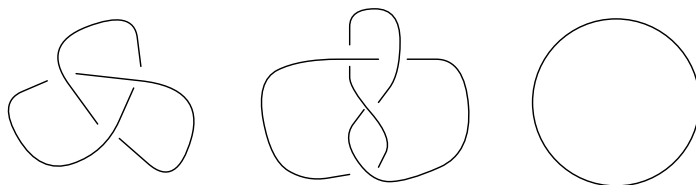


1 Basics I

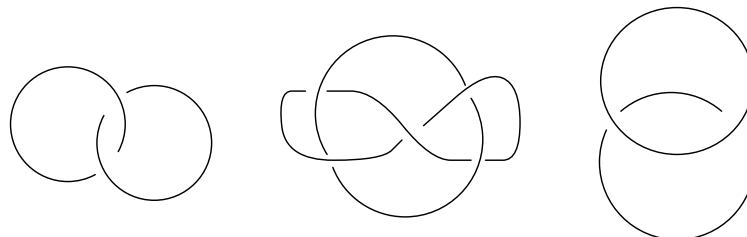
1.1 Introduction to Knots and Links

This course is about knots and links in 3-dimensional space. It's important to get one thing clear right at the start — a mathematical knot is a curve with no “free ends” such as you have in your shoelaces or your tie. We can think of knots as endless loops of string. A link is a configuration made up of several such loops. The loops are not allowed to have any self-intersections; also, we imagine them to have no thickness. Intuitively, two knots or links are regarded as being the same if one can be physically deformed into the other in 3-dimensional space. Of course, the strings are not allowed to pass through themselves or each other during the deformation. It is useful to make a few physical models of knots and links for yourself, as this will help you to appreciate better what is going on. My models are made of nylon rope from a hardware shop. The ends are easily joined by melting rope over a (very low!) gas flame. Adams recommends the use of an electric extension cable, which can be plugged into itself to form the knot.

Of course we don't want to have to carry sackloads of knots around with us in order to study them mathematically. The conventional way to represent knots, on paper or a blackboard, is by means of 2-dimensional diagrams. These diagrams show curves in the plane, and the curves are “broken” at certain points called crossings to show that one strand is supposed to pass behind the other in the actual 3-dimensional knot. Of course, many different diagrams can represent the same knot. Here are some examples of knot diagrams.



And here are some examples of link diagrams.

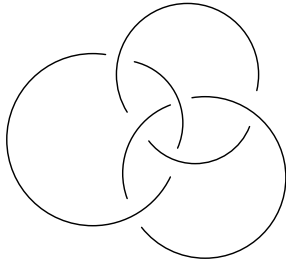


Let's look at linking first, in order to appreciate what some of the problems are. For this we can use simple curves such as circles, which are not themselves knotted.

If we consider all configurations consisting of pairs of circles in \mathbf{R}^3 , there are clearly two types — those which are linked as in the first diagram above, and those which are unlinked as in the third diagram. The second diagram illustrates the point that there are further possible ways to link two unknotted curves if they are not required to be circles.

Problem: How can we capture mathematically the difference between a linked and an unlinked pair of circles?

The next example, the *Borromean rings*, demonstrates a more subtle form of linking that is possible for three circles. Note that if any one of the circles is removed, the other two are unlinked, yet the set of three circles cannot be separated.



Of course we need precise definitions of knots and links, and of equivalence of knots and links, in order to set up a proper mathematical theory. This can be done, and such a theory does exist. Indeed, it is a very active area of research, with links to several other branches of mathematics as well as to other sciences such as physics, chemistry and biology. Some enthusiasts have been known to claim that knot theory may eventually become as “applicable” a branch of mathematics as calculus. It would be surprising, in fact, if such fundamental geometrical properties of curves in 3-dimensional space did not turn out to have good applications!

However, it is not the job of this course to set up the mathematical theory of knots in a fully rigorous way. Instead, we are going to treat certain concepts as intuitive. In this way we shall be able to reach the interesting calculations in a short time, and obtain a real “feel” for knots, without getting bogged down in generalities. This is the way you learned calculus at school. First you learn rules for differentiating and integrating standard functions, and only later do you get involved with the concept of what it really means for a function to be differentiable or integrable.

To illustrate what I mean by the intuitive approach, let's consider the simplest and most obvious *invariant* of a link, namely the number of *components* or

constituent curves it has. By an invariant we mean some property of the link that is not changed by deformations in 3-dimensional space. (We shall have a lot more to say about invariants later.) It is intuitively “obvious” that we cannot start with a link of, say, three components and conjure it into a link of two or four components without cutting and rejoining the strings. For example, we cannot make a component disappear.

