Lecture Notes
Knot Theory and Applications

Davide Michieletto

March 21, 2018

Contents

Chapter 1

Introduction

see Davide's notes + historical introduction, Tait, Maxwell and Thomson + knot diagram + links + knots as embedding or inclusions

Knots are very much part of our everyday life. In some cases, they are extremely useful: sailors and climbers need these for their own safety or to secure boats, and can tie a number of them with little thought. In other cases, knots may instead be quite a nuissance. You may have found this yourself when needing to reel in a long extension cable, after hoovering a socket-free area in your parents' attic, or when having to disentangle your headphones! In this lecture we will introduce the physics and maths of knots. We will first discuss the classification of knots into prime and composite, and how to distinguish the first few simple knots. We will also introduce the concept of knot chirality, of twist and torus knots, and of the unknotting number. These concepts will be used in the next lecture when we discuss the physics of DNA knots. Finally, we will discuss a knot invariant, the Alexander polynomial.

1.1 Historical Context

The theory of knots was pioneered by the man who gives the name of the road where this building is, Peter Guthrie Tait (1831-1901). Incidentally (or not), this building was named after J. C. Maxwell who was first a friend of Tait (they went to the same school) and then a competitor (the latter beat him to the chair of natural philosophy in 1859). The two corresponded at length regarding the theory of knots.

The year 1867 (recently seen the Origin of the species from Darwin and the first colour photograph from Maxwell) and Tait invites W. Thomson from Glasgow to Edinburgh to assist an experiment with smoke rings. Helmoltz had theorised that fluid vortexes within an ideal medium (no dissipation or friction) are stable configurations. The smoke rings in fact bounced off one another and never mixed or linked together. Thomson associates the smoke rings to atoms, and coined a new theory of atoms as knotted vortexes of aether. As the smoke rings did not combine to form differently shaped rings, so knotted vortexes would remain the same when combining, each defining a different element. Tait was sceptical, especially because one can create many many knots, many more than there are different elements.

Eventually, Tait was taken by the beauty of knots and set out to create "periodic tables" of knots. Before him, only Gauss and Listing had been interested in the mathematics of knots and links.

Tait set his own notation, he would label each crossing and then each knot could be associated with a word spelling out the order of the crossing encountered by following the curve. Obviously, different words might correspond to the same knot, or mirror ones or composites (see below) so Tait had to do an enormous work to find the genuinely unique words.

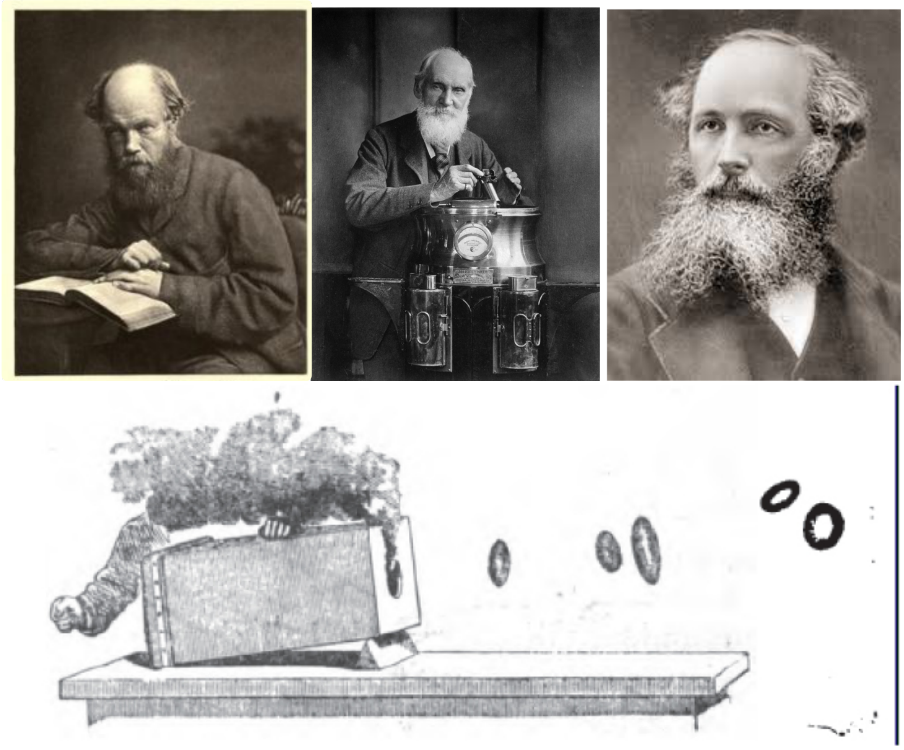


Figure 1.1: Tait's smoke machine.

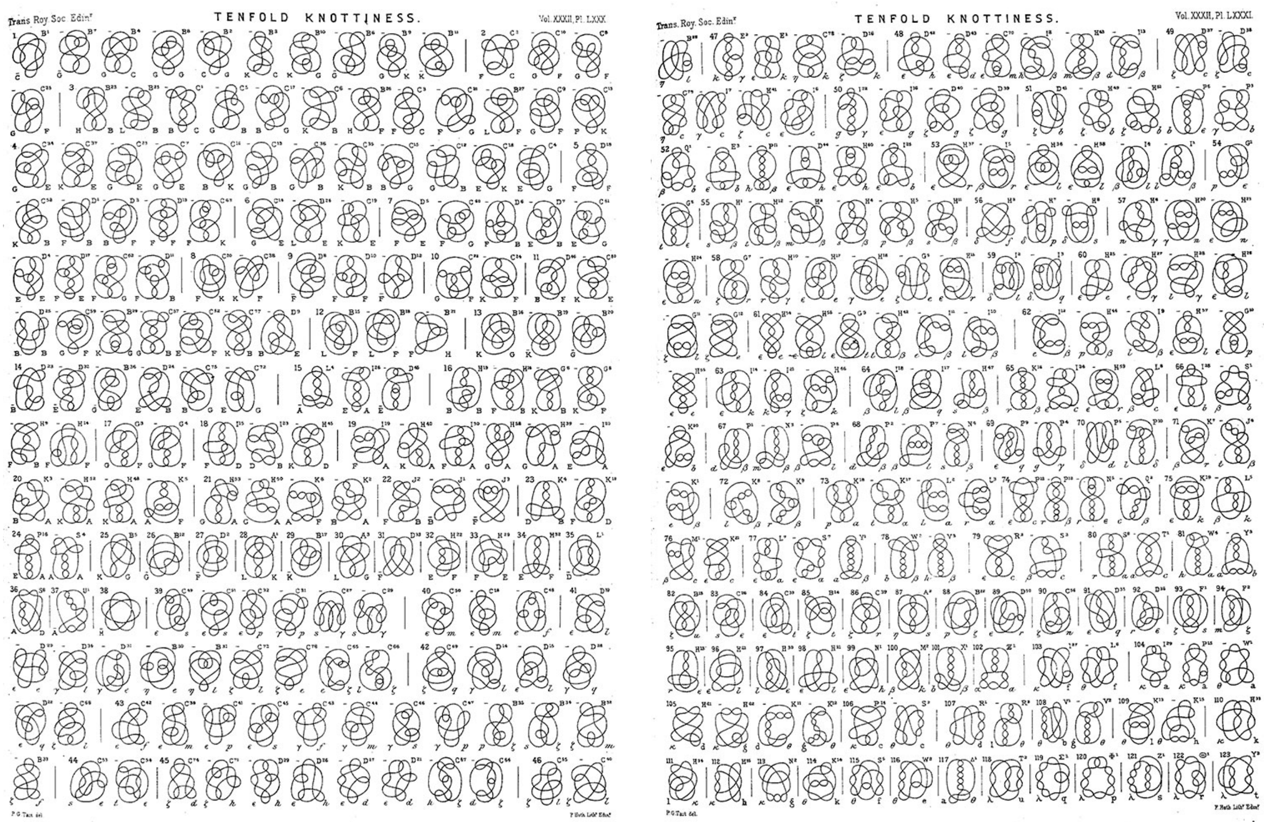


Figure 1.2: Little & Tait periodic table: knots with 10 crossings.

Tait's Conjectures

Tait's conjectures (which we will see later if true or false):

- A reduced alternating diagram has the smallest number of crossings

- An alternating knot with zero writhe is amphicheiral
- An alternating amphicheiral knot must have an even number of crossings

have been proven 100 years later they have been formulated (in 1980s by Kauffman and Jones).

There are 6 billion 22 crossing knots.

1.2 Simple Classification of Knots

Knot theory is the study of the "placement problem" [4]. In other words given the spaces X and Y, one wants to classify how X may be placed within Y. If X is S^1 (the circle) and Y is $\mathbb{R}^3$ then we have classical knot theory. The question of how to classify knots can then be reformulated into how can one distinguish different knots? There are two traditional ways to look at this problem. One studies the "embedding" $\alpha : S^1 \to \mathbb{R}^3$ (or $S^3 = \mathbb{R}^2 + \infty$) the other studies the "complementary space" $S^3 - \alpha(S^1)$. We will start with the former and then mention some techniques that can be applied to the latter.

An operational definition of a knotted curve can be given by resorting to our common experience of taking a piece of rope, tying a knot in it and finally gluing the two ends together. The knotted rope can be geometrically manipulated in three-dimensional space in countless ways but, unless we use scissors and glue to cut and rejoin the rope, it is impossible to turn it into a plain, unknotted ring. The above example builds on our intuitive notion that a rope has an approximately uniform thickness that forbids the rope self-crossing during the manipulations. Since all the configurations obtained by deforming the rope preserve the initial knot type, it appears natural to define a knot as the class of equivalence of configurations obtained by these manipulations in the three-dimensional embedding space. Going back to our initial question, how can we distinguish one knot from another? or, how can we classify knots? One popular way, dating back to Tait is to use knot diagrams.

1.2.1 Knot Diagrams

Ordinary mathematical knots in three-dimensions can be defined as embeddings of S^1 in $\mathbb{R}^3$ or inclusions in $\mathbb{R}^3$. Knotted one-dimensional curves only exist in 3 dimensions: they cannot form knots in d=1 or 2 because there are not enough degrees of freedom to tie a knot in these spaces (try it yourself with a lace flattened on your desk). At the same time, knotted 1D curves do not exist in d $\geq$ 4 because crossings can be undone in the 4-th (or higher) dimension, i.e. the crossings do not represent physical constraints.

Most commonly, knots in 3D are studied by using their projections onto a 2D surface. Hence we need a conventional way to draw these projections. The most important rules are two: (i) crossings must be denoted by over- and under-passing lines and (ii) they must be restricted to a point, i.e. the tangent of the contour at s_1 must not become parallel to that at any other $s \neq s_1$. A diagram with these two properties is called "regular".

1.2.2 Minimal Crossing Number

By smoothly deforming the three-dimensional knotted curve we start from, it is possible, in principle, to minimise the number of crossings in a knot diagram, compatibly with the topology of that knot type. This gives rise to a minimal knot diagram representation of that knot type having the smallest possible number of crossings n^{cr}_{min} (see also sec 2).

Figure 1.3: A regular crossing versus a non-regular crossing and a non-regular knot diagram versus a regular knot diagram.

Minimal crossing number is used to classify prime knots; the simplest topology has zero crossing, and is the unknot. The simplest non-trivial knot has three crossings, and is called the trefoil knot.

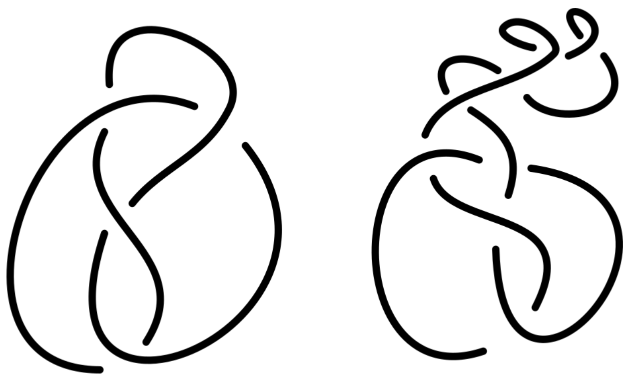


Figure 1.4: Two diagrams of the same knot with different number of crossings. The one on the left displays the minimum number of crossings.

1.2.3 Alexander-Briggs Notation

Although few knots have their own traditional names, such as “trefoil” or “figure-of-eight”, the vast majority of knots do not have a name. For this reason it is useful to assign a common notation and language that we will use when talking of knots and links.

There are many notations, but the simplest and perhaps most used is the Alexander-Briggs notation. It groups knots with same minimal crossing number, n^{cr}_{min}, and then assigns an arbitrary subscript that identifies different knots in the same family. For instance the family of 3-crossings knots has one member, the trefoil or 3_1. On the other hand, the family of 6-crossings knots has 3 members, the “Stevedore’s knot” 6_1, the 6_2 and the 6_3. Knot tables (such as the one made by Tait and Little in Fig. 1.2) are commonly used to keep track of this arbitrary numbering.

This notation can be extended for links, in which case a superscript denoting the number of components is shows, the Hopf link is 2^2_1. Composite knots are instead commonly identified with a hash, i.e. two trefoil knots tied on a rope make up a $3_1\#3_1$ knot.

The number of prime knots existing at a given value of n^{cr}_{min} can be shown to grow exponentially with n^{cr}_{min} (is there a simple proof?). This rapid growth is reflected in the fact that presently available exhaustive knot tables (minimal representations) exist only for knots with up to 16 crossings.

