Algebraic Knot Theory

Zsuzsanna Dancso
Damian Lin
Nelson Odins-Jones

March 28, 2025

Contents

Introduction **4**

Acknowledgements **6**

1 Algebra **7**
1.1 Algebraic Structures 7
1.2 Generators and Relations 13
1.3 Categories 17

2 Algebraic Structures for Knots and Tangles **22**
2.1 Knots and links 22
2.2 Link Diagrams 24
2.3 An Algebraic Structure on Knots 28
2.4 Tangles: a finitely presented monoidal category 29

3 Gradings and Filtrations **42**
3.1 Gradings 42
3.2 Filtrations 44
3.3 The Associated Graded Functor 45
3.4 The Vassiliev-Goussarov Filtration for Knots 48
3.5 Chord Diagrams 52
More Exercises for Chapter 3 55

4 Formality **56**
4.1 Formality Maps 56
4.2 Identifying Associated Graded Structures 57
More Exercises for Chapter 4 58

5 The Kontsevich Integral **59**
5.1 Prelude: a Complete Invariant for B_2 59
5.2 Definition 60
5.3 Calculation for the Positive Crossing 62
5.4 Invariance 63
More Exercises for Chapter 5 63

6 Drinfeld Associators **64**
6.1 Tangles as a finitely presented tensor category 64
6.2 The Kontsevich integral for tangles: parametrised tensor products 64
6.3 Parenthesised tangles 64
6.4 Formality for parenthesised tangles 64
6.5 A combinatorial Kontsevich Integral 64
6.6 Drinfel'd associators 64
More Exercises for Chapter 6 64

A Solutions to Selected Exercises **65**

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

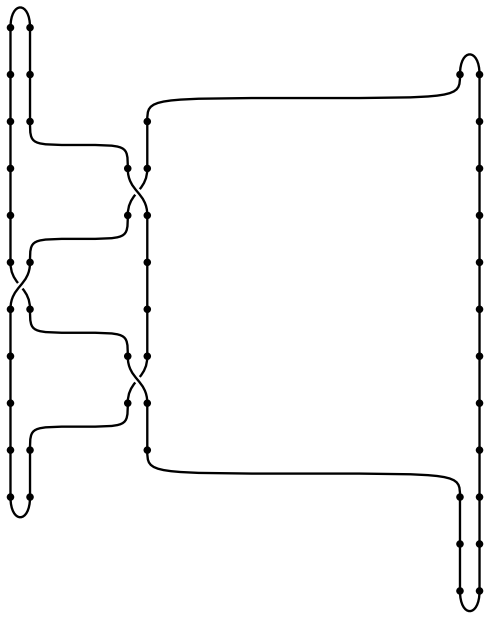
Drinfeld's work accomplishes the same goals via very different routes: Tamarkin, Bar-Natan and Fresse showed that *universal quantum invariants* of knots are in one-to-one correspondence with *Drinfeld associators*.

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:

$$\begin{array}{ccccc} G \times \{*\} & \xrightarrow{\mathrm{id}\times 1} & G \times G & \xleftarrow{1\times \mathrm{id}} & \{*\} \times G \\ & \searrow^{\cong} & \downarrow{\scriptstyle \cdot} & \swarrow_{\cong} & \\ & & G & & \end{array}$$

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.

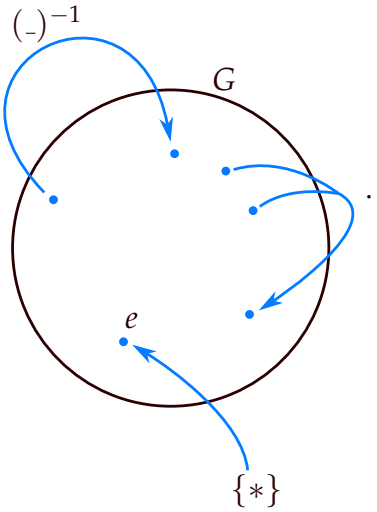


Figure 1: A schematic sketch of a group, with the arrows showing the operations.

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 the following conditions:

- 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).

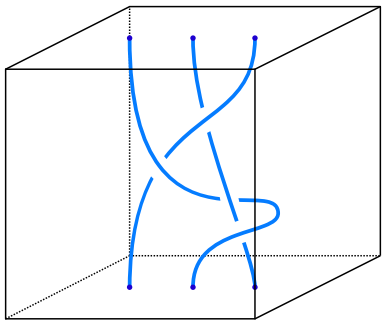


Figure 2: A braid on $n = 3$ strands.

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?

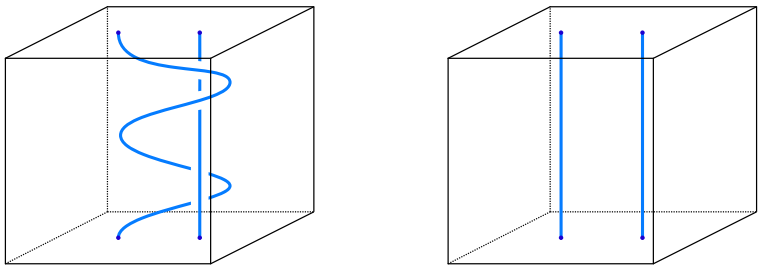


Figure 3: Two equivalent braids, as one can be deformed into the other without passing strands through each other.

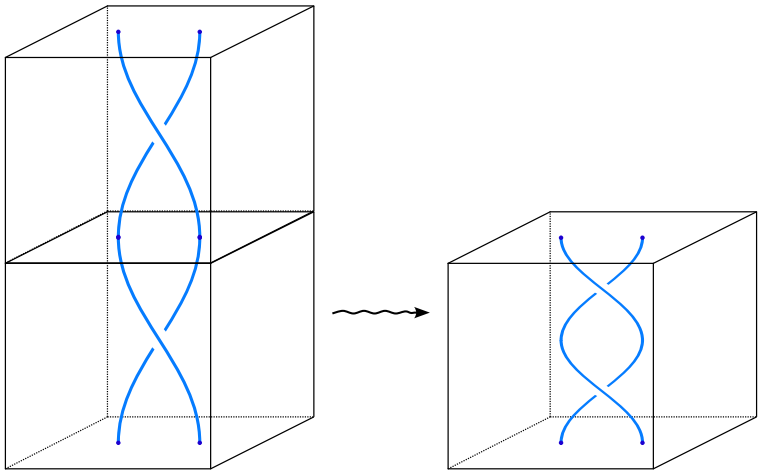


Figure 4: The composition operation is given by vertical stacking, followed by vertical compression.

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 $\mathbb{Q}[\sqrt{2}]$.

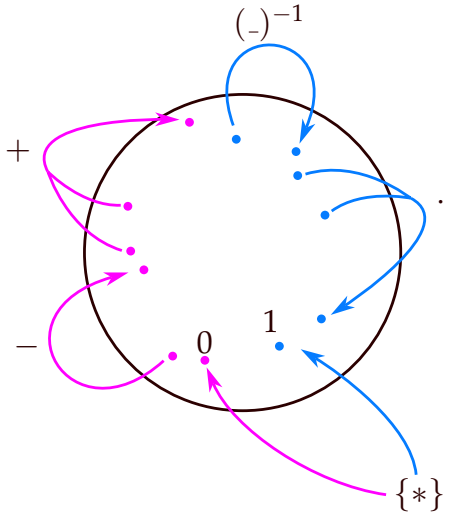


Figure 5: A schematic of the field structure.

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

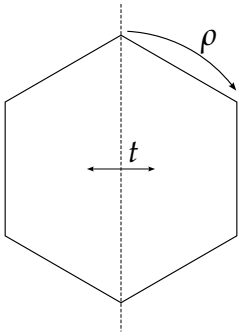
So far we've seen a variety of different algebraic structures. In many examples, one can describe the structure by pointing to an "atomic" set of elements, which are building blocks for all of the other elements, along with a set of rules which the products or combinations of atomic elements follow. 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:

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

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

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

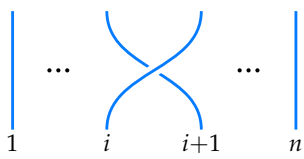
The following famous theorem – which we take on faith here – is due to Artin:

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

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

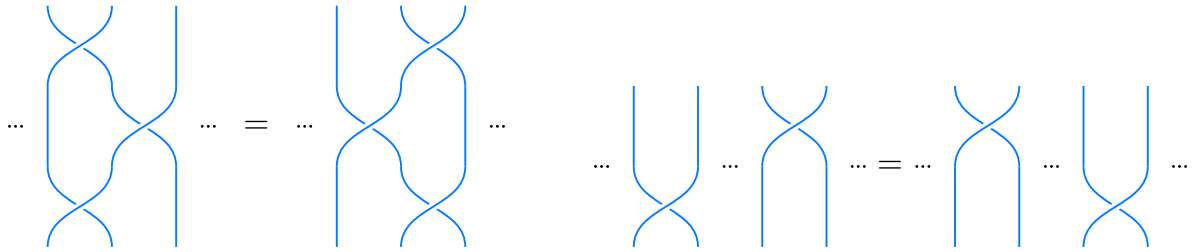
where

σ_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:

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

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 differ only 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 justification for 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.

Definition 1.3.2. An **isomorphism** is a morphism with an inverse – that is, a morphism $f : A \to B$ for which there exists another morphism $g : B \to A$ with $g \circ f = \mathrm{id}_A$ and $f \circ g = \mathrm{id}_B$.

[1]Famously, $\mathrm{ob}(\mathbf{Set})$ 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 too do the morphisms between them. The category of sets is not a small category. We ignore this issue from here on.

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.

Definition 1.3.3. 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 $\mathcal{C}$.

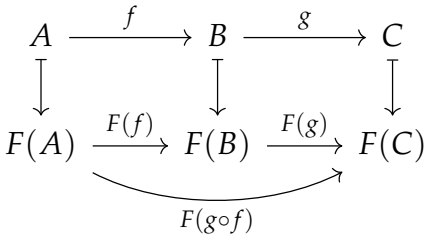


Figure 7: The composition condition in the definition of a functor, expressed in a commutative diagram.

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 foremost 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

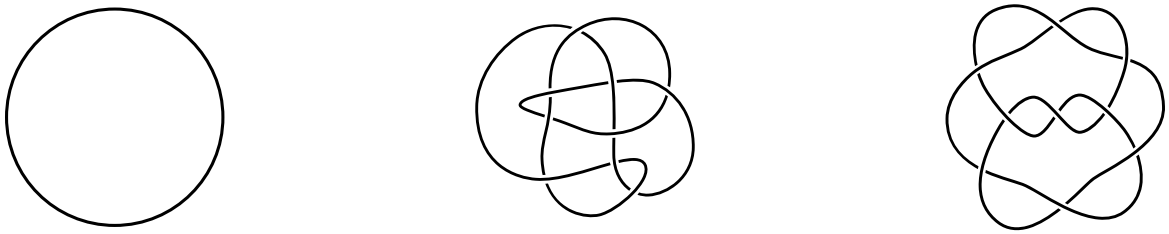
Of course, knot theory is concerned with knots, but from the algebraic point of view of this course, we will see that knots are not an ideal knotted object to study: they have little in the way of algebraic structure, and are not finitely presented in any sense.

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 – continuous injective mapping – of a circle in 3-dimensional space. Unfortunately continuous functions can be quite badly behaved, and this definition includes phenomena called *wild knots*. This can be avoided in various equivalent ways: by considering only differentiable embeddings, or only piecewise linear embeddings, or requiring that each part of the image is separated by a thin, constant width neighbourhood from other parts. We choose this third approach as it is conceptually simplest.

some picture or explanation of wild knots — Z.

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

[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.

from the circle S^1 into $\mathbb{R}^3$, or into the 3-sphere S^3. A knot K is *tame* if there exists a small number ε for which the image of K does not intersect an ϵ-width nighbourhood of itself. Otheriwse, K is *wild*.

Need a picture of a tame know with the epsilon tunnel around it. — Z.

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 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.

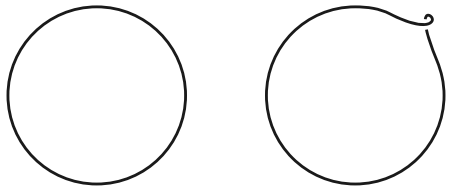


Figure 9: A small twist changes the embedding but not the knot.

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 sewing? 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 trick would be nice — Z.

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*.

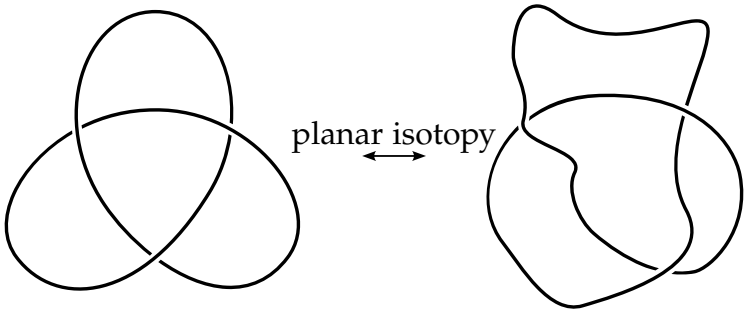
All of the images of knots you have seen on the previous pages were in fact knot diagrams: projections where at the crossings, the arc going under was drawn broken, and the arc going over was drawn solid. Your brain has likely interpreted these automatically as 3-dimensional objects.

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.

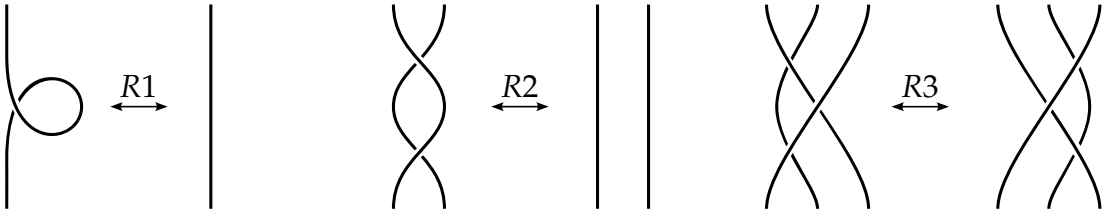


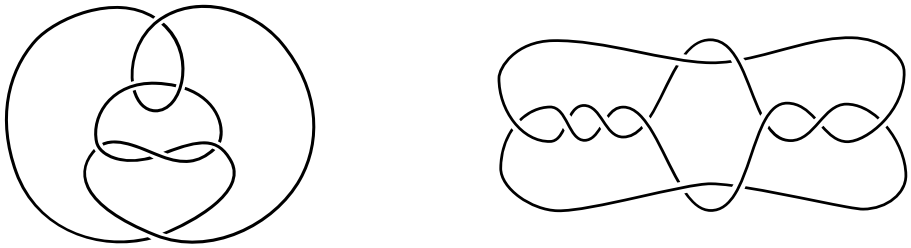
Figure 10: The Reidemeister moves.