Now consider one possible attempt to model knots mathematically. We might try to define a knot by a set of equations, and equivalence by continuous variation of parameters occurring in these equations. However, this requires care. Suppose that we start with the standard circle

$$x^2 + y^2 = r^2, \quad z = 0$$

lying in the (x, y) -plane. The centre is at the origin and the radius is r . Each circle is an example of the simplest knot, called the *unknot*. As we vary r continuously, we obtain a family of equivalent knots, parametrised by r . However, we are not allowed to put $r = 0$, because the equations then represent a single point, the origin, and not a circle. And $r < 0$ would give the empty set. So we clearly cannot allow a completely arbitrary continuous variation of parameters occurring in the equations. In fact, a similar “method” could be used to make any knot vanish: put the knot inside a sphere of sufficiently large radius R , and then let $R \rightarrow 0$.

The branch of mathematics which sorts out this kind of thing is called *point set topology*, and it needs a whole course to itself. The courses MT3101 on metric spaces and MT3121 on topology of surfaces cover some of these ideas. However, I am not suggesting that you need to have attended either of these courses in order to follow this one.

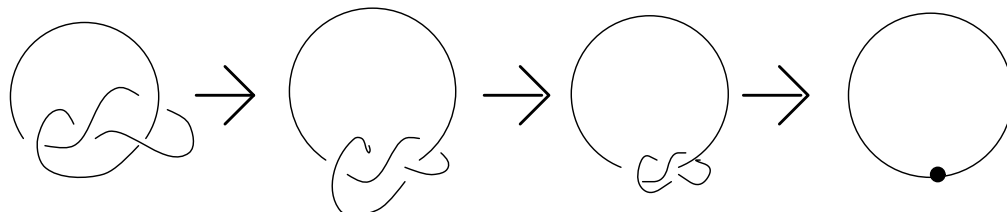
One term from topology will be useful to us, the idea of *homeomorphism*. A homeomorphism is simply a continuous function which has a continuous inverse. That is, the function must be a one to one correspondence (or *bijection*) h from its domain of definition X to its range Y , so that the inverse function is defined as a function $f^{-1} : Y \rightarrow X$, and both f and f^{-1} must be continuous. In our case, the domain X and range Y will normally be the whole of 3-dimensional space \mathbf{R}^3 . We visualise homeomorphisms as continuous but *reversible* deformations. The simplest homeomorphisms are the linear isomorphisms given by invertible matrices.

We shall also want to consider homeomorphisms where X is a circle, such as the standard circle given above. If Y is a plane curve and there is a homeomorphism $h : X \rightarrow Y$, then Y is usually called a *Jordan curve*. However if Y is a curve in \mathbf{R}^3 , then Y is a knot. Another way to say the same thing is that Y can be parametrised as $(x(\theta), y(\theta), z(\theta))$ for $0 \leq \theta \leq 2\pi$, where these coordinate functions x , y and z are continuous functions of the parameter θ which all take the same values for $\theta = 0$ and $\theta = 2\pi$, but which do not all take the same values for any other pair of distinct values of θ .

This gives a perfectly reasonable mathematical definition of a knot, but in fact it is not the one we are going to use. This is because it is too general, and hence too complicated for our needs. We shall call a curve which is homeomorphic to a circle a *topological circle*.

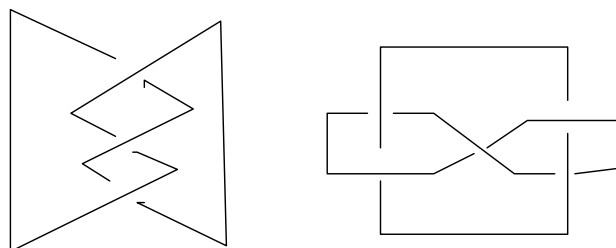
Problem: Why can't a Jordan curve be "knotted"? What exactly would this mean?

The next illustration shows a more subtle way in things can go wrong if we are not careful about definitions. Suppose we start with any knot and "pull it tight" as indicated by the diagrams below.



This certainly gives a continuously parametrised family of curves in \mathbf{R}^3 , and in this case every one of the curves in the family is a topological circle, in the sense that it can be continuously parametrised by a variable θ in the range $0 \leq \theta \leq 2\pi$. Clearly, every knot is equivalent to the unknot if this sort of move is allowed, so once again there would be no theory of knots.

Thus, even with an intuitive approach, we do have to make it clear what the rules of the game are. The easiest way is to assume that all our knots, and all components of links, are finite polygons in \mathbf{R}^3 .

