a slice doodle and NO otherwise.

Suppose that a diagram Δ is slice. Fix ψ as in Definition 1.2. Let $S(\psi) \subset H$ be the set of double points of ψ . The set $S(\psi)$ is a disjoint union of arcs and circles.

Let $\psi^{-1}(S(\psi))$ be the preimage of $S(\psi)$ on $\bigcup_{i=1}^{k} \mathbf{D}_{i}^{2}$. It is a disjoint union of arcs and circles on $\bigcup_{i=1}^{k} \mathbf{D}_{i}^{2}$.

If β is a connected component of the double point set $S(\psi)$ there are 4 possibilities:

(i) β is an arc with both ends on $M = \partial H$ (figure 2). Then $\psi^{-1}(\beta)$ is a pair of arcs on $\bigcup_{i=1}^{k} \mathbf{D}_{i}^{2}$.

(ii) β is an arc with one end on M and the other end in the interior of H (figure 3). The latter end is a branch point of ψ , and $\psi^{-1}(\beta)$ is an arc on $\bigcup_{i=1}^{k} \mathbf{D}_{i}^{2}$.

(iii) β is a circle (figure 4a). In that case $\psi^{-1}(\beta)$ is a pair of circles in the interior of $\bigcup_{i=1}^{k} \mathbf{D}_{i}^{2}$.

(iv) β is an arc with both ends in the interior of H. The ends of β are branch points of ψ , and $\psi^{-1}(\beta)$ is a circle inside $\bigcup_{i=1}^{k} \mathbf{D}_{i}^{2}$.

Note that a connected component γ of $\psi^{-1}(S(\psi))$ is a circle iff $\psi(\gamma)$ is either a circle (case (iii)) or an arc with both ends inside the handlebody H (case (iv)). In the latter case $\psi(\gamma)$ has two branch points.

Proposition 1.1. If Δ is slice, there is a mapping $\psi_1 : \bigcup_{i=1}^k \mathbf{D}_i^2 \to H$ such that $\psi_1^{-1}(S(\psi_1))$ is a disjoint union of arcs (no circles).

This is equivalent to $S(\psi_1)$ being the union of connected components of types (i) and (ii).

Proof. Start with an arbitrary ψ . If $\psi^{-1}(S(\psi))$ has no circles, we are done. Otherwise take an innermost circle c. It bounds a disc **D** inside one of the discs $\mathbf{D}_1^2, ..., \mathbf{D}_k^2$ with no points from $\psi^{-1}(S(\psi))$ inside **D**. There are two possibilities:

1) $\psi(c)$ is a circle (c is of type (iii)). Then we cut and paste the neighborhood of $\psi(\mathbf{D})$ in H (see figure 4) and arrive at ψ^* with fewer number of double point circles.

(ii) $\psi(c)$ is an arc with both ends inside the handlebody H (c is of type (iv)). We throw away **D** and its image in H, glue c to itself (as on figure 5) and obtain ψ^* with one less double point circle than ψ .

Repeat till we arrive at the required ψ_1 .

(*)

We call a map ψ_1 satisfying the conditions of Proposition 1.1 minimal.

The complexity of a diagram Δ of a doodle is the pair $(dp(\Delta), c(\Delta))$ where $dp(\Delta)$ is the number of double points of Δ and $c(\Delta)$ is the number of components of Δ . For example, figure 10 diagram has complexity (6,3). Introduce a partial ordering on the set of diagrams with $\Delta_1 \leq \Delta_2$ iff $dp(\Delta_1) \leq dp(\Delta_2)$ or $dp(\Delta_1) = dp(\Delta_2)$ and $c(\Delta_1) \leq c(\Delta_2)$.

Let Δ be a diagram of a doodle on M. Suppose there is an oriented 3-manifold $N, M \subset \partial N$ and a triple point free mapping $\psi : \bigcup_{i=1}^{k} \mathbf{D}_{i}^{2} \to M$ that makes diagram

$$\begin{array}{ccc} \cup_{i=1}^{k} \mathbf{D}_{i}^{2} & \stackrel{\psi}{\longrightarrow} & N \\ & & & & & & \\ & & & & & & & \\ & & & & & & & \\ & & & & & & & \\ & & & & & & & & \\ & & & & & & & & \\ & & & & & & & & \\ & & & & & & & & & \\ & & & & & & & & \\ & & & & & & & & \\ & & & & & & & & \\ \end{array}$$

commute. In this case the doodle (or the diagram) is called *weakly slice*. Again, this concept is independent from the choice of diagram. The difference from slice doodles is that N is not necessarily a handlebody. The next proposition shows that it does not matter.