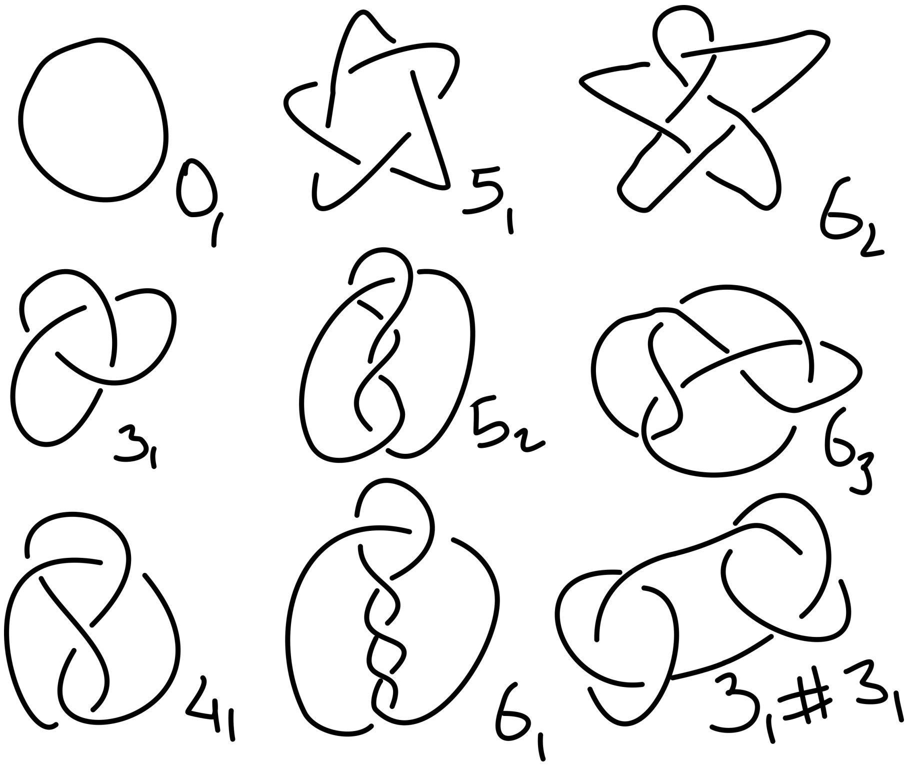


Figure 1.5: List of the first 8 prime knots up to six crossings and one non prime knot. The Alexander-Briggs notation is also shown.

1.2.4 Dowker Notation

This is another notation that can be used to distinguish knots and it descends directly from the one originally designed by Tait. Quite remarkably, it keeps track of non minimal conformations, i.e. even knot diagrams displaying a non minimal number of crossings can be reconstructed from the Dowker code.

To construct this code, consider the regular diagram of the knot and perform this sequence of steps:

- start from an arbitrary point along the contour and initialise a counter;
- every time you encounter a crossing, assign the value of the counter;
- if the counter is even, then revert its sign (from positive to negative) if you walk "over" the crossing. Do not change the sign if you walk "under" it;
- At the end you should have assigned $2n$ numbers (2 per each crossing, one even and one odd);
- the Dowker code for the knot can be compiled as the sequence of even numbers accompanying 1, 3, 5,..., $2n-1$.

As a practical example, see Fig. 1.6. From the picture one can notice that, as they are drawn, the minimal representation of the right-hand trefoil appears to have Dowker notation (-4,-6,-2) while the that of the left-hand trefoil (4,6,2). Yet, this notation cannot distinguish the chirality of knots and in fact both knots can be identified with the code (4,6,2). The simple proof of this fact is left as exercise.

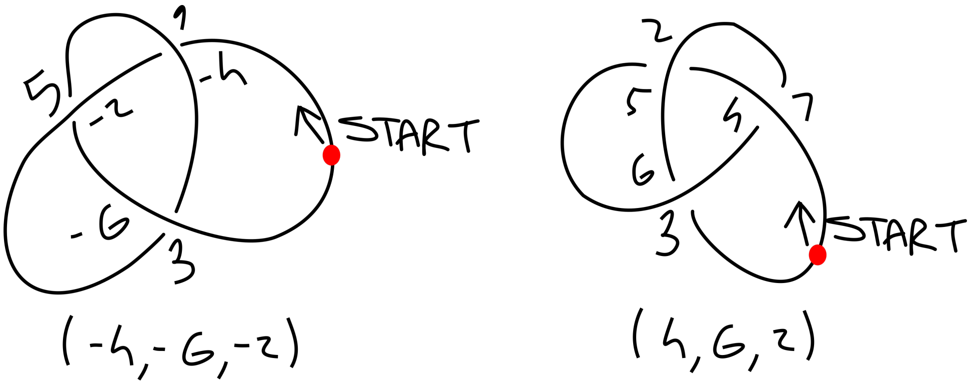
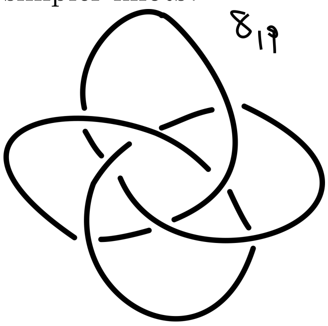


Figure 1.6: Constructing the Dowker code for the right-hand and the left-hand trefoil.

Exercise 1

(*) Prove that the Dowker code is insensitive to a knot chirality.
(**) Calculate the Dowker code for the 8_{19} knot. How does it differ from the Dowker codes of simpler knots?



1.2.5 Prime Knots

What is a prime knot? This is best defined as the opposite of a composite knot. As common experience teaches us, no knot τ can be untied by introducing another knot τ' in a portion of the closed curve. In jargon, we say that there exist no anti-knots. A further result is that the resulting knot, $\tau \# \tau'$, belongs to a knot type that is different either from τ or τ' and is called the composite knot of the two original knots. The two knots in the sum are called factors of $\tau \# \tau'$. Then, a prime knot is a non-trivial knot that cannot be decomposed as into a non-trivial connected sum. In other words, if a prime knot τ is equivalent to the connected sum $\tau' \# \tau''$ this implies that either τ' or τ'' are unknots. All standard knot tables, such as the one in Fig. 1.2, list only prime knots.

1.2.6 Chirality

An important notion in knot theory is chirality, or handedness. Suppose we look at a knot in a mirror. Is the mirror image equivalent to the original one or not ? As an example let us consider the trefoil knot and its mirror image in Fig. 1.8. The two cannot be continuously deformed into one another. For this reason, the trefoil knot is said to be topologically chiral.

For simple prime knots (which are not "alternating", see below), the handedness of chiral knots can be distinguished by computing the balance of left- and right-handed crossings in the minimal diagram. The handedness of each crossing is assigned by the right-hand rule applied to the oriented over- and under-passing segments as shown in Fig. 1.7.

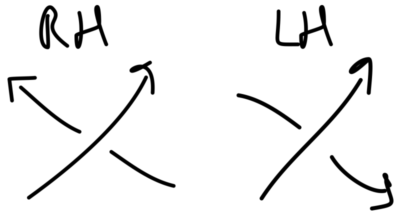


Figure 1.7: A right-hand and a left-hand crossing.

If a knot can be continuously deformed into its mirror image we say that the knot is topologically achiral, or amphichiral. The simplest amphichiral knot is the 4_1 knot in Fig.1.5 (see Ex. 2 for a proof).

As a curiosity, amphichiral knots are relatively rare. Indeed among the first 35 prime knots in the knot table less than a 1/4 are achiral (unknot, 4_1, 6_3, 8_3, 8_9, 8_{12}, 8_{17}, 8_{18}); all 49 9-crossings knots are chiral and only 13 of a total of 165 10-crossings knots are amphichiral.

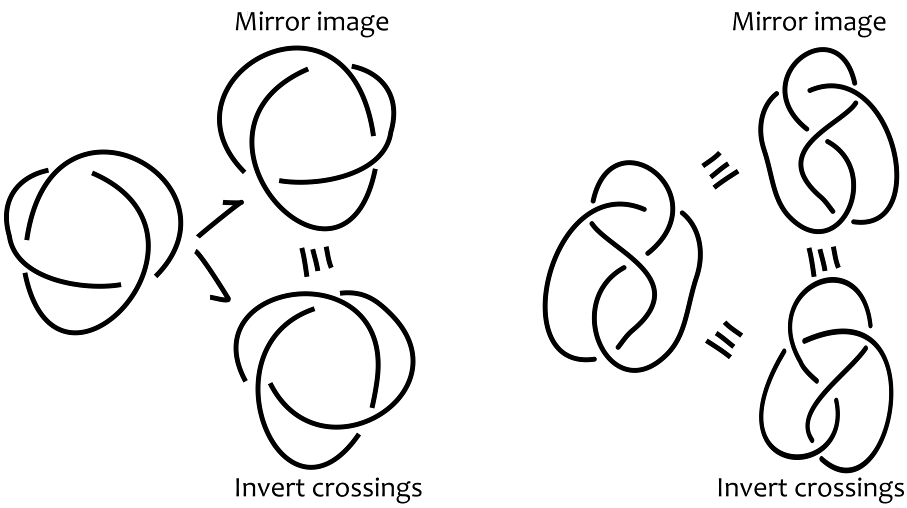


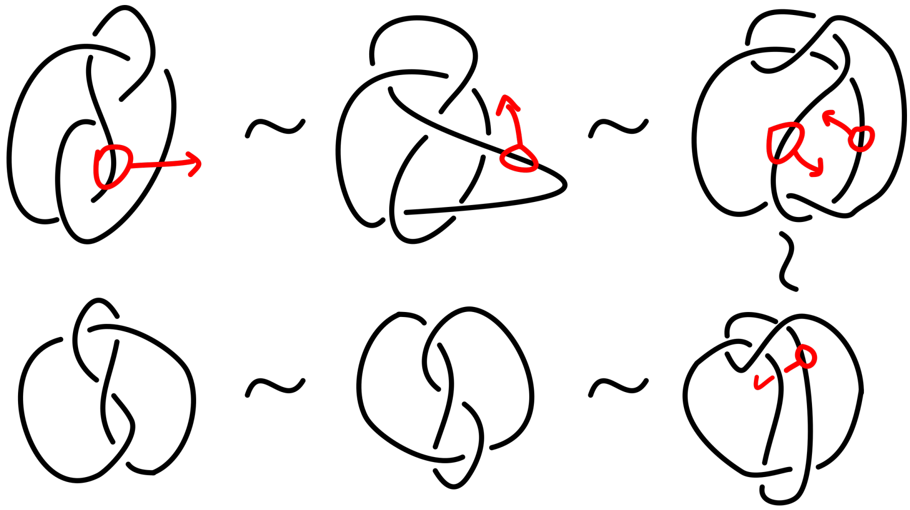
Figure 1.8: The trefoil is chiral, i.e. the mirror image of the knot is not equivalent to itself. The figure-of-eight is amphichrial, i.e. non-chiral. The mirror image of a knot is always equivalent to its image with inverted crossings.

Exercise 2

(*) Deform the unknot to show that it is not equivalent to a trefoil.

(*) Deform and rotate (but do not reflect) a trefoil to show that it cannot be deformed into its mirror image.

(**) Deform the figure-of-eight to show that it is amphichiral.



Here we have also introduced the concept of "continuous deformation". To be more precise, equivalence between knots can be defined up to homotopy, i.e. a smooth deformation of their contours which do not involve cutting the contour. We can indicate equivalence with the symbol $\sim$, as in the previous exercise. Deforming curves can be quite messy, for this reason there have been established a set of "rules" that can be methodically applied to smoothly deforming knotted curves. These are called "Reidemeister moves" and we will see them in the next Chapter.

1.2.7 Unknotting Number

Within the minimal knot diagram representation it is easy to realize that one can convert a knot into the unknot by reversing one or more crossings in the knot diagram. For example the minimal representation of the trefoil displays only 3 crossings: by reversing one of these, i.e. turning the under-passing strand into the overpassing one or vice-versa, we obtain the unknot. In general, for a given knot diagram it is always possible to find a set of crossings that can be switched over to obtain the unknot. On the other hand, for each knot diagram there can be several possible choices of crossings that can lead to the unknot. Moreover the number of crossings required might depend on the diagram. It is then natural to define the "unknotting number" as the minimum (taken over all possible knot diagrams) number of crossing reversals needed to turn a given knot into the unknot.

1.2.8 Alternating Diagrams

Prime knots can be partitioned in families according to salient topological indicators, such as the minimal number of crossings. An important one is the alternating knots, that is knots admitting a minimal diagrammatic representation where under and over crossings alternate along the path. In practice, most of the simplest types of knots, that is those with sufficiently small crossing number are alternating. In fact, all prime knots with crossing number smaller than 8 are alternating, with the simplest non-alternating instance being the 8_{19} knot.

You will remember that in an earlier exercise we have seen that the Dowker code for the 8_{19} knot is different from that of simpler knots. This is because it is non-alternating, and hence can display positive and negative numbers in the code, while alternating knots are characterised by codes with all positive (or negative) numbers.

1.3 Knot Families

There are three important families of knots that we ought to mention at this stage. These are the "torus", "twist" and "Lissajous" knots.

1.3.1 Torus Knots

With the exception of the unknot no other knot can be drawn on the surface of a sphere without self-intersections. Notably, there is a class of knots that can be drawn as a simple closed curve on the surface of a standardly embedded (i.e. unknotted) torus. This is the family of torus knots.

One of the most important properties of this family of knots is that there exists a generic analytical parametrisation for any knot belonging to this family. In general, one can exactly describe a curve that winds around the torus p times in one direction, say meridionally, and q times in the other, longitudinally (see Fig. 1.9) as

$$\begin{aligned} x(t) &= (r_0 + R\cos(pt))\cos(qt) \\ y(t) &= (r_0 + R\cos(pt))\sin(qt) \\ z(t) &= -R\sin(pt) \end{aligned}$$

where r_0 is distance from the centre of the torus to the centre of the tube forming the torus while R is the radius of the tube itself.

