Algebraic Knot Theory

Zsuzsanna Dancso
Damian Lin
Nelson Odins-Jones

January 14, 2025

Contents

Introduction

These lecture notes are based on the course *"Algebraic knot theory"* delivered at the Australian Mathematical Sciences Institute Summer School in January 2025, which consisted of a series of twenty lectures, supported by eight hours of problem-solving tutorials.

We present an algebraic perspective on quantum topology, more specifically, on *universal quantum invariants* of knots and related structures. This perspective is informed by the rational homotopy theory idea of *formality*. It explains developments in mathematical physics, Lie theory and quantum algebra, where solutions to independently important sets of equations (pentagon and hexagon equations, Kashiwara–Vergne equations) are shown to be equivalent to universal quantum invariants of various generalisations of knots.

History

In the 1990's, four Fields medals were awarded for results which later came to form the foundations of quantum topology:

- to Vladimir Drinfeld in 1990, for foundational work establishing structures called *quantum groups* and their properties;
- to Vaughan Jones in 1990, for the discovery of the Jones polynomial, an algebraic tool which helps distinguish knots from each other;
- to Edward Witten in 1990, for physical insight leading to new mathematical results, particularly in topology;
- to Maxim Kontsevich in 1998, for his discovery of the universal quantum invariant of knots.

It is noteworthy that none of these mathematicians worked primarily in knot theory, or topology: Drinfeld's primary interests lie in algebra and algebraic number theory; Jones was an analyst working on von Neumann algebras; Witten is a theoretical physicist; and Kontsevich is known primarily for his contributions in mathematical physics and geometry.

Over the following years in became clear that all of these groundbreaking results are deeply related to each other. Through this course you will learn how Kontsevich's and

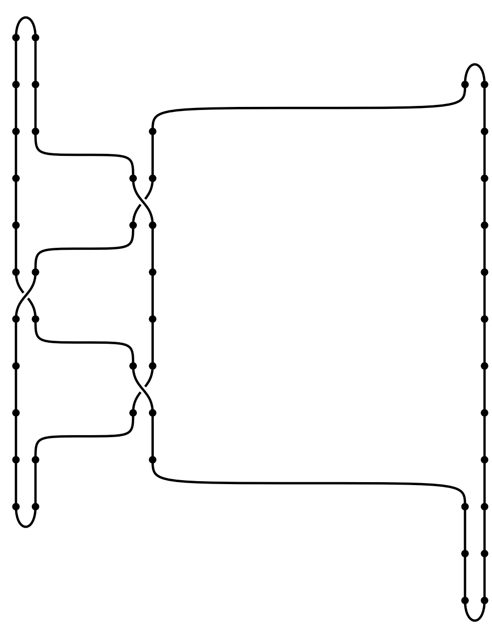
Drinfeld's work accomplishes the same goals via very different routes: Tamarkin, Bar-Natan and Fresse showed that *universal quantum invariants* of knots are in one-to-one correspondence with *Drinfeld associators*.

In fact, this phenomenon fits into a general pattern where universal quantum invariants of topological objects – or, in this course, *formality maps* – are in one-to-one correspondence with solutions to sets of equations in *graded algebras*. These sets of equations in key examples turn out to be of independent interest in Lie theory, quantum algebra of mathematical physics. The pentagon and hexagon equations which define Drinfeld associators is one such example.

This correspondence works best when the topological objects in question form a *finitely presented* algebraic structure. In this case the phenomenon fits into a general algebraic framework which illuminates the commonalities among the different examples.

Outline

This course approaches the subject of universal quantum invariants from this algebraic perspective: starting in Chapter 1 with a discussion of different algebraic structures, and finite presentations. In Chapter 2 we introduce knots and tangles as examples of algebraic structures and discuss whether they are finitely presented. (Spoiler: knots are not!) In Chapter 3 we move on to introduce the key algebraic ideas required: gradings, filtrations, and the associated graded functor. We discuss the key example of the Vassiliev filtration on the linearly extended space of knots, and its associated graded algebra. Chapter 4 introduce formality maps: the algebraic definition of a universal quantum invariant. Chapter 5 defines and proves the key properties of Kontsevich's univeral quantum invariant: this is of a different mathematical flavour than the other chapters, using primarily analysis rather than algebra. In Chapter 6, we translate Kontsevich's work to the algebraic setting of this course, and in the process discover Drinfeld associators.



Acknowledgements

We are grateful to James Morgan for his participation and helpful feedback during the development of these notes.

Chapter 1

Algebra

1.1 Algebraic Structures

In your first abstract and linear algebra courses you encounter a variety of algebraic structures: vector spaces, fields, groups, rings, and possibly more. All of these structures follow a general pattern: a set or multiple sets, with one or more operations on these sets, satisfying some axioms. Recall the definition of a group:

Definition 1.1.1. A *group* is a non-empty set G together with a binary operation – that is, a function $\cdot : G \times G \to G$ – called the *group multiplication*, satisfying the axioms:

- *Associativity*: For all a, b and $c \in G$, $(a \cdot b) \cdot c = a \cdot (b \cdot c)$.
- *Identity*: There exists an element $e \in G$ such that for every $a \in G$, $e \cdot a = a$ and $a \cdot e = a$.
- *Invertibility*: For each $a \in G$, there exists an element $a^{-1} \in G$ such that $a \cdot a^{-1} = e$.

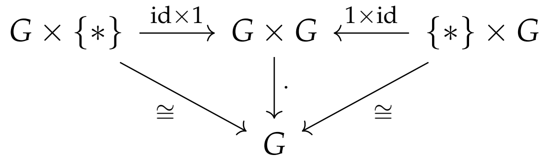
On first glance, a group is a set with one operation: the group multiplication. The first axiom, associativity, is convenient to write as a commutative diagram:

$$\begin{array}{ccc} G \times G \times G & \xrightarrow{\mathrm{id} \times \cdot} & G \times G \\ {\scriptstyle \cdot \times \mathrm{id}} \downarrow & & \downarrow {\scriptstyle \cdot} \\ G \times G & \xrightarrow{\quad \cdot \quad} & G \end{array}$$

Algebraists love commutative diagrams, as they are a clean, visual way to represent equality between different formulas. The second and third axioms, however, are logically much more complex, as they involve existential quantifiers: there exists an element $1 \in G$... Is there a way to express such an existential statement as a commutative diagram? In order to do that, we need to change how we think of this axiom, and in fact, the operations in a group. Instead of saying *there exists* an identity element, we could say there is an additional *operation* $1 : \{*\} \to G$, satisfying:

- *Identity*: For all $a \in G$, $1(*) \cdot a = a$ and $a \cdot 1(*) = a$.

In other words, the map 1 sends the element $*$ of the one-element set to the group identity e from the previous version of the axiom. This can now be written as a commutative diagram:



Here the isomorphism $G \times \{*\} \to G$ is given by $(g, *) \mapsto g$, and similarly on the other side $(*, g) \mapsto g$.