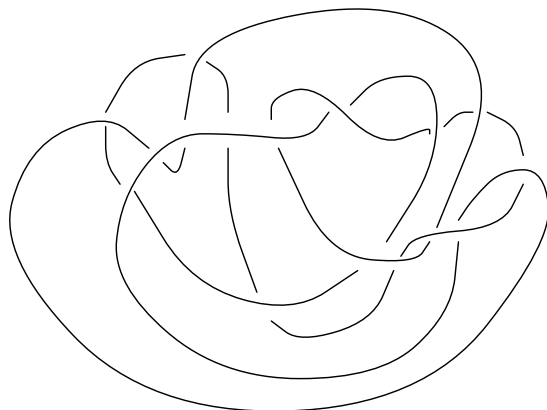


This will be our "official" definition of a knot or link, for use in proofs when it is necessary to be precise. However, in practice, we shall continue to draw knots using curves — it's too awkward to do anything else — but we'll always bear in mind that our pictures could be made into polygonal ones. We could for example imagine that the curves we draw are actually made up of a very large number of tiny straight line segments.

Later we'll discuss equivalence between such polygonal knots and links in more detail. The key point is that we need only a finite amount of data to specify a knot, for example the coordinates of the vertices of the polygons, and the list of pairs of vertices which are to be joined by line segments.

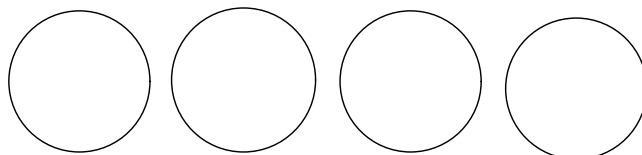
For the record, the appropriate definition of equivalence, for knots taken as topological circles, is that two of these are equivalent if there is an orientation-preserving homeomorphism of *the whole of* \mathbf{R}^3 which carries one topological circle on to the other one. However, just as we are going to work with the simpler, polygonal version of knots, we are also going to work with a simpler definition of equivalence. We'll come to all this in Chapter 2.

We have already mentioned that the simplest knot is the unknot. However, it is a non-trivial problem to decide whether a given knot is the unknot or not, *i.e.* whether it is “really knotted”. For example, you might like to try your hand at proving that the following diagram represents the unknot.

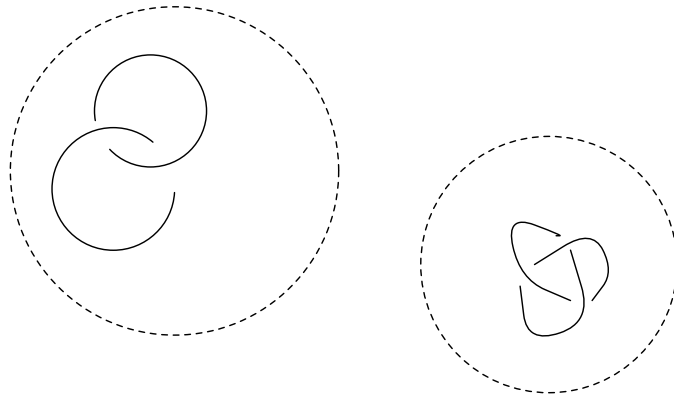


Every knot diagram can be transformed into a diagram for the unknot by making suitable crossing changes. It is intuitively obvious that we can untie any knot if we are allowed to pass the string through itself! However, it's not quite so easy to write out a proper proof, so we'll defer this job till later.

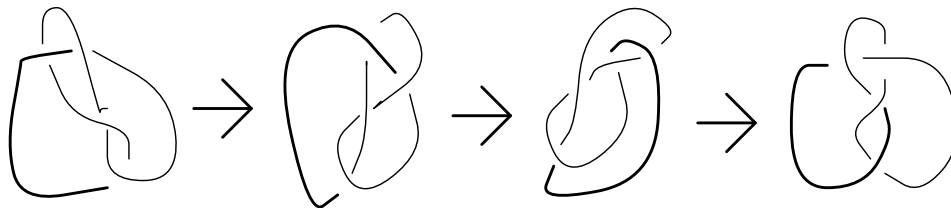
Along with the unknot, we mention unlinks. A link is called *split* if the components can be separated from each other, so as to form two links which could for example be enclosed in different soap bubbles in \mathbf{R}^3 . The simplest example is just a set of n circles such as we can easily draw in a plane so that they do not intersect.



Any link equivalent to this one is called a *split unlink*. In general the individual components of a split link can be knotted.



We end this section by mentioning *chirality*, or left and right handedness for knots. (This can be illustrated quite practically with the rope models.) The reflection of a knot K in a mirror is another knot K^* . Is K^* equivalent to K ? The answer is sometimes yes and sometimes no. For example the figure eight knot is equivalent to its mirror image — it is said to be *amphicheiral* or *achiral*. However, the trefoil is *chiral* — it exists in both a left-handed and a right handed version.