If $p/q \in \mathbb{Z}$ then these equations describe an unknotted curve lying on the surface of the torus whereas if p and q are prime they describe a knotted curve. To see this consider a "topological" representation of the torus: this can be drawn as a flat square or rectangle, where the sides are identified with one another as in fig. 1.9. This figure indicates that one can join the green sides to form a cylinder and then the red sides to make a torus. Viceversa, one can image to cut the torus along the red and green lines to obtain a flattened out representation. Curves that lie on the torus can also be drawn in this flat representation. In particular, p and q can now be easily visualised as the number of times these curves span the rectangle along the two directions. The identification of the sides also imply that curves that exit the rectangle at a certain point re-enter from the same point on the other side.

In Fig. 1.9 I show an example of a torus knot (the trefoil) explicitly drawn on the surface of a torus. This is the simplest torus knot and corresponds to the case with $p = 3$ and $q = 2$ (The same knot also can be drawn as $p = 2$ and $q = 3$, but it would appear different to the one in Fig. 1.9). In the figure I also show the flattened-out representation of this curve along the torus (with same colour scheme): this starts from the bottom-left corner, exits the top side at 2/3 of its length, re-enters from the bottom at the same position, exits the right side at 1/2 of its length, re-enters from the left, exits from the top at 1/3, re-enters from the bottom and finally ends up at the top-right corner, which is identified with the left-bottom one thus closing the curve.

An interesting properties of torus knots is that their unknotting number can be computed knowing the values of p and q defined above as $((p-1)(q-1))/2$. is there a simple proof?

Exercise 3

(*) Use the following code in Mathematica to visualise the trefoil drawn on the surface of a torus:

```
filledtorus[u_, t_] :=
```

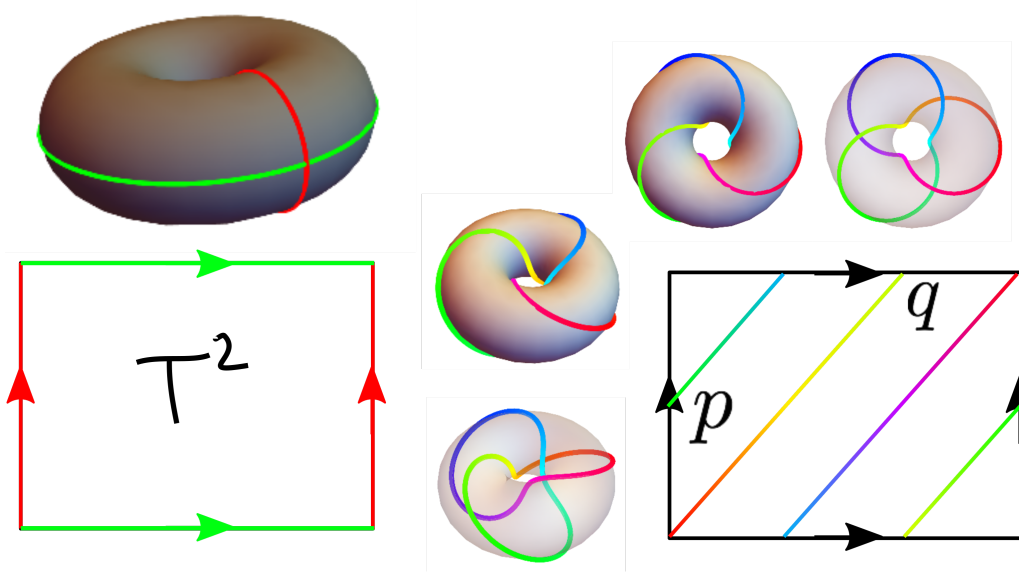


Figure 1.9: Torus knots can be drawn on the surface of a torus (T^2). If you imagine to cut a torus along the green and red circles, you will obtain a rectangle as in the figure. The green and red sides are identified with one another, as they are really the same line on the torus. Curves that go through one of the sides reappear on the other, because of the torus periodicity. Accordingly, a (p, q) torus knot can be thought as a line traversing the rectangle p times in one direction and q times in the other, and ending up at the starting point. Here I show a trefoil $(3, 2)-$torus knot drawn on the surface of a solid torus and on the flattened-out representation using the same colour scheme to aid its visualisation.

```
{Cos[t] (1.5 + 0.95*Cos[u]), Sin[t] (1.5 + 0.95*Cos[u]), Sin[u]}
pqtorusknot[t_] :=
{(1.5 + Cos[p t]) Cos[q t], (1.5 + Cos[p t]) Sin[q t], -Sin[p t]}
Show[ParametricPlot3D[filledtorus[u, t], {t, 0, 2 Pi}, {u, 0, 2 Pi},
PlotStyle -> Directive[LightGray, Opacity[1]], Mesh -> None,
Boxed -> False, Axes -> False],
ParametricPlot3D[pqtorusknot[t] /. {p -> 3, q -> 2}, {t, 0, 2 Pi},
ColorFunction -> Function[{x, y, z, t}, Hue[t]],
PlotStyle -> Directive[Black, Thickness[0.02]]]]
```

(*) Modify the parameters p and q to draw other (torus) knots.

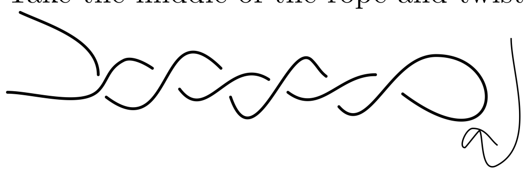
(*) Try the formula to find the unknotting number of few torus knots: $3_1 = (3, 2)$, $5_1 = (5, 2)$, $7_1 = (7, 2)$.

1.3.2 Twist Knots

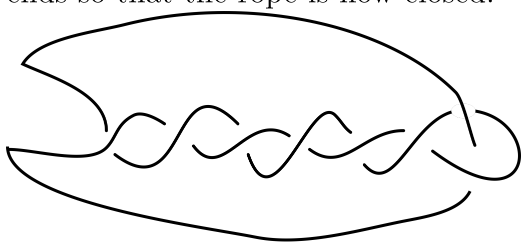
Another important family of knots is that of "twist" knots. The simplest way to define these is by describing how to tie them:

- Take a rope and hold it fixed at its two extremities while shaping a long "U";

- Take the middle of the rope and twist the whole loop, an integer number of times;



- Finally, take one of the ends and pass it through the top hoop of the rope, and link the ends so that the rope is now closed.



Increasing the number of twists you can create increasingly complex knots: trefoil (which, notably, is not only a torus, but a twist knot too), figure-of-eight (4_1), 5_2, 6_1, 7_2, 8_1 (in the figure above), etc.. One interesting property of these knots is that they all have unknotting number one: reverting one of the two crossings at the top loop is enough to undo the knot. Another interesting fact is that the only amphichiral twist knots are the unknot and the 4_1.

Although these knots have a straightforward way to be tied, they do not have equally simple analytical parametrisation. Indeed, only the figure-of-eight can be described through parametric equations, thanks to the fact that can be drawn on the surface of a punctured torus (or a torus with 2 holes). The knots that share this feature are called double torus knots.

One possible parametrisation of the 4_1 knot is

$$\begin{aligned} x(t) &= (2 + \cos(2t)) \cos(3t) \\ y(t) &= (2 + \cos(2t)) \sin(3t) \\ z(t) &= \sin(4t) \end{aligned}$$

1.3.3 Lissajous Knots

Lissajous curves in 2D have been studied in the past for their use in electronics. Their extension in 3D can describe some knotted curves. This family of knots is described by the simple equations

$$\begin{aligned} x(t) &= \cos(n_x t + \delta_x) \\ y(t) &= \cos(n_y t + \delta_y) \\ z(t) &= \cos(n_z t + \delta_z) \end{aligned}$$

where n_x, n_y and n_z are integers and the phase shifts δ_x, δ_y and δ_z may be reals. As knotted curves can never self-intersect there are constraints on the values of these 6 parameters. An important one is that the 3 integers n_x, n_y and n_z must be pairwise relatively prime.

Examples of Lissajous knots are:

- $5_2 = \{(n_x, n_y, n_z), (\delta_x, \delta_y, \delta_z)\} = \{(3, 2, 7), (0.7, 0.2, 0)\}$
- Stevedore's knot: $6_1 = \{(3, 2, 5), (1.5, 0.2, 0)\}$
- Square knot: $3_1\#3_1 = \{(3, 5, 7), (0.7, 1.0, 0)\}$
- $8_{21} = \{(3, 4, 7), (0.1, 0.7, 0)\}$

1.4 Simple Classification of Links

Although the theory of links considers generic multi-component links, we will primarily focus here on two-components links which are made by a pair of chains. Consider two simple closed curves, C_1 and C_2 that are disjoint, i.e. with no points in common, and embedded in three-dimensional space. The curves are said to be topologically linked or, more simply, linked if no smooth deformation exists by which they can be pulled apart so that they lie in two different half-spaces separated by a plane.

As mentioned before, the Alexander-Briggs notation for links is made by using the minimal crossing number with a superscript denoting the number of components and subscript indexing the member of the family. A pair of unlinked rings is denoted as 0^2_1 while the simplest link, i.e. the Hopf link, as 2^2_1.

1.4.1 Linking Number

One important number that can be used to classify knots has first been proposed by Gauss and it is the “linking number” (it is actually a “topological invariant”, see next chapter). Loosely speaking it corresponds to the effective algebraic number of times the one curve winds around the other, and viceversa. If the link can be approximated by a series of points in 3D $\boldsymbol{r}_1$ and $\boldsymbol{r}_2$ for the two components C_1 and C_2 then their linking number can be computed as

$$Lk(C_1, C_2) = \frac{1}{4\pi} \int_{C_1} \int_{C_2} \frac{(\boldsymbol{r}_1 - \boldsymbol{r}_2) \cdot (\boldsymbol{dr}_1 \times \boldsymbol{dr}_2)}{|\boldsymbol{r}_1 - \boldsymbol{r}_2|^3} = \tag{1.1}$$

$$= \frac{1}{4\pi} \int_{S^1 \times S^1} ds\, dt \frac{(\gamma_1(s) - \gamma_2(t)) \cdot (\dot{\gamma}_1(s) \times \dot{\gamma}_2(t))}{|\gamma_1(s) - \gamma_2(t)|^3} \tag{1.2}$$

where the latter equation holds if the curves C_1 and C_2 can be analytically described by parametric curves $\gamma_1(s)$ and $\gamma_2(t)$ with $s, t \in [0, 2\pi)$ and tangents $\dot{\gamma}$. [Notice that this integral formulation has strong connections to the Biot-Savart law of electromagnetism. xxx]

Alternatively, if the curves cannot be simply described analytically or approximated numerically (for instance when you draw them on a piece of paper) their linking number can be conveniently evaluated as follows:

- Choose an orientation for each curve;
- To each crossing assign +1 (right-hand) or -1 (left-hand) value according to its handedness;
- The linking number $Lk(C_1, C_2)$ is the sum of all these values divided by 2.

It is important to notice that the value of Lk does depend on how the two curves are oriented. In the case that the curves possess an intrinsic orientation based on their chemical or physical properties then it is meaningful to consider the signed value of Lk. Otherwise in abstract or general contexts, where the choice of orientation is subjective, it is more appropriate to consider only the absolute value $|Lk|$ which is independent of the curves orientation.

The linking number is a useful descriptor but suffers of serious limitations. Consider for instance Fig. 1.10. The 5^2_1 link is called the “Whitehead link” and has $|Lk| = 0$. Nonetheless, this link cannot be taken apart, i.e. it is a non-trivial link. [This link also appears in Thor’s hammer]. This ambiguity has made necessary the introduction of the concept of homologically linked curves which is used to denote two curves $C1$ and $C2$ having $|Lk(C1, C2)| = 0$. Homologically-linked curves are also referred to as being algebraically linked, in contrast to topological or geometrically linked ones. Indeed, pairs of curves that are homologically linked are guaranteed to be topologically linked but, as shown by the previous example, the converse is not true. Another example of the same type is the case of "Borromean rings", depicted as the 6_2^3 link in Fig. 1.10. This link is made of three rings, none of which is topologically linked to any other and yet, the components cannot be taken apart. On the other hand, by cutting any one of the rings one can release the other two. [This link was used as a symbol of strength and unity in the coat of arms of the noble family of Borromeo in Milan around '300, from where it takes the name.]

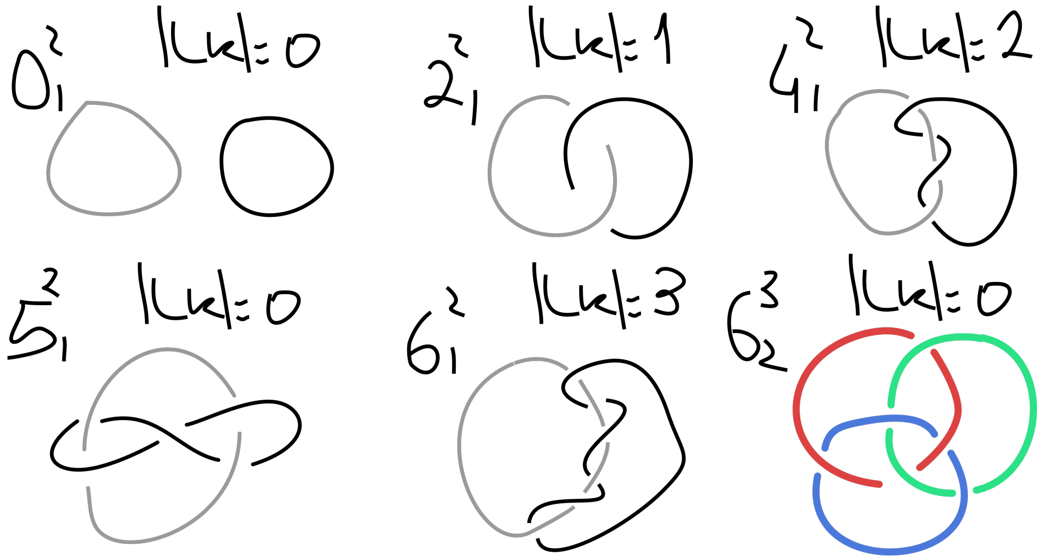


Figure 1.10: Some 2-component links and one interesting 3-component link.

