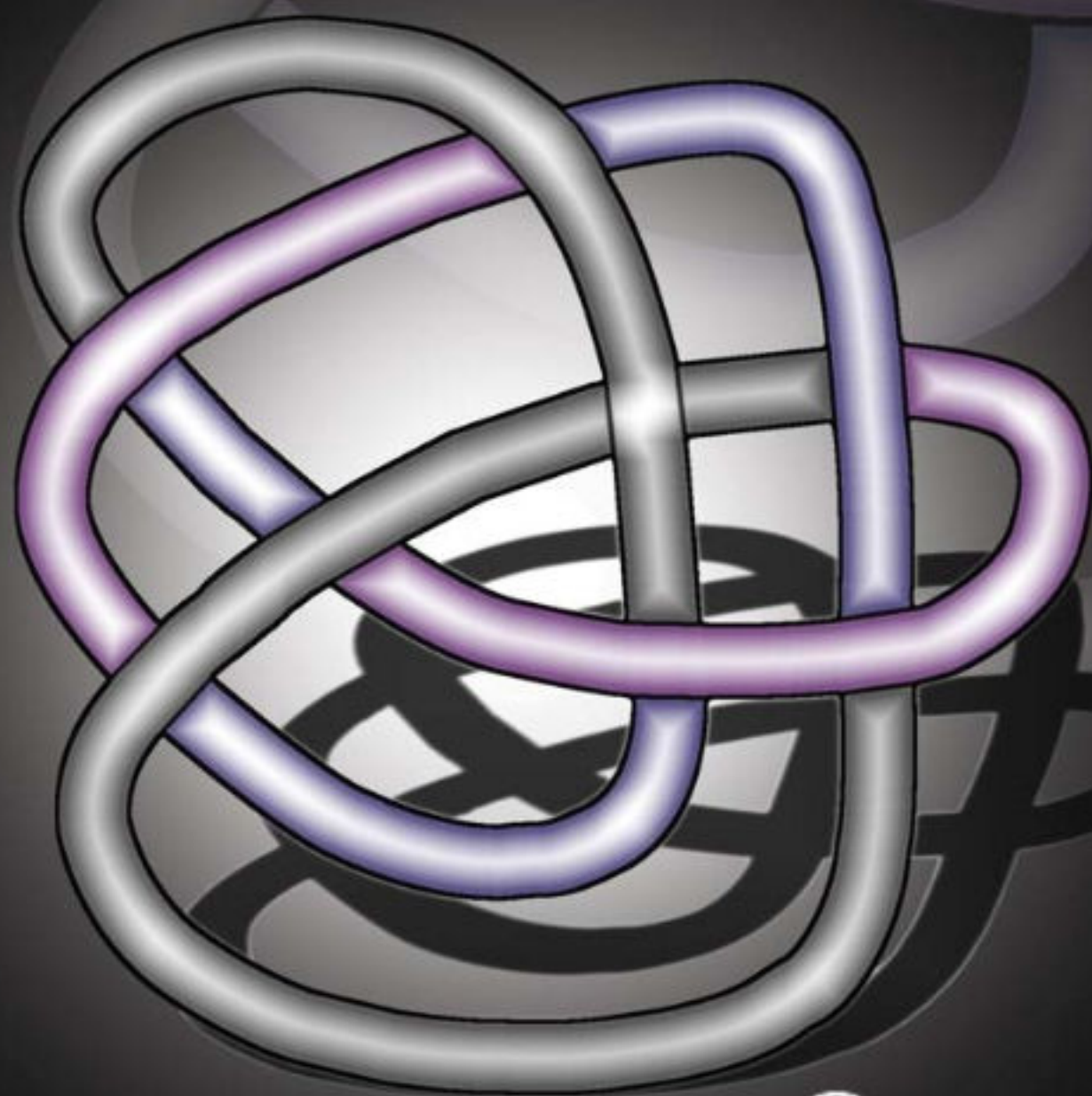


STUDENT MATHEMATICAL LIBRARY
Volume 74

Quandles

An Introduction to
the Algebra of Knots

Mohamed Elhamdadi
Sam Nelson



Quandles

An Introduction to
the Algebra of Knots

STUDENT MATHEMATICAL LIBRARY
Volume 74

Quandles

An Introduction to the Algebra of Knots

Mohamed Elhamdadi
Sam Nelson



American Mathematical Society
Providence, Rhode Island

Editorial Board

Satyan L. Devadoss
Erica Flapan

John Stillwell (Chair)
Serge Tabachnikov

2010 *Mathematics Subject Classification*. Primary 57M25, 55M25, 20N05, 20B05, 55N35, 57M05, 57M27, 20N02, 57Q45.

For additional information and updates on this book, visit
www.ams.org/bookpages/stml-74

Library of Congress Cataloging-in-Publication Data

Elhamdadi, Mohamed, 1968–

Quandles: an introduction to the algebra of knots / Mohamed Elhamdadi, Sam Nelson.

pages cm. – (Student mathematical library ; volume 74)

Includes bibliographical references and index.