We can do the same for the invertibility axiom, by introducing an *inverse operation* $(\_)^{-1} : G \to G$, which satisfies

- *Invertibility*: For all $a \in G$, $a \cdot a^{-1} = 1(*)$ and $a^{-1} \cdot a = 1(*)$.

Exercise 1.1.1. Write the inverse axiom as a commutative diagram.

In this style of thinking, a group can be illustrated as two sets, G and $\{*\}$, with operations drawn as arrows indicating the number of inputs, as in Figure 1.

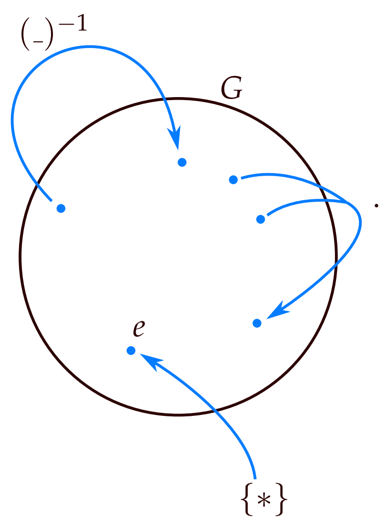


Figure 1: A schematic sketch of a group, with the arrows showing the operations.

A group is called *abelian* or *commutative* if, on top of the previous three axioms, there is an additional axiom:

- *Commutativity*: For each $a, b \in G$, $ab = ba$.

Exercise 1.1.2. Write the commutativity axiom as a commutative diagram.

Example 1.1.1. The following are some examples of groups that you would have encountered in a first abstract algebra course. For each of these examples, write down the identity and inverse maps.

1. The integers, $\mathbb{Z}$, under addition form a group.

2. The integers modulo n, $\mathbb{Z}_n$, under addition modulo n form a group. This is isomorphic to the quotient group $\mathbb{Z}/n\mathbb{Z}$.

3. The rotational and reflectional symmetries of a regular n-gon forms a group under composition, denoted D_n.

4. The set of non-zero complex numbers $\mathbb{C} \setminus \{0\}$ forms a group under multiplication.

5. The set of bijective funtions $f : \{1, \cdots, n\} \to \{1, \cdots, n\}$, that is, permutations of size n, forms a group under composition. This is known as the symmetric group and denoted S_n.

This course is ultimately about knot theory, so we take the opportunity here to introduce our first knotted structure, the braid group.

Definition 1.1.2. A *braid* on n strands – as in Figure 2 – is an embedding (injective map) of n intervals into the unit cube, connecting a set of n collinear marked points on the bottom face to n collinear endpoints directly above on the top face, satisfying:

- Each interval starts at a marked point on the bottom face, and ends at a marked point on the top face.
- The strands are strictly ascending (monotone increasing functions with respect to height in the cube).

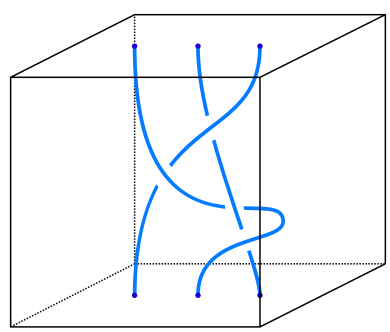


Figure 2: A braid on $n = 3$ strands.

Two braids are *equivalent* or *isotopic* if one can be continuously deformed into the other without passing the strands (intervals) through each other, such as the two braids in Figure 3. (We will learn the formal definition of isotopy later.)

If two braids have the same number of strands, we can stack them vertically, then vertically compress the stack of cubes into a single unit cube, as shown in Figure 4. This operation produces a new braid, is associative, and has corresponding identity and inverse operations, so it makes the set of n-strand braids into a group:

Example 1.1.2. The *braid group* B_n on n strands is the group of equivalence classes of n-strand braids, the operation given by vertical stacking followed by vertical compression, called *composition* (see Fig. 4).

Exercise 1.1.3. Sketch a proof that equivalence classes n-strand braids form a group under the stacking composition:

(a) Sketch a proof that the stacking composition is associative (up to braid equivalence).

(b) What is the identity braid with respect to the stacking composition? What is the inverse of a given braid? Justify why your proposed identity and inverse satisfy the corresponding axioms.

(c) For which n is B_n commutative?

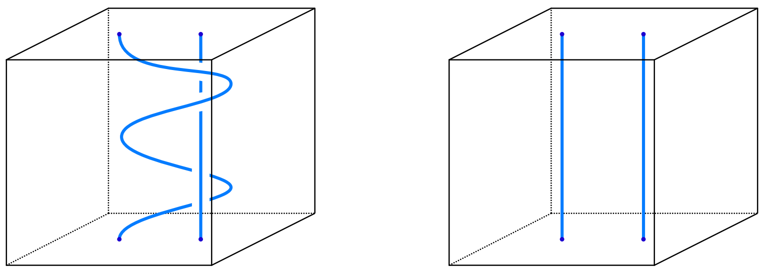


Figure 3: Two equivalent braids, as one can be deformed into the other without passing strands through each other.

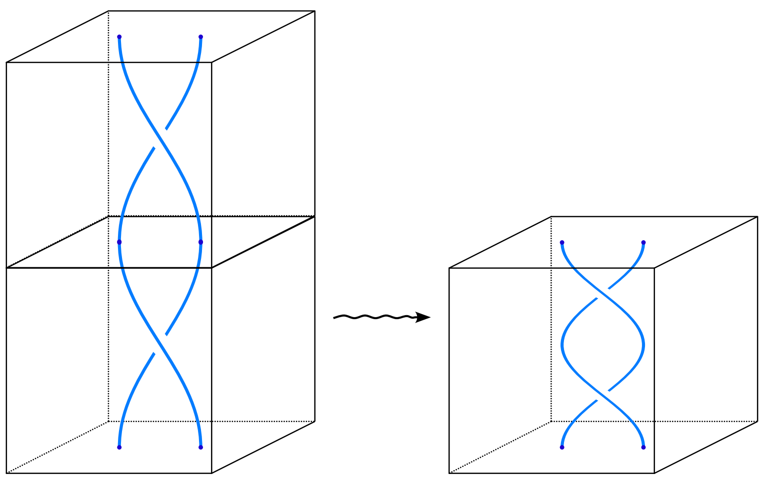


Figure 4: The composition operation is given by vertical stacking, followed by vertical compression.

From now on we refer to elements of the braid group as braids, rather than equivalence classes of braids. In figures we will no longer draw the boxes.

There is an important group homomorphism $\pi : B_n \to S_n$, which sends a braid to the permutation it induces on the set $\{1, \ldots, n\}$. Namely, for a braid β, number the bottom endpoints and top endpoints from left to right. Then $\pi(\beta)$ is the permutation which maps each bottom point to the top end of its braid strand.

Exercise 1.1.4. (a) What is $\pi(\beta)$ for the braid β in Figure 2? What is the order of $\pi(\beta)$?

(b) Prove that π is a homomorphism, and find its kernel. The kernel of this homomorphism is a subgroup of B_n called the *pure* braid group, denoted PB_n.

(c) Let β be the braid in Figure 2. What is the order of the element β? What is the order of the group B_3?