1.5 Linking, Twisting and Writhing

The substance that carries our genes, i.e. the DNA, is a long double-stranded helix that can be knotted and linked [2, 1]. DNA is far from being a 1-dimensional curve. Indeed, it is more similar to a twisted ribbon. For this reason, the simple theory of knotting and linking cannot capture the full complexity of DNA topology. To this end, we need to develop a more sophisticated theory for the possible topological states of a ribbon.

The two edges of a ribbon can be thought of as two curves. When we close the ribbon to make a strip, it makes sense to ask whether these two curves are linked or not.

Strictly speaking, when we close the ribbon we need to make sure that the top edge on one end matches the top edge on the other: therefore, while the ribbon may be twisted prior to closure, it has to be twisted by an angle which is a multiple of 2π. If we were to twist the ribbon by π or an odd multiple of π before joining the ends, we would generate a non-orientable surface (so-called Möbius strip for π rotation). A non-orientable surface is a surface which cannot be unambiguously oriented. For instance, imagine to be a small ant living on the middle strip drawn in Fig. 1.11. Start walking from the red dot and circle the strip once. When you get back to the red spot you will realise that your journey has led you on the wrong side of the strip. You soon realise that the strip needs to be traversed twice in order to get back to the original tarting point. [There are famous paintings by M. C. Escher which artistically represent these types of surfaces.] You can also realise that the Möbius strip does not have to independent edges, but one edge connects into the other forming a unique contour.

Given a closed ribbon where the edges define two distinct curves, it then makes sense to introduce the concept of linking number of the curves corresponding to the two edges, say C_1 and C_2, $Lk(C_1, C_2)$. In practice, the linking number measures the number of turns n introduced

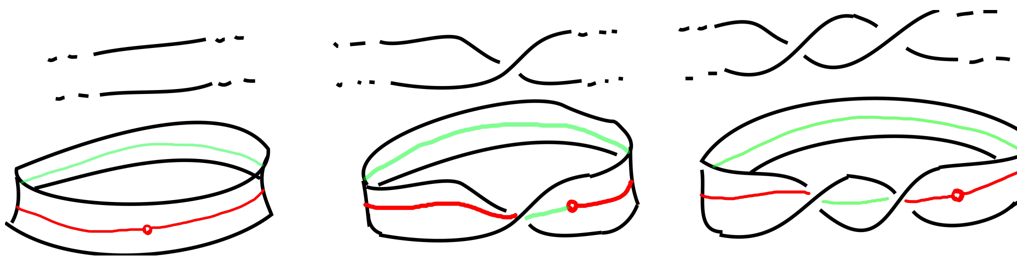


Figure 1.11: A non-twisted ribbon, a Möbius (half-twisted) strip and a once-twisted ribbon. The paths drawn on the ribbons should be understood as being red when lying on the side facing the reader while green when lying on the side hidden from the reader. This eases the visualisation of the transition between sides.

in the ribbon before the two ends are joined. In this case, the convention is to give the two edges of the ribbon the same orientation, so that it now makes sense to consider the linking number as a signed quantity, with positive linking number corresponding to a right-handed ribbon.

In the context of DNA, the linking number measures the number of double helical turns. For a DNA molecule in equilibrium, i.e. on which no stress or torque is applied, this number equals its contour length, L, divided by the double helical pitch, $p = 3.5$ nm or 10.5 base-pairs [2].

It is intuitively clear that, once the ribbon is closed, the linking number $Lk(C_1, C_2)$ is a globally conserved quantity, i.e., whatever the dynamics of the ribbon, it cannot change provided the ribbon is not cut. However, there are many possible conformations compatible with the constraint of having a fixed linking number. It is useful to characterise these conformations via two geometric quantities: the twist, Tw, and the writhe, Wr, exemplified in Fig. 1.12.

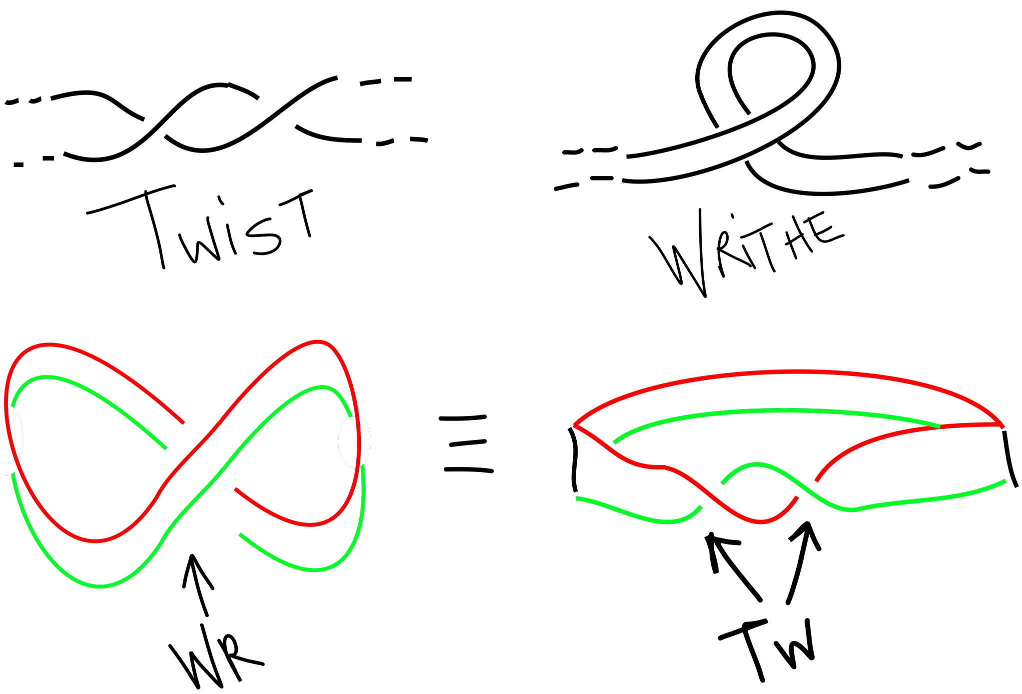


Figure 1.12: Examples of twist (local crossing of ribbon's edges) and writhe (non-local crossing of ribbon centreline). A ribbon displaying on unit of writhe (self-intersection of ribbon centreline) is topologically equivalent to a ribbon displaying one unit of twist (2π rotation or 2 half-twists). The linking number $Lk = Tw + Wr$ is conserved.

The "twist" of the ribbon tells how many full turns should be removed in order to have it flat on a plane. Whereas the "writhe" tells how many times the ribbon crosses over itself. Both these quantities have a sign: the convention is that a right-handed twist is positive, while the sign convention for the writhe is the same as for the crossings (see Fig. 1.7).

Empirically, we can convince ourselves that if we increase the linking number by twisting up the ribbon many times before closure, then a simple circular structure may not be stable and the ribbon may coil onto itself and display writhe. This coiling exchanges the local torsional stress with configurational one. In other words, the ribbon sacrifices entropy to reduce the energy stored under the form of local twist. This interplay between local stress and non-local response is peculiar of elastic curves and can be observed in DNA [2].

The observation that twist and writhe are geometrically linked leads to a very important result, known as White-Fuller-Călugăreanu's theorem [5, 3]:

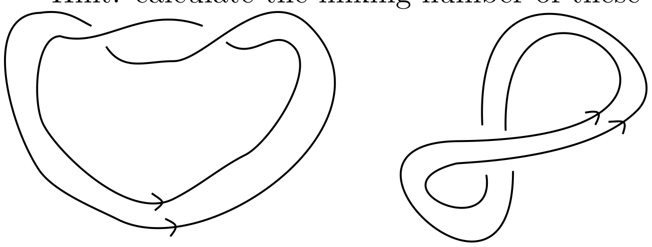
$$Lk = Tw + Wr\,. \tag{1.3}$$

Therefore, as the linking number is conserved, in possible physical conformations one can convert twist into writhe, provided that their sum is constant. This result is central for the physics of DNA, and DNA supercoiling.

Exercise 4

(*) Prove the White-Fuller-Călugăreanu's theorem using the definition of linking number Lk between two curves.

Hint: calculate the linking number of these two configurations:



Chapter 2

Topological Invariants

2.1 The Need of Invariants

How can you decide if two knots are equivalent? This is a question that has been around for a long time and no one has found a formal answer yet. The most useful resource we have to classify knots is based on "topological invariants". These are objects (in general numbers) that do not change when the contour of the knots is continuously deformed, i.e. up to isotopy.

There are several choices for topological invariants, but some are better than others. Here's a list of some commonly used ones:

- Linking number: it is the number of times one of the link components winds around the other;
- Minimal crossing number: it is the **minimum** number of self-intersections that can be seen in a regular knot diagram;
- Minmax number: it is the **minimum** number of local minima (or maxima) that can be found in a knot embedding in a fixed Cartesian direction;
- Genus: it is the number of "handles" of the minimal surface spanning a link or a knot;
- Conway Polynomial: it is constructed as a recursive algorithm on a knot diagram (see below);
- Alexander Polynomial: it is generated by constructing the associated matrix from a knot diagram (see below);
- Knot Group: it is the fundamental group of the knot complement (see next chapter);
- Braid Group: it is the group of the braid associated with a given knot (see next chapter).

In this Chapter we will see how to construct more sophisticated invariants with respect to the ones used in the previous chapter. One key point will remain valid though, that two knots can be classified as equivalent if they can continuously deformed into one another. So, let's first define some more precise ways to continuously deform knot diagrams.

2.2 Reidemeister Moves

In the previous Chapter we have seen some cases in which it was required to deform knotted curves, for instance to see whether they looked like their mirror image. In those cases we did not follow precise indications on how to deform them and this may have led to some confusion! We now introduce some prescriptions on how to formally deform knotted diagrams so that we can find out whether two knots are equivalent by checking that they look the same after having performed a well-defined series of moves, called "Reidemeister moves". These moves are the essential ones to produce all kinds of ambient isotopies between knot diagrams. [An "ambient isotopy" is a continuous deformation of the knot contour within the embedding space, i.e. "ambient".] Hence, two knot diagrams that differ by a number of Reidemeister moves are equivalent, or "ambient isotopic".

The Reidemeister moves are three:

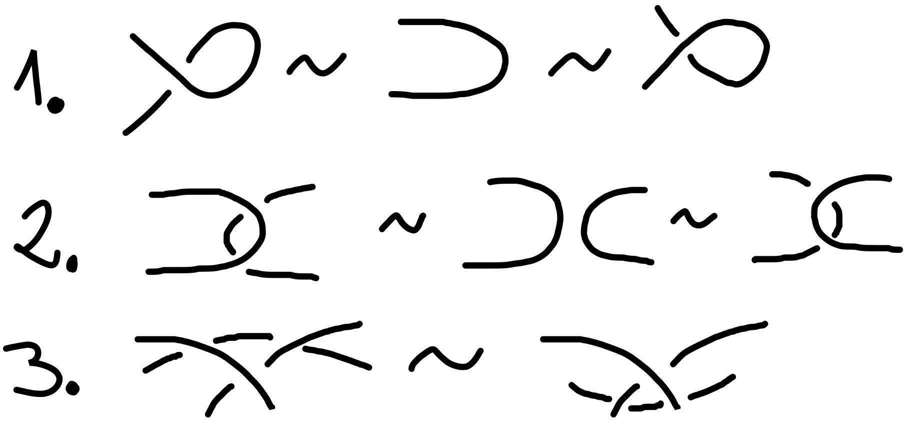
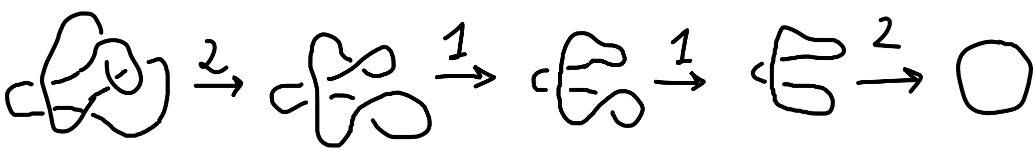


Figure 2.1: Reidemeister moves.

Exercise 5

(*) Simplify the following diagram using Reidemeister moves.



(**) Simplify the following diagram using Reidemeister moves.

(*) Prove that the following process will always lead to unknots: start drawing; every time you cross a previously drawn segment, underpass it; repeat until you return to the start.

(**) Two equivalent links have same linking number Lk. This means that Lk is invariant under ambient isotopy. Show it using Reidemeister moves.

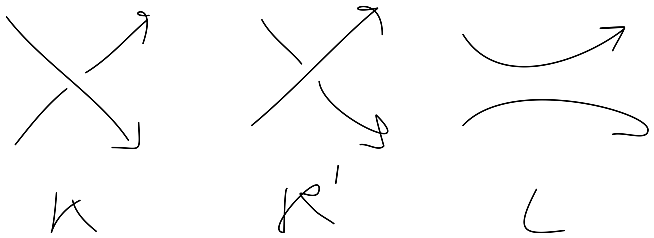
2.3 Conway Polynomial

In order to introduce the concept of polynomial invariant we need to start from some axioms. The first two are general to all polynomial invariants, i.e. equivalent knots give identical polynomials and that there is an "identity" element (the unknot). The third axiom is more specific to the polynomial under consideration.