ISBN 978-1-4704-2213-4 (alk. paper)

1. Knot theory. 2. Low-dimensional topology. I. Nelson, Sam, 1974– II. Title. III. Title: Algebra of Knots.

QA612.2.E44 2015

514'.2242–dc23

2015012551

Copying and reprinting. Individual readers of this publication, and nonprofit libraries acting for them, are permitted to make fair use of the material, such as to copy select pages for use in teaching or research. Permission is granted to quote brief passages from this publication in reviews, provided the customary acknowledgment of the source is given.

Republication, systematic copying, or multiple reproduction of any material in this publication is permitted only under license from the American Mathematical Society. Permissions to reuse portions of AMS publication content are handled by Copyright Clearance Center's RightsLink® service. For more information, please visit: <http://www.ams.org/rightslink>.

Send requests for translation rights and licensed reprints to reprint-permission@ams.org.

Excluded from these provisions is material for which the author holds copyright. In such cases, requests for permission to reuse or reprint material should be addressed directly to the author(s). Copyright ownership is indicated on the copyright page, or on the lower right-hand corner of the first page of each article within proceedings volumes.

© 2015 by the authors. All rights reserved.

Printed in the United States of America.

∞ The paper used in this book is acid-free and falls within the guidelines established to ensure permanence and durability.

Visit the AMS home page at <http://www.ams.org/>

10 9 8 7 6 5 4 3 2 1 20 19 18 17 16 15

Contents

Preface	vii
Chapter 1. Knots and Links	1
§1. Knots and Links	1
§2. Combinatorial Knot Theory	9
§3. Knot and Link Invariants	16
Chapter 2. Algebraic Structures	27
§1. Operation Tables and Isomorphisms	27
§2. Quotient Sets and Equivalence Relations	36
§3. Modules	47
§4. Groups	53
§5. Cohomology	66
Chapter 3. Quandles	73
§1. Kei	73
§2. Quandles	87
§3. Alexander Quandles and the Alexander Polynomial	96
Chapter 4. Quandles and Groups	107
§1. Fundamental Group	107
§2. Braid Groups	110

§3. Knot Groups	120
§4. Knot Quandles	125
§5. Augmented Quandles	128
§6. Quandles and Quasigroups	135
Chapter 5. Generalizations of Quandles	147
§1. Racks	147
§2. Bikei	156
§3. Biracks and Biquandles	163
Chapter 6. Enhancements	173
§1. Basic Enhancements	173
§2. Structure Enhancements	181
§3. Quandle Polynomials	188
§4. Quandle Cocycle Enhancements	192
Chapter 7. Generalized Knots and Links	207
§1. Colorings of Tangles and Embeddings	207
§2. Surface Knots	213
§3. Virtual Knots	220
Bibliography	237
Index	243

Preface

Quandles and their kin (kei, racks, biquandles and biracks) are algebraic structures whose axioms encode the movements of knots in space in the same way that groups encode symmetry and orthogonal transformations encode rigid motion. Quandle theory thus brings together aspects of topology, abstract algebra and combinatorics in a way that is easily accessible using pictures and diagrams.

The term “quandle” was coined by David Joyce in his PhD dissertation, written in 1980 and published in 1982 [**Joy82**]. Previous work had been done as far back as 1942 by Mituhisa Takasaki [**Tak42**], who used the term “kei” for what Joyce would later call “involutory quandles”. In the 1950s Conway and Wraith [**CW**] informally discussed a similar structure they called “wracks” from the phrase “wrack and ruin”. At the same time Joyce was writing about quandles, Sergey V. Matveev [**Mat82**] was writing behind the iron curtain about the same algebraic structure, using the more descriptive term “distributive groupoids”. Louis Kauffman [**Kau91**] used the term “crystals” for a form of the quandle structure. In the mid 1980s a generalized form of the quandle idea was independently discovered by Brieskorn [**Bri88**], who chose the descriptive term “automorphic sets”.

In 1992 Roger Fenn and Colin Rourke [**FR92**] wrote a seminal work reintroducing the quandle idea and a generalization; they chose to use the Conway/Wraith term “wracks” while dropping the “w”

to obtain the term “racks”, canceling the “w” along with the writhe independence. In subsequent work [FRS95] they suggested a further generalization known as “biracks” with a special case known as “biquandles”. Biquandles were explored in detail in 2002 by Louis Kauffman and David Radford [KR03], with later work by others [CES04, FRS95, NV06].

Fenn, Rourke and Sanderson introduced in [FRS95] a cohomology theory for racks and quandles, analogous to group homology. This ultimately led to the current popularity of quandles, since it allowed Scott Carter, Daniel Jelsovsky, Seiichi Kamada, Laurel Langford and Masahico Saito in [CJK⁺03] to define an enhancement of the quandle counting invariant using quandle cocycles, leading to new results about knotted surfaces and more. It was this and subsequent work that led the present authors to study quandles, and ultimately led to this book.