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

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 [**knot-theory-and-its-applications**].

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.
Note: The unknot on the left is easier, from [**the-knot-book**]; the one on the right is called the Goeritz unknot, first appearing in [**bemerkungen-zur-knotentheorie**].)

2.2.1 Knot symmetries

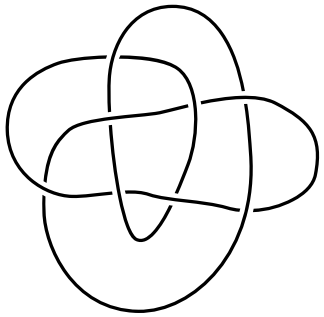
The parametrisation of S^1 – for example, the counterclockwise direction along the unit circle in the complex plain – induces an orientation (direction) along a knot, and more generally, along each component of a link. An ambient isotopy will always preserve this orientation. A weaker notion of equivalence 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 for many of the simplest knots (for a suitable definition of simplicity that we don't state), K^* does happen to be ambient isotopic to K. In fact, the simplest knot that isn't reversible is the knot 8_{17}, shown below.



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 $K_1 \# K_2$.

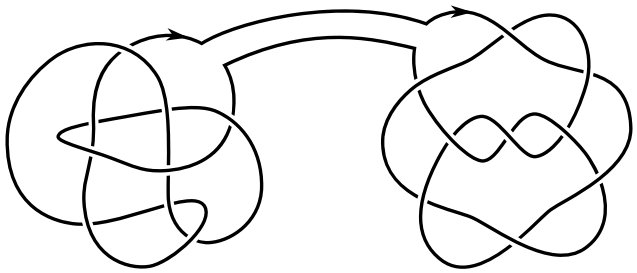


Figure 12: The knot $9_{33} \# 9_{35}$.

Many arbitrary choices are made in this definition, so it's not at all clear that the connected sum operation is well defined. If the knots are un-oriented, then indeed it isn't: the resulting knot depends which of the ends of K_1 is connected to which of the ends of K_2. But for oriented knots, the gluing of ends is uniquely determined.

We also made choices as to where along K_1 and K_2 to remove small arcs. It turns out that the resulting connected sum knot is independent of this choice:

Exercise 2.3.1. Give a topological argument showing that $K_1 \# K_2$ does not depend on where along each knot K_1 and K_2.

It is clear that the connected sum operation is associative, as long as it is well-defined. Since the definition doesn't treat K_1 any differently to K_2, the connected sum is also commutative. In other words, knots form a *commutative monoid*, denoted $\mathcal{K}$, under the connected sum operation #.

A commutative monoid you know well is the monoid of nonnegative integers under multiplication. Integers which cannot be obtained as products of other (non-unit) integers are called *prime*, and it is a famous elementary result of number theory that there are infinitely many prime numbers. Hence, the monoid $(\mathbb{Z}_{\geq 0}, \times)$ is not finitely generated: any generating set will include all prime numbers. The news is not all bad, though: we know that every integer has a unique prime decomposition, up to permutation of factors.

There is a striking similarity between the monoid of knots and the monoid of integers in this respect:

Definition 2.3.2. A prime knot is a non-trivial knot which cannot be written as the connected sum of two non-trivial knots.

Theorem 2.3.1 (Euclid's Theorem for Knots)**.** There are infinitely many prime knots.

There are many proofs of this theorem by finding some infinite family of prime knots, see for example [**knots**].

Theorem 2.3.2 (Schubert's Theorem)**.** Up to permutation of factors, every knot can be uniquely decomposed as a connected sum of prime knots.

A proof of Schubert's theorem can be found in [**knots**].

Corollary. The set of knots forms a free commutative monoid generated by prime knots under the connected sum operation.

The fact that there are infinitely many prime knots means that the commutative monoid of knots is not finitely generated, and hence not finitely presented. We remedy this by instead thinking about other knotted objects with more powerful algebraic structures, where each object can be assembled from finitely many 'atomic' parts. The braids that we saw in Example 1.1.2 (as a group) and Example 1.3.4 (as a category) are one option. A more powerful way to achieve finite generation is to extend our world to include *tangles*.

2.4 Tangles: a finitely presented monoidal category

In Example 1.3.4, we defined a category **Braid**, with objects indexed by natural numbers, and morphisms between n and n given by equivalence classes of braids on n strands. The category of *tangles* is a simultaneous extension of the braid category and the set of links, where interval strands are not required to be monotonic, circle components are allowed, and there exist morphisms between distinct m and n.

Definition 2.4.1. A *tangle* T with m and n endpoints is an embedding of some number of intervals and some number of circles – that is, a compact one dimensional manifold, possibly with boundary – into the unit cube $[0,1]^3$, such that the boundary maps to m collinear, unit-spaced marked points on the bottom face and n such points on the top face. Tangles are considered up to ambient isotopy fixing the boundary of the cube. We denote the set of ambient isotopy classes of tangles with n bottom and m top endpoints by $\mathcal{T}_{m,n}$.

Tangles may be multiplied, by composing them with the same associative stacking operation as for braids. Of course, two tangles may only be composed when the number of top endpoints of the first matches the number of bottom endpoints of the second.

Tangle figure needed — Z.

Definition 2.4.2. The monoidal category **Tangle**, also denoted $\mathcal{T}$ consists of:

- *Objects*: nonnegative integers $\mathbb{Z}_{\geq 0}$
- *Morphisms:* The morphisms are ambient isotopy classes of tangles with the appropriate number of bottom and top endpoints, that is, $\mathrm{Hom}(m,n) := \mathcal{T}_{m,n}$. Composition is vertical stacking.

Exercise 2.4.1. Make Definition 2.4.2 precise by resolving the subtleties around reparametrising cubes as they are stacked. Show that $\mathcal{T}$ satisfies the axioms of a category.

Our goal is to write a finite presentation for tangles: importantly, a finite number of morphisms which generate all others. With only vertical stacking as the operation, this is not possible: essentially, morphisms between few endpoints don't generate morphisms between many endpoints.

Luckily, this problem can be solved by enriching the algebraic structure of tangles with an additional, quite natural, operation: stacking side by side. Given a tangle T_1 in $\mathrm{Hom}(m_1, n_1)$ and a tangle T_2 in $\mathrm{Hom}(m_2, n_2)$, their side-by-side disjoint union, denoted $T_1 \otimes T_2$, is an element of $\mathrm{Hom}(m_1 + m_2, n_1 + n_2)$.

an illustrative

This side by side product of tangles satisfies the categorical properties of a tensor product, hence the notation $\otimes$. Unlike composition, it is defined for any pairs of tangles; it is clearly associative, and unital thanks to the empty tangle in $\hom(0, 0)$. The percise way to summarise the above is that the $\otimes$ operation upgrades **Tangle** to a *monoidal category*. Monoidal categories provide a bridge between knot theory and quantum algebra: thus, we make a brief detour back to the world of abstract category theory to define them.

tensor product figure — Z.

Monoidal categories

To define tensor products in categories, we need a notion of a *bifunctor*: a functor which has two inputs rather than one, and one output as usual.

Definition 2.4.3. Given a categiry $\mathcal{C}$, a *bifunctor* $F : \mathcal{C} \times \mathcal{C} \to \mathcal{C}$ assigns an object $F(A, B)$ to any pair of objects (A, B) of $\mathcal{C}$. For each pair of morphisms $f : A \to C$ and $g : B \to D$, F assigns a morphism $F(f, g)$ in $\mathrm{Hom}\left((A, B), (C, D)\right)$, and these assignments respect component-wise composition and identities:

$$(F(f', g') \circ F(f, g) = F(f' \circ f, g' \circ g),$$

$$\mathrm{id}_{F(A,B)} = F(\mathrm{id}_A, \mathrm{id}_B).$$

A *tensor product* $\otimes$ in a category $\mathcal{C}$ is a bifunctor $\otimes : \mathcal{C} \times \mathcal{C} \to \mathcal{C}$. For objects, we denote $\otimes(A, B)$ by $A \otimes B$, and similarly for morphisms $f : A \to C$ and $g : B \to D$ we write $f \otimes g : A \otimes B \to C \otimes D$. A tensor product is, in essence, associative and unital, however, some terminology is required in order to make this precise.

The formal associativity requirement states that all pairs of objects $(A \otimes B) \otimes C$ and $A \otimes (B \otimes C)$ to be **naturally isomorphic**. In particular, part of the structure of a monoidal category is a specific isomorphism $\alpha_{A,B,C} : (A \otimes B) \otimes C \to A \otimes (B \otimes C)$, for each triple of objects A, B and C in $\mathcal{C}$. We will define what it means for these component isomorphisms to assemble into a *natural* isomorphism; for now you may read that informally as "canonical".

It will be useful to think of tensor components in parentheses as being "closer together", in which case the isomorphism $\alpha_{A,B,C}$ can be visualised as below:

Next, for the tensor product $\otimes$ to behave as tensor products do in the wild, it need to come with an identity or *unit* object I: this is why monoidal categories are named after

monoids. As an example, think of vector spaces over $\mathbb{C}$: given any vector space V, we have $V \otimes \mathbb{C} \cong V \cong \mathbb{C} \otimes V$, making the field a unit object for the tensor product.

The formal unit requirement is that there exists an object I in $\mathcal{C}$, such that $A \otimes I$ and $I \otimes A$ are naturally isomorphic to A, via a set of given iomorphisms $\lambda_A : I \otimes A \to A$ and $\rho_A : A \otimes I \to A$.

Lastly, the isomorphisms α, λ, and ρ must satisfy some compatibility conditions, which ensure that they work together as you would expect.

The first such condition is called the *triangle axiom*: for all objects A and B in $\mathcal{C}$, it shouldn't matter if you re-associate before taking a tensor product with the identity, thus the triangle below should commute.

$$\begin{array}{ccc} (A \otimes I) \otimes B & \xrightarrow{\alpha_{A,I,B}} & A \otimes (I \otimes B) \\ & {\scriptstyle \rho_A \otimes \mathrm{id}_B} \searrow \quad \swarrow {\scriptstyle \mathrm{id}_A \otimes \lambda_B} & \\ & A \otimes B & \end{array}$$

It is important to note that the content in this statement is not that $(A \otimes I) \otimes B$ is isomorphic to $A \otimes B$: of course it is, this follows from the natural isomorphisms already built in. The issue is that the isomorphism $(A \otimes I) \otimes B \cong A \otimes B$ can be constructed from the natural isomorphisms *two different ways*, and if these ways differed – if the triangle above didn't commute – then this isomorphism would no longer be canonical.

The second and last compatibility condition is called the *pentagon axiom*. Consider the entirely left-branching parenthesisation of four objects, $((A \otimes B) \otimes C) \otimes D$. There are two possible sequences of re-parenthesisation isomorphisms α that take us from this to the entirely right-branching parenthesisation $A \otimes (B \otimes (C \otimes D))$. These are shown in the upper and lower paths in Figure 14, which together pass through all five possible parenthesisations of the four objects A, B, C, and D.

$$\begin{array}{ccc} & (A \otimes B) \otimes (C \otimes D) & \\ {\scriptstyle \alpha_{A\otimes B,C,D}} \nearrow & & \searrow {\scriptstyle \alpha_{A,B,C\otimes D}} \\ ((A \otimes B) \otimes C) \otimes D & & A \otimes (B \otimes (C \otimes D)) \\ \Big\downarrow {\scriptstyle \alpha_{A,B,C} \otimes \mathrm{id}_D} & & \Big\uparrow {\scriptstyle \mathrm{id}_A \otimes \alpha_{B,C,D}} \\ (A \otimes (B \otimes C)) \otimes D & \xrightarrow{\alpha_{A,B\otimes C,D}} & A \otimes ((B \otimes C) \otimes D) \end{array}$$

Figure 14: The associativity pentagon for four components.

The isomorphism $((A \otimes B) \otimes C) \otimes D \cong A \otimes (B \otimes (C \otimes D))$ ought to be canonical, so we must require that both ways of producing it be the same. This is the same as saying that the pentagon diagram of Figure 14 commutes, for every choice of four objects A, B, C, and D. This is known as the *pentagon axiom*, or *pentagon equation* when expressed as an equality in the various isomorphisms, as below:

$$\alpha_{A,B,C\otimes D} \circ \alpha_{A\otimes B,C,D} = \mathrm{id}_A \otimes \alpha_{B,C,D} \circ \alpha_{A,B\otimes C,D} \circ \alpha_{A,B,C} \otimes \mathrm{id}_D \,.$$

You can also visualise the pentagon axiom as two ways of gradually rearranging left-grouped objects to the right, as shown in Figure 15.

Summarising all these ingredients, we arrive at the definition of a monoidal category:

Definition 2.4.4. A *monoidal category* $(\mathcal{C}, \otimes, I, \alpha, \lambda, \rho)$ is a category $\mathcal{C}$, with

- a tensor product $\otimes : \mathcal{C} \times \mathcal{C} \to \mathcal{C}$,
- with an associativity constraint α, consisting of component isomorphisms $\alpha_{A,B,C} : (A \otimes B) \otimes C \to A \otimes (B \otimes C)$
- and a distinguished unit object $I \in \mathrm{ob}(\mathcal{C})$, and left and right unit constraints λ and ρ with respect to I, with component isomorphisms $\lambda_A : I \otimes A \to A$, and $\rho_A : A \otimes I \to A$;
- satisfying the *triangle* and *pentagon* axioms.

The monoidal category $\mathcal{C}$ is called *strict* if all the isomorphisms furnished by α, λ, and ρ are identity morphisms in $\mathcal{C}$.

Example 2.4.1. The motivating example of a monoidal category is that of $\mathbf{Vect}_{\mathbb{F}}$, the category of vector spaces over a given field $\mathbb{F}$. The tensor product $\otimes$ is the usual tensor product of vector spaces, and for all vector spaces U, V, and W, simply shifting the brackets will do the trick:

$$\alpha_{U,V,W} : (u \otimes v) \otimes w \mapsto u \otimes (v \otimes w).$$

The unit object is the ground field $\mathbb{F}$ itself, and the left and right unit constraints respectively send $a \otimes v$ and $v \otimes a$ to av, for all a in $\mathbb{F}$.