Proposition 1.2. If a doodle is weakly slice, it is slice.

Proof is by induction on the complexity of some diagram Δ of doodle.

Step 1: If $dp(\Delta) = 0$, Δ is a union of simple closed pairwise disjoint curves on M. Then Δ is slice because there is a handlebody H, $\partial H = M$ and a collection of properly embedded pairwise disjoint discs in H whose boundaries are Δ .

Step 2: Suppose that proposition is true for any diagram that is less than Δ w.r.t. \leq . Let us prove it for Δ . Suppose that Δ is weakly slice. We have commutative diagram (*). By Proposition 1.1 we can assume that discs intersect (and selfintersect) only along arcs with ends on the boundaries of discs. More precisely, we need the analogue of Proposition 1.1 for weakly slice doodles and it is immediate.

Next, take an outermost arc α on some disc \mathbf{D}_{j}^{2} (figure 6). Let l be the segment on the boundary on that disc such that $\alpha \cup l$ bound a subdisc with no arcs inside. Consider two cases.

Case 1: $\psi(\alpha)$ is of type (i), i.e. $\psi^{-1}(\psi(\alpha))$ is a union of two arcs.

Take $L = M \times [0, 1]$, $\Delta \in M \times \{0\}$. Consider a diagram (Δ, α) that coinsides with Δ outside a small neighborhood U of l(figure 7a depicts U). and is as on figure 7c inside U.

The complexity of (Δ, α) is less than the complexity of Δ because Δ has two more double points than (Δ, α) . Also, it is easy to see that (Δ, α) is weakly slice. By induction hypothesis, (Δ, α) is slice. Thus, there is a handlebody H and a set R of triple point free discs in H with $\partial R = (\Delta, \alpha)$.

Let Q be the following surface in $M \times [0, 1]$

(i) $\partial Q = \Delta \cup (\Delta, \alpha),$

(ii) outside $U \times [0,1]$ (U is a neighborhood of l on M) Q is $(\Delta \setminus (\Delta \cap U)) \times [0,1]$,

(iii) inside $U \times [0, 1]$ Q has one double point arc. Slices $Q \cap (U \times \{t\}), t = 0; 0.5; 1$ are shown on figure 7.

Then $R \cup_{(\Delta,\alpha)} Q$ is a collection of triple point free discs inside the handlebody obtained by gluing H and $M \times [0,1]$ through the identification $\partial H \cong M \times \{1\}$.

Therefore, Δ is slice.

Case 2: $\psi(\alpha)$ is of type (ii). This is treated similarly. Intersections of Q with $U \times \{t\}, t = 0; 1$ are depicted on figure 8.

For a diagram Δ of a doodle, denote the double points of Δ by different letters X, Y, \ldots . The doodle is composed of curves, say, C_1, \ldots, C_k . Pick up an orientation of each curve. Also, we fix an orientation of the surface M. For each $i = 1, \ldots, k$, start at some point of C_i , travel along C_i and read off the letters assigned to double points as we cross them. For each i, we get a sequence S_i (up to the cyclic order) of elements from the set $\{X, Y, \ldots\}$.

Again, for each i = 1, ..., k start at the same point of C_i , travel along C_i and compose the sequence L_i of +'s and -'s: each time we cross a double point from left to right, we add +, if we cross a double point from right to left we add - (see figure 9).

For each *i* take a disk \mathbf{D}_i . Let n_i be the number of elements in the *i*-th sequence. Mark n_i points on the boundary of \mathbf{D}_i and assign a letter to each marked point so that if we read off letters as we go along $\partial \mathbf{D}_i$ in clockwise order, we get the sequence S_i . Also, each marked point has a *polarization* + or - (from L_i).