If one restricts oneself to the most important quandle axiom, namely self-distributivity, then one can trace this back to 1880 in the work of Pierce [Pei80] where one can read the following comments: *“These are other cases of the distributive principle These formulae, which have hitherto escaped notice, are not without interest.”* Another early work fully devoted to self-distributivity appeared in 1929 by Burstin and Mayer [BM29] dealing with distributive quasigroups: binary algebraic structures in which both right multiplication and left multiplication are bijections, and with the extra property that the operation is left and right distributive on itself (called also Latin quandles).

As quandle theorists, we have found quandle theory not only intrinsically interesting but also very approachable for undergraduates due to its unique mix of geometric pictures and abstract algebra. This book is intended to serve as a text for a one-semester course on quandle theory which might be an upper division math elective or as preparation for a senior thesis in knot theory.

This book assumes that the reader is comfortable with linear algebra and basic set theory but does not assume any previous knowledge of abstract algebra, knot theory or topology. The reader should be

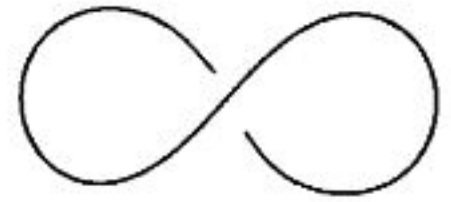
familiar with sets, unions, intersections, Cartesian products, functions between sets, injective/surjective/bijective maps as well as vector spaces over fields, linear transformations between vector spaces, and matrix algebra in general. Readers should also be familiar with the integers \mathbb{Z} , rationals \mathbb{Q} , reals \mathbb{R} and complex numbers \mathbb{C} .

The book is organized as follows.

Chapter 1 introduces the basics of knot theory; advanced readers may opt to skip directly to Chapter 2. Chapter 2 introduces important ideas from abstract algebra which are needed for the rest of the book, including introductions to groups, modules, and cohomology assuming only a linear algebra background. Chapter 3 gives a systematic development of the algebraic structures (quandles and kei) arising from oriented and unoriented knots and links, including both theory and practical computations. Chapter 4 looks at important connections between quandles and groups and introduces the basics of algebraic topology, including the fundamental group and the geometric meaning of the fundamental quandle of a knot. In Chapter 5 we look at generalizations of the quandle idea, including racks, bikei, biquandles and biracks. Chapter 6 introduces enhancements of representational knot and link invariants defined from quandles and their generalizations. In Chapter 7 we conclude with applications to generalizations of knots including tangles, knotted surfaces in \mathbb{R}^4 , and virtual knots.

The authors wish to thank our many students, colleagues and friends without whom this book would not have been possible.

Chapter 1



Knots and Links

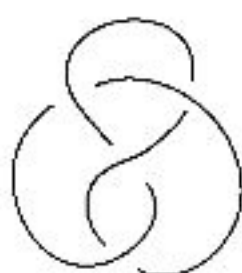
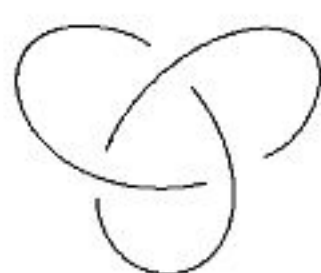
1. Knots and Links

A *knot* is a simple closed curve, where “simple” means the curve does not intersect itself and “closed” means there are no loose ends. We usually think of knots in three-dimensional space since simple closed curves in the line and plane are pretty boring and, perhaps surprisingly, simple closed curves in 4 or more dimensions are also boring, as we will see.

Two knots K_0 and K_1 have the same *knot type* if we can move K_0 around in space in a continuous way, i.e. without cutting or tearing the knot (or the space in which the knot lives!) to match up K_0 with K_1 . Formally, K_0 is *ambient isotopic* to K_1 if there is a continuous map $H : \mathbb{R}^3 \times [0, 1] \rightarrow \mathbb{R}^3$ such that $H(K_0, 0) = K_0$, $H(K_0, 1) = K_1$ and $H(x, t)$ is injective (one-to-one) for every $t \in [0, 1]$. Such a map is called an *ambient isotopy*; if you think of t as a time variable, then H is a movie showing how to continuously deform K_0 onto K_1 . If there exists an ambient isotopy H taking K_0 to K_1 we write $H : K_0 \xrightarrow{\sim} K_1$.