Example 2.4.2. Given two representations (V_1, ρ_1) and (V_2, ρ_2) of a group G, it is a fundamental fact of representation theory that $(V_1 \otimes V_2, \rho_1 \otimes \rho_2)$ is a representation of the direct product $G \times G$. It is less obvious that we can make a representation of G itself[2], by defining

$$\rho_{V_1 \otimes V_2} : G \to \mathrm{End}(V_1 \otimes V_2), \quad g \mapsto \rho_1(g) \otimes \rho_2(g).$$

The associativity constraint and unit object $\mathbb{F}$ are inherited from the monoidal category **Vect**, with the latter upgraded to a representation of G trivially:

$$\rho_{\mathbb{F}}(g)(a) = a \quad \text{for all } g \text{ in } G \text{ and } a \text{ in } \mathbb{F}.$$

It's a good exercise to check that $\rho_{V_1 \otimes V_2}$ is a group homomorphism, and hence $(V_1 \otimes V_2, \rho_{V_1 \otimes V_2})$ is a representation of G.

[2]The underlying structure making this possible is a *coproduct* on the group algebra $\mathbb{F}G$. In general, the category A-**Mod** of modules over any bialgebra A is monoidal. This is an example of *Tannaka-Krein duality*.

The name *monoidal category* stems from the fact that the binary operation $\otimes$ makes $\mathrm{ob}(\mathcal{C})$ into a monoid, with I serving as the identity element. Of course, we are sneakily overlooking that in general, $\otimes$ is only associative up to isomorphism, and $\mathrm{ob}(\mathcal{C})$ is not a set. In the case that $\mathcal{C}$ is small and strict, $\mathrm{ob}(\mathcal{C})$ is an honest monoid. A map F between monoidal categories $(\mathcal{C}, \otimes_{\mathcal{C}}, I_{\mathcal{C}})$ and $(\mathcal{D}, \otimes_{\mathcal{D}}, I_{\mathcal{D}})$ should thus be a functor resembling a morphism of monoids, with

$$F(A \otimes_{\mathcal{C}} B) = F(A) \otimes_{\mathcal{D}} F(B) \quad \text{and} \quad F(I_{\mathcal{C}}) = I_{\mathcal{D}}.$$

This actually defines a *strict monoidal functor*. To define a general *monoidal functor*, relax the equalities above to isomorphisms

$$\varphi_{A,B} : F(A \otimes_{\mathcal{C}} B) \xrightarrow{\sim} F(A) \otimes_{\mathcal{D}} F(B) \quad \text{and} \quad \varphi_1 : F(I_{\mathcal{C}}) \xrightarrow{\sim} I_{\mathcal{D}},$$

which will form part of the data of our functor.

In the discussion above, we suppressed what exactly it means for isomorphisms to be *natural*, and didn't explain why the axioms are sufficient. The next two short vignettes aim to fill these gaps.

Natural transformations

Categories have objects and morphisms, and categories themselves form a category (modulo set theoretic subtleties), whose morphisms are functors. This game continues: functors too are the objects in their own category, and *natural transformations* are the morphisms of functors. In other words, natural transformations are "maps" sending one functor to another.

Let $\mathcal{C}$ and $\mathcal{D}$ be two categories, and F and G two functors from $\mathcal{C}$ to $\mathcal{D}$. Given two objects A and B of $\mathcal{C}$, and a morphism f in $\mathrm{Hom}(A, B)$, we can apply the functors to both the objects and the morphism:

$$\begin{array}{ccc} F(A) & & G(A) \\ {\scriptstyle F(f)}\downarrow & & \downarrow{\scriptstyle G(f)} \\ F(B) & & G(B) \end{array}$$

A natural transformation is a collection of morphisms in $\mathcal{D}$ relating F to G, compatible with the images of morphisms of $\mathcal{C}$:

Definition 2.4.5. A *natural transformation* η from F to G is a collection of morphisms $\eta_X : F(X) \to G(X)$ in $\mathcal{D}$, for each object X in $\mathcal{C}$, satisfying that the diagram below commutes *for all* choices of A and B in $\mathrm{ob}(\mathcal{C})$, and all f in $\mathrm{Hom}(A, B)$.

$$\begin{array}{ccc} F(A) & \xrightarrow{\eta_A} & G(A) \\ {\scriptstyle F(f)}\downarrow & & \downarrow{\scriptstyle G(f)} \\ F(B) & \xrightarrow{\eta_B} & G(B) \end{array}$$

If the maps η_X are isomorphisms for all X, then η is called a *natural isomorphism*, and the objects $F(X)$ and $G(X)$ in $\mathcal{D}$ are said to be *naturally isomorphic*.

Mathematicians throw the words *natural* and *canonical* around often, usually in reference to maps which arise without a dependence on choices. The above definition is the fruit of category theory's efforts to make precise what this really means.

Example 2.4.3. A classic example of unnatural versus natural isomorphisms is found in the category $\mathbf{Vect}_{\mathbb{F}}$ of finite dimensional vector spaces over F.

Every finite-dimensional vector space V over $\mathbb{F}$ is isomorphic to its dual V^*, the space of linear functionals $\varphi : V \to \mathbb{F}$. The isomorphism, however, depends on a choice of a basis $\{e_1, \ldots, e_d\}$ of V, which in turn defines a dual basis $\{e^1, \ldots, e^d\}$ of V^* by setting $e^i(e_j) = \delta^i_j$. This basis-dependence means that the isomorphism $V \cong V^*$ is not canonical, not *natural*.

Every finite-dimensional vector space V is also isomorphic to its double dual V^{**}, which consists of linear functionals from V^* to $\mathbb{F}$. Here, we could use a repeated version of the above isomorphism, but there is a better, more *natural* one! Define the *evaluation map* $\eta_V : V \to V^{**}$ by $v \mapsto \widehat{v}$, where $\widehat{v}$ sends a linear functional φ on V to its evaluation at v:

$$\widehat{v}(\varphi) := \varphi(v).$$

Exercise 2.4.2. Check that the $\eta = \{\eta_V\}$ of Example 2.4.3 is a natural isomorphism between the identity functor, and the double dual functor $V \mapsto V^{**}$, which are both *endofunctors* (functors from a category to itself) on the category $\mathbf{Vect}_{\mathbb{F}}$.

Exercise 2.4.3. The functor GL_n sends a commutative ring R to the group $\mathrm{GL}_n(R)$ of invertible $n \times n$ matrices with coefficients in R. The functor and $(\cdot)^\times$ sends a ring R to its group of units $R^\times$. Both of these are functors from the category of commutative rings **CRing** to the category of groups **Grp**. How do these functors act on morphisms? Can you find a natural transformation between them?

Exercise 2.4.4. Let G be a group, and BG be its single-object category. Recall from Exercise 1.3.6 that a functor $F : BG \to \mathbf{Vect}_{\mathbb{F}}$ is exactly the same as a representation (V, ρ) of G. Given two representations (V_1, ρ_1) and (V_2, ρ_2) of G, a *G-equivariant map* is a linear map $\varphi : V_1 \to V_2$ satisfying

$$\rho_2(g) \circ \varphi = \varphi \circ \rho_1(g)$$

for all g in G. In fact, G-equivariant maps are the morphisms in the category of representations of G. Show that G-equivariant maps of representations correspond exactly to natural isomorphisms between the corresponding functors $BG \to \mathbf{Vect}_{\mathbb{F}}$.

Example 2.4.4. We can now parse what it means for the associativity and unit constraints in a monoidal category to assemble into a natural isomorphisms. Given a category $\mathcal{C}$ with a tensor product $\otimes$, we have two endofunctors on $\mathcal{C} \times \mathcal{C} \times \mathcal{C}$:

$$\otimes \circ (\otimes \times \mathrm{id}_{\mathcal{C}}) : (A, B, C) \mapsto (A \otimes B) \otimes C \quad \text{and} \quad \otimes \circ (\mathrm{id}_{\mathcal{C}} \times \otimes) : (A, B, C) \mapsto A \otimes (B \otimes C).$$

An *associativity constraint*, then, is a natural isomorphism

$$\alpha : \otimes \circ (\otimes \times \mathrm{id}_{\mathcal{C}}) \to \otimes \circ (\mathrm{id}_{\mathcal{C}} \times \otimes).$$

Naturality means that for every pair of triples of objects and corresponding triple of morphisms in $\mathcal{C}$,

$$\begin{array}{ccc} A & B & C \\ \downarrow f & \downarrow g & \downarrow h \\ D & E & F \end{array}$$

the isomorphisms $\alpha_{A,B,C}$ and $\alpha_{D,E,F}$ make the following diagram commute:

$$\begin{array}{ccc} (A \otimes B) \otimes C & \xrightarrow{\alpha_{A,B,C}} & A \otimes (B \otimes C) \\ \downarrow{\scriptstyle (f \otimes g) \otimes h} & & \downarrow{\scriptstyle f \otimes (g \otimes h)} \\ (D \otimes E) \otimes F & \xrightarrow{\alpha_{D,E,F}} & D \otimes (E \otimes F) \end{array}$$

Similarly, the *left unit constraint* with respect to I is a natural isomorphism

$$\lambda : \otimes \circ (I \times \mathrm{id}) \to \mathrm{id}_{\mathcal{C}},$$

and the *right unit constraint* with respect to I is a natural isomorphism

$$\rho : \otimes \circ (\mathrm{id} \times I) \to \mathrm{id}_{\mathcal{C}}\,.$$

Exercise 2.4.5. Check that the associativity and unit constraints in Examples 2.4.1 and ex:MonoidalReps are natural, and that the triangle and pentagon axioms are satisfied.

The Pentagon Axiom and Coherence

We discussed how the pentagon axiom is necessary to ensure that the isomorphisms between differently parenthesised tensor products of four objects are canonical. But what about more than four objects? Surely, we must impose an axiom that all compositions of isomorphisms getting us from, for example, $(((A \otimes B) \otimes C) \otimes D) \otimes E$ to $A \otimes (B \otimes (C \otimes (D \otimes E)))$ be equal too, giving a very large number of pentagon-style axioms?

Exercise 2.4.6. How many parenthesisations are there of n objects?

Remarkably, we do not need any more reassociation axioms: the pentagon is sufficient, according to the following powerful result from one of the co-founders of category theory, which we state without proof:

Theorem 2.4.1 (Mac Lane's Coherence Theorem)**.** Let $X_1, X_2, \ldots, X_n$ be n objects in a category $\mathcal{C}$, with a tensor product $\otimes$ and an associativity constraint α. Let P_1 and P_2 be two parenthesised tensor products of all the X_i, and f and g be two isomorphisms from P_1 to P_2, built from isomorphisms α_{X_i,X_j,X_k}, their inverses, and identity morphisms, by means of composition and tensor products. If $\otimes$ and α satisfy the pentagon axiom, then $f = g$.

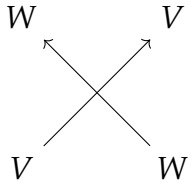
Braided Monoidal Categories

We add one last piece of structure to a monoidal category: a means of flipping from $A \otimes B$ to $B \otimes A$. In the monoidal category of vector spaces $\mathbf{Vect}_{\mathbb{F}}$, simply permuting the tensor factors gives an isomorphism between $V \otimes W$ and $W \otimes V$,

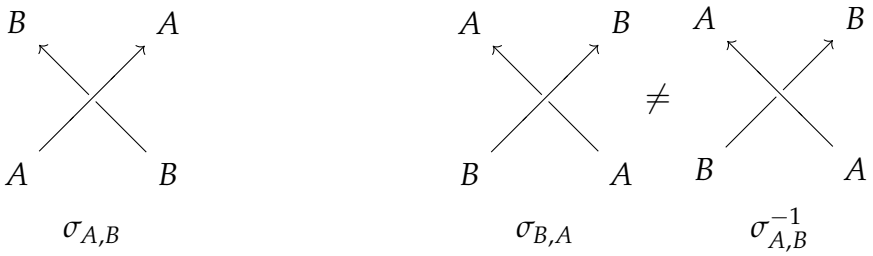
$$\sigma_{V,W} : v \otimes w \mapsto w \otimes v.$$

Note that the isomorphisms $\sigma_{V,W}$ assemble into a natural isomorphism between the bifunctor $\otimes$ and the *opposite* tensor product bifunctor $\otimes^{op} : (V, W) \mapsto W \otimes V$. We can also write $\otimes^{op}$ as a pre-composition of $\otimes$ with the *transposition* functor $\tau : \mathcal{C} \times \mathcal{C} \to \mathcal{C} \times \mathcal{C}$ given by $\tau(V, W) = (W, V)$.

In $\mathbf{Vect}_{\mathbb{F}}$, the flip $\sigma_{W,V} : w \otimes v \mapsto v \otimes w$ is clearly the inverse of $\sigma_{V,W}$. Such flip isomorphisms are called *symmetric*, inspired by the similar behaviour of the transpositions σ_{ij} in a symmetric group. This situation can be portrayed diagrammatically as below:



However, this symmetry need not always be the case! In some monoidal categories, flipping B and A is *not* equivalent to the inverse of flipping A and B. We have encountered such non-trivial flips before, in the braid groups. Drawing inspiration from this, we will depict such maps with crossings.



A braided monoidal category is a monoidal category where natural flipping isomorphisms $\sigma_{A,B} : A \otimes B \to B \otimes A$ are guaranteed by the structure. When there are more natural isomorphisms, there are more coherence conditions too!

Each hexagon in Figure 16 shows two ways to combine flips and associativity isomorphisms to arrive at the same result: in each, the top path re-associates, then flips, then re-associates; the bottom path flips, then re-associates, then flips. Imposing the compatibility condition that both hexagons commute and reading the equalities from the diagrams, we obtain the two *hexagon equations*. This condition on the associativity constraint and the flipping map is known as the *hexagon axiom*:

$$\alpha_{B,C,A} \circ \sigma_{A,B\otimes C} \circ \alpha_{A,B,C} = \mathrm{id}_B \otimes \sigma_{A,C} \circ \alpha_{B,A,C} \circ \sigma_{A,B} \otimes \mathrm{id}_C$$

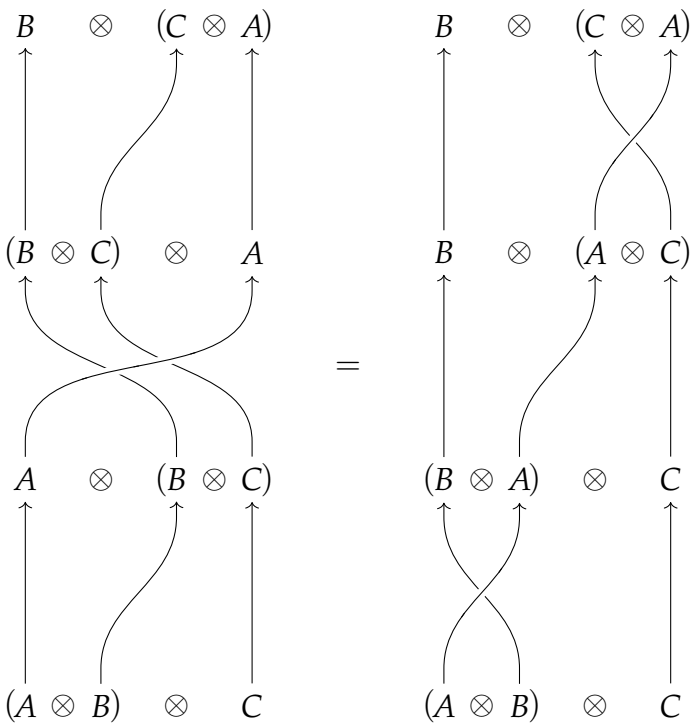


Figure 17: Visual representation of the first hexagon equation.