Exercise 1.1.5. Does any element of B_n other than the identity have a finite order?

Another algebraic structure you have likely encountered – perhaps in a linear algebra course even before you knew about groups – is a field. A field structure consits of two abelian group structures "stuck together":

Definition 1.1.3. A *field* is a set $\mathbb{F}$ with operations $+$ and $\cdot$, such that $(\mathbb{F}, +)$ is an abelian group with additive identity 0 and additive inverse denoted by $-$; and $(\mathbb{F} \setminus \{0\}, \cdot)$ is an abelian group with multiplicative identity 1 and inverse denoted $(\_)^{-1}$. In addition to the abelian group axioms for $+$ and $\cdot$, these operations must satisfy:

- *Distributivity*: For all x, y and z in $\mathbb{F}$, $x(y+z) = xy + xz$.

Distributivity is very important here: if the two group structures had nothing to do with each other, fields would be far less interesting, and far less useful. Distributivity describes how the two structures interact in a way that encodes phenomena that occur all over mathematics. The field structure is illustrated in Figure 5

Example 1.1.3. Common examples of fields include $\mathbb{Q}, \mathbb{R}, \mathbb{C}, \mathbb{Z}_p$, and their field extentions, such as $\mathbb{Q}[\sqrt{2}]$.

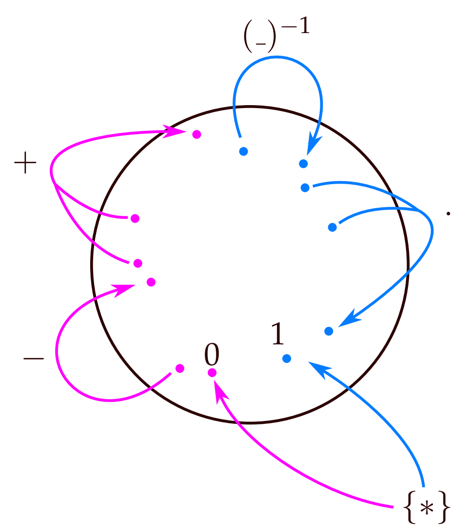


Figure 5: A schematic of the field structure.

A vector space adds another layer of complexity as it involves structures on multiple sets:

Definition 1.1.4. A *vector space* over a field $\mathbb{F}$ is an abelian group $(V, +)$ along a map $\cdot : \mathbb{F} \times V \to V$, called *scalar multiplication*, such that in addition to the field and abelian group axioms, the following properties are satisfied:

- *Distributivity*: For all λ and μ in $\mathbb{F}$, and all v, w in V, $(\lambda + \mu)v = \lambda v + \mu v$ and $\lambda(v + w) = \lambda v + \lambda w$
- *Compatibility of scalar and field multiplications*: $(\lambda\mu)v = \lambda(\mu v)$.
- *Compatibility with Identity*: For all v in V and the multiplicative identity $1 \in \mathbb{F}$, $1v = v$.

Exercise 1.1.6. Draw a schematic picture illustrating the structure of a vector space.

Other examples of algebraic structures you may have encountered are *monoids* (like groups but without inverses), rings (like fields but without multiplicative inverses, and not necessarily commutative multiplication), and algebras (vector spaces which are also a rings in a compatible way). If you have taken abstract algebra courses, you will have noticed that a seemingly small difference, like ommitting the multiplicative inverse requirement to obtain a ring from a field, can lead to dramatic differences in how the structures behave. We review the definitions of these structures as they will play a role in the topology we study later.

Definition 1.1.5. A *monoid* is a set M, such that there is a map $\cdot : M \times M \to M$, and a map $1 : * \to M$ satisfying the follwing axioms:

- *Associativity*: For all a, b and c in M, $(a \cdot b) \cdot c = a \cdot (b \cdot c)$
- *Identity*: For all $m \in M$, $1(*) \cdot m = m$ and $m \cdot 1(*) = m$.

Definition 1.1.6. A *ring* is a set R such that $(R, +)$ is an abelian group, $(R, \cdot)$ is a monoid, and in addition to the abelian group and monoid axioms, the Distributivity axiom holds:

- *Distributivity*: For all a, b and c in R, $a \cdot (b + c) = a \cdot b + a \cdot c$ and $(b + c) \cdot a = b \cdot a + c \cdot a$.

Definition 1.1.7. An *algebra* over a field $\mathbb{F}$ is a vector space A over $\mathbb{F}$ with scalar multiplication $\cdot : \mathbb{F} \times A \to A$, and an additional *algebra product* $* : A \times A \to A$ such that $(A, +, *)$ is a ring, and the product $*$ is bilinear.

Exercise 1.1.7. The bilinearity requirement for the algebra product ensures that the ring structure and the vector space structure on A interact well, rather than forming two separate structures. Write down the bilinearity requirement as a formula.

Matrices with elements in $\mathbb{F}$, with scalar multiplication and matrix multiplication are a prominent example of an algebra. Polynomials with scalar and polynomial multiplication are another. One example which will become an important prototype for the structures we study in this course is an algebra constructed from an arbitrary group, called the group algebra:

Definition 1.1.8. Given a group G and a field $\mathbb{F}$ the *group algebra FG* of G over $\mathbb{F}$, as a vector space, is given by formal linear combinations of elements of G: that is, the vector space with basis G. The algebra multiplication given by the bilinear extension of multiplication in G.

Example 1.1.4. The following are elements of the group algebra of $\mathbb{Q}S_3$:

$$\frac{4}{3}(123) - \frac{2}{3}(12)$$

and

$$(23) + \frac{1}{2}\,\mathrm{id}\,.$$

Multiplying them gives

$$\begin{aligned}\left(\frac{4}{3}(123) - \frac{2}{3}(12)\right)\left((23) + \frac{1}{2}\,\mathrm{id}\right) &= \frac{4}{3}(123)(23) - \frac{2}{3}(12)(23) + \frac{2}{3}(123) - \frac{1}{3}(12) \\ &= \frac{4}{3}(12) - \frac{2}{3}(123) + \frac{2}{3}(123) - \frac{1}{3}(12) \\ &= (12).\end{aligned}$$

Exercise 1.1.8. Is it true that if α and β in $\mathbb{Q}B_3$, then $(\alpha-\beta)(\alpha+\beta) = \alpha^2 - \beta^2$? What about in $\mathbb{Q}B_2$?

1.2 Generators and Relations

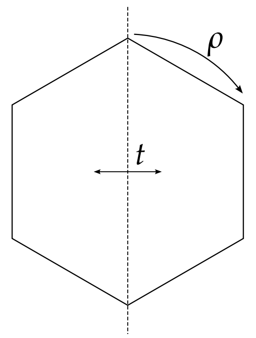
So far we've seen a variety of different algebraic structures. In many examples, one can describe the structure by pointing to an "atomic" set of elements, which are building blocks for all of the other elements, along with a set of rules which the products or combinations of atomic elements follow. This section formalises this idea: the building blocks are called *generators*, and the rules are called *relations*. We introduce these concepts in the context of groups, but the same principles apply to all of the algebraic structures we have seen.