An example is provided on figure 10. We have a diagram with 6 double points. Denote them by X, Y, Z, W, P, Q. Denote the components by C_1, C_2, C_3 . Then S_i, L_i 's, up to an obvious ambiguity, are

$$S_1 = \{X, Y, P, Q\}, S_2 = \{Q, W, Y, Z\}, S_3 = \{X, Z, P, W\},$$

$$L_1 = \{-, -, +, +\}, L_2 = \{-, -, +, +\}, L_3 = \{+, -, -, +\}.$$

Suppose now that the doodle is slice. Consider a minimal ψ of Δ (i.e. ψ satisfies the same conditions as ψ_1 in Proposition 1.1). Take a double point set $S(\psi)$ and its preimage $\psi^{-1}(S(\psi))$ on $\bigcup_{i=1}^{k} \mathbf{D}_i^2$. Then our marked points on the boundary of $\bigcup_{i=1}^{k} \mathbf{D}_i^2$ are pairwise connected by arcs inside discs. The following conditions hold: (a) Each letter is assigned to exactly two marked points,

(b) each marked point is connected by an arc to exactly one other point,

(c) arcs do not intersect,

(d) If a marked point with a letter, say, X, is connected to a marked point with another letter, say, $Y, Y \neq X$, then the other two of the marked points with letters X, Y are connected with each other,

(e) ends of each arc have different polarization (one is +, the other is -).

Vice versa, if we have a diagram Δ of a doodle and it is possible to satisfy conditions (a)-(e), the doodle is slice. We leave it to the reader to verify (*Hint:* Prove first that the doodle is weakly slice and then use Proposition 1.2 to deduce that it is slice).

§2. Reduction of 3-SATISFIABILITY to SLICE DOODLE.

We reduced SLICE DOODLE to the following combinatorial problem:

There is an alphabet (a finite set of cardinality n) of letters: X, Y, Z, ..., W and k discs D(1), ... D(k). Each disc has several marked points on its boundary and to each marked point one of the letters is assigned and + or - is assigned (the latter assignment is called *polarization*) such that

(i) Each letter in the alphabet is associated to exactly two marked points and these two points have different polarizations,

(ii) Each disc has at least one marked point (that implies $k \leq 2n$).

Problem. Is there a way to draw n straight line segments on discs D(1), ..., D(k) so that

 (a^*) ends of each segment are marked points of different polarizations,

 (b^*) each marked point is the end of exactly one segment,

(Note that (a^*) , (b^*) imply that the set of marked points is divided into sets of cardinality two – boundaries of segments)

 (c^*) if two marked points are connected and the letters assigned to them are different, then the two other marked points with the same letters are connected,

 (d^*) no two segments intersect.

Figure 11 depicts an example.

Note that a segment is allowed to connect two points with the same letters assigned to them.

We call this problem COMBINATORIAL SLICE DOODLE, or CSD, because it is a combinatorial reformulation of SLICE DOODLE.

Later in this chapter letters of an alphabet for CSD will always be capital letters with indices.

CSD is, obviously, an NP problem. Next we prove that CSD is NP-complete by *log-* space reducing 3-SATISFIABILITY to CSD.

An instance of 3-SATISFIABILITY (3SAT) is a finite set of variables $\{A, B, ...\}$ and a finite set of clauses, each clause containing no more than 3 elements from the set $\{A, \neg A, B, \neg B, ...\}$. The answer to an instance of 3SAT is YES if we can assign values *true/false* to the variables such that in each clause at least one of the elements is *true*. Otherwise the answer is NO. Variables of 3SAT have no indices, but letters of CSD in the reduction from 3SAT to CSD will always have indices (placed in brackets), so no misunderstanding should arise.

Now for the reduction. Pick an instance of 3SAT. We first construct a variable gadget. Each of the variables A, B, \ldots can take two values *true/false*. Pick one variable, say, X.

Let D be a disc (figure 12) with 4 marked points X(1), X(2), X(3), X(4). Let X(1), X(3) have polarization +, X(2), X(4) – polarization -. There are two ways to pairwise connect these points without segments intersecting. If X(1) is connected to X(4) and, consequently, X(2) to X(3), variable X is *true*, otherwise X is *false*.