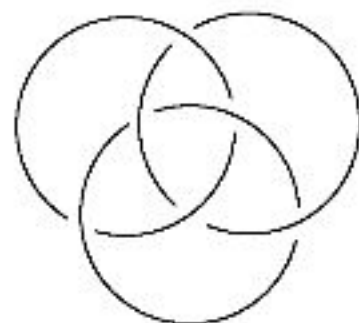
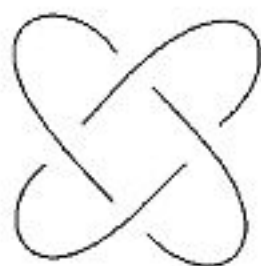
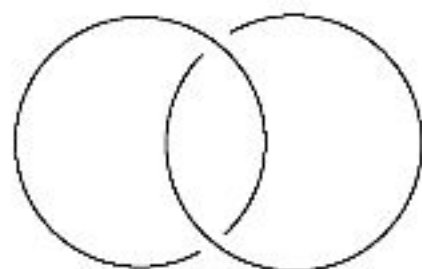
To specify a knot K we could make a physical model by tying the knot in a rope or cord; a nice trick suggested by Colin Adams in [Ada04] is to use an extension cord, so you can join the ends together by plugging the plug into the outlet end.

To specify knots in a more print-friendly format, we could give a parametric function $f(t) = (x(t), y(t), z(t))$ where $0 \leq t \leq 1$ and $f(0) = f(1)$. This approach is required in order to study *geometric knot theory*, where the exact positioning of K in space is important. In *topological knot theory*, however, we only care about the position up to ambient isotopy; thus, a simpler solution is to draw pictures or *knot diagrams*. Formally, a *knot diagram* is a projection or shadow of a knot on a plane where we indicate which strand passes over and which passes under at apparent crossing points by drawing the understrand broken.



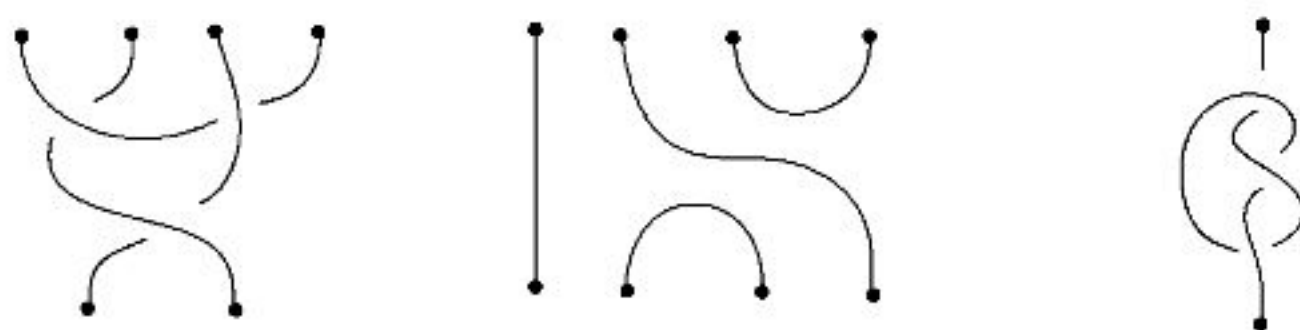
A knot is *tame* if it has a diagram with a finite number of crossing points; knots in which every projection has infinitely many crossing points are called *wild knots*. We will only deal with tame knots in this book.

Links, Tangles and Braids (oh my!) There are many kinds of objects related to knots. A *link* consists of several knots possibly linked together; each individual simple closed curve is a *component* of the link. A knot is a link with only one component.

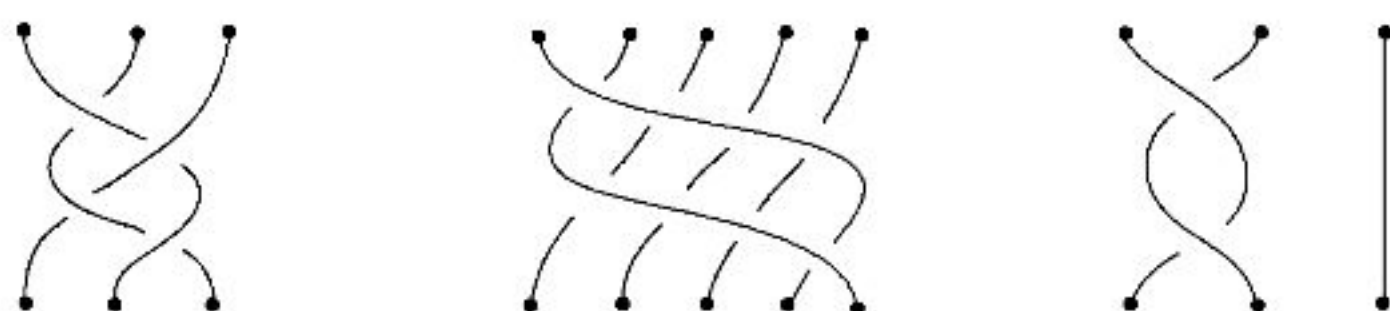


A *tangle* is a portion of a knot or link with fixed endpoints we can think of as inputs and outputs. If there are n inputs and m outputs,