Definition 1.2.1. If G is a group, and S is a subset of G, then the *subgroup generated by S* is the set of all elements of G which can be obtained as finite products of elements of S and their inverses. We denote this subset by $\langle S\rangle$, or simply list the elements of S inside of angle brackets, for example $\langle x, y\rangle$.

Exercise 1.2.1. Show that $\langle S\rangle$ is indeed a subgroup of G, and that it is the smallest subgroup which includes S.

Definition 1.2.2. If $\langle S\rangle = G$, then we say that S *generates* G, or that S is a *generating set* for G. The elements of S are called the *generators*. We say G is *finitely generated* if there exists a finite set which generates it.

Example 1.2.1. Recall that D_n is the set of symmetries of a regular n-gon. D_n is generated by the elements ρ and t, where ρ is a rotation by $2\pi/n$ and t is any reflectional symmetry of the n-gon, of which there are n to choose from.



Definition 1.2.3. A *relation* (between generators of G) is an equality which they satisfy in G.

Exercise 1.2.2. Prove that in D_n, the following relations hold:

$$\rho^n = 1 \qquad\qquad t^2 = 1 \qquad\qquad \rho t = t\rho^{-1}$$

There are only three equations listed above, but these are certainly not all the equations in D_n. For instance, $t = t^{-1}$ and $\rho^{n+1}t^{-1} = t^{101}\rho^{-1}$, and many more, can be deduced from the three relations above. In fact, any equality satisfied by the generators ρ and t can be deduced from the these three relations! This gives a complete description of D_n as the group generated by two elements – ρ and t – subject to the three rules which describe how products of these two generators work in the group. Such a description is called a *presentation*. This is an informal definition: to formally define group presentations, one needs to delve into the notion of free groups. For our purposes, this will suffice:

Definition 1.2.4. A *presentation* for a group G is a generating set S along with a set of relations R between elements of S, which imply all true equalities in G. When S and R form a presentation for G, we write

$$G = \langle S \mid R \rangle \quad \text{or} \quad G = \langle s_1, s_2, \cdots \mid r_1, r_2, \cdots \rangle.$$

If S and R are finite sets, we say that G is *finitely presented*.

A presentation is a convenient way of specifying a group – or other algebraic structure, as we will see later in the course. For instance,

$$H = \langle a, b \mid a^n = 1, b^2 = 1, ab = ba^{-1} \rangle$$

means that the group H is generated by a and b, subject to the three relations (and any further relations implied by these). In other words, the elements of H are products of the letters a and b and their inverses, and we can manipulate these products using the three relations.

However, a presentation not a unique way to describe a group: the same group can have many different presentations. For example, the relations in the definition of H above, are the same equations that we saw D_n satisfied, just with ρ renamed to a and t

renamed to b. This should make us strongly suspect that G is isomorphic to D_n. Unfortunately, it's not always easy to tell when two presentations represent *isomorphic* groups (from the perspective of group theory, essentially the same). For example,

$$\langle a, b \mid aba^{-1} = b^2, bab^{-1} = a^2 \rangle$$

is a fancy presentation of the trivial group! In general, the problem of deciding whether two groups given by finite presentations are isomorphic – called the *group isomorphism problem* – is *undecidable*, meaning there is no possible computer algorithm that could give a yes/no answer in finite time. It's also undecidable to determine whether an arbitrary finite presentation gives an abelian group, or the trivial group.

Another famous problem in this area is known as the *word problem*, which asks whether two words (products) made up of the generators and their inverses are equal in a finitely presented group. This is also undecidable in general, though it is decidable for many specific groups.

Exercise 1.2.3. Prove that the word problem is decidable in

$$G = \langle \rho, t \mid \rho^n = 1, t^2 = 1, \rho t = t\rho^{-1} \rangle.$$

That is, construct an algorithm which can decide whether two arbitrary words in ρ and t (and their inverses) are equal in G.

If we're to look for isomorphisms between finitely presented groups, we'd better be able to construct homomorphisms. This depends on the following critical observation:

Observation 1.2.1. To define a homomorphism ϕ from a group defined by a presentation $G = \langle S \mid R \rangle$ into another group H, it's enough to specify where the generators are mapped. The rest is determined by the multiplicativity of ϕ, as every element g of G can be written in terms of elements of S and their inverses. For example, if $g = s_1^5 s_2 s_1^{-3}$, then $\phi(g) = \phi(s_1)^5 \cdot \phi(s_2) \cdot \phi(s_1)^{-1}$.

However, we cannot freely "make up" a homomorphism by arbitrarily choosing the images of the generators: these images must respect the relations! For example if $s_1^5 s_2 = 1$ is a relation in R, then we need $\phi(s_1)^5 \cdot \phi(s_2) = 1_H$ to be true in H, as $\phi(1_G) = 1_H$ for any group homomorphism. However, since the relations in R imply all equalities between elements of G, it is sufficient for the ϕ-images of the generators to satisfy these relations.

This train of thought is summarised by the following theorem, a cornerstone of this course:

Theorem 1.2.1 (Von Dyck's Theorem)**.** *Given a group presentation $G = \langle S \mid R \rangle$ and a group H, a group homomorphism $\phi : H \to G$ is determined by the values $\phi(s_i)$ for $s_i \in S$. Furthermore, ϕ is a group homomorphism if and only if $\phi(s_i)$ satisfy the equations $\phi(r_i)$ in H, for all relations $r_i \in R$.*

Exercise 1.2.4. Using Von Dyck's theorem, prove that there is a group isomorphism between

$$H = \langle a, b \mid a^n = 1, b^2 = 1, ab = ba^{-1} \rangle$$

and

$$G = \langle \rho, t \mid \rho^n = 1, t^2 = 1, \rho t = t\rho^{-1} \rangle.$$

Exercise 1.2.5. Using Von Dyck's Theorem, construct a group F with a surjective but not injective homomorphism $\phi : F \to G$, with G as above.

Exercise 1.2.6. We stated without proof that the finite presentation

$$G = \langle \rho, t \mid \rho^n = 1, t^2 = 1, \rho t = t\rho^{-1} \rangle.$$

is in fact a finite presentation of D_n, the group of symmetries of a regular n-gon. Prove this theorem, i.e., prove that $G \cong D_n$.
Hint: Exercise 1.2.3 might help.

Example 1.2.2. Let's count how many homomorphisms there are from D_6 to $\mathbb{Z}_{12}$. Before you read on, try to count them yourself!

Note that D_6 is a multiplicative group, while $\mathbb{Z}_{12}$ is written additively. So, the defining property of a group homomorphism $\phi : D_6 \to \mathbb{Z}_{12}$ is that for all $x, y \in D_6$, we have $\phi(xy) = \phi(x) + \phi(y)$.