Definition 2.4.6. Let $(\mathcal{C}, \otimes, I.\alpha, \lambda, \rho)$ be a monoidal category. Let τ be the transposition endofunctor on $\mathcal{C} \times \mathcal{C}$, that is, $\tau(A, B) = (B, A)$ for A and B in $\mathrm{ob}(\mathcal{C})$. A *braiding* on $\mathcal{C}$ is a natural isomorphism $\sigma : \otimes \to \otimes \circ \tau$, satisfying the hexagon axiom. A *braided monoidal category* is a monoidal category with a braiding σ. A *symmetric monoidal category* is a monoidal category with a braiding σ satisfying $\sigma_{B,A} = \sigma_{A,B}^{-1}$.

Exercise 2.4.7. Show that the second hexagon equation can be obtained from the first by using the inverse flips $\sigma_{A,B}^{-1}$, and draw a picture of it in the style of Figure 17.

A map between two braided monoidal categories $(\mathcal{C}, \otimes, \tau)$ and $(\mathcal{D}, \otimes, \tau)$ is a monoidal functor (F, φ) which respects the braidings: for all objects V and W in $\mathcal{C}$, we have that

$$\begin{array}{ccc} F(V) \otimes F(W) & \longrightarrow & F(V \otimes W) \\ \downarrow & & \downarrow \\ F(W) \otimes F(V) & \longrightarrow & F(W \otimes V) \end{array}$$

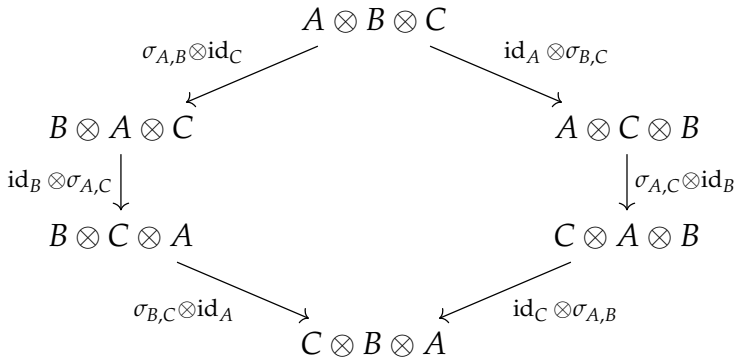
commutes.

Exercise 2.4.8. a) Recall that a monoidal category is *strict* if all the associativity isomorphisms are identities: for all objects A, B, and C, $(A \otimes B) \otimes C = A \otimes (B \otimes C)$. In this

setting, we are justified in ditching the brackets and writing $A \otimes B \otimes C$. In a strict monoidal category $\mathcal{C}$, the two hexagons of Figure 16 reduce to the two triangles

$$\begin{array}{ccc} A \otimes B \otimes C & \xrightarrow{\sigma_{A,B\otimes C}} & B \otimes C \otimes A \\ {\scriptstyle \sigma_{A,B}\otimes \mathrm{id}_C}\downarrow & \nearrow_{\mathrm{id}_B \otimes \sigma_{A\otimes C}} & \\ B \otimes A \otimes C & & \end{array} \quad \text{and} \quad \begin{array}{ccc} A \otimes B \otimes C & \xrightarrow{\sigma_{A\otimes B,C}} & C \otimes A \otimes B \\ {\scriptstyle \mathrm{id}_A \otimes \sigma_{B\otimes C}}\downarrow & \nearrow_{\sigma_{A,C}\otimes \mathrm{id}_B} & \\ A \otimes C \otimes B & & \end{array} .$$

Show that the following new and improved hexagon commutes for all objects A, B, and C:

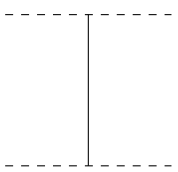


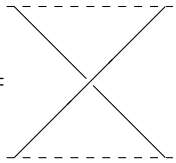
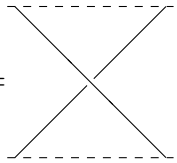
b) Hence, for a fixed object A, find an equation satisfied by $R \coloneqq \sigma_{A,A}$. This is called the **quantum Yang-Baxter equation**. Representing this equality of morphisms diagrammatically, as in Figure 17, which old friend do you recognise?

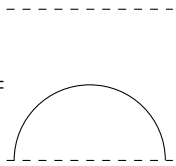
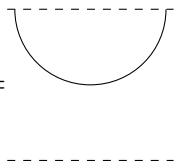
You will note that in the pictorial representation of Figure 17, the two equivalent maps of the first hexagon equation are shown as isotopic braids. This is no mistake! Morphisms in braided monoidal categories behave much like braids and tangles, and indeed, provide a well-suited algebraic framework to describe them.

Tangles as a Finitely-Presented Braided Monoidal Category

Consider the category whose objects are finite sequences of points $\bullet$, which we identify with the natural numbers $\mathbb{N}$. Setting $m \otimes n = m + n$, we are on our way to defining a strict monoidal structure on this category, with unit object 0. As morphisms we have the following diagram elements, called *simple tangles*,

$\mathbb{1} \coloneqq \mathrm{id}_\bullet =$  $\in \mathrm{Hom}(1,1),$

$x_+ :=$  $\in \mathrm{Hom}(2,2)$, $x_- :=$  $\in \mathrm{Hom}(2,2)$,

$t :=$  $\in \mathrm{Hom}(2,0)$, $b :=$  $\in \mathrm{Hom}(0,2)$,

and all diagrams obtained from them by means of horizontal juxtaposition and vertical stacking, which represent the tensor product and composition respectively. Any morphism constructed in this way represents an unoriented tangle as we have previously defined it, and so we have a generating set for our tensor category. But now we have the same problem we faced for knot diagrams: when do two such morphisms depict isotopic tangles? In this finitely presented case, our problem is to find the relations satisfied by these generators. Recall that two knot diagrams depict isotopic knots if and only if one may be obtained from the other through planar isotopies and a finite sequence of Reidemeister moves. Let us first construct representations of the Reidemeister moves from our building blocks:

Reidemeister moves:$T1, T2, T3$.

The following four diagram moves depict planar isotopies:

planar isotopies:$T0, T4, T5, T5'$.

Together, $\{T0, T1, T2, T3, T4, T5, T5'\}$ are known as the *Turaev relations*. In close analogy to the knot case, the remarkable fact that this small set of very believable relations is actually all we need, is one we shall state without proof.

Theorem 2.4.2 (Turaev)**.** The category **Tangle** is equivalent (as a monoidal category) to the monoidal category generated by $\{x_+, x_-, b, t\}$, modulo the Turaev relations.

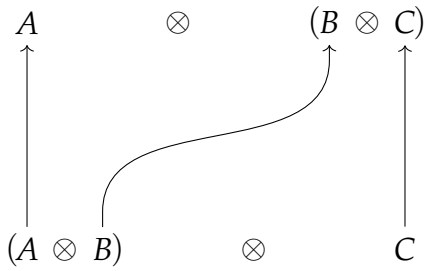


Figure 13: Visualisation of the associativity isomorphism $\alpha_{A,B,C}$.

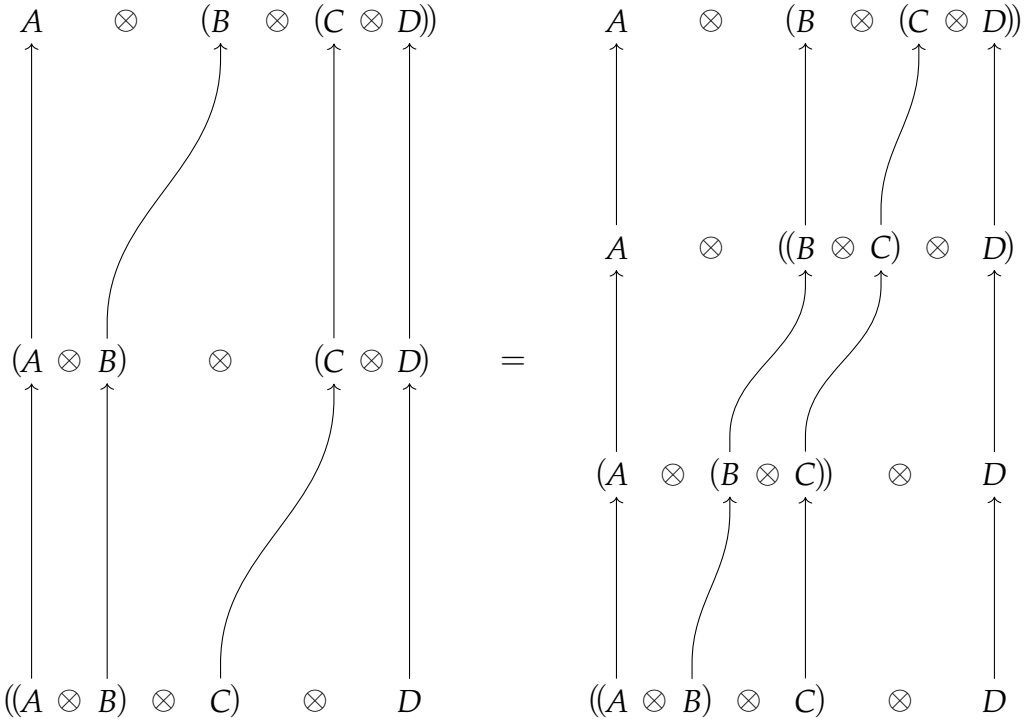


Figure 15: The pentagon equation, diagrammatically.

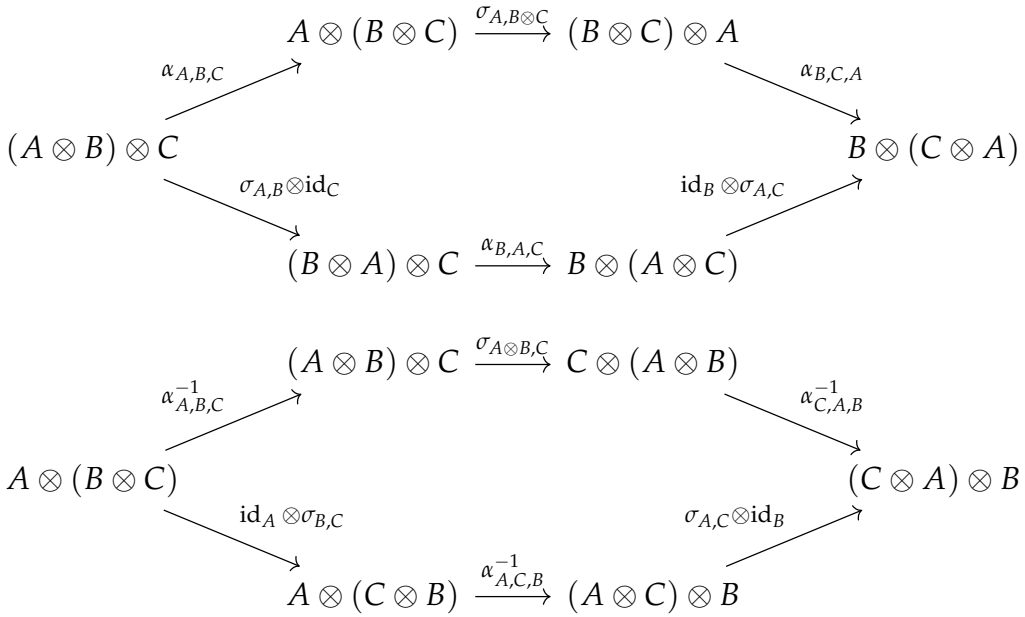


Figure 16: The two hexagons, the compatibility conditions for α and σ.

Chapter 3

Gradings and Filtrations

Algebras are easier to work with when they come with a decomposition into building blocks which behave in a predictable way. Here we will consider two such notions: first *gradings*, and then the less powerful *filtrations*.

3.1 Gradings

Definition 3.1.1. Let A be an algebra over a field $\mathbb{F}$. A *grading* on A is a linear decomposition into $\mathbb{F}$-vector spaces A_n,

$$A = \bigoplus_{n=0}^{\infty} A_n, \quad \text{or} \quad A = \prod_{n=0}^{\infty} A_n,$$

such that if $a \in A_i$ and $b \in A_j$, then $ab \in A_{i+j}$. The elements of A_n are called *homogeneous of degree n*. A *graded algebra* is an algebra equipped with a grading.

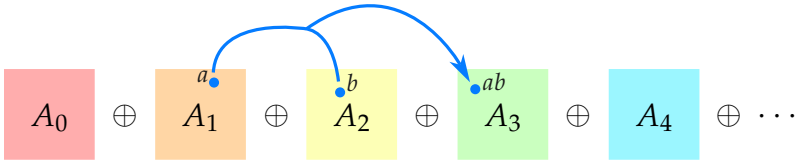


Figure 18: A graded algebra A.

Example 3.1.1. The classic example to keep in mind is that of polynomials. Let $A = \mathbb{C}[x_1, \ldots, x_r]$ denote the algebra of complex polynomials in r variables. In an intentional synergy of terminology, the homogeneous elements of degree n are exactly the homogeneous polynomials of degree n: those for which the exponents in each monomial sum to n. For example, $x^2y + \frac{22}{7}xy^2 + \sqrt{2}iy^3$ is a homogeneous polynomial in two variables, while $\pi + x + 2xy$ is not homogeneous. Since any polynomial can be written as a finite sum of homogeneous polynomials, we have $A = \bigoplus_{n=0}^{\infty} A_n$, where $A_n \subseteq A$ is the set of homogeneous polynomials of degree n.

Exercise 3.1.1. Show that $A = \mathbb{C}[x_1, \ldots, x_r]$ satisfies the remaining defining property of a graded algebra, namely, if p and q are homogeneous polynomials of degrees i and j respectively, then pq is homogeneous of degree $i + j$.