Axiom 1: For each knot and link K there exist a corresponding polynomial $C(K)$. Equivalent knots (or links) are given identical polynomials, i.e. $K \sim K' \Rightarrow C(K) = C(K')$.

Axiom 2: The trivial knot is associated with the "identity" $C_K = 1$.

Axiom 3: Suppose that 3 knots or links differ at the site of one crossing as:



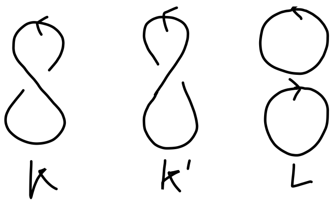
then

$$C(K) - C(K') = zC(L)\,. \tag{2.1}$$

The last axiom defines a relationship between knot diagrams in which one of the crossings is changed and the associated polynomials. By iterating this last relationships we end up either to the trivial knot or to cases for which we know the associated polynomial (see exercise below). At the end of this procedure we obtain a polynomial of the form $C_K = a_0 + a_1 z + a_2 z^2 + \ldots$. The coefficients $a_n \in \mathbb{Z}$ have well-defined physical meaning that we will discover later.

Exercise 6

(*)Show that any split link (i.e. a link that can be drawn as disjoint components) has zero Conway polynomial.

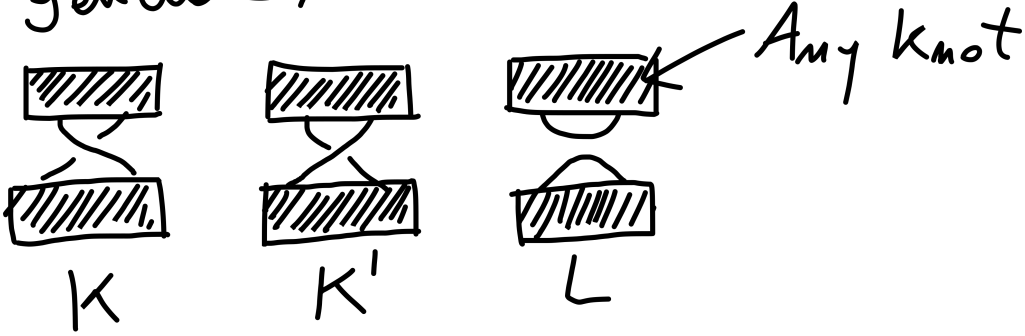


$\Rightarrow C(K) - C(K') = zC(L)$

but $K \sim K'$ $\Rightarrow C(K) = C(K')$

$\boxed{C(L) = 0}$

In general:

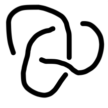
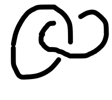


(*)Find the Conway polynomial of the trefoil.

$C(K') = 1$

$C(L') = 0$ $C(W) = 1$

Hence:

C() = 1 + C() z

$= 1 + z(0 + z)$

$= 1 + z^2$

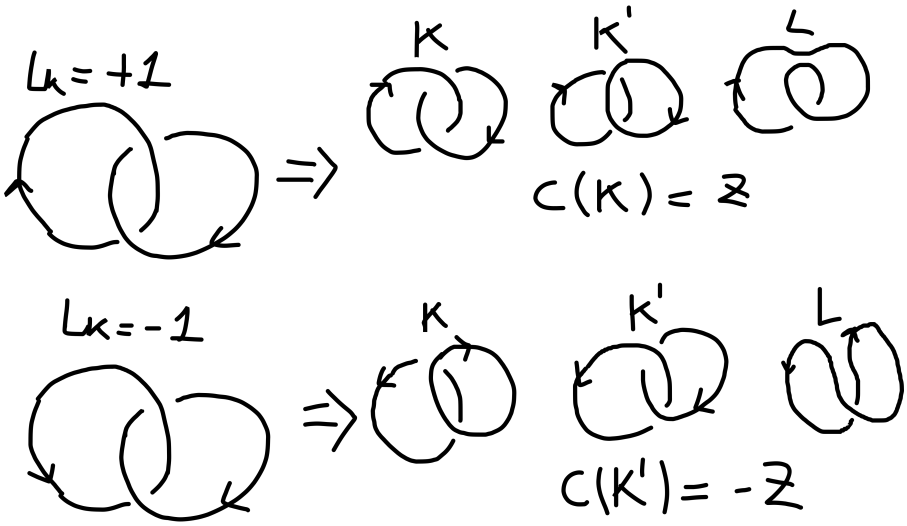
(*)Find the Conway polynomial of the figure-of-eight.

2.3.1 Why Polynomials Are Better

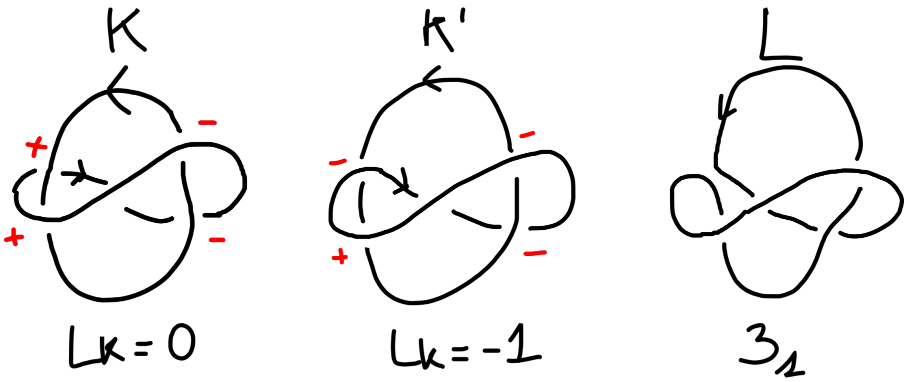
Up to here we had looked at examples which do not reflect the true power of polynomial invariants. For instance, the minimal crossing number can already capture the knottedness of the trefoil. Now, we will have a look at two cases in which the Conway polynomial outperforms other invariants.

Exercise 7

(*) Find the Conway polynomial for the positive and negative Hopf link.

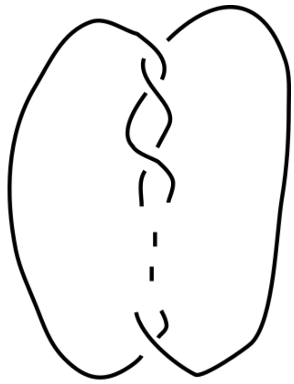


(*) Find the Conway polynomial for the Whitehead link.



$$C(K) = zC(3_1) + C(2^2_1) = z(1+z^2) - z = z^3$$

(*) Find the Conway polynomial for the the family of 2n-crossings link:



So, we have seen that the Conway polynomial can detect handedness of the Hopf link (but not of the trefoil, try it!) and can even detect the linkedness of the Whitehead link.

2.3.2 Meaning of the Polynomial Coefficients

With the examples we have examined we can then see that a_0 is 0 if the link has more than one component and 1 if it is a knot (1 component). At the same time a_1 is the linking number between the components (if there are 2) and it is 0 otherwise (see exercise in the previous section). By computing the Conway polynomial for the trefoil one can see that a_2 is a sort of self-linking from splicing crossings on K. One can iterate this reasoning:

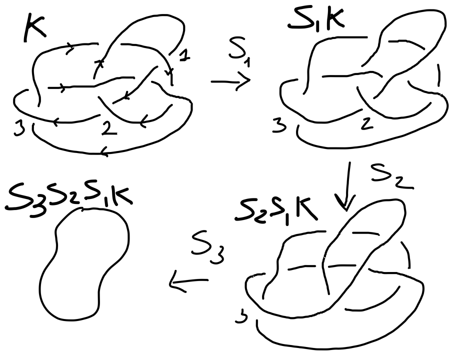
$$a_0(K) - a_0(\bar{K}) = 0 \tag{2.2}$$
$$a_1(K) - a_1(\bar{K}) = a_0(L) \tag{2.3}$$
$$a_2(K) - a_2(\bar{K}) = a_1(L) \tag{2.4}$$
$$\ldots \tag{2.5}$$

In general:

$$a_{n+1}(K) - a_{n+1}(S_iK) = \epsilon_i(K)a_n(E_iK) \tag{2.6}$$

where a_n is the coefficient of the term z^n in the polynomial, ϵ_i is the sign of $i-$th crossing, S_i denotes switching crossing i and E_i denotes eliminating crossing i.

As last exercise, calculate the a_2 of the following knot K:



Hence:

$$\partial_2(K) - \partial_2(S_1K) = \varepsilon_1\partial_1(E_1K) \qquad (1)$$
$$\partial_2(S_1K) - \partial_2(S_2S_1K) = \varepsilon_2\partial_1(E_2S_1K) \qquad (2)$$
$$\partial_2(S_2S_1K) - \partial_2(S_3S_2S_1K) = \varepsilon_3\partial_1(E_3S_2S_1K) \qquad (3)$$

$$\boxed{\text{With } \varepsilon_1=-1,\ \varepsilon_2=+1,\ \varepsilon_3=-1 \text{ and } \partial_1 = lk}$$

$$\Downarrow$$

$$(1)+(2)+(3)$$
$$\|$$
$$\partial_2(K) - \partial_2(\underbrace{S_3S_2S_1K}_{\text{unknot}})$$
$$\|$$
$$-lk(E_1K) + lk(E_2S_1K) - lk(E_3S_2S_1K)$$

$$lk(E_1K)=0 \qquad lk(E_2S_1K)=1 \qquad lk(E_3S_2S_1K)=0$$

$$\boxed{\partial_2(K)=1}$$

2.4 Alexander Polynomial

One of the best known knot polynomial invariant is the Alexander polynomial. The latter is defined in terms of a single variable, t and is computed, starting from a given diagram according to the following general prescriptions:

1. attach an orientation to the diagram and establish the sign of each crossing using the right-hand rule, as in Fig. ??
2. assign a progressive numbering index to the n arcs of the diagrams and (separately) to the n crossings.
3. define an $(n \times n)$ matrix M. The elements of the x-th row of M are calculated by considering the x-th crossing in the diagram and the three arcs, i, j and k taking part to the crossing. For definiteness we shall assume that the i-th arc passes over arcs j and k. All elements of the x-th row of M are set to zero except for $M(x,i)$, $M(x,j)$, $M(x,k)$. These three entries are calculated as follows:

   (a) if the crossing x is positive then $M(x,i) = 1-t$, $M(x,j) = -1$ and $M(x,k) = t$,

   (b) if the crossing x is negative then set $M(x,i) = 1-t$, $M(x,j) = t$ and $M(x,k) = -1$.

Iterating the procedure for all crossings the matrix is completely defined.

Deleting any one of these columns and any one row yields a $(n-1)\times(n-1)$ matrix. This is the Alexander matrix associated to a given diagram. The determinant of the Alexander matrix (which is therefore a minor of M) is the desired Alexander polynomial, $\Delta(t)$. Strictly speaking, the Alexander polynomial is not uniquely defined for a given knot type. This is evident because the size of the matrix, M, and hence the determinant, depend of the number of crossings and hence on the details of a given diagrammatic representation. It turns out that the Alexander polynomials of two different diagrammatic representations of the same knot type can differ only by a multiple of $\pm tm$, with $m \in \mathbb{Z}$. This ambiguity is immaterial if one computes $|\Delta(-1)|$, which is indeed a topological invariant, and is normally computed in applications.

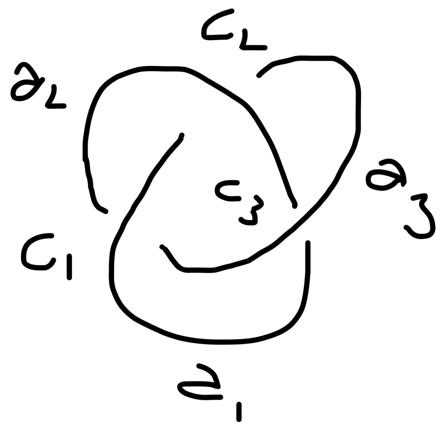


Figure 2.2: A possible labelling of crossings (c1, c2, c3), and arcs (a1, a2, a3) for the knot diagram of the trefoil to compute its Alexander polynomial.

As an exercise, we can compute the Alexander polynomial of a trefoil (see Fig. 2.2 for a possible labelling of crossings and arcs). The Alexander matrix is

$$\begin{pmatrix} 1-t & t & -1 \\ -1 & 1-t & t \\ t & -1 & 1-t \end{pmatrix} \tag{2.7}$$

and the corresponding Alexander polynomial is $\Delta(t) = t^2 - t + 1$, so that the invariant quantity $|\Delta(-1)| = 3$ for the trefoil knot.

Chapter 3

Knot Groups

In this chapter we will introduce the concept of group and see how it applies to knot theory.

3.1 A brief recap of groups

In mathematics, a group is an algebraic structure equipped with an operation that can associate any two elements to form a third, together with identity, inversion, closure and associativity. Well known examples of groups are the integers with the addition operation and symmetry or rotations in d-dimensions.

More specifically a set G, together with an operation $\cdot$ that combines two elements of G, is a group, denoted by $(G, \cdot)$ if

(closure) if $a, b \in G$ then $a \cdot b \in G$;

(associativity) $a \cdot (b \cdot c) = (a \cdot b) \cdot c$;

(identity) $a \cdot id = a$;

(inversion) $a \cdot a^{-1} = id$.

Groups in which the operation is commutative are called Abelian and Non-Abelian otherwise.

Exercise 8

(*) Show that $\mathbb{Z}$ with addition is a group.

3.1.1 Dihedral groups

A useful example of groups are the symmetry groups. A specific example of these are the finite, dihedral groups. These are composed by polygonal shapes, together with symmetry operations. The notation is that for a n-sided polygon, the associated group with 2n operations is denoted D_n.