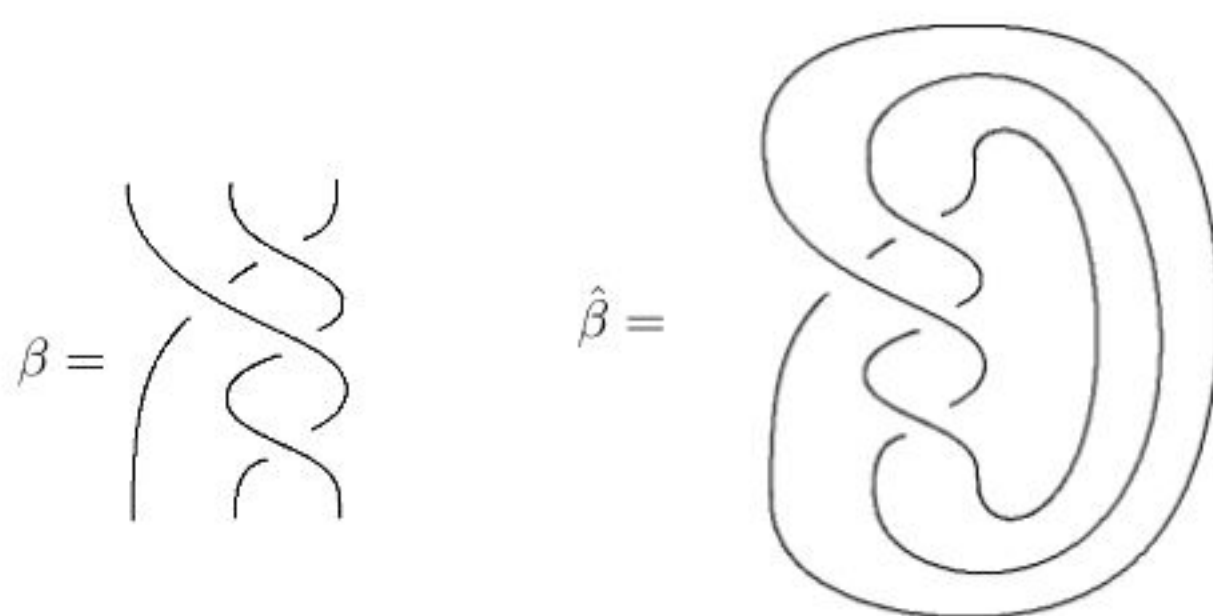
we have an (n, m) -tangle.



A *braid* is a tangle which has no maxima and no minima in the vertical direction, i.e., a tangle whose strands do not turn around. Note that in any braid, the number of inputs must equal the number of outputs, unlike more general tangles.

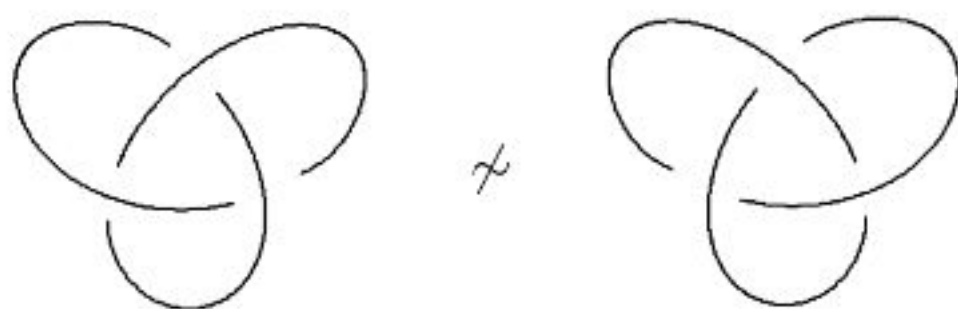


For any braid β there is a knot or link $\hat{\beta}$ called the *closure* of the braid, obtained by joining the top strands to the bottom strands. The converse is also true – every knot or link can be put into braid form, a fact known as *Alexander's Theorem*.