In Definition 3.1.1 we allowed a graded algebra to be either a direct sum, or a direct product. As vector spaces, the only difference is that the former consists of *finite* linear combinations of homogeneous elements, whilst in the latter, we also have *infinite* linear combinations: think power series versus polynomials. Power series (and, in general, infinite graded sums) are multiplied using the Cauchy formula: if $a, b \in A = \Pi_{n=0}^{\infty} A_n$, and the homogeneous degree n components of a and b are denoted a_n and b_n, respectively, then the degree n component of the product ab is given by

$$(ab)_n = \sum_{i=0}^{n} a_i b_{n-1}. \tag{3.1}$$

should turn more improtant displayed formulas into numbered equations — Z.

Remark 3.1.1. Technically, $\Pi_{n=0}^{\infty} A_n$ is the *degree completion* of $A = \bigoplus_{n=0}^{\infty} A_n$ is called the *degree completion* of A, often denoted $\widehat{A}$. There is a sequence of projections

$$\sum_{n=0}^{0} A_n \leftarrow \sum_{n=0}^{1} A_n \leftarrow \sum_{n=0}^{2} A_n \leftarrow \cdots,$$

and the graded completion is, by definition, the inverse limit

$$\widehat{A} = \varprojlim_{N} \sum_{n=0}^{N} A_n.$$

In these notes, we will work mainly with direct products, and omit the hats to make notation less cluttered.

should we define inverse limits here? probably? — Z.

Need to fix the numbering so that everything (remarks, examples, defs, thms etc) comes in one sequence — Z.

Exercise 3.1.2. There is a clear linear inclusion $A = \bigoplus_{n=0}^{\infty} A_n \hookrightarrow \Pi_{n=0}^{\infty} A_n$. Show that the algebra multiplication of A extends *uniquely* to $\Pi_{n=0}^{\infty} A_n$, and the unique extension is given by the Cauchy formula, *if* we require that in $\Pi_{n=0}^{\infty} A_n$ the product of elements a and b up to degree N depends only on the up to degree N parts of a and b. (In other words, the image of ab in the degree quotient $\Pi_{n=0}^{\infty} A_n / \Pi_{n=N+1}^{\infty} A_n$ depends only on the images of a and b in the same degree quotient.)

Graded algebras $A = \bigoplus_{n=0}^{\infty} A_n$ form a category **GrAlg**, whose morphisms are degree-preserving algebra homomorphisms:

Definition 3.1.2. Let $A = \bigoplus_{n=0}^{\infty} A_n$ and $B = \bigoplus_{n=0}^{\infty} B_n$ be two graded algebras. A *graded algebra homomorphism* is an algebra homomorphism $f : A \to B$ such that $f(A_n) \subseteq B_n$.

Degree-completed graded algebras form their own category $\widehat{\mathbf{GrAlg}}$, with graded homomorphisms as above. In fact, *degree completion* is a functor from **GrAlg** to $\widehat{\mathbf{GrAlg}}$.

Graded algebras are nice! Degrees provide a natural notion of complexity; one may even choose to blissfully ignore the complicated higher degree terms, and perform calculations *up to degree n*, that is, in a degree quotient. A familiar example of this is approximating functions by the first few terms of their Taylor series. Similarly, a grading unlocks the power of induction: one may prove things one degree at a time. These points will prove useful later in the course, when calculating knot invariants, and proving theorems about them.

3.2 Filtrations

Let A be an algebra over $\mathbb{F}$. Recall that *ideals* are to algebras and rings what normal subgroups are to groups: the sub-structure you can quotient by.

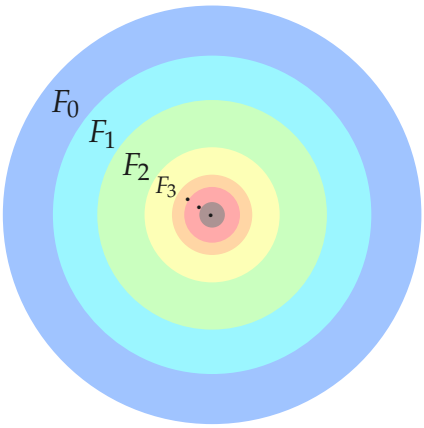
Definition 3.2.1. An ideal in an algebra A is a subset $I \subseteq A$ which is not only closed under linear combinations and the algebra multiplication (a subalgebra), but also under multiplication with *all* elments of $AI \subseteq I$ and $IA \subseteq I$. This definition is for a *two-sided* ideal, and in non-commutative algebras one may also define *left* and *right* ideals, which are only closed under left (respectively, right) multiplication by elements of A. For two-sided ideals we write $I \lhd A$.

Exercise 3.2.1. Given two ideals I and J in A, the product IJ is – by definition – the linear span of all products xy with x in I and y in J. Show that IJ is an ideal A.

Definition 3.2.2. A (descending) *filtration* on A is a decreasing sequence of nested ideals

$$A = F_0 \supseteq F_1 \supseteq F_2 \supseteq \cdots,$$

such that if $a \in F_i$ and $b \in F_j$, then $ab \in F_{i+j}$. A *filtered algebra* is an algebra A equipped with a descending filtration.



(a) A filtered algebra viewed from "above"...



(b) ...and from "the side". This picture makes sense if you've played with a Matryoshka doll: the big ones include all the smaller ones within.

Figure 19: A filtered algebra A.

Filtered algebras are useful, but not as useful as graded algebras. For example, given $x \in A_2$ and $y \in A_3$ in a graded algebra $A = \bigoplus_{n=0}^{\infty} A_n$, we know exactly which component

their product ab lives in: A_5. If $A = F_0 \supseteq F_1 \supseteq F_2 \supseteq \cdots$ is merely filtered, and $x \in F_2$ and $y \in F_3$, then we still know that $xy \in F_5$. However, since $F_5 \supseteq F_6 \supseteq F_7 \supseteq \cdots$, the product xy could live anywhere down this line of filtered components.

Nonetheless, filtered algebras form a category **FiltAlg**, whose morphisms are algebra homomorphisms which respect filtrations:

Definition 3.2.3. Let $A = F_0 \supseteq F_1 \supseteq \cdots$ and $B = G_0 \supseteq G_1 \supseteq \cdots$ be two filtered algebras. A *filtered algebra homomorphism* is an algebra homomorphism $f : A \to B$ such that $f(F_n) \subseteq G_n$.

While this definition looks identical to Definition 3.1.2, the notion of a filtered homomorphism is much weaker than a graded one. As before, B_n includes within it all the B_i for $i > n$. So, for some $x \in A_3$, it could be that $f(x) \in B_{274} \subseteq B_3$. Graded algebras, in contrast, come as linear decompositions into their graded components, so anomalies like this cannot happen. The following exercise makes this distinction explicit.

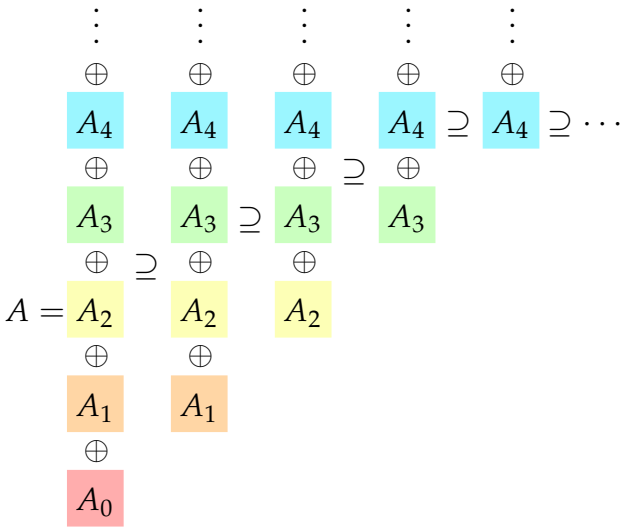


Figure 20: Filtering a graded algebra.

Exercise 3.2.2. Graded algebras can always be considered filtered: given a graded algebra $A = \bigoplus_{n=0}^{\infty} A_n$, verify that the subspaces

$$F_k = \bigoplus_{n=k}^{\infty} A_n$$

are ideals, and give a decreasing filtration on A. See Figure 20 for an illustration of the filtered components. Note, of course, that the graded algebra A and its filtered version are isomorphic as algebras. Verify that this process of "weakening" a grading defines a *forgetuful* functor **GrAlg** $\to$ **FiltAlg**: namely, graded homomorphisms are also filtered.

3.3 The Associated Graded Functor

Exercise 3.2.2 showed how to make filtered algebras out of graded ones. Since gradings are more useful than filtrations, however, one might prefer the other direction: to make

graded algebras out of filtered ones. Happily, there is a functor which does just that, called the *associated graded* functor. Better yet, we will soon see that in many cases, a tool called a *formality map* enables one to use the associated graded algebra to learn about the original filtered one.

Definition 3.3.1. If $A = F_0 \supseteq F_1 \supseteq \cdots$ is a filtered algebra, the (completed) *associated graded* algebra of A with respect to this filtration is the linear direct product

$$\operatorname{gr} A = \prod_{n=0}^{\infty} F_n \Big/ F_{n+1},$$

with the multiplication defined on the degree components by

$$\overline{a} \in F_i \Big/ F_{i+1} \quad \text{and} \quad \overline{b} \in F_j \Big/ F_{j+1} \quad \longmapsto \quad \overline{a} \cdot \overline{b} := \overline{ab} \in F_{i+j} \Big/ F_{i+j+1}.$$

The multiplication is extended to all of $\operatorname{gr} A$ via the Cauchy product formula (3.1).

The associated graded algebra is illustrated in Figure 21

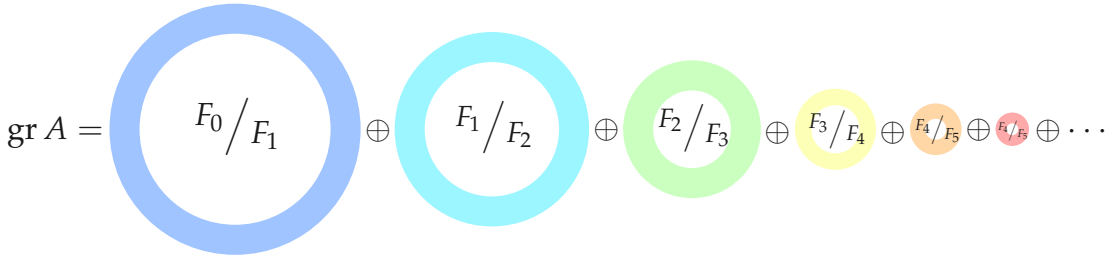


Figure 21: The associated graded of the filtered algebra from Figure 19.

Given two filtered algebras $A = F_0 \supseteq F_1 \supseteq \cdots$ and $B = G_0 \supseteq G_1 \supseteq \cdots$, a filtered algebra homomorphism $f : A \to B$ induces an *associated graded homomorphism*

$$\operatorname{gr} f : \operatorname{gr} A \to \operatorname{gr} B$$

between the associated graded algebras. The associated graded map is defined as follows: given an element x in F_n, denote the coset of x in F_n / F_{n+1} by $\overline{x}$. Then $\operatorname{gr} f(\overline{x})$ is defined to be the coset $\overline{f(x)}$ of $f(x)$ in G_n / G_{n+1}.

Exercise 3.3.1. Verify that the above is a well-defined homomorphism of graded algebras, and $\operatorname{gr}(f \circ g) = \operatorname{gr} f \circ \operatorname{gr} g$. Therefore, gr is a functor from **FiltAlg** to **GrAlg**. Check that gr is a one-sided inverse to the forgetful functor considered in Example 3.2.2.

3.3.1 The Pro-Unipotent Filtration

Let G be a group, and $\mathbb{F}G$ the group algebra over a given field $\mathbb{F}$. In this section we define a famous filtration on $\mathbb{F}G$, called the *pro-unipotent* filtration, studied extensively in group theory and algebraic geometry.

Definition 3.3.2. Let $\iota : \mathbb{F}G \to \mathbb{F}$ denote the *augmentation map* sending a linear combination of group elements to the sum of its coefficients:

$$\iota\left(\sum_{g\in G} c_g g\right) = \sum_{g\in G} c_g,$$

where $c_g \in \mathbb{F}$ and $g \in G$.
The *augmentation ideal* $I \lhd \mathbb{F}G$ is the kernel of the augmentation map, that is, the set of elements of $\mathbb{F}G$ whose coefficients sum to zero:

$$I = \ker \iota = \left\{ \sum_{g\in G} c_g g \;\middle|\; \sum_{g\in G} c_g = 0 \right\}.$$

Exercise 3.3.2. Verify that the augmentation map is indeed an algebra homomorphism, and hence, the augmentation ideal is indeed an ideal.

For $k > 0$, write I^k for the product of I with itself k times. In other words, I^k is linearly spanned by all elements of the form $x_1 x_2 \dots x_k$ with all x_i in I. By convention, $I^0 = \mathbb{F}G$. You can verify that the I^k are nested ideals: $I^k \supseteq I^{k+1}$ for all k.

Definition 3.3.3. The *pro-unipotent filtration* on $\mathbb{F}G$ is given by

$$\mathbb{F}G = I^0 \supseteq I^1 \supseteq I^2 \supseteq \cdots.$$

Exercise 3.3.3. The inverse limit algebra

$$\widehat{\mathbb{F}G} = \varprojlim_n \mathbb{F}G \Big/ I^n$$

is called the *pro-unipotent completion* of the group G. The algebra $\widehat{\mathbb{F}G}$ is also filtered, with the filtration fiven by

$$F_k = \varprojlim_{n>k} I^k \Big/ I^n.$$

(a) Prove that $\widehat{\mathbb{F}G} \supseteq F_1 \supseteq F_2 \supseteq \cdots$ is indeed a sequence of nested ideals, and describe the elements of F_k explicitly.

(b) Use the universal property of inverse limits to show that there is a canonical homomorphism $i : \mathbb{F}G \to \widehat{\mathbb{F}G}$. Given $x \in I^k$, describe $i(x) \in \widehat{\mathbb{F}G}$ explicitly.

(c) Show that i is a filtered map: $i(I^k) \subseteq F_k$.

(d) Show that i is injective if and only if $\bigcap_n I^n = \emptyset$.

(e) Show that $\operatorname{gr} i$ is an isomorphism.

Example 3.3.1. Consider the "baby example" $G = \mathbb{Z}$, which in this case is more conveniently written multiplicatively as the free group on one generator: $\mathbb{Z} = \langle x \rangle$. The group algebra over the complex numbers is

$$\mathbb{C}\mathbb{Z} = \left\{ \sum_{i=-\infty}^{\infty} c_i x^i \;\middle|\; c_i \in \mathbb{C};\, c_i = 0 \text{ for all but finitely many } i \right\},$$

otherwise known as the algebra of complex Laurent polynomials in a single variable. The augmentation map ι is the evaluation homomorphism $\iota = \text{eval}_1 : \mathbb{C}\mathbb{Z} \to \mathbb{C}$, which maps a Laurent polynomial to its value at $x = 1$.