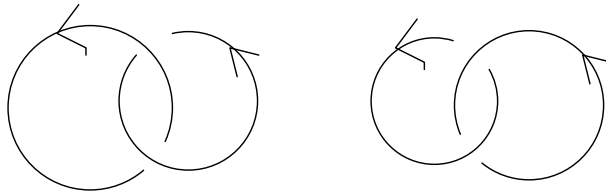
In practice we're going to simplify things by representing knots (as in the drawings) by *knot diagrams* in \mathbf{R}^2 . A formal definition of a knot diagram will be given later, but for now note that it must at least record which is the “overpass” and which the “underpass” at each crossing point. In this way we avoid ever having to deal with the “actual” knot in \mathbf{R}^3 , because all the manoeuvres we can do in \mathbf{R}^3 can be represented by sequences of “moves” which transform one knot diagram into another. These moves are called *Reidemeister moves* and we'll come to them in Chapter 2.

1.2 Problems of Knot Theory

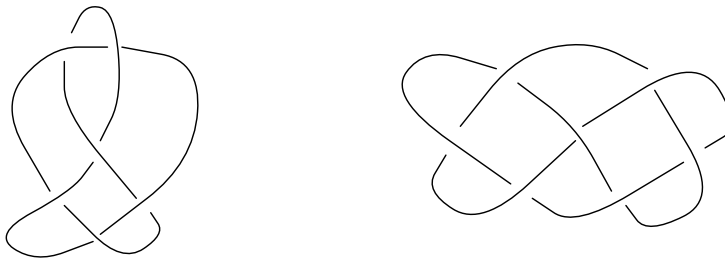
Much of knot theory is concerned with telling when two knots are the same and when they are different. As we have already mentioned, the basic idea for telling knots apart is to evaluate a suitable *knot invariant*, and show that it takes different values for the two knots being considered. The invariant may be a number, or it may be a more complicated mathematical object such as a group or a polynomial. In this course we shall study two important knot invariants which are polynomials. They are called the *Conway polynomial* and the *Jones*

polynomial. Although knot theory had its origins in the 19th century, these two invariants were discovered only in the last 20 years. (This is not quite true in the case of the Conway polynomial, as it is really an “improved version” of the Alexander polynomial discovered in the 1920’s.)

Actually there is not just one classification problem for knots, but several different problems. An important concept is that of an *oriented* knot or link. This means that an arrow (to indicate the “positive” direction) is placed on each component. In this situation, equivalences have to match up the arrows as well. For example, it is easy to convince yourself that the following two links are different as *oriented* links, although they are equivalent as ordinary links if the arrows are removed.



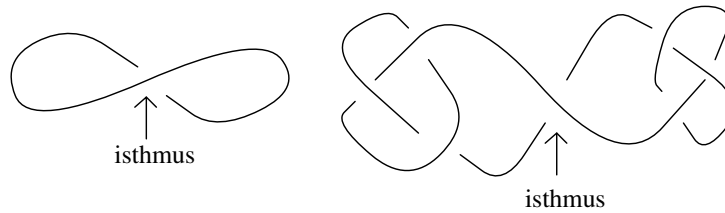
We shall also be considering different types of knots. An important special class is that of *alternating knots*. These are knots which have diagrams in which the overpasses and underpasses alternate as we go round the knot.



Every knot diagram can be made into the diagram of an alternating knot by changing crossings suitably. Experimenting with small diagrams should soon convince you of this, and it’s a good exercise to try to find a method for doing this which will always work.

A long standing conjecture about alternating knots was proved recently (1986). This says if we have an alternating diagram for a knot, then that knot cannot have any diagram with fewer crossings than there are in the alternating diagram. In particular, it follows from this that the number of crossings in an alternating diagram is an invariant of the knot. This gives a means to tell for example that the trefoil knot and the figure-eight knot are different, and that both of them are different from the unknot.

Actually we have to be a bit more careful about the definition of an alternating diagram for this theorem to apply. We can easily alter a knot diagram by introducing a loop or twisting a part of the knot through 180° .