By Exercise 1.2.6, we know that $D_6 = \langle \rho, t \mid \rho^6 = 1, t^2 = 1, \rho t = t\rho^{-1} \rangle$. By Von Dyck's Theorem, this means that homomorphisms ϕ are given by solutions $(\phi(\rho), \phi(t))$ in $\mathbb{Z}_{12}$ of the system of equations

$$\{6\phi(\rho) = 0, \quad 2\phi(t) = 0, \quad \phi(\rho) + \phi(t) = \phi(t) - \phi(\rho)\}.$$

Notice that the last equation simplifies to $2\phi(\rho) = 0$, which also implies $6\phi(\rho) = 0$. So, the set of three equations above is equivalent to

$$\{2\phi(\rho) = 0, \quad 2\phi(t) = 0\}$$

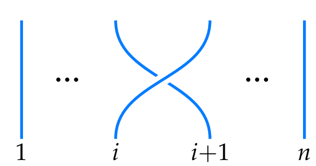
The following famous theorem – which we take on faith here – is due to Artin:

Theorem 1.2.2. *The braid group, B_n has a finite presentation*

$$B_n = \left\langle \sigma_1, \sigma_2, \cdots, \sigma_{n-1} \;\middle|\; \begin{array}{ll} \sigma_i\sigma_j\sigma_i = \sigma_j\sigma_i\sigma_j & \text{if } j = i \pm 1 \\ \sigma_i\sigma_j = \sigma_j\sigma_i & \text{if } |j-i| > 1 \end{array} \right\rangle,$$

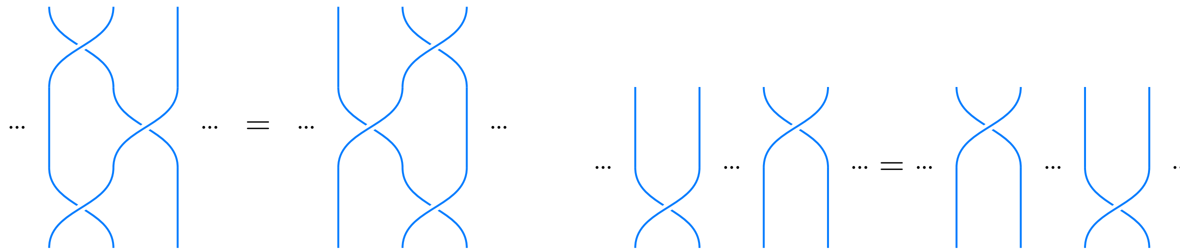
where

$\sigma_i =$



.

Proving that this is a presentation for B_n includes proving that the relations in this presentation, as shown below, are equivalent to ambient isotopy.



(a) $\sigma_i\sigma_{i+1}\sigma_i = \sigma_{i+1}\sigma_i\sigma_{i+1}$ (b) $\sigma_i\sigma_j = \sigma_j\sigma_i$

Exercise 1.2.7. How many relations are there in B_4? How about B_n?

There is a presentation for the symmetric group, which demonstrates how it is a quotient of the braid group:

Theorem 1.2.3. *The following is a finite presentation of the symmetric group:*

$$S_n = \left\langle s_1, s_2, \cdots, s_{n-1} \;\middle|\; \begin{array}{ll} s_is_js_i = s_js_is_j & \text{if } j = i \pm 1 \\ s_is_j = s_js_i & \text{if } |j-i| > 1 \\ s_i^2 = 1 & \text{for all } i \end{array} \right\rangle,$$

Example 1.2.3. Another important example of finitely presented groups is finitely generated *free groups*, which have a presentation with finitely many generators and *no relations*.

For example, the free group on one generator, $FG\langle x \rangle$, is isomorphic to the group $\mathbb{Z}$ of integers. In general the elements of the free group are *words* (non-commutative monomials) in the generators and their inverses.

Finite presentations can be constructed for any other algebraic structure you know, not just groups: monoids, rings, vector spaces, algebras, and more. The notion of *generation* needs to be modified to whatever makes sense in a given structure: for example, in a vector space, generation is via linear combinations; in a ring, it is via addition, subtraction and multiplication. A finitely generated free structure is one with a finite generating set, and no relations in the presentation. To test your understanding, think about the following questions:

Exercise 1.2.8. All vector spaces are free! Which fundamental theorem in linear algebra tells you this?

Exercise 1.2.9. What is the difference between:

- the group algebra over $\mathbb{Q}$ of the free group on k generators, $\mathbb{Q}FG\langle s_1, \ldots, s_k \rangle$, and
- the free $\mathbb{Q}$-algebra on k generators, $FA\langle s_1, \ldots, s_k \rangle$?

1.3 Categories

When encountering a new kind of mathematical object, we usually immediately learn about the maps between them that respect their defining structures. Groups have group homomorphisms, which play nice with their binary operations. Vector spaces have linear maps, which cooperate with addition and scalar multiplication. Topological spaces have continuous functions, which send nearby points to nearby points. And sets, with no additional structure to speak of, are happy as long as one element is sent to only one element.

Indeed, a lot may be learned from studying the structure-preserving maps. For example, studying the homomorphisms from a group to matrix groups, or an *endomorphism groups* of vector spaces, is the premise of *representation theory*. Often, the properties of the structure preserving maps are more important than the elements of the objects themselves. For example, you may recall that most of linear algebra considers linear maps, rather than elements in vector spaces.

This line of thinking is formalised with the notion of a **category**: a class of mathematical objects, together with the structure-preserving maps between them. It is perhaps best understood as the *context* within which we are viewing the objects.

Definition 1.3.1. A *category* $\mathcal{C}$ consists of:

- a set of *objects*, $\mathrm{ob}(\mathcal{C})$,
- for each A and B in $\mathrm{ob}(\mathcal{C})$, a set of *morphisms* from A to B, $\mathrm{Hom}(A, B)$,
- for each A in $\mathrm{ob}(\mathcal{C})$, an element 1_A of $\mathrm{Hom}(A, A)$, called the *identity* on A,
- for each A, B, and C in $\mathrm{ob}(\mathcal{C})$, an operation called *composition*

$$\begin{aligned} \mathrm{Hom}(B, C) \times \mathrm{Hom}(A, B) &\to \mathrm{Hom}(A, C) \\ (g, f) &\mapsto g \circ f, \end{aligned}$$

satisfying the following axioms:

- *Associativity*: Whenever f, g and h are morphisms that can be composed, we have $(h \circ g) \circ f = h \circ (g \circ f)$.
- *Identity*: For each $f \in \mathrm{Hom}(A, B)$, we have $f \circ 1_A = f = 1_B \circ f$.

Exercise 1.3.1. Can you draw a schematic picture, like those in the previous sections, illustrating the structure of a category? Hint: there will be many sets.

Example 1.3.1. The following are some of the motivating examples for categories:

- The category of groups, **Grp**, whose objects are groups, and whose morphisms are group homomorphisms. Composition is the composition of homomorphisms.
- The category **Vect** of vector spaces, whose objects are vector spaces, and morphisms are linear maps. Composition is composition.
- The category **Top**, whose objects are topological spaces, and whose morphisms are continuous functions. Composition is composition.

- The category **Set** whose objects are sets[1], and whose morphisms are set functions. Composition is composition.