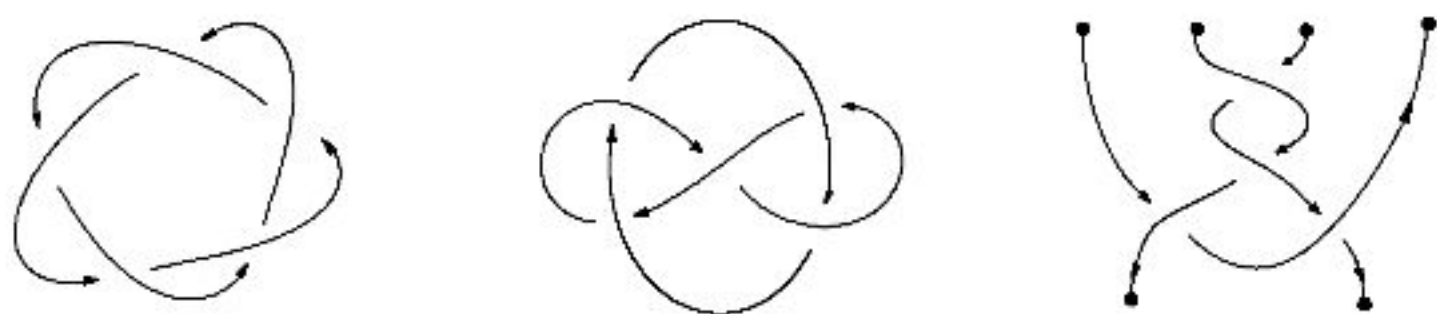


The *obverse* of a knot K is the mirror image of K , denoted \bar{K} . A knot may or may not be equivalent to its obverse – the trefoil knot comes in distinct left- and right-handed varieties, for instance. Knots which are different from their obverses are called *chiral*, while knots

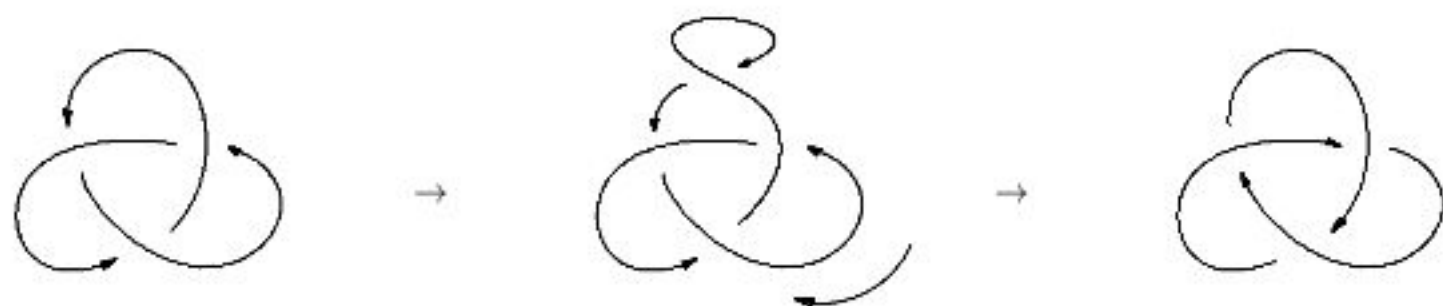
which are ambient isotopic to their obverses are called *amphichiral*.



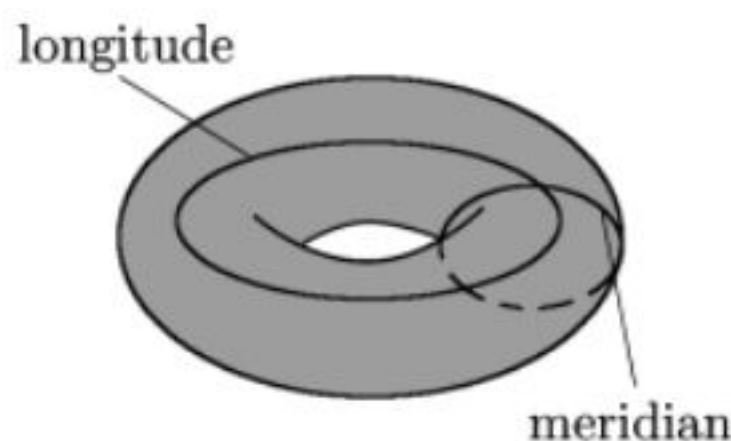
Oriented Knots. For each strand in a knot, link, tangle or braid, we can make a choice of *orientation* or preferred direction of travel. Knots described by a parametrization have an implied orientation in the direction of increasing t value; braids also have an implied orientation of all strands oriented in the same direction (up or down depending on the author's choice of convention). For generic oriented knots, links and tangles, we specify the orientation of each strand with an arrow.



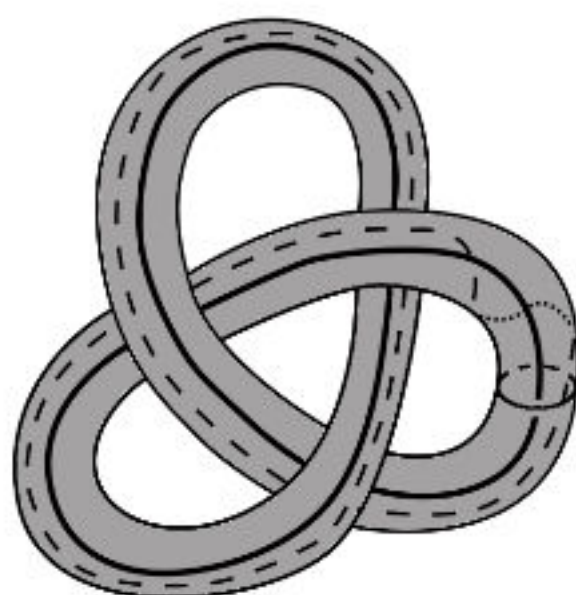
Reversing the orientation of an oriented knot K yields a possibly different oriented knot called the *inverse* or *reverse* of K , denoted $-K$. For two oriented knots K_0 and K_1 to be equivalent, we need an ambient isotopy $H : K_0 \xrightarrow{\sim} K_1$ which respects the orientation of K_1 . For example, the trefoil knot K below is equivalent to its inverse $-K$ as illustrated:

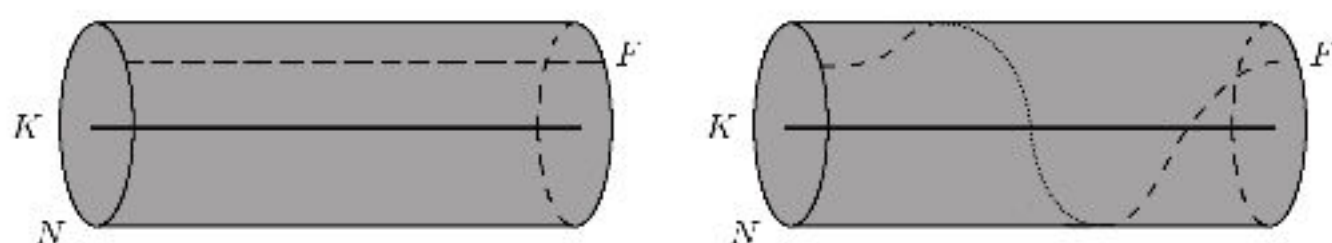


Framed Knots. Like a choice of orientation, a *framing* of a knot is a choice of extra structure we can give to a knot which then must be preserved by an ambient isotopy for two framed knots to be equivalent. Start by inflating the knot K like an inner tube, so we have a knotted solid torus N with K as its core. This solid torus is called a *regular neighborhood* of the knot. A circle on the torus which goes around the torus with the knot is called a *longitude*, while a circle going around a disk slice of the solid torus with the knot at its center is called a *meridian*.



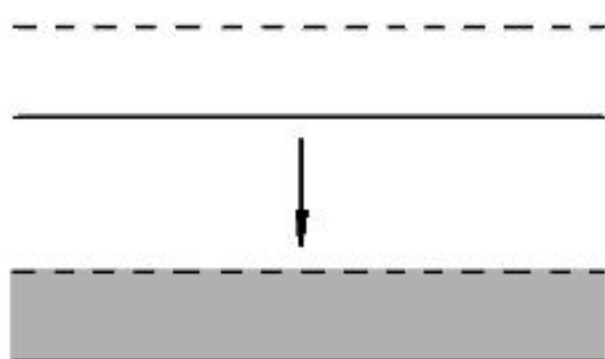
A *framing curve* F is a simple closed curve on the surface of the torus which projects down onto the original knot K in an injective (one-to-one) way, i.e., a longitude of the torus. While F goes around the torus with K exactly once in the longitudinal direction, it can wrap around the meridional direction of the torus any integer number of times.





Let K be a knot and F a framing curve. A *framed isotopy* of (K_0, F_0) to (K_1, F_1) is an ambient isotopy which carries K_0 to K_1 and carries F_0 to F_1 . For a given knot K , with framing curve F , the number of times F wraps meridianally around K (with counterclockwise wraps counted with a positive sign and clockwise twists counted with a minus sign) is called the *framing number* of the framed knot (K, F) . For a fixed knot K , two framed knots (K, F_0) and (K, F_1) are framed isotopic only if the framing numbers are equal.

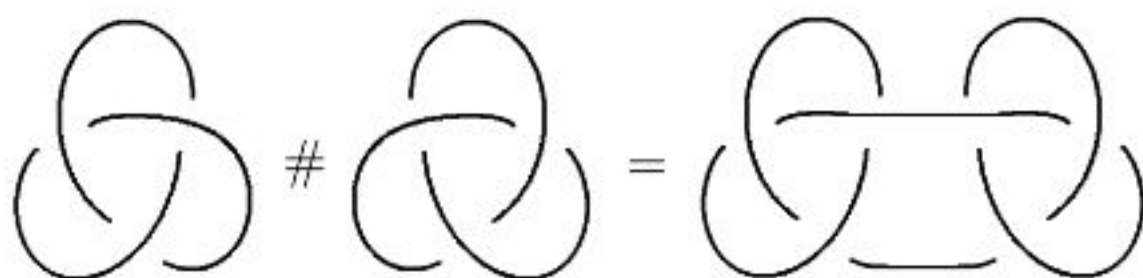
We can think of a framed knot as a 2-component link with the knot and its framing curve forming two sides of a ribbon.