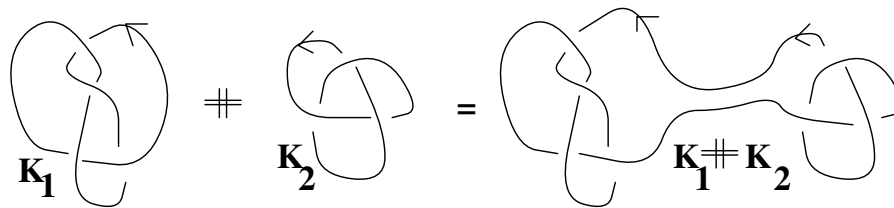


An alternating diagram is called *reduced* if there are no crossings of this trivial kind. This special kind of crossing is easily recognised, as two of the four plane regions of the diagram which meet at the crossing are part of the same global region in the whole plane. Such a trivial crossing, called an *isthmus*, is easily eliminated by twisting a part of the diagram through 180° . In the next section, we shall see that we do not really lose anything by restricting to diagrams without isthmuses. Those with isthmuses can be obtained by a process called *composition* from the others.

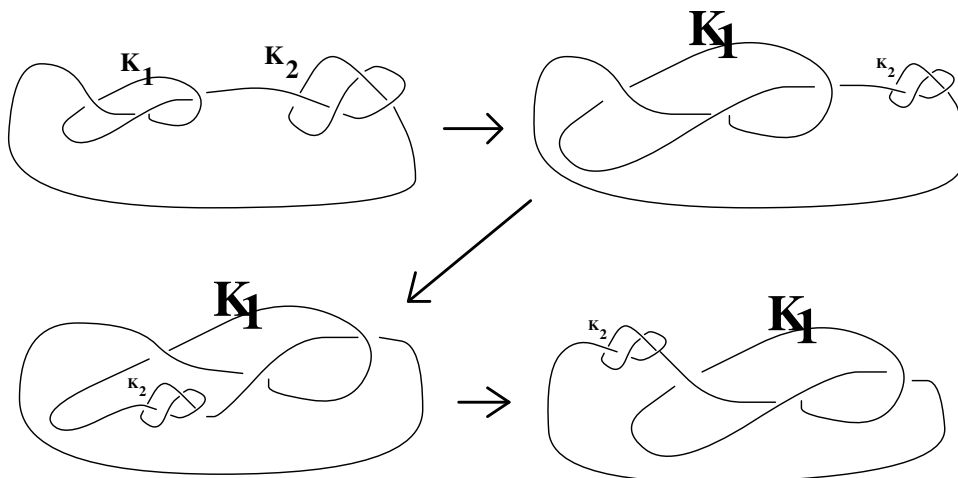
1.3 Composition of Knots

The fundamental problem in knot theory is the *classification* of knots. This asks for a “list” of all possible knot types or equivalence classes, so that every knot is equivalent to one and only one knot on the list. Further, to be of practical use we need a method for deciding which member of the list corresponds to a given knot. These are very hard problems! Tables of knots are arranged according to the “crossing number”, *i.e.* the smallest number of crossing points for any knot in the equivalence class. For example these tables will list only one knot type with three crossings (the trefoil) and only one with four crossings (the figure eight knot). For three up to ten crossings the number of knot types is 1, 1, 2, 3, 7, 21, 49, 165 respectively.

Informally, we define the *composition* of two knots K_1 and K_2 to be the knot obtained by tying K_1 followed by K_2 on the same piece of string.



We denote this new knot by $K_1 \# K_2$. The following sequence of diagrams explains why the operation of composition of knots is commutative.

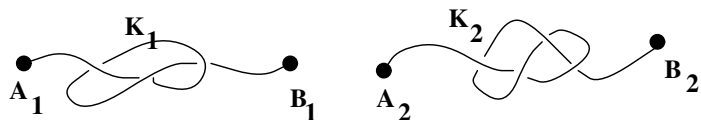


The strategy is to make one knot small and slide it along the other one. Composition is also associative (obviously). Also, the unknot 0 functions as an identity element, *i.e.* $K \# 0 = K = 0 \# K$.

A very important fact about the composition of knots is that knots can't "cancel each other out". That is, if the composition $K_1 \# K_2$ is the unknot, then both knots K_1 and K_2 must themselves be unknots. It's worth pausing to reflect a little on this result. If it were not true, that is to say if there were some remarkable knots K_1 and K_2 which combined to give the unknot, then *every* knot would be a composite knot. This would follow from the trivial observation that every knot is the composition of itself with the unknot.

We define a knot to be *prime* if it is not the unknot and is not the composition of two non-trivial knots $K_1 \neq 0$, $K_2 \neq 0$. This definition depends on the result mentioned in the preceding paragraph. It was proved by Schubert in 1949 that every knot is uniquely (up to order of the terms) the composition of prime knots. Tables of knots are actually tables of prime knots.

This brief discussion of composition of knots has passed over a possible ambiguity in the definition. We can think of the composition as being made by tying K_1 and K_2 in separate pieces of rope with end points A_1, B_1 and A_2, B_2 . Thus K_1 will really mean the knot obtained by joining A_1 to B_1 , and similarly for K_2 .