But note that not all categories' morphisms are homomorphisms. In fact, not all categories' morphisms are functions:

Example 1.3.2. There is a category $\mathbf{Order}(\mathbb{Z})$, whose objects are integers, and whose morphisms are true propositions $n \leq m$. Composition of morphisms is given by

$$(n \leq m) \circ (m \leq k) = n \leq k.$$

In othe words, for any given two objects $n, m \in \mathbb{Z}$, the category $\mathbf{Order}(\mathbb{Z})$ has a single morphism $n \to m$ in $n \leq m$, and no morphisms $n \to m$ otherwise. Verify that this category has identity morphisms, explain what composition is (there is only one possibility), and verify that composition is associative.

Exercise 1.3.2. Is it possible for two categories to have the same objects but different morphisms (and composition)? What about the other way, can two categories only differ in their objects?

Example 1.3.3. Given an arbitrary group G, there is a category BG with one object, $*$, and morphisms given by elements of G. Composition of objects is given by multiplication in G. Check that this is an example of a category.

The last two examples can be seen as evidence of the category theory mantra that the morphisms of a category (and the operations that tell you how they compose) are more important than the objects. The names **Grp**, **Ring**, **Top**, are counterintuitive in this sense, but they have stuck, and it's easier to write **Top** than **ContinuousMap**.

Exercise 1.3.3. Let G and H be two groups, and BG and BH their single-object categories as in Example 1.3.3. Let $\mathbf{Vect}_{\mathbb{F}}$ be the category of vector spaces over a field $\mathbb{F}$. Show that functors $BG \to BH$ correspond exactly to group homomorphisms $\varphi : G \to H$.

Example 1.3.4. The braid category, $\mathbf{Braid} = \sqcup_n BB_n$ has objects indexed by positive integers, and for $n \in \mathbb{Z}_{>0}$, $\mathrm{Hom}(n, n) = B_n$. There are no morphisms between different objects: $\mathrm{Hom}(n, m) = \emptyset$ for $n \neq m$. Composition of morphisms is given by braid multiplication in the appropriate braid group. While this example is a little silly, you will see how it can be meaningfully extended with more morphisms to produce a category of tangles.

Is there a category of categories? To answer this question, we ought to follow the mantra above, and understand structure preserving morphisms between categories. These are called *functors*, and map objects to objects and morphisms to morphisms in a way preserving all the structure of a category.

[1]Famously, ob(**Set**) is not a set like our definition seems to require. To oversimplify, there are too many sets! Technically, Definition 1.3.1 is the definition of a *small category*, where the objects form a set and so to the morphisms between them. The category of sets is not a small category. We ignore this issue from here on.

Definition 1.3.2. Let $\mathcal{C}$ and $\mathcal{D}$ be two categories. A *functor* $F : \mathcal{C} \to \mathcal{D}$ is a map which

- maps each object $A \in \mathrm{ob}(\mathcal{C})$ to an object $F(A) \in \mathrm{ob}(\mathcal{D})$, and
- maps each morphism $f \in \mathrm{Hom}(A, B)$ to a morphism $F(f) \in \mathrm{Hom}(F(A), F(B))$.
- These mappings respect identities and composition. That is, $F(1_A) = 1_{F(A)}$ for all objects A in $\mathcal{C}$, and $F(g \circ f) = F(g) \circ F(f)$ whenever the two morphisms g and f may be composed in $\mathcal{C}$.

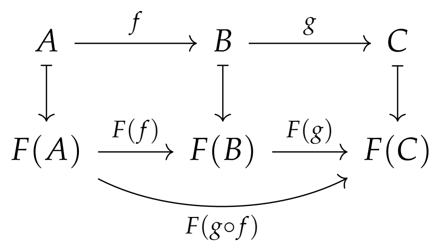


Figure 7: The composition condition in the definition of a functor, expressed in a commutative diagram.

With functors as the morphisms, categories do indeed form a (large) category.

Example 1.3.5 (Forgetful functors)**.** Many algebraic structures, such as groups, monoids and fields, are sets endowed with some additional structure. Similarly, morphisms between them are first and set functions, which also respect the additional structure. The *forgetful functors* from the categories of these structures to the category send such objects to their underlying sets, and maps between them to their underlying set maps. In some sense, they do nothing: they simply tell us to view objects and maps in a different context, and *forget* about the extra structure.

Exercise 1.3.4. Consider the forgetful functor **Grp** $\to$ **Set**. Is it surjective on morphisms?

Exercise 1.3.5. There is a forgetful functor **Vect** $\to$ **Set**. Can you construct a functor **Set** $\to$ **Vect**?

Example 1.3.6 (Free functors)**.** If you solved Exercise 1.3.5, you have most likely constructed an example of a *free functor*. Any set S can be used as a generating set for a free group, a free algebra, or as a basis for a vector space. By Von Dyck's theorem, a homomorphism from a free structure is uniquely and freely determined by choosing the images of the generators. For example, any set map $f : S \to T$ lifts to a homomorphism between the free groups $\tilde{f} : FG\langle S\rangle \to FG\langle T\rangle$. In other words, this method gives functors from **Set** to **Grp**, **Vect**, and so on: these are called *free functors*.

Exercise 1.3.6. Recall that the set of endomorphisms of a vector space V forms a group under composition, called $\mathrm{End}(V)$. A *representation* of a group G is a vector space V together with a group homomorphism $\rho : V \to \mathrm{End}(V)$. Show that functors $BG \to \mathbf{Vect}_{\mathbb{F}}$ correspond exactly to group representations (V, ρ).

Chapter 2

Algebraic Structures for Knots and Tangles

"Knots are the wrong objects to study in knot theory." – Dror Bar-Natan

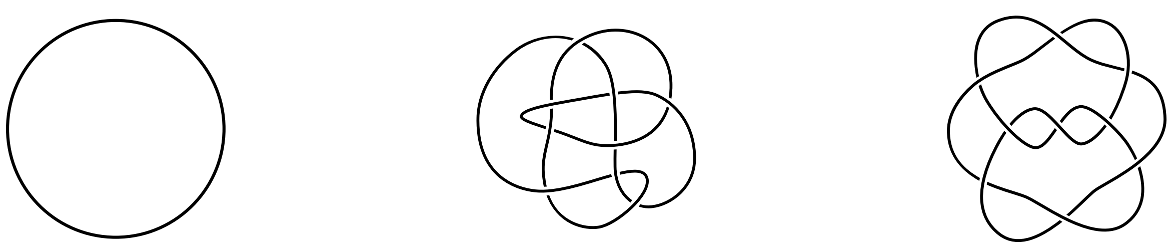
Of course, knot theory concerned with knots, but from the algebraic point of view of this course, we will see that knots are not an ideal knotted object to study: they have little in the way of algebraic structure, and are not finitely presented in any sense.

They do, however, have many better-behaved cousins. We have already seen Artin's Theorem, which gives a finite presentation for each braid group. In this chapter we also get to know tangles, which can be thought of as the atomic building blocks of knots.

To honour the tradition of the field, we do start out with good old knots.

2.1 Knots and links

A knot, intuitively, is a (potentially) tangled up circular piece of string, such as the examples shown in Fig. 8.



(a) 0_1 (b) 9_{33} (c) 9_{35}

Figure 8: Some examples of knots.