Dihedral group D_2 It is the dihedral group of an planar object that has 2 sides. It is also called the Klein 4-group and it is associated to liquid crystals and biaxial nematics. D_2 is Abelian.

Dihedral group D_4 This group is that of an object with 4 sides (a square). For a square there are 8 possible operations that leave the square unchanged, 4 rotations (0,$\pi/2$, π, $3\pi/4$) and 4 reflections (up-down, left-right and 2 diagonals). Compositions of these operations give a new element which is identical to the original square, i.e. satisfying closure. The group is not Abelian (as one can notice by depicting a non-symmetrical face within the square) draw pic.

3.2 The representation of groups

A group can be represented by the set S of generators of the group and the relations (if any) between them. The generators are "minimal words" that can be used to compose all the elements of a group. For instance, the group of discrete rotations can be composed by powers of one element, ρ being the smallest rotation. Finite dihedral groups are composed by the generators of rotations ρ and reflections σ. In these cases, a certain number of rotations are equivalent to no rotation at all, hence there exists a relation $\rho^n = 1$, that defines the limits of the group actions. Take D_4 as an example, the element ρ describing a $\pi/2$ rotation can be iterated to form the others ρ^2, ρ^3 but $\rho^4 = 1$, no rotation at all. The same holds for reflections, σ for which $\sigma^2 = 1$. Hence its representation is

$$D_4 = \langle \sigma, \rho | \rho^4 = \sigma^2 = 1, (\sigma\rho)^2 = 1 \rangle \, , \tag{3.1}$$

the last relation is a combination of generators which independently lead to the identity.

Exercise 9

(*) Try it to see if the relations of D_4 work

3.3 The fundamental group

Given a topological space X, its fundamental group ($\pi_1(X, x)$) is generated by all classes of loops based at x.

To visualise this group, imagine to be embedded in a space, say $\mathbb{R}^3$ to start with, and image to draw a point x somewhere. Now image to draw many loops that start from x and come back to x after meandering in $\mathbb{R}^3$. Now ask yourself the following question: are all these loops equivalent or can they be divided into distinct classes? It turns out that for $\mathbb{R}^3$, all loops you can draw are equivalent and are null-homotopic. Hence, the fundamental group of $\mathbb{R}^3$ is trivial, i.e. the identity.

Imagine now to sit on S^1 (a circle), decide a point x and draw all loops you can draw. The simplest one is the loop that circles the circle once. The second simplest is the one that circles it twice. The third .... Then, can all these loops be deformed one into the other? The answer is no! Because the circle is a 1D line in a 2D space, there is no room for deforming the loops one into the other. Hence, these loops you have drawn all belong to distinct classes. The fundamental group of the topological space S^1 is homeomorphic to a group that we know very well (see Section 3.1)! Imagine to draw a loop that winds n times the origin (call it a), and then one that winds it m times (call it b) the composition of $a \cdot b = c$ is a loop that winds $n+m$ times, i.e. $\pi_1(S^1, x) = (\mathbb{Z}, +)$.

Theorem 1 *If the topological space is path-connected, then $\pi_1(X, x) \equiv \pi_1(X)$.*

Theorem 2 *If the topological space is simply connected, then $\pi_1(X, x) \equiv Id$.*

Exercise 10

(**) Explain why the above statements are true

The distinct classes of loops form the elements of the group. The elements can be represented by the generators. For instance, $\mathbb{R}^3$ has trivial generators, S^1 has one generator that produces all possible loops.

3.4 The fundamental group of the knot complement

First, we introduce the concept of the complement of a knot. Consider a curve $\mathcal{K}$ embedded in $D = \mathbb{R}^3$ or $D = S^3 = \mathbb{R}^3 + \infty$, its complement is defined as $D - \mathcal{K}$ (the embedding space minus the knotted curve). Then, we define the group of a knot as the fundamental group of the knot complement. Then we can state that

Theorem 3 *The groups of equivalent knots are equal.*

One can use the complement of a knot to determine whether two knots are the same knot. In other words, the knot complement is, up to homeomorhpisms, a **topological invariant**.

Exercise 11

(***) Show that link complement is **not** a topological invariant

---

3.4.1 The Wirtinger Representation

In order to compute the group of knotted curves it is often useful (bur not necessary) to start from the Writinger representation. This can be constructed in the following way:

1. Draw a regular knot diagram and orient the curve;
2. Label the arcs with α_i, $i = 1, ...$;
3. Draw an arrow passing under each arc in a right-handed fashion. This indicates a section of the loop starting from the eye of a distant observer (you) and passing under the arc according to the arrow direction and coming back to the eye.
4. Label the arrows x_i, $i = 1, ...$;
5. At each crossing of the knot diagram, there are two intersecting arcs and three labels: α_i, α_{i+1} (for the arc passing under) and α_k for the arc passing over. Thus, there are three arrows, x_i, x_{i+1} and x_k. Depending on the sign of the crossing there must be one of the two relations:
   - $x_k x_i = x_{i+1} x_k$ if positive
   - $x_i x_k = x_k x_{i+1}$ if negative

6. For a diagram with n crossings there are n relations with n unknowns, hence the system is fully determined. Because we only need relations, it is enough to analyse $n-1$ crossings and then obtain the minimal generators and relations to represent the group.

Example: Unknot

Draw an uknot and all loops with base point in $\mathbb{R}^3$. It is clear that it is equivalent to S^1.

Example: trefoil

[draw] Relations:

- $x_1x_2 = x_2x_3$
- $x_1x_2 = x_3x_1$
- $x_3x_1 = x_2x_3$

Use two of these relations to get

$$\pi_1(3_1) = \langle x_1, x_2 | x_1x_2x_1 = x_2x_1x_2 \rangle \tag{3.2}$$

Exercise 12

(*) Show that this is equivalent to

$$\pi_1(3_1) = \langle a, b | a^2 = b^3 \rangle \tag{3.3}$$

Exercise 13

(***) Show this group

$$\pi_1(3_1) = \langle a, b | a^2 = b^3 \rangle \tag{3.4}$$

is the one for the trefoil by using the fact that it is a torus knot. In other words, draw it on the surface of a torus and then study the topological space $T^2 - 3_1$, how can you describe it via generators?

Exercise 14

(**) Generalise the finding of the previous exercise to $T_{p.q}$ torus knots.

Example: figure-of-eight

[draw] Relations:

- $x_1x_4 = x_3x_1$
- $x_3x_2 = x_1x_3$
- $x_1x_2 = x_2x_4$
- $x_3x_4 = x_4x_2$

Use three of these relations to get

$$\pi_1(4_1) = \langle x_1, x_3 | x_1^{-1} x_3 x_1 x_3^{-1} x_1 x_3 = x_3 x_1^{-1} x_3 x_1 \rangle \tag{3.5}$$

Exercise 15

(**) [draw] The square knot and the granny knot are **not** equivalent. Show that the fundamental group the granny and the square knot are the same.

Easy way:
The knot complement does not capture handedness. Hence, because
square=$3_1^r \# 3_1^l$
granny=$3_1^r \# 3_1^r$
the two fundamental groups must be the same.

Difficult way:
Compute and compare the knot groups.

$$\pi_1(square) = \langle x, y, z, w | xyx = yxy, wzw = zwz, x = w \rangle \tag{3.6}$$

$$\pi_1(granny) = \langle x, y, z, t | xyx = yxy, tzt = ztz, t = x \rangle \tag{3.7}$$

Theorem 4 *The fundamental group of the knot complement is not a complete topological invariant.*

Exercise 16

(**) [draw]
Compute the knot group of twist knots with n crossings.
[xx todo xx]

Chapter 4

Seifert Surfaces

Definition: A **Seifert surface** of a knot or link K is a connected, orientable, compact 2-manifold M with $\partial M = K$, i.e. its boundary is the knot or link K.

Theorem 5 *Every closed orientable connected 2-manifold is homeomorphic to one of the ones in this table and it is classified by its genus $g \geq 0$ (or Euler characteristic χ).*

Manifold	S^2	T^2	$T^2 \# T^2$	...	$T^2 \# \ldots \# T^2$
genus	*0*	*1*	*2*	...	g
χ	*-2*	*0*	*2*	...	*2-2g*

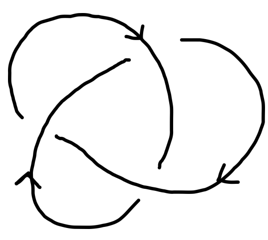
The Euler characteristic can be computed from the simplicial representation of the manifold. The genus is effectively the number of "holes" in the manifold.

Definition: The genus of a knot or link is the least genus of all possible Seifert surfaces.

Theorem 6 *Every knot or link is the boundary of a Seifert surface.*

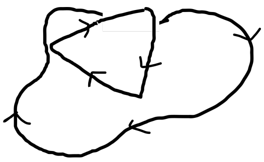
The proof of the theorem is the actual construction of such a surface for a generic knot or link. In order to construct it, one follows these steps:

- Draw a regular knot (or link) diagram

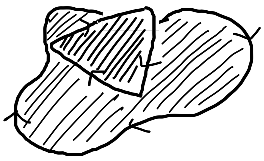


- "reconnect" all crossings according to the orientation of the crossing arcs. The resulting curves form disjoint discs.

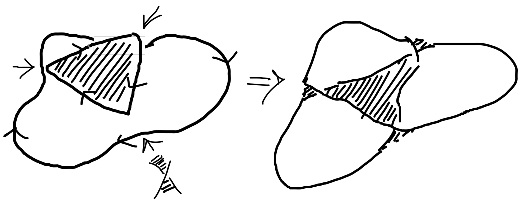
- Push the nested discs upward to make them stack like a wedding cake



- Orient the discs according to the direction of the boundary curve

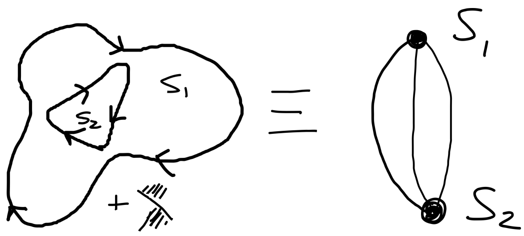


- Replace the crossings with half-twists joining the discs



At the end of this procedure, you have created an **oriented** surface whose boundary is the knot or link.

The disjoint discs created in this procedure are called Seifert circles. One useful way to compute the genus of a surface is by first computing the Euler characteristic of the corresponding simplicial complex. This is done by noticing that, topologically, the Seifert construction generates discs connected by strips, which can shrunk to points connected by edges.



Then the Euler characterstic χ can be simply found as $\chi = \#(points) - \#(edges)$ and the genus can then be calculated reminding that $\chi = 2 - 2g - k$ ($g = 1 - (\chi + k)/2$) and k is the number of components of the knot or link), hence

$$g = 1 - \frac{\chi + b}{2} = 1 - \frac{s - c + b}{2} \tag{4.1}$$

where s is the number of Seifert circles (points), c the number of crossings (edges) and b the number of boundary components.

Exercise 17

(*)Calculate the genus of the Hopf Link.

(*) Calculate the genus of the following link.

(**) Show that the genus of a torus knot is $geq (p - 1)(q - 1)/2$.

Chapter 5

Minimal Surfaces

What is a minimal surface

Surface bounded by a boundary whose mean curvature is zero.

Surface Tension

At a liquid-air interface, the liquid molecules experience a larger attraction to other liquid molecules, due to cohesion, than air molecules. This means that the net force on a molecule at the boundary of a film is inward, towards the bulk. While a molecule in the bulk does not feel a net force (see Fig. 5.1). Thus the surface of a liquid can be seen as an elastic layer, which is under tension due to imbalanced forces. Chemical details are important to determine the surface tension of liquids. For instance water has a high surface tension due to hydrogen bonds, whereas soapy liquids display a much lower surface tension. For this reason, any given body of liquid tries to minimises its surface, and hence form droplets rather than streams of fluids.

The Young-Laplace equation describes the difference in pressure between two fluids when there is a thin wall with surface tension γ and mean curvature H in the middle:

$$\Delta p = -\gamma \nabla \cdot \hat{n} = 2\gamma H = \gamma \left(\frac{1}{R_1} + \frac{1}{R_2} \right) \tag{5.1}$$

A minimal surface can be defined through the Young-Laplace equation by saying that it is a surface whose pressure difference is zero.

The Plateau–Rayleigh instability (that for which any stream of liquid is destined to break up into droplets, or beads of liquid) is driven by surface tension.

Why is a soap film a minimal surface?

Soap films are fluids suspended from a solid frame. They are thin, typically a few hundred nanometres thick whilst spanning a frame with lateral dimensions of several centimetres, so that surface effects dominate many of their properties, both static and dynamic. Surface tension causes soap films to minimise their surface area, leading to the well-known characterisation of their shapes as minimal surfaces. This also accounts for a striking feature of soap film dynamics, namely that they exhibit topology changing transitions, triggered by a fundamental instability of the area functional. The canonical example of this is the collapse of the catenoid minimal surface, spanning a frame of two round circles, to a pair of disconnected discs.