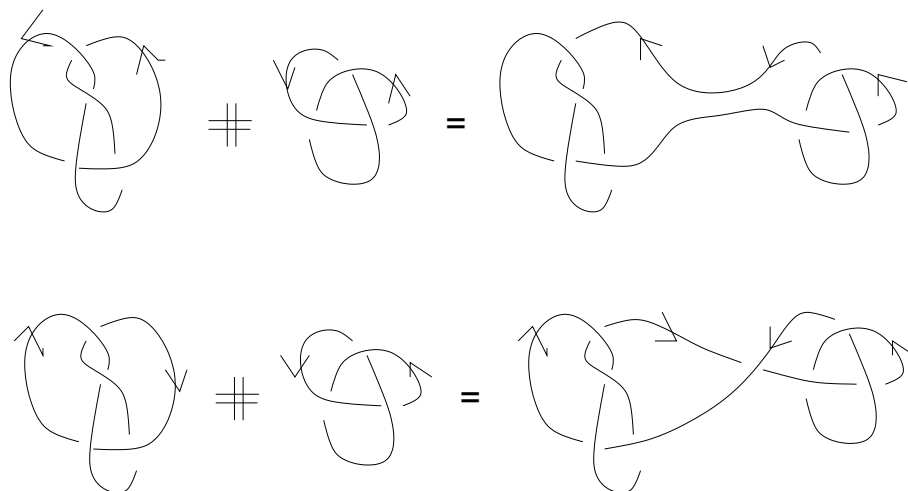


Now we can form a composition by joining A_1 to A_2 and B_1 to B_2 . But we could equally well join A_1 to B_2 and B_1 to A_2 . Unless we know which end of the strings is which, we have two possibilities. Do these always give the same result for the composition?

This turns out to be quite a subtle question! If you experiment with a few small knots, the chances are that it will not matter which way you join the ends.

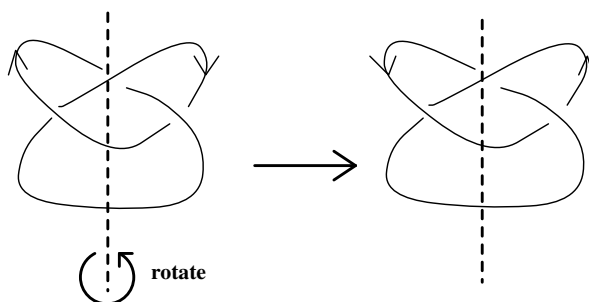
But in fact there do exist knots for which it *does* matter. To explain this point, we have to go back to the idea of orienting a knot. A moment's reflection should convince you that labelling the ends of the "open" string as A and B is equivalent to putting an orientation on the knot.

Thus when we form the composition of two *oriented* knots, we will always do so in such a way that the orientations match up correctly, so that the composition itself receives an orientation. The diagrams below illustrate this.



In this way, we can define the composition of *oriented* knots unambiguously.

We have already met one type of symmetry that a knot can have, namely that of being equivalent to its mirror image. For oriented knots a second symmetry question arises. Is the knot equivalent to its *inverse*, that is, to the same knot with the opposite orientation? Of course, the deformation in \mathbf{R}^3 which carries one oriented knot to the other must match up the orientations as well. It is easy to see that a trefoil knot is invertible by rotating it through 180° , as shown below.



However, it was shown in 1964 by H. F. Trotter that non-invertible knots do exist. The first non-invertible knot in the table is the knot 8_{17} .

If at least one of the knots K_1 and K_2 is invertible, then the compositions obtained from the two different choices of matching the ends of the strings are equivalent knots. However, in the case where both K_1 and K_2 are non-invertible,

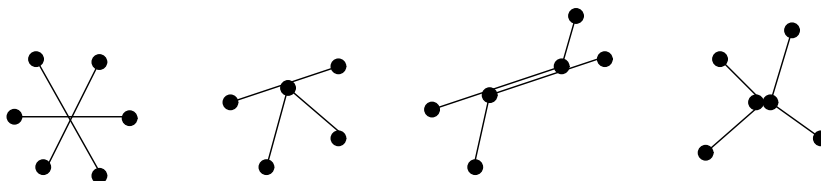
it is possible that the two compositions will be inequivalent. To be sure of forming a composition unambiguously in all cases, therefore, we must work with *oriented* knots.