A variable and its negation may appear several times in different clauses. Thus, we need a device to make several copies of X. This is achieved by a configuration of 3 discs **D1**, **D2**, **D3** on figure 13. Marked point on the boundary of **D1** has polarization + iff its first index is even. Marked point on the boundary of **D2** or **D3** has polarization + iff its index is odd.

Each of the 4 points $X(1,j), X(2,j), X(3,j), X(4,j), 1 \le j \le s$ is a "copy" of X(1), X(2), X(3), X(4) as the following lemma shows.

Lemma. If X is true all segments connecting marked points on figure 13 discs are horizontal. If X is false all segments are vertical.

In other words, if X(1) is connected to X(4), (X = true) there is only one choice to connect the rest of the marked points on figure 13 according to the rules (a^*) - (d^*) :

X(2) is connected to X(3), P(1) to P(4), P(2) to P(3), Q(1) to Q(4), Q(2) to Q(3), X(1,j) to $X(4,j), 1 \le j \le s$, X(2,j) to $X(3,j), 1 \le j \le s$. Similarly, if X = false, X(1) is connected to X(2), X(3) to X(4) and X(1,j) to $X(2,j), 1 \le j \le s$, X(3,j) to $X(4,j), 1 \le j \le s$. The fact that the set of th

Therefore, for each j = 1, ..., s, quadruple X(1, j), X(2, j), X(3, j), X(4, j) is connected exactly as X(1), X(2), X(3), X(4). In that sense it is a copy of X. If, for some j, we want X(1, j), X(2, j), X(3, j), X(4, j) to behave as the negation of X, we exchange X(2, j) and X(3, j) on figure 13.

Proof of the lemma is straighforward. \Box

We choose s to be the number of times X and its negation appear in the clauses of our instance of 3SAT.

Let us construct a clause gadget. Some clauses of 3SAT may have just one or two variables. In that case we change them into pseudo-three variable clauses:

$$(X) \to (X \cup X \cup X),$$
$$(X \cup Y) \to (X \cup X \cup Y).$$

Now, we are given a clause $(X \cup Y \cup Z)$, where X, Y, Z are some variables/their negations.

From variable gadgets we get a "copy" of each X, Y, Z:

$$egin{aligned} X(1,a), X(2,a), X(3,a), X(4,a), \ Y(1,b), Y(2,b), Y(3,b), Y(4,b), \ Z(1,c), Z(2,c), Z(3,c), Z(4,c) \end{aligned}$$

for some a, b, c.

Assign two discs depicted on figures 14-15 to the clause $(X \cup Y \cup Z)$. Here T(1) - T(6), W(1) - W(6) are new letters that appear only on the discs assigned to that clause.

Marked points with polarization + on the boundary of figure 14 disc are

X(1, a), X(3, a), W(1), W(3), W(5), T(2), T(4), T(6).

Marked points with polarization - on the boundary of figure 15 disc are

Y(1,b), Y(3,b), Z(1,c), Z(3,c), W(2), W(4), W(6), T(1), T(3), T(5).

We claim that $(X \cup Y \cup Z)$ is *true* if and only if there is a way to divide marked points on discs of figures 14-15 into pairs so that the conditions (a^*) - (d^*) are satisfied. To prove that, consider separate cases.

Case 1: X is true. Then X(1) is connected to X(4), X(2) to X(3). Now connect T(i) to W(i), $1 \le i \le 6$. We are left with 8 marked points Y(1) - Y(4), Z(1) - Z(4). We connect them in the way defined by values true/false of Y, Z. Conditions (a^*) - (d^*) hold.

Case 2a: X is false, Y is true. Connect X(1,a) to X(2,a), X(3,a) to X(4,a). Connect Y(1,b) to Y(4,b), Y(2,b) to Y(3,b), W(1) to W(6), W(4) to W(3), W(2) to W(5), T(1) to T(6), T(4) to T(3), T(2) to T(5). Four marked points Z(1,c), Z(2,c), Z(3,c), Z(4,c) are left. We connect them according to the value of Z. Again no two segments intersect.

Case 2b: X is false, Z is true. This case is equivalent to the case 2a with roles of Y and Z exchanged.