Two of the most striking features of soap films are their beautiful morphology, and the fact that they are inherently unstable. Both of these may be understood on the basis that their

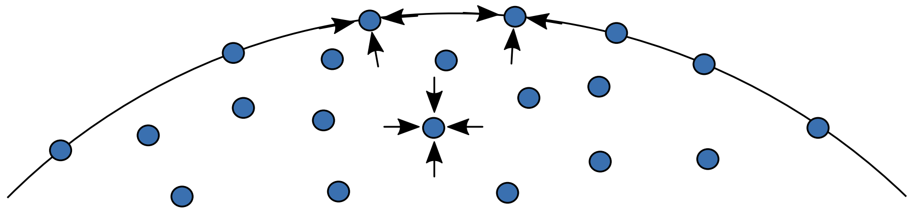


Figure 5.1: surface tension

energetics are dominated by interfacial tension so that an appropriate free energy is simply the total surface area of the soap film (or of the surface approximating it). The shape of the film is then given by the condition that it be a critical point of the area functional, while the issue of (in)stability comes from the Hessian at those critical points

By the Young-Laplace eq. the difference in pressure between the two sides of a soap film has to be zero. This turns out to be eual to the mean curvature $H = (k_1 + k_2)/2$. So soap films and minimal surfaces have zero mean curvature. This means that they locally look like saddles everywhere.

Minimal Surfaces in Linked Rings

Perhaps the simplest example of two surfaces that are homeomorphic but not isotopic are the two distinct soap films spanning the Hopf link (n = 2) frame, which differ in the relative linking number that they induce between the two boundary circles

Instabilities and Topological Transitions

Such instabilities and topological transitions are generic; a well-known theorem in the mathematics literature states that the only stable complete minimal surface embedded in R3 is the plane.

The transition from a Moebius strip to a disc involves a singularity that occurs at the boundary of the soap film [xx], whereas the singularity of the catenoid occurs in the bulk.

Chapter 6

Surface Evolver

6.1 What is Surface Evolver

Surface evolver is a software developed by K. Brakke (Susquehanna university) to study the shape of surfaces driven by surface tension and other forces. The simplest algorithms implemented in the software numerically minimise the free energy of the system by steep gradient method. As seen in the previous chapter, any surface holds tension and hence energy. Minimising the free energy in the case that only surface tension is acting means minimising the area of the surface, given its boundary. This is done by triangulating the surface and attempting moves which reduce the number of triangles.

The software is highly flexible especially in terms of the surface boundaries. Any well-defined boundary, **even knotted or linked ones**, are handled by the evolver (see examples in following sections). In addition, volume, surface or boundary constraints can be added.

The Evolver can be used, and has been used, to obtain insights into a number of problems: shape of cell membranes[xx], soap films[xx], instabilities of minimal surfaces[xx], threadings in solutions of ring polymers [xx], etc.

Installation

Surface Evolver is available for Windows, Linux and MacOS. Detailed informations on how to install are given on Brakke's web page: `http://www.susqu.edu/brakke/evolver/evolver.html`. Make sure to acknowledge his effort by citing his work every time you will use surface evolver in the future!

6.2 Basics

The basic elements used to define a surface are vertices, edges, facets and bodies. Vertices are points in Euclidean space. Edges are straight lines joining points. Facets are **flat** triangles bounded by three edges. A surface is a union of facets. A body is defined by its bounding surface.

Importantly, there are no limitations on how many edges are originating from a vertex or on how many facets are bound by one edge. This means that arbitrary topologies are possible, including triple junctions in surfaces of soap films.

There is no restrictions on orientation of edges and facets, thus non-orientable surfaces are possible (e.g. Moebius strip).

A surface that is constructed as the union of facets has an area and thus an energy arising from the various contributions chosen by the user: surface tension, gravity, constraints, etc.

In order to start the evolver, one needs an initial datafile `data.fe` which contain the list of faces, edges and vertices (see below).

One an initial surface is defined, the key commands to obtain the final result are:

g n: performs n iterations to lower the free energy

r: refines triangulation

s: shows surface on screen

P: graphics output (print on file)

q: quit

Every time an iteration is performed, the new are is calculated as follows: first, the force on each vertex is calculated from the gradient of the total energy as a function of the position of that vertex. Second, the force gives the direction of motion unless constraints are applied, in this case the force is made to conform to the constraint. Finally, the motion of the vertex is found by multiplying the force by a global scale factor.

6.3 Working Examples

6.3.1 Soap Cube into a Soap Bubble

We are all familiar with soap bubbles. Because they are made of soap (as the name suggests) they are shaped by surface tension. At this point the eagle-eyed diligent student should raise his/her hand and ask whether there is some problem with the maths, as we showed in the previous chapter that minimal surfaces have zero mean curvature everywhere, while a sphere of radius R has principal curvatures equal to $1/R$ everywhere or mean curvature $2/R$. The problem is that for a freely floating soap bubble there is no boundary and hence the spherical shape is the only one that minimises its finite area.

In this section we will make a soap bubble starting from a soap cube. Let's start by defining the cube using vertices, edges and faces. Start with the vertices. They do not necessarily have to appear in any order. One can just list them as:

```
vertices
1 0.0 0.0 0.0
2 1.0 0.0 0.0
3 1.0 1.0 0.0
4 0.0 1.0 0.0
5 0.0 0.0 1.0
6 1.0 0.0 1.0
7 1.0 1.0 1.0
8 0.0 1.0 1.0
```

Now the edges are defined using the "id" of the vertices and one has to list the source and target vertices for each new edge as:

```
edges
1 1 2
2 2 3
3 3 4
4 4 1
```

```
5 5 6
6 6 7
7 7 8
8 8 5
9 1 5
10 2 6
11 3 7
12 4 8
```

Finally, using the edges, one can define the faces, which are defined as the surface spanned by a closed loop of edges. The orientation of the face is given by the orientation of the loop using the usual right-hand rule

```
faces
1 1 10 -5 -9
2 2 11 -6 -10
3 3 12 -7 -11
4 4 9 -8 -12
5 5 6 7 8
6 -1 -4 -3 -2
```

Lastly, one can define the body as the volume contained by the faces as

```
bodies
1 1 2 3 4 5 6 volume 1
```

The first thing to do is to check out your cube. Start the evolver load the file with the instructions to make a cube and type ‘s’. If you then hit ‘h’ while focusing on the graphicss window you will get all the possible graphics functions/options, like rotation, translation, spinning, etc. Hit ‘q’ to go back to the command line for the evolver. To print your cube, hit ‘P’, choose PostScript (option 3) and follow instructions.

To minimise the area of the cube, hit ‘g’, you will see that a line is printed with the iteration number, area, energy, and current scale factor. Also, the vertices in the middle of the faces are now popping-out. If you keep hitting ‘g’, not much happens after this step. This is because the algorithm has minimised the area, given a certain triangulation. Of course, the triangulation of the surface you see in Fig. 6.1 is very coarse. We can make a better triangulation by hitting ‘r’. The evolver will tell you how many vertices, edges and facets are now in the surface. Evolve your surface hitting ‘g 10’, you will see that you have reached another minimum and you are stuck! Again, refine and repeat the iteration until you find that further refining does not decrease the area further.

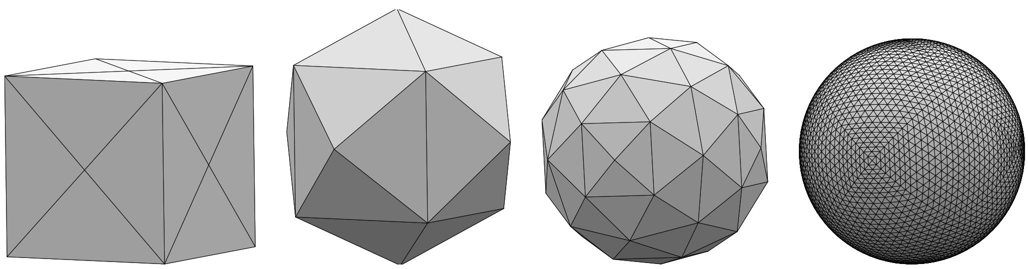


Figure 6.1: A soap cube becoming a soap bubble.

As exercise, try to remove the volume constraint from the “bodies” section (this is done by deleting “volume 1”) and see what happens to our soap cube. Can you explain why?

6.4 Catenoid Collapse

In this section we will test the true power of surface evolver, i.e. performing changes of topology of the surface. In this example we will start from a minimal surface that is formed between two parallel rings that are close enough, i.e. a catenoid. "Catenoid" comes from "catena" or "chain" because it is the surface obtained by spinning the shape of a chain hanging from two poles under its own weight around a central axis. This surface is thus a surface of revolution and has an analytical representation:

$$\begin{aligned} x =& c \cosh \frac{v}{c} \cos u \\ y =& c \cosh \frac{v}{c} \sin u \\ z =& v \end{aligned} \tag{6.1}$$

where $u = [-\pi, \pi)$ and $v \in \mathbb{R}$. The catenoid exists as a minimal surface only when the two parallel rings are placed near enough. When they are parted more than $d^* = 0.66R$, then the catenoid collapses into two disjoint disks (themselves minimal surfaces). See also a home-made movie at `https://www.youtube.com/watch?v=XqKDZB9nxDI`. Let's see if surface evolve can capture this behaviour.

Let's start by defining our frame. Here's a problem! How can we define circles using vertices and edges. We would need an infinite number of them. Luckily one can also give exact parametric expressions for curves in 3D to surface evolver and select vertices on these curves.

We will start from a cylinder expressed as rectangles formed by edges belonging to the two rings. The parameter we will need are (1) the separation between the rings d and (2) the radius of the rings R. Unfortunately surface evolver does not like parameters in single letters so we will use dz and Rc. We can thus specify parameters, curves, vertices, edges and faces as follows:

```
PARAMETER Rc = 1
PARAMETER dz = 0.5

boundary 1 parameters 1 // upper ring
x1: Rc * cos(p1)
x2: Rc * sin(p1)
x3: dz

boundary 2 parameters 1 // lower ring
x1: Rc * cos(p1)
x2: Rc * sin(p1)
x3: -dz

vertices
//they are fixed because otherwise will slide along boundary
1 0.0 boundary 1 fixed
2 pi/3 boundary 1 fixed
3 2*pi/3 boundary 1 fixed
4 pi boundary 1 fixed
5 4*pi/3 boundary 1 fixed
6 5*pi/3 boundary 1 fixed

7 0.0 boundary 2 fixed
8 pi/3 boundary 2 fixed
```

```
9 2*pi/3 boundary 2 fixed
10 pi boundary 2 fixed
11 4*pi/3 boundary 2 fixed
12 5*pi/3 boundary 2 fixed

edges
//first the ones in each ring
//they are fixed because the new vertices cannot slide
1 1 2 boundary 1 fixed
2 2 3 boundary 1 fixed
3 3 4 boundary 1 fixed
4 4 5 boundary 1 fixed
5 5 6 boundary 1 fixed
6 6 1 boundary 1 fixed

7 7 8 boundary 2 fixed
8 8 9 boundary 2 fixed
9 9 10 boundary 2 fixed
10 10 11 boundary 2 fixed
11 11 12 boundary 2 fixed
12 12 7 boundary 2 fixed

//now the ones from ring 1 to ring 2
13 1 7
14 2 8
15 3 9
16 4 10
17 5 11
18 6 12


faces
1 1 14 -7 -13
2 2 15 -8 -14
3 3 16 -9 -15
4 4 17 -10 -16
5 5 18 -11 -17
6 6 13 -12 -18
```

Starting from this surface, which should mimic a discretised cylinder, refine the triangulation once or twice and minimise the surface. Once you are happy with the result, hit 'A' and the program will tell you which parameters you defined and which ones you want to change 'on the go' by pressing the corresponding number and value. Press 2 and 0.6, this brings the distance of the two boundary circles from 0.5 to 0.6. Now minimise again, you will notice that the catenoid has a "slimmer waist". Because this procedure accumulates "defects" in the triangulation, it is often a good idea to equi-triangulate (redistribute the area among the triangles) by hitting 'u' and to remove the edges that have become shorter than a certain (small) value by hitting 't 0.05'.

For instance, follow this sequence from the start to get a nice catenoid: s, q, r, r, g 1000. From here, separate the rings: A, 2 0.6, g 1000, u, g 1000, u, g 1000. And once more: A, 2 0.66, g 1000, u, g 1000, u, g 1000. Now the catenoid is ready to undergo the collapse: A, 2 0.665, g

1000, u, g 1000, u, t 0.05, g 1000, t 0.05, g 1000, t 0.05, g 1000.

As an exercise, from the start position yourself beyond the collapse point and observe the neck shrinking from the top. Use a long iteration such as 'g 10000'. You will see that the neck stops, but after equitriangulation and removing short edges, it start collapsing again.

Another exercise is to remove the fixed condition from the boundaries and see what happens.

One more exercise, the catenoid instability can be induced when $z < z^*$. This can be done by defining the catenoid to be the boundary of a body, and adjusting the body volume with the b command to get zero pressure. This effectively generates a catenoid "implosion". Try it.

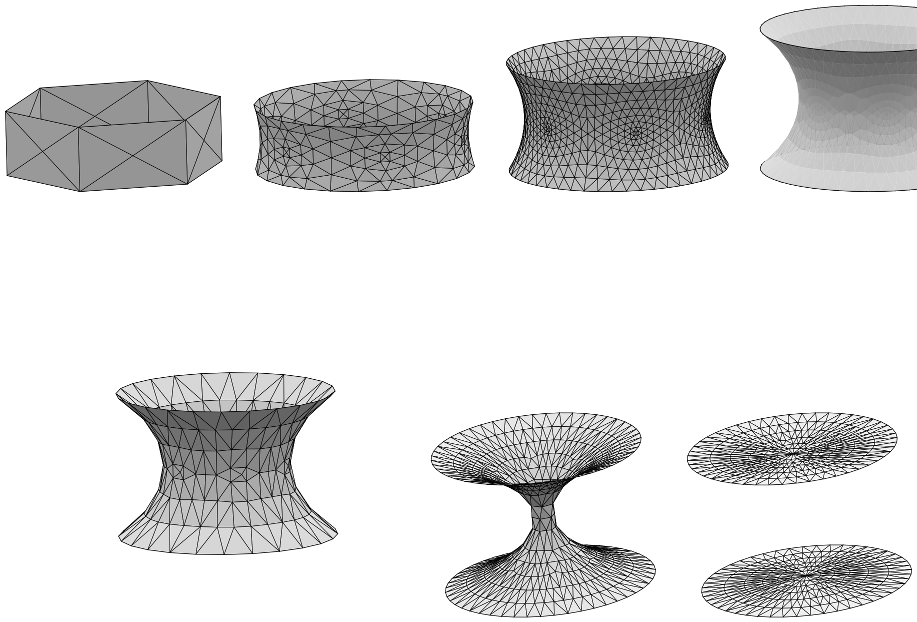


Figure 6.2: A catenoid collapse.