1.4 Knot Diagrams and Invariants

Intuitively, we can think of a diagram of a knot or link in \mathbf{R}^3 as its shadow cast on some plane, which we then identify with \mathbf{R}^2 . If the knot is polygonal, then the shadow will be a polygon with self-intersections. To reconstruct the knot from the shadow, we have to add information to say which is the overpass and which is the underpass at each crossing. Also, we want to avoid projecting the knot in a direction where “accidental” intersections occur. We assume

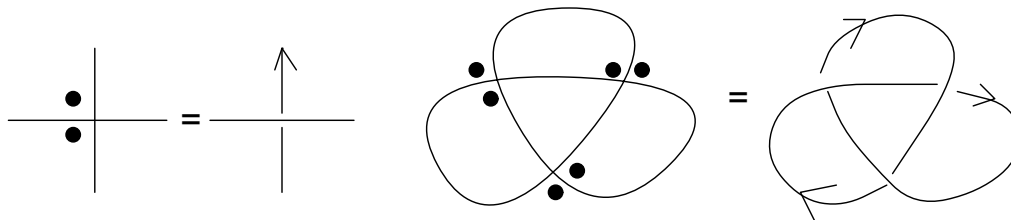
1. at each crossing, exactly two line segments cross,
2. line segments are not allowed to touch without crossing.

The following pictures show examples of outlawed configurations.



We shall regard it as intuitively obvious that every knot has lots of projections satisfying the required conditions. In fact, by starting with a projection in any direction in \mathbf{R}^3 and moving the direction of projection a small amount, we can avoid the bad cases indicated above.

By convention, we draw a knot diagram by breaking the curve to show the underpass at each crossing. This of course is just an attempt to give a more formal description of the diagrams such as we draw to represent knots on paper or a blackboard. Any way of indicating which of the two strands at a crossing is nearer the viewer will do as a knot diagram. For example in many older papers on knot theory the “dot convention” is used instead. For this, two of the four local regions at a crossing are marked by dots, in such a way that an (imaginary) insect crawling along an *oriented* knot will have the dots on its left hand side as it moves along the *underpass* at each crossing. The following example should make this clear.



In this course we shall use the modern method of breaking the curve and showing the orientation by means of arrows placed along it. Note that the orientation must be indicated for each component of a link diagram by at least one arrow on each component.

Two knot or link diagrams will be called *equivalent* if the knots or links that they represent are equivalent. Of course, it would be nice to have a criterion for equivalence of diagrams that is expressed entirely in terms of the diagrams themselves, without reference to the “actual” knots and links in \mathbf{R}^3 . This would mean that the whole theory of knots and links could be reduced to questions in two dimensions!

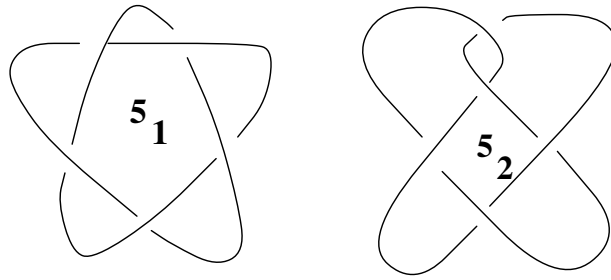
Actually, this can be done, and it will be discussed in Chapter 2 when we deal with the *Reidemeister moves*. These are changes we can make to a diagram locally which yield equivalent diagrams, and have the property that in a sense all diagrams equivalent to a given one can be obtained by using just these moves. However, there is a price to be paid: there is no known effective procedure for deciding which moves to apply when we want to know whether two given knot diagrams are equivalent.

A special case of this problem is that of deciding whether a particular knot diagram represents the unknot. In 1961 the German mathematician W. Haken came up with an algorithm for solving this special case. However, the algorithm is so complicated that nobody has yet succeeded in turning it into a computer program.

1.5 Some knot invariants

We conclude this chapter with a brief discussion of some invariants for knots that can be defined using diagrams. Our discussion will necessarily be very incomplete, as we do not at this stage have a proof that there are any knots at all that are not equivalent to the unknot! In fact, it is by defining and evaluating invariants that we aim to prove this. So in this section we shall *define* some invariants, but we shall not be able to *evaluate* them.

The first invariant is the *crossing number*. This is the smallest number of crossings that occur in any diagram for the knot (or link). This invariant was already discussed briefly in Sections 1.2 and 1.3. For example, the unknot has crossing number 0, the trefoil knot has crossing number 3 and the figure eight knot has crossing number 4. This invariant could be used to distinguish between these three knots. There are two (equivalence classes of) knots with crossing number 5.



As these examples suggest, knots are traditionally labelled in tables in the form c_n , where c is the crossing number and this is the n th knot in the table with crossing number c . This isn't really a very helpful notation! There is in fact a more systematic notation, again due to John Conway and based on his theory of "tangles". If this interests you, you can find out about it by reading *The Knot Book*, but I won't get around to dealing with it in the course.