Exercise 3.3.4. Show that the augmentation ideal $I \lhd \mathbb{C}\mathbb{Z}$ is linearly generated by the set $\{x^i - 1 \mid i \in \mathbb{Z}\}$. Hence or otherwise, show I is the (principal) ideal generated by $(x-1)$. Then show that I^k is generated, as an ideal, by $(x-1)^k$.

We will now identify the associated graded algebra of $\mathbb{C}\mathbb{Z}$ with respect to the pro-unipotent filtration $\mathbb{C}\mathbb{Z} = I^0 \supseteq I^1 \supseteq I^2 \supseteq \cdots$.

To warm up, let us find the degree-0 component $\mathbb{C}\mathbb{Z}/I$. The augmentation map is clearly surjectivem as any complex number z has at least $x - 1 + z$ in its pre-image under eval_1. So, by the first isomorphism theorem,

$$\mathbb{C}\mathbb{Z} \Big/ I \cong \text{im}(\text{eval}_1) = \mathbb{C}.$$

Exercise 3.3.5. Show that in $I \,/\, I^2$, we have $(x-1) = -(x^{-1} - 1)$.

Proposition 3.3.1. For all $k \geq 0$, the degree-k component I^k / I^{k+1} of $\text{gr}\,\mathbb{C}\mathbb{Z}$ is isomorphic to $\mathbb{C}$.

Proof. This can be achieved with a slight upgrade of the argument above. First note that I^k is the ideal generated by $(x-1)^k$, and so any element $p(x)$ of I^k may be written in the form $q(x)(x-1)^k$, for some $q(x) \in \mathbb{C}\mathbb{Z}$. Define a generalised evaluation map $\iota_k : I^k \to \mathbb{C}$, given by $\iota_k(p) = q(1)$. Once again ι_k clearly surjective. Furthermore, $p(x) = q(x)(x-1)^k$ is mapped to zero if and only if $(x-1)$ divides $q(x)$. This is the case if and only if $(x-1)^{k+1}$ divides $p(x)$, that is, if and only if $p(x)$ is in I^{k+1}. Thus, by the first isomorphism theorem, we have $I^k / I^{k+1} \cong \mathbb{C}$. □

Corollary. There is an isomprphism $\text{gr}\,\mathbb{C}\mathbb{Z} \cong \mathbb{C}[[\alpha]]$, via the identification $\overline{(x-1)} \mapsto \alpha$.

corollaries should be numbered too — Z.

We were lucky this time to be able to conclusively identify the associated graded of an interesting filtered algebra via such an elemntary argument. In general, we will find that this is not usually so easy.

3.4 The Vassiliev-Goussarov Filtration for Knots

So far we have defined filtrations and gradations in an abstract algebra setting. Now, we are ready to study a specific example of a filtration, on a space of knots, getting into proper algebraic knot theory territory!

Definition 3.4.1. Let $\mathbb{C}\mathcal{K}$ denote the complex vector space with basis $\mathcal{K}$. By bilinearly extending the knot connected sum #, $\mathbb{C}\mathcal{K}$ is elevated to an algebra: the *algebra of knots*.

Note that this is in fact the monoid algebra of the knot monoid. The same construction would work over any other field of characteristic zero, such as the rational numbers or the real numbers. We use the complex numbers as this will be required later for the construction of the Kontsevich integral.

Next, we the Vassiliev – a.k.a. Vassiliev–Goussarov – filtration on $\mathbb{C}\mathcal{K}$. As usual, the 0-th component of the filtration is the entire algebra $\mathbb{C}\mathcal{K}$. Here and in general, when we extend a structure linearly, such as $\mathbb{C}\mathcal{K}$, we abbreviate its filtered components by omitting the field form the notation: we write the 0-th filtered component as $\mathcal{K}_0 = \mathbb{C}\mathcal{K}$, and denote the i-th filtered component by $\mathcal{K}_i$ rather than $\mathbb{C}\mathcal{K}_i$. The first and further components are defined diagrammatically, as below.

Definition 3.4.2. In an oriented knot diagram we call a crossing *positive* if the over strand followed by the under strand forms a positive local coordinate system for $\mathbb{R}^2$ (↗), and *negative* otherwise (↖).

Definition 3.4.3. In Vassiliev theory, a *double point* or *singular point* of a knot diagram is used to denote a positive crossing minus a negative crossing, as in Figure 22. In other words, a knot with a double point is a difference of the two knots where the double point is *resolved* as a positive or as a negative crossing. An *n-singular knot* is a knot with n double points.



Figure 22: A double point and its resolution.

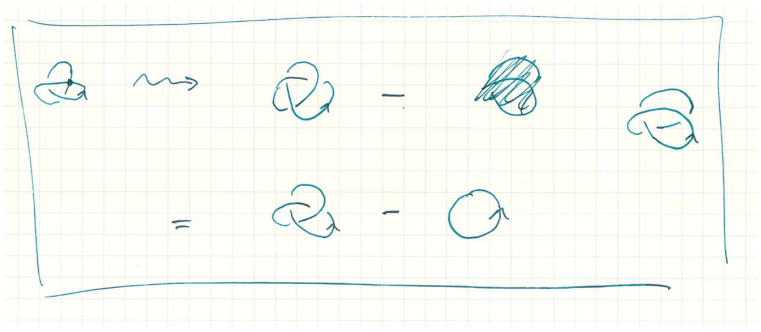


Figure 23: A one-singular knot and its resolution.

A *crossing change* on a knot diagram is the operation of changing an over-crossing to an under-crossing or vice versa. The two knots coming from the resolution of a double point always differ by a crossing change at the crossing that was resolved. Every singular knot resolves to some element of $\mathbb{C}\mathcal{K}$, as in Figure 23 and Figure 24.

A singular knot can be easily resolved into an element of the knot algebra. But to determine for an element of $\mathbb{C}\mathcal{K}$ whether it can be written as some singular knot, and if so, how, is a lot more difficult.

When resolving a knot with n double points, it doesn't matter in what order we resolve them. We will end up with a signed sum of knots with 2^n terms at first, though this may simplify if some of them turn out the be the same knot, as in Figure 24.

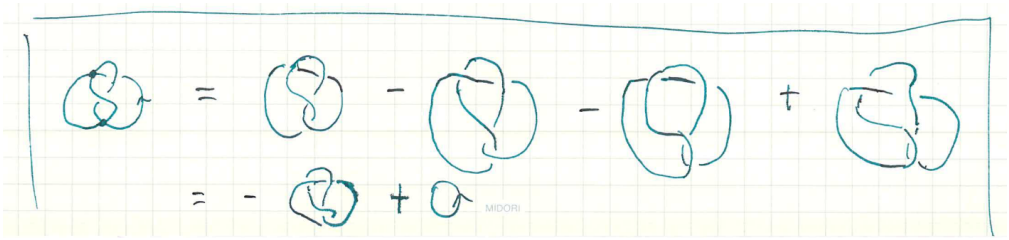


Figure 24: A two-singular knot and its resolution.

need real figures — Z.

Definition 3.4.4. Let $\mathcal{K}^n$ denote the linear span of resolutions of knots with at least n double points. The *Vassiliev Filtration* on the $\mathbb{C}\mathcal{K}$ is the filtration

$$\mathbb{C}\mathcal{K} = \mathcal{K}^0 \supseteq \mathcal{K}^1 \supseteq \mathcal{K}^2 \supseteq \mathcal{K}^3 \supseteq \cdots$$

Exercise 3.4.1. Show that the connected sum operation is compatible the Vassiliev filtration: that is, if $K \in \mathcal{K}^i$ and $L \in \mathcal{K}^j$, then $K \# L \in \mathcal{K}^{i+j}$. Thus, the Vassiliev filtration makes $\mathbb{C}\mathcal{K}$ into a filtered algebra.

Following the previous chapter, we study this filtered algebra by looking at the associated graded algebra:

$$\mathcal{A} := \operatorname{gr} \mathbb{C}\mathcal{K} = \bigoplus_{n=0}^{\infty} \mathcal{K}^n \Big/ \mathcal{K}^{n+1} \tag{3.2}$$

Let's start by looking at a the first few degrees. In degree zero we have $\mathcal{A}^0 = \mathcal{K}^0/\mathcal{K}^1$. Elements of this algebra are represented by elements of $\mathcal{K}^0 = \mathbb{C}\mathcal{K}$, and two knots represent the same element of the quotient if their difference is an element of $\mathcal{K}^1$. For any two knots that are related by a crossing change, their difference is a knot with one double point, hence, this quotient consists of knots, considered up to crossing changes.

Exercise 3.4.2. Prove that every knot diagram can be transformed into the unknot via crossing changes and Eeidemeister moves.

It follows from Exercise 3.4.2 the degree zero component of $\mathcal{A}$ is of dimension 1:

$$\mathcal{A}^0 = \mathcal{K}^0 \Big/ \mathcal{K}^1 \cong \mathbb{C}\langle 0_1 \rangle,$$

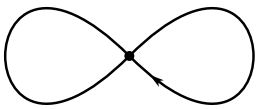
that is, the $\mathbb{C}$-linear span of the equivalence class of the unknot 0_1.

In degree one, $\mathcal{A}^1 = \mathcal{K}^1/\mathcal{K}^2$. In other words, degree one is spanned by knots with one double point, which are equivalent is they differ by a knot with two double points.

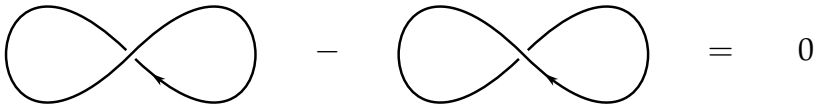


Figure 25: Two two-singular knots with numbered double points. The cyclic order of traversal for the knot on the left is $(1,1,2,2)$, while on the right it is $(1,2,1,2)$.

This means that as long as a knot already has a double point, we can change any other crossings arbitrarily. This means that anything one-singular knot diagram then be transformed via crossing changes and Reidemeister moves into the following:



In turn, this singular knot stands for the difference



Therefore, the first degree of $\mathcal{A}$ is of dimension 0:

$$\mathcal{A}^1 = \mathcal{K}^1 \Big/ \mathcal{K}^2 \cong \{0\}$$

In degree two, things get a little more complicated. Again, $\mathcal{A}^2$ is spanned by two-singular knots, and three-singular knots are set to vanish. However, unlike in lower degrees, there isn't a single knot diagram with two double points which every other knot with two double points can be transformed into.

To demonstrate this, number the double points of a knot diagram, and observe that crossing changes and Reidemeister moves don't change the cyclic order one encounters the double points while traversing around the knot. Consider for example the two knots in Figure 25: it is not possible to transform one into the other, as the cyclic ordering of the double point encounters are different, and this remains true even if we allow arbitrary re-labeling.

As in degree 1, the singular knot on the left hand side of Figure 25 of these resolves to zero in the knot algebra. It turns out that the one on the right hand side does span $\mathcal{A}^2$:

$$\mathcal{A}^2 = \mathcal{K}^2 \Big/ \mathcal{K}^3 \cong \mathbb{C} \left\langle \right\rangle$$

In this figure, can we number the double points 1 and 2, for both knots? — Z.

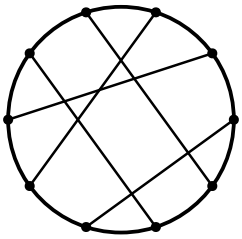
Instead of proving this explicitly, we will develop the notion that will let us attack all of the degrees of $\mathcal{A}$ at once: chord diagrams.

3.5 Chord Diagrams

So far, we have thought of elements of $\mathcal{K}^n \,/\, \mathcal{K}^{n+1}$ as equivalence classes of knots with n double points, modulo crossing changes. However, our study of degree two hints that perhaps a nicer way would be simply to record the order the n double points are encountered as one traverses the knot. We will record exactly this information as its own mathematical object, so as not to have to make an arbitrary, specific choice of an n-singular knot for each cyclic ordering. This new mathematical object is called a chord diagram:

Definition 3.5.1. A *chord diagram* of degree (or order) n is a collection of $2n$ marked points along an oriented circle, partitioned into n pairs, and considered up to orientation-preserving diffeomorphism of the circle. This is to say, the role of the circle is only to provide a cyclic ordering on the n points. In figures, chord diagrams are drawn by connecting the pairs of points with chords. We denote the vector space spanned by chord diagrams $\mathcal{C}$, and the homogeneous degree n component by $\mathcal{C}_n$.

Below is an example of a chord diagram:



Chord diagrams of order n can describe the spanning elements of $\mathcal{A}_n$, in other words, there is a surjective map

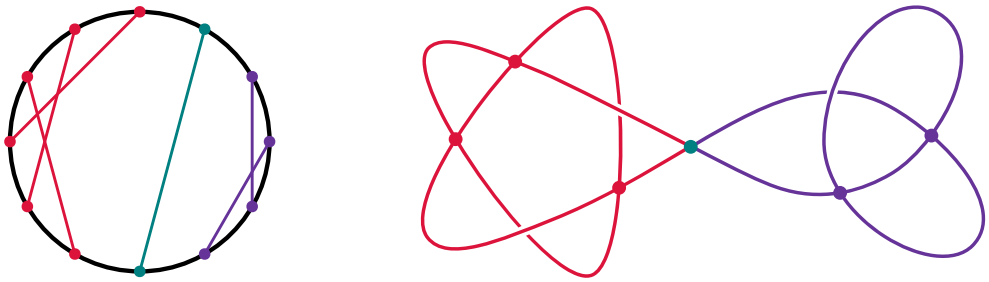
$$\pi : \mathcal{C} \to \mathcal{A}$$

given by choosing a representative n-singular knot. This can be done explicitly as follows: imagine that the cords of the chord diagram are "rubber bands" connecting pairs of points along the circle, and contract these rubber bands to create double points. Of course, this might create additional crossings, which we choose arbitrarily. The class of this representative is well-defined in $\mathcal{A}^n = \mathcal{K}^n / \mathcal{K}^{n+1}$, as it has n double points, and hence rest of the crossings can be arbitrarily changed. Because of this intuitive description, we refer to π as the *rubber band map.*

Can we have a picture of the rubber band map for some simple singular knot, like a 3-singular trefoil or figure 8? I'm thinking: on the left, chord diagram, $\mapsto$ on the right, singular knot. — Z.