To formalise this mathematically, we describe a knot as an embedding of a circle in 3-dimensional space:

Definition 2.1.1. A knot is an *embedding* (injective continuous[1] function) $K : S^1 \to \mathbb{R}^3$, from the circle S^1 into $\mathbb{R}^3$, or into the 3-sphere S^3.

Two knots are considered equivalent, or "the same" if one can be moved around in space, without tearing the string, to look exactly like the other. For example the two

[1]In general, a topological embedding also needs to be a homeomorphism onto its image. Since S^1 is compact, this follows from being injective and continuous.

embeddings in Fig. 9 are not the same, but since the former can be deformed by a small twist into the latter, we want to say they represent the same knot. In order to make this notion of equivalence precise, we need to recall some basic notions from algebraic topology.

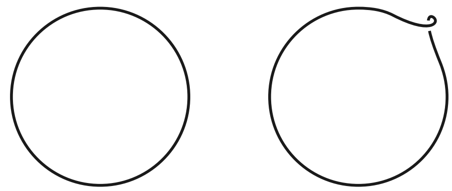


Figure 9: A small twist changes the embedding but not the knot.

The first such notion is a *homotopy*, which is the precise notion for a continuous deformation of a function.

Definition 2.1.2. If $f : X \to Y$ and $g : X \to Y$ are continuous functions between topological spaces, then a *homotopy* between f and g is a continuous function $H : X \times [0,1] \to Y$ such that $H(x,0) = f(x)$ and $H(x,1) = g(x)$. If there exists a homotopy between f and g, we say f is *homotopic* to g.

Since the intuitive notion of homotopy is a gradual deformation over time, we refer to the interval $[0,1]$ as *time*, and elements of the interval as *moments* of the homotopy.

Exercise 2.1.1. Show that homotopy is an equivalence relation on the set of continuous functions $X \to Y$.

Is homotopy a good notion of equivalence for knots? Say H is a homotopy between two knots K_1 and K_2: the issue is that while K_1 and K_2 are embeddings, there is no requirement that all the other moments during the homotopy are. Intuitively, a homotopy will potentially pass parts of the knot through each other, and thus all knots are equivalent under homotopy.

Exercise 2.1.2. (a) Give an example of two non-homotopic functions.

(b) Given any two knots K_0 and K_1, write down a homotopy between them.

To more faithfully describe our intuitive understanding of knot equivalence, we need to consider homotopies which restrict to embeddings at every moment:

Definition 2.1.3. An *isotopy* is a homotopy through embeddings, that is, a homotopy $H : X \times [0,1] \to Y$ such that for all $t \in [0,1]$, $H(x,t)$ is an embedding of X in Y.

Unfortunately, knots are still all equivalent under isotopy: this is less obvious, and known as the "lazy seamstress trick". Have you ever got a pesky knot on your thread when sawing? Were you tempted to – rather than spending the time to untangle it – just pull it as tight, until it practically disappears? This is an isotopy: given any knot, you can pull all the knotty part to a point while extending the rest to a plain circle: the unknot.

a figure illustrating this

So, we want isotopies for knot equivalence, but need to make the seamstress trick illegal. There are multiple equivalent ways of adjusting the notion of isotopy to achieve this. Here we'll go with *ambient ispotopy*: the word *ambient* means relating to the surroundings of something. Indeed, ambient isotopy is a continuous deformation of not just the knot, but the entire 3-dimensional space in which the knot is embedded. Such a deformation of $\mathbb{R}^3$ "takes the knot with it", and it is a homeomorphism at every moment, hence won't collapse the knot to a point:

Definition 2.1.4. Given two knots K and L, an *ambient isotopy* between K and L is a continuous function $F : \mathbb{R}^3 \times [0, 1] \to \mathbb{R}^3$ satisfying

- $F_0 : \mathbb{R}^3 \times \{0\} \to \mathbb{R}^3$ is the identity map,
- at each moment $F_t : \mathbb{R}^3 \times \{t\} \to \mathbb{R}^3$ is a homeomorphism, and
- $F_1 \circ K = L$.

Exercise 2.1.3. Show that ambient isotopy is an equivalence relation on knots.

Exercise 2.1.4. Given an ambient isotopy taking K to L, derive a corresponding (plain) isotopy between them.

A natural extension of the concept of a knot is a *link*, which involves possibly more than one loop, possibly tangled together. Ambient isotopy is also the right notion to capture the intuitive equivalence of links:

Definition 2.1.5. A *link* is a continuous embedding of a finite disjoint union of circles into $\mathbb{R}^3$, up to ambient isotopy.

2.2 Link diagrams

Link diagrams help translate knot theory from continuous functions – the realm of analysis and geometry – to diagrams that you can describe combinatorially, and even feed to a computer: the realm of algebra and discrete mathematics. In order to define knot diagrams, we need to talk about projections.

Definition 2.2.1. A projection π of a link L onto a plane $P \subseteq \mathbb{R}^3$ is *regular*, or *generic*, if:

- All but finitely many points of the image of πL have only one pre-image in $\sqcup S^1$.
- The points which have multiple preimages (called the *singular points*) are all *transverse double points*. That is, singular points in the image have two pre-images each, and lie at the intersection of two projected arcs (which cross each other, rather than meeting tangentially). These transverse double points are called the *crossings*.

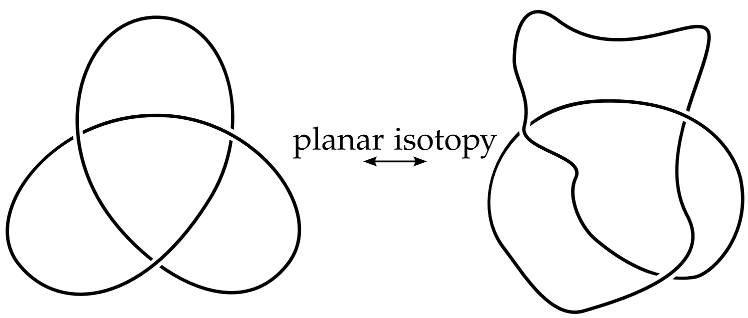
All of the images of knots you have seen on the previous pages were in fact knot diagrams: projections where at the crossings, the arc going under was drawn broken, and the arc going over was drawn solid. Your brain has likely interpreted these automatically as 3-dimensional objects.

Definition 2.2.2. A link diagram of a link L is a is a regular projection of L onto a plane (a copy of $\mathbb{R}^2$), with additional over/under strand information at the double points. When link diagrams are drawn, the under strand at each crossing is indicated by breaking the arc.

Exercise 2.2.1. Show that a link diagram determines the link up to ambient isotopy.

What about the converse? Does the knot uniquely determine the diagram? The answer is clearly no: if you move the projection plane around, or apply ambient isotpoy to the knot, the diagram will move around and change. Some of this movement is captured by the notion of *planar isotopy*:

Definition 2.2.3. Two knot diagrams D and D' are *planar ispotopic* is there is an ambient isotopy of the plane $F : \mathbb{R}^2 \times [0, 1] \to \mathbb{R}^2$, such that F_0 is the identity and $F_1(D) = D'$.