6.5 Minimal Surfaces of Links and Knots

6.5.1 Hopf Link

Links and knots can be thought as boundaries of surfaces, for instance Seifert surfaces! On the other hand, the Seifert surface is not necessarily minimal. As far as I know, there is no general analytic description for minimal surfaces bound by knots and links. But, we have the evolver. If you dip a Hopf link in a soap solution you do not immediately get a surface, but a film with some triple junctions. [xx draw pic xx]. By popping the right part of the film one gets a surface, with no triple junctions.

A ring in 3D space centred in zero and radius R can be parametrised as

$$\begin{aligned} x_1(\theta) &= R\cos\alpha\cos\theta \\ y_1(\theta) &= R\sin\theta \\ z_1(\theta) &= R\sin\alpha\cos\theta \end{aligned}$$

where the angle α represent a rotation around the y-axis. This can be see as the first boundary of a film. Now we add a second, linked to the first,

$$\begin{aligned} x_2(\theta) &= R\cos\theta \\ y_2(\theta) &= R + R\sin\theta \\ z_2(\theta) &= 0 \end{aligned}$$

This ring lies flat in the x-y plane with normal directed as $\hat{z}$ and is linked to the first, when $\alpha \neq n\pi$ and $n \in \mathbb{Z}$. Both are parametrised by $\theta = [0 : 2\pi)$.

Let's use these curves as boundaries for the input script of the evolver. Now we want to draw a surface in between these boundaries. If you think about it, you will see that you can do this in two ways. One, you can circle the vertices in one boundary in one direction, say clockwise, while connecting the vertices on the other boundary with edges and circling them either clockwise or counter-clockwise. By doing this you will obtain two different surfaces which are mirror symmetric [xx to check xx].

Let's try. We can write the following instructions for the first case, clockwise[2]

```
PARAMETER Rc = 10
PARAMETER atilt = pi/2

boundary 1 parameters 1 // first ring
x1: Rc * cos(atilt) * cos(p1)
x2: Rc * sin(p1)
x3: Rc * sin(atilt) * cos(p1)

boundary 2 parameters 1 // lower ring
x1: Rc * cos(p1)
x2: Rc + Rc * sin(p1)
x3: 0

vertices
//they are fixed because otherwise will slide along boundary
//use increments of pi/3 and start from pi/2
1 pi/2 boundary 1 fixed
2 5*pi/6 boundary 1 fixed
3 7*pi/6 boundary 1 fixed
4 3*pi/2 boundary 1 fixed
5 11*pi/6 boundary 1 fixed
6 13*pi/6 boundary 1 fixed

7 pi/2 boundary 2 fixed
8 5*pi/6 boundary 2 fixed
9 7*pi/6 boundary 2 fixed
10 3*pi/2 boundary 2 fixed
11 11*pi/6 boundary 2 fixed
12 13*pi/6 boundary 2 fixed

edges
//they are fixed because the new vertices cannot slide
1 1 2 boundary 1 fixed
2 2 3 boundary 1 fixed
3 3 4 boundary 1 fixed
4 4 5 boundary 1 fixed
5 5 6 boundary 1 fixed
6 6 1 boundary 1 fixed

7 7 8 boundary 2 fixed
8 8 9 boundary 2 fixed
```

```
9 9 10 boundary 2 fixed
10 10 11 boundary 2 fixed
11 11 12 boundary 2 fixed
12 12 7 boundary 2 fixed

//join the two rings using shortest path
//start from vertices on the same side
13 1 7
14 2 8
15 3 9
16 4 10
17 5 11
18 6 12

faces
1 1 14 -7 -13
2 2 15 -8 -14
3 3 16 -9 -15
4 4 17 -10 -16
5 5 18 -11 -17
6 6 13 -12 -18
```

If you want the other case, clockwise in the first ring and counter-clockwise in the second, one just has to change the vertices section into:

```
7 pi/2 boundary 2 fixed
8 pi/6 boundary 2 fixed
9 -pi/6 boundary 2 fixed
10 -pi/2 boundary 2 fixed
11 -5*pi/6 boundary 2 fixed
12 -7*pi/6 boundary 2 fixed
```

so that the second ring is circled in the opposite direction with respect to the first. At voila. Two surfaces for the Hopf link, the first identifying a +1 link whereas the second identifying the -1 link (fig. 6.3).

As exercise, check that if both rings are circled counter-clockwise, the obtained surface is identical to the first clockwise^2 case.

As exercise, modify the tilt angle alpha and examine how the minimal surface changes.

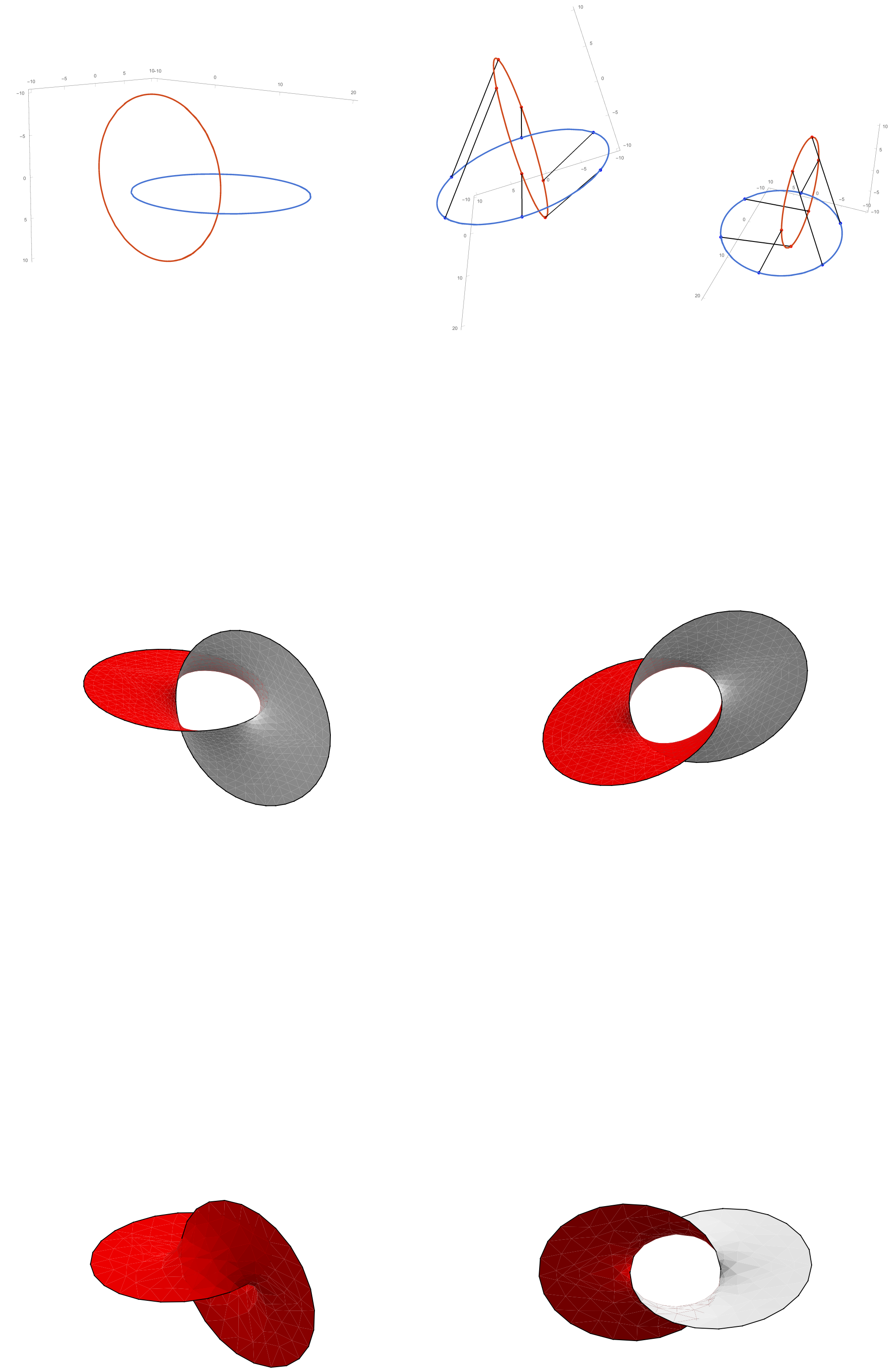


Figure 6.3: Constructing the minimal surfaces of a Hopf Link

6.5.2 Trefoil

One parametrisation of torus knots is

$$\begin{aligned} x_1(t) &= (r\cos qt + R)\cos pt \\ x_2(t) &= (r\cos qt + R)\sin pt \\ x_3(t) &= r\sin qt \end{aligned}$$

Where R is the radius of the torus centreline while r is the radius of the disk. The right-handed trefoil is obtained by setting $p = 2$ and $q = 3$, the left-handed one by setting $p = -2$ and $q = 3$. A good way to construct the surface in the trefoil is by setting a large R and a small r. In this way it is easier to visualise the polygons making up the surface. Once it is made for the right hand trefoil, the left-handed one follows easily just by changing the sign of p.

For a right-handed trefoil we can thus write the following instruction:

```
PARAMETER rmin = 0.5
PARAMETER rmax = 2

boundary 1 parameters 1
x1: (rmin * cos(3 * p1) + rmax) * cos(2 * p1)
x2: (rmin * cos(3 * p1) + rmax) * sin(2 * p1)
x3: rmin * sin(3 * p1)

vertices
1 0*pi boundary 1 fixed
2 0.2*pi boundary 1 fixed
3 0.4*pi boundary 1 fixed
4 0.6*pi boundary 1 fixed
5 0.8*pi boundary 1 fixed
6 1.0*pi boundary 1 fixed
7 1.2*pi boundary 1 fixed
8 1.4*pi boundary 1 fixed
9 1.6*pi boundary 1 fixed
10 1.8*pi boundary 1 fixed

edges

1 1 2 boundary 1 fixed
2 2 3 boundary 1 fixed
3 3 4 boundary 1 fixed
4 4 5 boundary 1 fixed
5 5 6 boundary 1 fixed
6 6 7 boundary 1 fixed
7 7 8 boundary 1 fixed
8 8 9 boundary 1 fixed
9 9 10 boundary 1 fixed
10 10 1 boundary 1 fixed

11 1 6
12 2 7
```

```
13 3 8
14 4 9
15 5 10

faces

1 11 6 -12 -1 color 4 backcolor 15
2 12 7 -13 -2 color 4 backcolor 15
3 13 8 -14 -3 color 4 backcolor 15
4 14 9 -15 -4 color 4 backcolor 15
5 15 10 11 -5 color 4 backcolor 15
```

Try to evolve this surface and refine it. Can you tell the difference with previous surfaces? What is happening to the colour of the surface? You can also try to deform the boundaries by playing with the parameters rmin and rmax (remember that this can be done on the fly by hitting 'A').

For the committed student, try different torus knots and check the orientability of the resulting minimal surface.

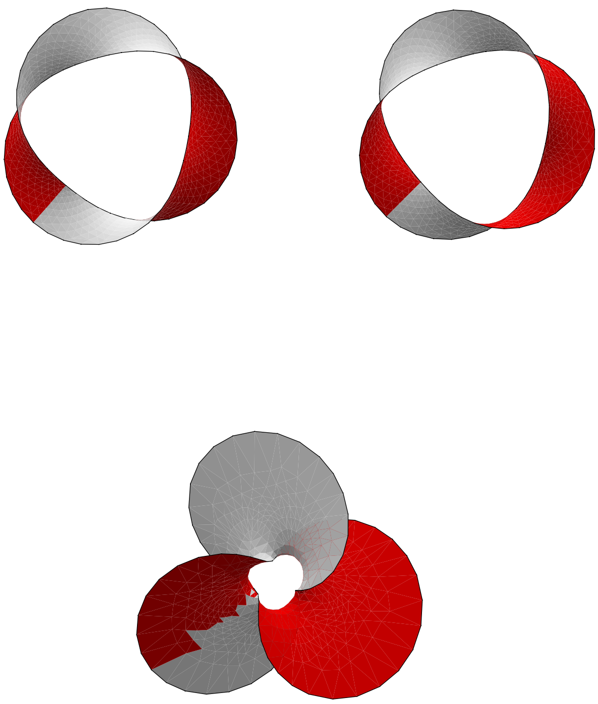


Figure 6.4: Trefoils

Bibliography

[1] AD Bates and A Maxwell. *DNA topology.* Oxford University Press, 2005.

[2] C R Calladine, H Drew, F B Luisi, A A Travers, and Eleanor Bash. *Understanding DNA: the molecule and how it works*, volume 1. Elsevier Academic Press, 1997.

[3] M.R Dennis and J.H Hannay. Geometry of Călugăreanu's theorem. *Proc. R. Soc. A*, 461(2062):3245–3254, 2005.

[4] L.H. Kauffman. *On Knots.* Annals of mathematics studies. Princeton University Press, 1987.

[5] J White. Self-Linking and the Gauss Integral in Higher Dimensions. *Am. J. Math*, 91(3):693–728, 1969.