We have already seen in in the example of degrees 1 and 2 that π is not injective. In both of these examples, the chord diagrams had an isolated chord — a chord which does

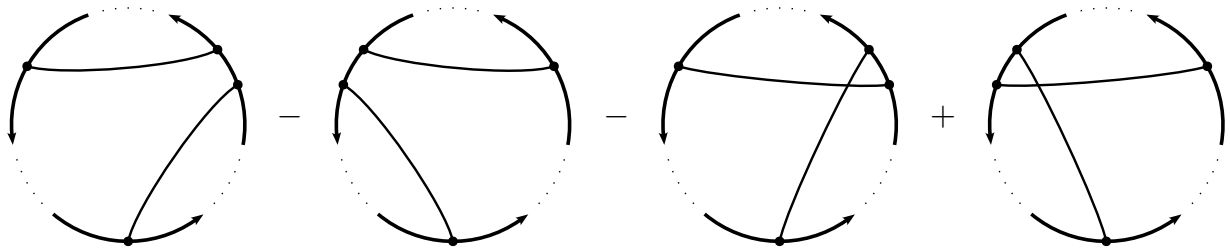
not touch (or can be deformed so as not to touch) any other chord on the chord diagram, just like the green chord below, which divides the chord diagram into a red section and a purple section.



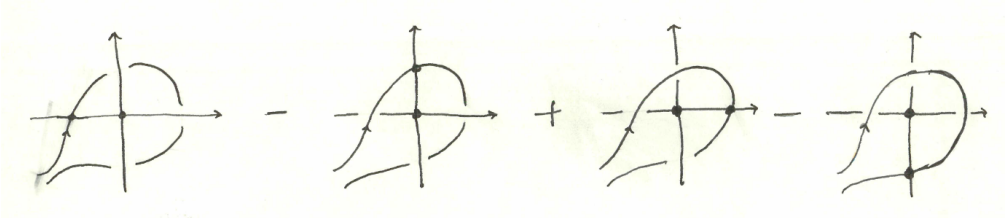
This is the same as colouring the knot string blue from the green double point, until the green double point is reached again, and then colouring it red from there till it is reached yet again. Since it's isolated, there are no chord diagrams between any of the red and blue distinguished points. That means that there are no double-points between red and blue parts of the knot, and so a representative can always be chosen in which the red and blue parts of the knot are totally disjoint (if there are red-blue crossings, they can be changed for free). So, the green double point can be resolved to give some knot K minus itself, which is 0.

Proposition 3.5.1 (One-Term (1T) or Framing Independence (FI) Relation)**.** Any chord diagram with an isolated chord resolves to $0 \in \mathcal{A}$. In other words, chord diagrams with isolated chords are in the kernel of π.

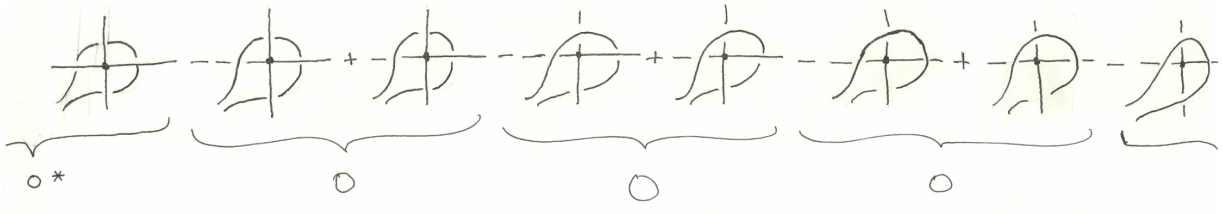
Is there anything else in the kernel of π? As it turns out, yes. Take the following element of $\mathcal{C}$:



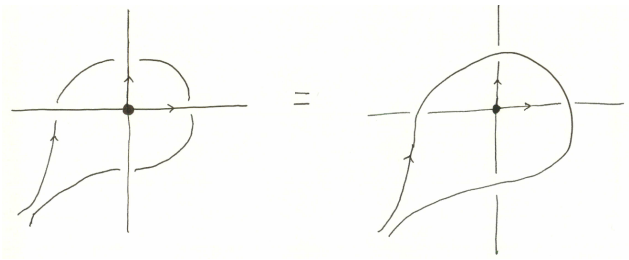
in which the dotted lines represent part of the chord diagram that are consistent between all four terms. Mapping this via π into A gives the following sketch,



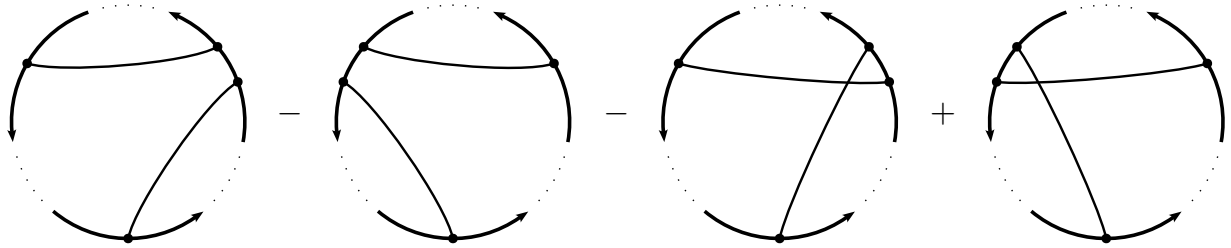
and note that since there are already $n-2$ double points in the parts we can't see and 2 in the parts we can see, we have n of them, so this is only defined up to crossing-change. To show that this is indeed in the kernel we will show that the sum of knots resolves to $0 \in \mathcal{A}^n$. The crossings have been chosen such that the sum actually resolves to 0, but in general it could resolve to an element of $\mathcal{K}^{n+1}$. Indeed, resolving the double point on the outer loop in each of the above four terms gives



in which all of the non-starred pairs of terms cancel out to give zero, and the starred pair of terms cancels out by that fact that



Proposition 3.5.2 (Four-Term (4T) Relation)**.** Any element of the form



is in the kernel of π.

As it turns out, (and we will state here and see why in the next chapter), the elements of the forms in propositions 3.5.1 and 3.5.2 span the kernel of π. By the first isomorphism theorem, we can quotient out by these elements (or rather, the relations that sets these elements to zero) and get a pictorial representation for elements of $\mathcal{A}$ in terms of $\mathcal{C}$.

Theorem 3.5.1. The associated graded space of knots (relative to the filtration $\mathcal{K}$) is isomorphic to the space of chord diagrams modulo the four-term and one-term relations.

$$\operatorname{gr}\mathcal{K} = \mathcal{A} \cong \mathcal{C} \Big/ 4T, 1T \,.$$

Exercise 3.5.1. Find all chord diagrams of degree 3.

Exercise 3.5.2. By finding all relations between chord diagrams of degree 3, find the dimension of and a basis for $\mathcal{A}^3$.

Associated graded of the connected sum induces algebra structure om chord diagrams... — Z.

More Exercises for Chapter 3

Exercise 3.5.3. Let $K \in \mathbb{C}\mathcal{K}$, and

$$K = 5_{1\text{ (r.h.)}} - 3_{1\text{ (r.h.)}}.$$

Find a singular knot that resolves to K, and hence show that K is in the first degree of the Vassiliev filtration.

Exercise 3.5.4. Let $K \in \mathbb{C}\mathcal{K}$, and

$$K = 5_{1\text{ (r.h.)}} - 2 \cdot 3_{1\text{ (r.h.)}} + 0_1.$$

Show that $K \in \mathcal{K}^2$.

Chapter 4

Formality

We saw in the previous chapter that the associated graded space of the knot algebra is chord diagrams modulo the 4T and 1T relations. Rejoice! Chord diagrams are much easier objects to deal with than knots, and in this chapter we will learn how to study knots (and more generally, other objects) through their associated graded spaces.

4.1 Formality Maps

If the associated graded construction of the previous chapter is going to be useful, then we should be able to work with it rather than with the filtered space, so there should be some filtered homomorphism from $f : A \to \operatorname{gr} A$, where we are treating $\operatorname{gr} A$ as a filtered structure by, that is, the ith filtered component consists of all of the graded components of degree i or higher.

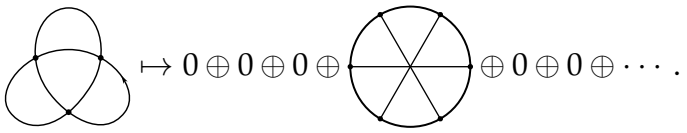
Recall that a filtered homomorphism is a homomorphism such that

$$f(A_i) \subseteq B_i.$$

In this special case of $B = \operatorname{gr} A$, this becomes

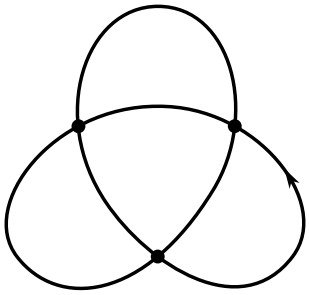
$$f(A_i) \subseteq \{0\} \oplus \cdots \oplus \{0\} \oplus \overbrace{\underbrace{A_i \big/ A_{i+1}}_{(\operatorname{gr} A)\text{'s } i\text{th graded component}} \oplus \underbrace{A_{i+1} \big/ A_{i+2}}_{(\operatorname{gr} A)\text{'s } i+1\text{th graded component}} \oplus \cdots}^{(\operatorname{gr} A)\text{'s } i\text{th } \textit{filtered} \text{ component}}$$

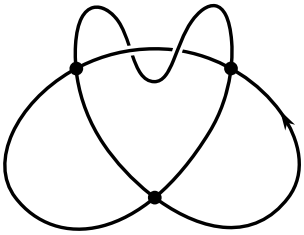
Perhaps the simplest non-trivial map that satisfies the above condition is one that sends an element of filtered degree i to its projection in $A_i \,/\, A_{i+1}$ and 0 everywhere else. In the case of knots and the Vassiliev filtration this would look like the following:



$$\mapsto 0 \oplus 0 \oplus 0 \oplus \quad \oplus 0 \oplus 0 \oplus \cdots .$$

However, in order to get the most out of A that we can from $\operatorname{gr} A$, we also want f to lose as little information as possible. In this example this clearly isn't the case as – we wouldn't be able to tell these two knots apart after this map, as they have the same chord diagram.





We want our map to retain more information than this, so we have to store more information about our input in higher degree components. We can encode this property as

$$\operatorname{gr} Z = \mathrm{id}_{\operatorname{gr}\mathcal{A}}.$$

Definition 4.1.1. A *formality map* is a filtered homomorphism satisfying

$$\operatorname{gr} Z = \mathrm{id}_{\operatorname{gr}\mathcal{O}}.$$

Formality maps are also called *homomorphic expansions* or *universal finite type invariants*.

Lemma 4.1.1 (Characterisation of formality maps)**.** A filtered homomorphism $\mathcal{O} \to \mathcal{A}$ is a formality map if and only if it has the form

$$(\operatorname{gr} Z)(K) = \sigma(K) + \text{higher degree terms.}$$

4.2 Identifying Associated Graded Structures

In general, identifying $\operatorname{gr} A$ is tricky. Often, especially in knot theoretic contexts, it is fairly easy to come up with a nice *candidate* C for what $\operatorname{gr} A$ may be – maybe we can easily give a generating set D_n for each A_n, and some relations R necessarily satisfied by their images in $A_n \,/\, A_{n+1}$. We then set $C = \prod_{n=0}^{\infty} D_n \,/\, R$, and immediately we have a surjection $\pi : C \to \operatorname{gr} A$. Identifying $\operatorname{gr} A$ is now reduced to showing π is an isomorphism. But perhaps it isn't! Perhaps our guess C is still too big, and there are some additional relations lurking in $\operatorname{gr} A$ that we have not captured in R. This is exactly the situation we find ourselves in for knots, as recalled in the second example above.
On the other hand, identifying $\operatorname{gr} A$ is worthwhile, as alluded to in (crossref): in $\operatorname{gr} A$, we may compute, and prove, things degree by degree. It would be wonderful if we had a tool for studying A via the user-friendly $\operatorname{gr} A$. This would take the form of a map $Z : A \to \operatorname{gr} A$ that lost as little information as possible.

For example we have stated but not proved the following examples of associated graded spaces:

Example 4.2.1.

- At the end of (that section), we showed directly that if one filters the group algebra $\mathbb{C}\mathbb{Z}$ of the integers - also known as the algebra of (complex) Laurent polynomials - by powers of its augmentation ideal, the associated graded is isomorphic to $\mathbb{C}[[x]]$, the algebra of (complex) power series.

- (where?) we placed the Vassiliev filtration on the algebra $\mathcal{K}$ of knots, and introduced the algebra $\mathcal{C}$ of chord diagrams on the circle, modulo the $4T$ and framing invariance relations. We found a surjection $\pi : \mathcal{C} \to \operatorname{gr}\mathcal{K}$, and conjectured that it is an isomorphism.

Everything can be neatly rephrased as follows

Theorem 4.2.1. Let $\mathcal{K}$ be some filtered algebraic structure, and $\mathcal{A}$ be some graded algebraic structure. Since $\mathcal{A}$ is graded, it is also filtered. The following are equivalent:

1. There exists a filtered homomorphism $Z : \mathcal{K} \to \mathcal{A}$ such that
$$\operatorname{gr} Z = \mathrm{id}_{\operatorname{gr}\mathcal{K}}.$$
2. There exists a filtered homomorphism $Z : \mathcal{K} \to \mathcal{A}$ such that
$$(\operatorname{gr} Z)(K) = \sigma(K) + \text{higher degree terms.}$$
3. $\mathcal{A}$ is the associated graded version of $\mathcal{K}$,
$$\operatorname{gr}\mathcal{K} \cong \mathcal{A}.$$

More Exercises for Chapter 4

Chapter 5

The Kontsevich Integral

In this chapter we construct the universal finite type invariant $Z : \mathcal{K} \to \mathcal{C}$, named the Kontsevich integral for its creator, Maxim Kontsevich, who received a Fields medal for his efforts. We will embed knots and tangles in $\mathbb{C} \times \mathbb{R} \cong \mathbb{R}^3$, so we can use complex analysis to do knot theory. As a warm-up, we will see how we can count the number of twists of a two-strand braid in this way, using an integral to count the rotations of a chord with endpoints on the braid. Essentially, the idea of the Kontsevich integral is to count the rotations of m chords, for all ways of placing m chords on the strands of a knot, for all values of m. It is believable that such a vast amount of information would yield a powerful knot invariant. Indeed it does, but proving so will take some work.