Suppose that we are given a knot K , maybe handed to us as a rope model, and we are asked to find its crossing number. What do we have to do? First, we have to fiddle about with it so as to project it as a plane diagram. By counting the number of crossings, we get a number which is an upper bound for the crossing number. We can fiddle about some more to try to reduce this estimate by finding another projection with fewer crossings. When we get tired of this, we should try to prove that our best estimate is correct.

Now the real problems start. Let's suppose we are trying to prove that the crossing number of our knot K is 7. Then we need to know that K can not be represented by any diagram with 6 crossings or less. To do this, we need a table of all knots with 6 or less crossings, together with suitable invariants by means of which we can distinguish K from each of the knots in this table. You can see why the crossing number is hard to calculate! In fact, the supposed table only exists at present up to 16 crossings — nobody has listed the knots of 17 crossings yet. For the special case of alternating knots, we now have the amazing breakthrough mentioned in Section 1.2, but in the general case, a method for calculating crossing numbers seems completely out of reach.

The next invariant is the *unknotting number*. Look carefully at the diagrams of the knots 5_1 and 5_2 above. If we switch the overpass and underpass at any of the crossings in the diagram for 5_1 , we get a trefoil knot. Hence we need *two* crossing switches to reduce to a diagram for the unknot. On the other hand, switching either of the upper two crossings in the 5_2 diagram gives an unknot at once. We say that the 5_1 diagram has unknotting number 2, and the 5_2 diagram unknotting number 1.

So far so good, but remember that to get a knot invariant, we have to take account of all possible diagrams for the knot. What if there is another diagram for 5_1 in which a single crossing switch would be sufficient to undo the knot? The *unknotting number* of a knot K is defined as the minimum number of crossing changes required to convert *some* diagram of K into an unknot diagram.

You might guess that in trying to calculate the unknotting number of K , we could confine the search to diagrams of K with the minimum number of crossings. Then at least we could reduce the problem of calculating the unknotting number to that of checking a finite number of diagrams. Unfortunately, this guess is wrong! In 1983, an example was found of a knot K with crossing number 10 with the property that 3 crossing switches are required to unknot any diagram of K with 10 crossings, yet K has a diagram with 12 crossings which can be unknotted by making only 2 crossing changes.

As this situation suggests, the unknotting number is an extremely difficult invariant to calculate, even harder than the crossing number. To make progress in knot theory, we need invariants which can be calculated by practical methods. Such invariants are provided by the knot polynomials, the Alexander polynomial and the Jones polynomial, that will form the principal topics of this course. As we shall see, the evaluation of these invariants, for many knots, is quite easy. But again there's a price to be paid. The hard work is largely shifted to the problem of proving that the polynomials we shall calculate really are invariants!

We conclude this section by proving a result we have been tacitly assuming throughout this discussion of unknotting numbers.

Proposition 1.1 *Every knot diagram can be changed into a diagram for the unknot by switching a suitable subset of the crossings.*

Proof Choose a point p on the knot diagram that is not a crossing point, and if the diagram is not already oriented, choose an orientation. Start an imaginary insect to walk along the knot starting at p and taking the chosen direction. The *first* time the insect arrives at a crossing, if it is on the underpass, then switch the crossing so that it is on the overpass. If the insect was originally on the overpass, do nothing. Continue doing this until the insect returns to p .

We shall prove that this set of switching operations has resulted in the diagram of an unknot. To do this, we shall construct an unknot in \mathbf{R}^3 which has the switched diagram as a projection on a plane in \mathbf{R}^3 .

To do this, we time the progress of our imaginary insect using an equally imaginary clock. We start the clock at $t = 0$, and assume that the insect crawls round the diagram at constant speed, returning to p at $t = T$ say. Using (x, y) as coordinates in \mathbf{R}^2 , plot the curve C given parametrically by $(x(t), y(t), t)$ for $0 \leq t \leq T$ in \mathbf{R}^3 . Finally join the end points $(p, 0)$ and (p, T) of the curve C by a straight line segment parallel to the t -axis, to obtain a closed curve K in \mathbf{R}^3 . Now look at K from a view point on the negative t -axis, *i.e.* project K orthogonally on to the plane $t = T$. Since the overpass at each crossing always has a smaller t value than the underpass, this projection is precisely our switched knot diagram. To see that K is unknotted, project it on to any plane containing the t -axis. Since no two points of C have the same t -coordinate, and since p was not a crossing point, the projection is a simple closed curve with no crossings. Hence K is an unknot, and the proof is complete.