Case 3: All 3 variables X, Y, Z are *false*. It is easy to check that there is no way to satisfy (a^*) - (d^*) in this case.

Thus, starting from an instance of 3SAT with n variables and m clauses, we get an instance of COMBINATORIAL SLICE DOODLE with 3n + 2m discs and less that 100nm marked points.

We constructed a *log*-space reduction from 3SAT to COMBINATORIAL SLICE DOODLE. Therefore, CSD is NP-complete and SLICE DOODLE is NP-complete.

Our reduction from 3SAT to CSD is not parsimonious but the following slightly more sophisticated reduction is:

The variable gadget remains the same except that s (number of "copies" of a variable) might be bigger than the number of appearances of the variable and its negation in the clauses of the instance of 3SAT.

To each clause $(X \cup Y \cup Z)$ we associate two discs with marked points (in clockwise order, up to the cyclic permutation):

Disk 1:

$$\begin{split} &X(1,a0), X(4,a0), T(1), X(4,a1), X(3,a1), T(2), X(4,a2), X(3,a2), T(3), X(4,a3), \\ &X(3,a3), T(4), X(4,a4), X(3,a4), T(5), X(4,a5), X(3,a5), T(6), X(3,a), X(2,a), \\ &W(6), X(2,a5), X(1,a5), W(5), X(2,a4), X(1,a4), W(4), X(2,a3), X(1,a3), W(3), \\ &X(2,a2), X(1,a2), W(2), X(2,a1), X(1,a1), W(1). \end{split}$$

Marked points with polarization + are X, W's with odd first index and T's with even index.

Disk 2:

$$\begin{split} &W(1),T(1),Y(4,b2),Y(3,b2),T(4),W(4),Y(4,b3),Y(3,b3),Z(4,c),Z(1,c),W(5),\\ &T(5),T(2),W(2),Z(2,c),Z(3,c),Y(2,b3),Y(1,b3),Y(3,b1),Y(2,b1),W(3),\\ &T(3),Y(2,b2),Y(1,b2),T(6),W(6),Y(1,b1),Y(4,b1). \end{split}$$

Marked points with polarization + are Y, Z, T's with odd first index and W's with even index.

a0 - a5 are indices for 6 "copies" of X and b1 - b3 are indices for 3 "copies" of Y.

We define #CSD as a function problem that, given an instance of COMBINA-TORIAL SLICE DOODLE, says in how many ways we can split marked points into pairs with conditions (a^*) - (d^*) satisfied. The latter reduction is parsimonious, thus, it is also a reduction from #3SAT to #CSD. Therefore, #CSD is #P complete.

Remark. Suppose that we were in the smooth category. The difference is that we have only figure 1a local move of doodles and discs cannot have branched points. Conbinatorially, that means that an arc is forbidden to join two marked points with the same letter assigned to them. We get a variation of our original problems, that we call SMOOTH SLICE DOODLE and COMBINATORIAL SMOOTH SLICE DOODLE. It is immediate that our reduction from 3SAT to COMBINATORIAL SLICE DOODLE such as an arc and, therefore, SMOOTH SLICE DOODLE is NP-complete.

Question: Is SLICE DOODLE NP-complete if we are restricted to the case of

(i) one component doodles

or to the case of

(ii) doodles on the two-sphere.

We do not know the answers. The existence of many components was very important because we "localized" each variable and clause by assigning different components to them.

Acknowledgements.

I would like to thank my advisor Igor Frenkel for his support and Tanya Khovanova for her interest.

References

^{1.} Carter, J.S., Extending immersions of curves to properly immersed surfaces, Topology and its Applications 40 (1991), 287-306.

- 2. Garey, M.R., Johnson, D.S., Computers and Intractability A Guide to the Theory of NPcompleteness, 1979.
- 3. Fenn, R., Taylor, P., Introducing doodles, Lect. Notes in Math. 722, 37-43.
- 4. Khovanov, M., Doodle groups, to appear.
- 5. Papadimitriou C.H., Computational complexity, Addison-Wesley, 1994.













Figure 6







•

Figure 7







•

Figure 10











Figure 13

•