5.1 Prelude: a Complete Invariant for B_2

Our main complex analytical tool will be the following special case of Cauchy's integral formula, which says that the number of times a curve $\gamma : [0,1] \to \mathbb{C}$ rotates anticlockwise around a point $a \in \mathbb{C}$ is given by

$$\frac{1}{2\pi i} \int_\gamma \frac{dz}{z-a},$$

which is not necessarily an integer when γ is not a closed curve. Recall that the set of ambient isotopy classes of braids on two strands, along with the operation of sticking braids together, forms a group called B_2. Suppose we take four points $(a, t_{\min})$, $(a, t_{\max})$, $(b, t_{\min})$, and $(b, t_{\max})$ in $\mathbb{C} \times \mathbb{R}$, and tie two pieces of string between them to form a braid.

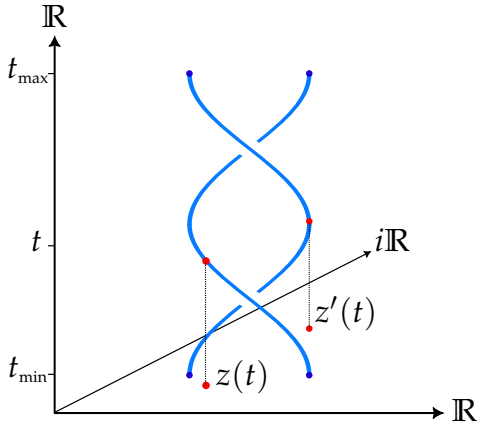


Figure 28: A two-strand braid living in $\mathbb{C} \times \mathbb{R}$.

How can we determine which element of the braid group our sculpture represents? B_2 has a single generator σ, a positive half-twist, and no relations. This means all elements of B_2 are uniquely determined by their (signed) number of half-twists, and all we have to do is count. Since we required braid strands to be monotonically ascending, we can write the complex horizontal coordinates along the two strands as functions of the vertical coordinate, as $z(t)$ and $z'(t)$. Then the two strands twisting around each other is equivalent to the chord with endpoints $z(t)$ and $z'(t)$ rotating. This in turn equivalent to the difference $z(t) - z'(t)$ rotating around the origin, and so

$$2 \cdot \frac{1}{2\pi i} \int_{t_{\min}}^{t_{\max}} \frac{dz - dz'}{z - z'}.$$

computes the number of half-twists, which is guaranteed to be an integer.

5.2 Definition

5.2.1 Setup

Begin with an **oriented** knot $K \subset \mathbb{C} \times \mathbb{R}$. We can measure how high we are on the knot by sending a point $(z, t) \in K$ to t. Call its minimum and maximum on the knot $t_{\min}$ and $t_{\max}$ respectively. A point where the differential of this function vanishes is called a critical point, and the value of the height function at a critical point is called a **critical value**. These are the heights t_c at which the number of intersections of our knot with a horizontal plane changes. We can, through judicious jiggling, adjust our knot so that:

- There are finitely many critical values, and they are distinct. We will label them $t_{\max} > t_{c_1} > t_{c_2} > \cdots > t_{c_r} > t_{\min}$.
- Between critical points, the height function is monotonic on each strand. That is, each slice between critical points $K \cap (\mathbb{C} \times (t_{c_{i+1}}, t_{c_i}))$ looks like a braid on some number of strands, except some strands may be going down.

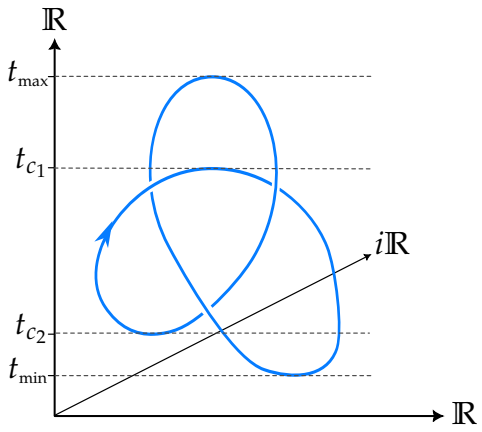


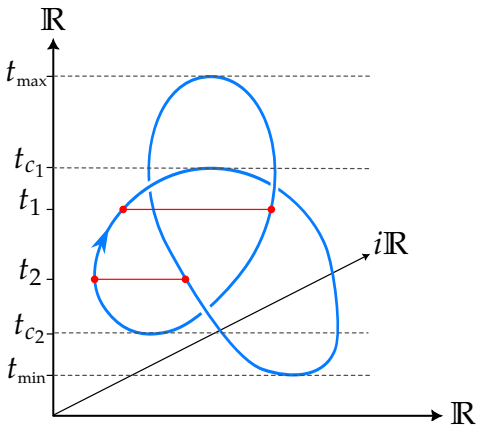
Figure 29: A strict Morse trefoil knot living in $\mathbb{C} \times \mathbb{R}$.

An embedding satisfying these conditions is called a **strict Morse knot**, and an example is given in Figure 29. We have satisfied the strict Morse requirement by slightly breaking the symmetry of the picture of the trefoil we know and love, and tugging the bottom right lobe slightly downward. It has four critical values, $t_{\max}$, t_{c_1}, t_{c_2}, and $t_{\min}$. With strict Morse knots, we are ready to cook up the Kontsevich integral.

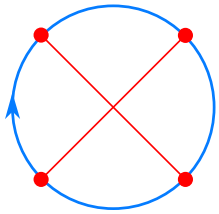
5.2.2 Ingredients

Let $m \geq 1$, and take m numbers $t_{\max} > t_1 > t_2 > \cdots > t_m > t_{\max}$, none of which are critical values of the height function on K.

- At each level $t = t_j$, choose an unordered pair of strands, and write their horizontal coordinates as functions z_j and z'_j of height as before. The set $\left\{(z_j, z'_j) \,|\, 1 \leq j \leq m\right\}$ is called a **pairing**. Think of each pair as a chord with endpoints on the knot, as in Figure 30.



(a) A pairing P on the trefoil, for $m = 2$.



(b) The corresponding chord diagram D_P.

Figure 30

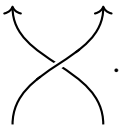
- $\downarrow_P$ is the number of strands in the pairing P that are going down if you follow the knot in the direction of its orientation. For the pairing in Figure 30, $\downarrow_P = 2$.
- D_P is the chord diagram obtained by unwrapping the knot, with the chords from P, into an oriented circle. Obviously this unwrapping is not an ambient isotopy, as we pass strands through each other at will.
- the iterated integral in m variables which we will write with the wedge

5.2.3 Procedure

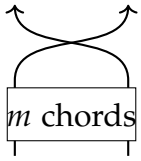
$$Z(K) = \sum_{m=0}^{\infty} \frac{1}{(2\pi i)^m} \int\limits_{\substack{t_{\min} \leq t_m < \cdots < t_1 \leq t_{\max} \\ t_j \text{ are noncritical}}} \sum_{\substack{\text{pairings} \\ P}} (-1)^{\downarrow_P} D_P \bigwedge_{j=1}^{m} \frac{dz_j - dz'_j}{z_j - z'_j}$$

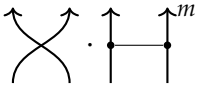
5.3 Calculation for the Positive Crossing

Let us perform our first, and likely only, explicit calculation of the Kontsevich integral. We will calculate it for a single positive crossing,

$T =$  .

There are no critical points, and for each m, there is only one pairing P, with $\downarrow_P = 0$, and

$D_P =$ 

$=$  .

Thus our integral reduces to

$$Z\left(\times\right) = \times \sum_{m=0}^{\infty} \uparrow\cdots\uparrow^{m} \frac{1}{(2\pi i)^m} \int_{t_{\min}\le t_m<\cdots<t_1\le t_{\max}} \bigwedge_{j=1}^{m} \frac{dz - dz'}{z - z'}$$

$$= \times \sum_{m=0}^{\infty} \uparrow\cdots\uparrow^{m} \int_{t_{\min}}^{t_{\max}} \int_{t_{\min}}^{t_1} \cdots \int_{t_{\min}}^{t_{m-1}} f(t_m)\cdots f(t_2) f(t_1)\, dt_m \cdots dt_2 dt_1,$$

where in an attempt to maintain some semblance of brevity, we write

$$f(t) := \frac{1}{2\pi i} \frac{\frac{d}{dt}\left(z(t) - z'(t)\right)}{z(t) - z'(t)}.$$

All that remains is to evaluate the iterated integral. We know that for $m = 1$ we are simply counting the braiding number, and

$$\int_{t_{\min}}^{t_{\max}} f(t)\, dt = \frac{1}{2}.$$

We can evaluate the iterated integral using the following general lemma.

Lemma 5.3.1.

$$\int_a^b \int_a^{t_1} \cdots \int_a^{t_{m-1}} f(t_m)\cdots f(t_2) f(t_1)\, dt_m \cdots dt_2 dt_1 = \frac{1}{m!}\left[\int_a^b f(t)\, dt\right]^m. \tag{5.1}$$

Proof. The proof is by induction. Equality for $m = 1$ comes for free. Then, applying the inductive hypothesis to the innermost $m - 1$ integrals, the left hand side of (5.1) is equal to

$$I := \frac{1}{(m-1)!} \int_a^b \left[\int_a^{t_1} f(t)\, dt\right]^{m-1} f(t_1)\, dt_1.$$

Integration by parts and the chain rule then reveal

$$I = \frac{1}{(m-1)!}\left[\int_a^{t_1} f(t)\,dt\right]^m - (m-1)I,$$

and so I is equal to the right hand side of (5.1). □

Applying this to the problem at hand, we finally arrive at

$$Z\left(\;\times\;\right) = \times \sum_{m=0}^{\infty} \frac{1}{m!}\left(\frac{1}{2}\;\uparrow\!\!\bullet\!\!-\!\!\bullet\!\!\uparrow\;\right)^m$$
$$= \times \exp\left(\frac{1}{2}\;\uparrow\!\!\bullet\!\!-\!\!\bullet\!\!\uparrow\;\right). \tag{5.2}$$

Exercise 5.3.1. Slightly modify the above argument to show that

$$Z\left(\;\times\;\right) = \times \exp\left(-\frac{1}{2}\;\uparrow\!\!\bullet\!\!-\!\!\bullet\!\!\uparrow\;\right).$$

Later we will be able to show this using only (5.2) and the properties of the Kontsevich integral.

5.4 Invariance

More Exercises for Chapter 5

Chapter 6

Drinfeld Associators

I've put in some section titles to help break this up — Z.

6.1 Tangles as a finitely presented tensor category

6.2 The Kontsevich integral for tangles: parametrised tensor products

6.3 Parenthesised tangles

6.4 Formality for parenthesised tangles

6.5 A combinatorial Kontsevich Integral

6.6 Drinfel'd associators

More Exercises for Chapter 6

Appendix A

Solutions to Selected Exercises

Chapter 1

Solution (Ex. 1.2.5)**.** For a surjective but non-injective homomorphism, we define $G = \langle x, y, z \mid x^n = 1, y^2 = 1 \rangle$, and $\phi(x) = a$, $\phi(y) = b$ and $\phi(z) = 1_H$.

Solution (Ex. **??**)**.** *Hint:* Can you delete a strand and be left with a single crossing (the generator of B_2? No, if you delete a strand what you get isn't even a braid. Can you do something to β to fix this problem?

Solution (Ex. 1.1.5)**.** *Hint:* Consider all pairs of braid strands in B_n and find a homomorphism into B_2.

Chapter 2

Chapter 3

Solution (Ex. 3.3.2)**.** Actually, we will prove a much more general and useful result, that the kernel $K = \ker \varphi \subseteq R$ of any ring homomorphism $\varphi : R \to S$ is a two-sided ideal of R. We need to show that:

a) K is an additive subgroup of R. It suffices to check that for all x and y in K, $x - y$ is also in K. This is true by linearity: $\varphi(x - y) = \varphi(x) - \varphi(y) = 0 - 0 = 0$, and hence $x - y \in K$.

b) For all $r \in R$ and $k \in K$, rk and kr are also in K. This is also a simple consequence of φ being a ring homomorphism: $\varphi(rk) = \varphi(r)\varphi(k) = \varphi(r) \cdot 0 = 0$ and so $rk \in K$. The proof for kr proceeds the same way, and so K is a two-sided ideal.

The special case of applying this to $\phi = \iota$ is the result we were after. In fact, we can go one better again: the converse is also true! As the quotient by a two-sided ideal I is again a ring, the projection $\pi : R \to R/I$ is a ring homomorphism, which clearly has kernel I. So, the two-sided ideals of R are exactly the kernels of ring homomorphisms from R.

Solution (Ex. 3.3.4)**.** Take $r \in I$. Then r is a Laurent polynomial of the form $\sum_{i=-N}^{N} c_i x^i$, with $\sum_{i=-N}^{N} c_i = 0$. We want to show that it can be written as a linear combination

of elements of the set $\{x^i - 1 \mid i \in \mathbb{Z}\}$, and do so by subtracting zero, in the form of $\sum_{i=-N}^{N} c_i$:

$$r = \sum_{i=-N}^{N} c_i x^i = \sum_{i=-N}^{N} c_i x^i - \sum_{i=-N}^{N} c_i = \sum_{i=-N}^{N} c_i (x^i - 1).$$

So the $x^i - 1$ span I. We will now show that $x - 1$ generates I as an ideal by showing that each of these elements is a product of $x - 1$ and something in $\mathbb{CZ}$. Let $i \geq 1$. It is easy to show by expanding the brackets that $x^i - 1 = (x-1)(x^{i-1} + x^{i-2} + \cdots + x + 1)$. Then as $x^{-i} - 1 = -x^{-i}(x^i - 1)$, we are done.

Chapter 4

Solution (Ex. 3.4.2)**.** ¡proof that every knot is crossing-change equivalent to the unknot¿

Solution (Ex. 3.5.3)**.** We take K and massage the second term by ambient isotopy to find something that is a crossing change away from the first term. This is shown in Fig. 31.

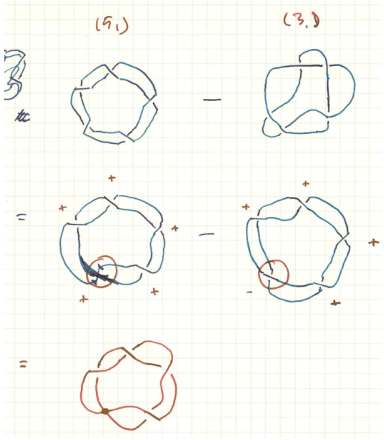


Figure 31: After ambient isotopy (equivalent to a Reidemeister II move on the diagram), the second term differs from the first by only a crossing change.

Solution (Ex. 3.5.4)**.** We take K and do some ambient isotopy to the last two terms, getting them to look like crossing changes of the cinquefoil.