However, since each moment of a planar isotopy is a homeomorphism of $\mathbb{R}^2$, planar isotopy will never change the number of – create or eliminate – crossings. Up to planar isotopy, a knot diagram is simply a 4-valent plane graph: a graph with four edge endings at each vertex, embedded in the plane. Since the vertices represent crossings, the incident edge endings are equipped with a cyclic orientation, and two opposite edge endings are marked "over". This information is easily encodable for computational purposed, and when we talk about knot diagrams, this is the information we care about. We say that two knot diagrams are the same if they differ only by planar isotopy. In other words, a knot diagram is really an equivalence class of diagrams under planar isotopy.

There is more complication, however: we have already seen multiple different diagrams representing the same knot, and it is not usually easy to tell when they do. In fact, there are infinitely many diagrams for each knot, for example by adding arbitrary numbers of little twists.

Reidemeister's Theorem gives a characterisation of when two diagrams represent ambient isotopic links, in terms of a short list of local "moves". This complete combinatorial description of links opened up countless possibilities for their algebraic study. It is easy enough to see that the moves in the theorem, known as the Reidemeister moves and shown below, don't change the ambient isotopy type. It is much more substantial to prove that the Reidemeister moves are sufficient to generate all isotopies.

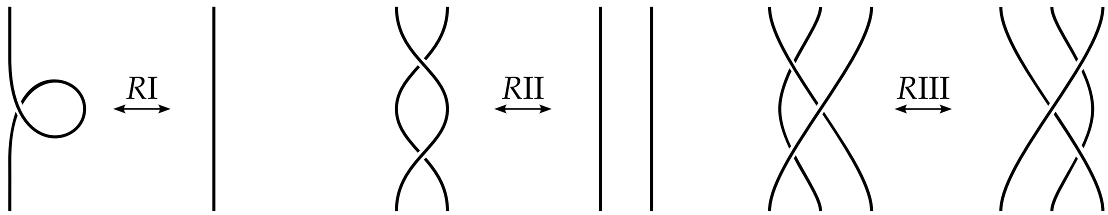


Figure 10: The Reidemeister moves.

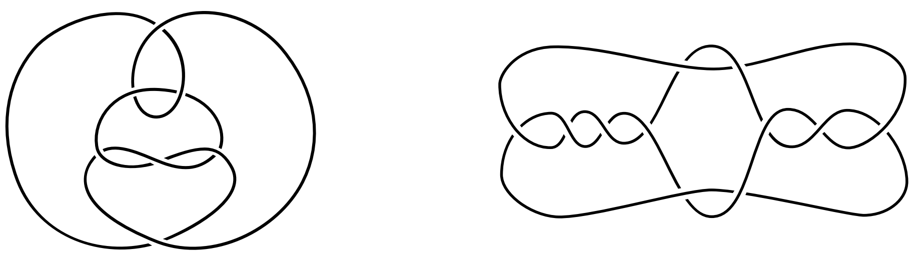
What is harder to see is that these are all the moves that are necessary:

R1 needs the + and - kinks. Minor but I'd prefer 1,2,3 to I, II, III — Z.

Theorem 2.2.1 (Reidemeister)**.** *Two knot diagrams D and D' represent the same knot if and only if D and D' are related by a finite sequence of Reidemeister moves (and planar isotopy).*

A proof of Reidemeister's theorem can be found in Chapter 4 of [Mur96].

Reidemeister's theorem is an existence result: it doesn't tell us how long or how complicated the sequence of moves relating two equivalent diagrams may be. It is not even easy to tell whether a given knot diagram is equivalent to the unknot.



Exercise 2.2.2. Untangle the knots above on the using Reidemeister moves and planar isotopies. (The one on the left is easier, from [Ada94]; the one on the right is famously tricky and called the Goeritz unknot.)

cite Ben et al? — Z.

2.2.1 Orientations and Mirror Images

The parametrisation of S^1 – for example, the counterclockwise direction along the unit circle in the complex plain – induces an orientation (direction) along a knot, and more generally, along each component of a link. An ambient isotopy will always preserve this orientation. A weaker notion of equivalence would be to allow both ambient isotopies and parameterisation-reversals of S^1. Knots under this more permissive equivalence are called *unoriented* knots.

In this course, we are working primarily with oriented knots and links. In pictures, it is often useful to indicate the direction (corresponding to counterclockwise in the pre-image) with arrows along the link diagram.

Definition 2.2.4. The *reverse* is a unary knot operation taking a knot K to the knot K pre-composed with a parameterisation-reversal of S^1. Knots for which $K = K^*$ are called *reversible*. The reverse of K is denoted by K^*.

That is, the reverse of K the same knot, with the opposite orientation: in a knot diagram this amounts to simply reversing the arrow.

In general, K^* is not ambient isotopic to K, but there are many reversible knots, which can be can be continuously moved back to their original orientations. In fact, the simplest knot that isn't invertible is the knot 8_{17}, shown below.

Add figure of 8_17. — Z.

Exercise 2.2.3. Prove that the unknot and the trefoil are reversible by finding ambient isotopies to their reverses.

Definition 2.2.5. The *mirror image* is a unary knot operation taking a knot K to K post-composed with a reflection of $\mathbb{R}^3$. The mirror image of K is denoted $\overline{K}$. Knots for which $K = \overline{K}$ are called *amphichiral*, and knots for which $K \neq \overline{K}$ are called *chiral*.

Diagramatically, the mirror image can be achieved by swapping the over and under strands at each crossing of the knot diagram. This corresponds to reflecting along a plane parallel to the projection plane.

Exercise 2.2.4. Prove that the figure-8 knot is amphichiral.

Both the mirror image and the reverse are *involutions*: doing either operation twice returns the original knot.

Definition 2.2.6. Knots for which $K = \overline{K^*}$ are called *minus-amhichiral*. Knots for which $K = \overline{K} = K^* = \overline{K^*}$ are called *fully symmetric*. Knots for which K, K^*, $\overline{K}$ and $\overline{K^*}$ are all different are called *totally asymmetric*.

We summarise the symmetry types of knots in the table below:

Name	**Symmetry**			
totally asymmetric	K	K^*	$\overline{K}$	$\overline{K^*}$
invertible but chiral	$K = K^*$		$\overline{K} = \overline{K^*}$	
(plus-)amphichiral, noninvertible	$K = \overline{K}$		$K^* = \overline{K^*}$	
minus-amphichiral, noninvertible	$K = \overline{K^*}$		$K^* = \overline{K}$	
fully symmetric	$K = K^* = \overline{K} = \overline{K^*}$			

Exercise 2.2.5. Show that the figure-eight knot is fully symmetric.

2.3 An Algebraic Structure on Knots

Knots carry an algebraic structure, with respect to the binary *connected sum* operation.

Definition 2.3.1. The *connected sum* of two (oriented) knots K_1 and K_2 is the knot produced by removing a small arc from each, and connecting the four ends in a way compatible with the orientations. The connected sum of K_1 and K_2 is denoted $K_1 \# K_2$.