Then, framed isotopy can be understood as movement of the ribbon through space. Similarly a framed link of n components can be understood as an ordinary link of $2n$ components where the components come in parallel pairs forming the two sides of n ribbons, with each of the original n components having its own framing curve and framing number. Similarly, in a framed braid or framed tangle, each strand has its own framing curve and framing number.

We can think of framed isotopy as a mathematical model for knotted 3-dimensional ropes or tori, where ambient isotopy is the model for knotted 1-dimensional curves.

Connected Sums. Knots and links have an operation known as *connected sum* where two knots are joined into a single knot by cutting the knots and joining the loose ends to form a single knot. We write $K_0 \# K_1$ for the connected sum of K_0 and K_1 .

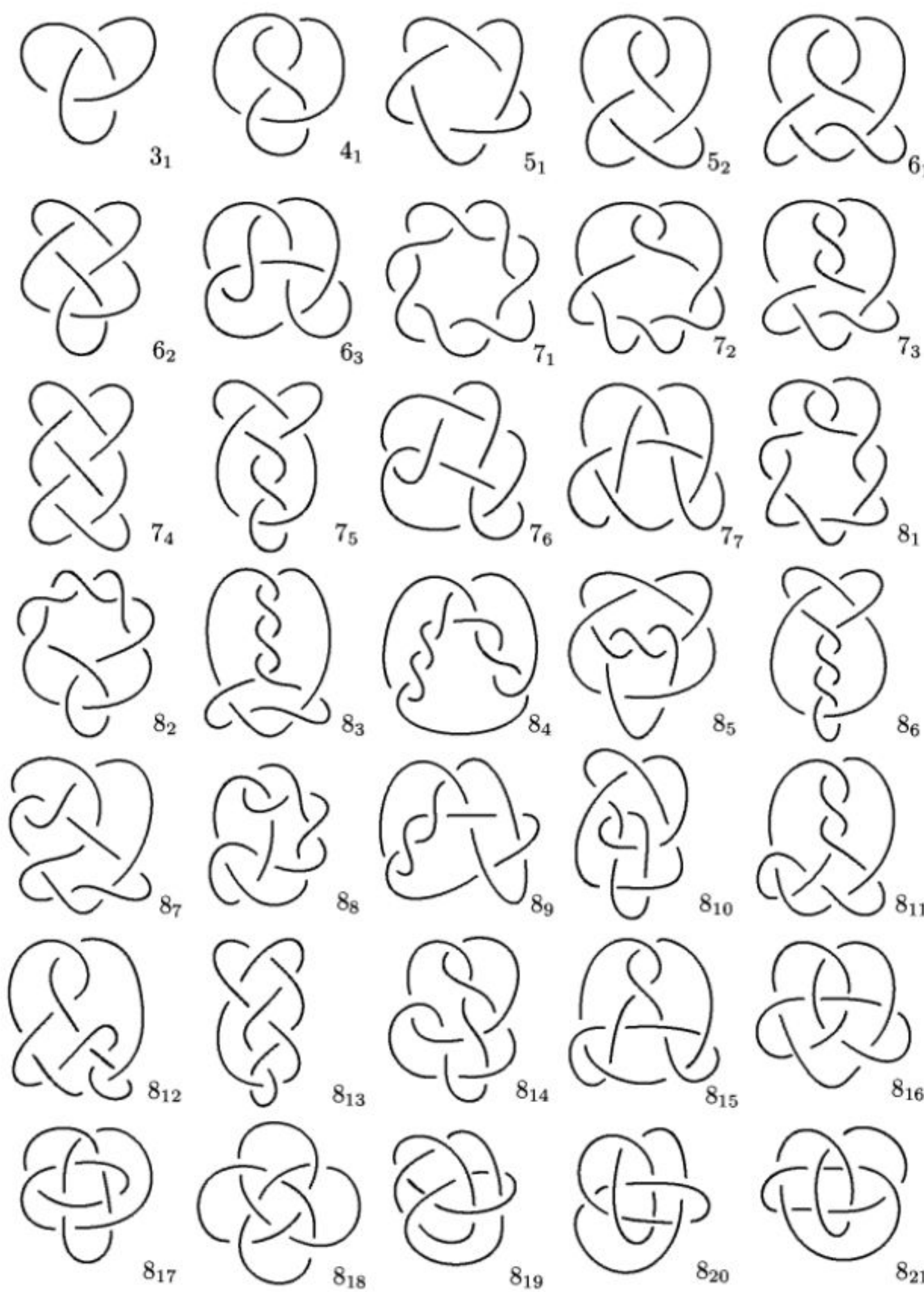


A connected sum $K_0 \# K_1$ can also be understood as the result of first tying K_0 in a piece of string, then tying K_1 before joining the ends.

A knot K is *prime* if the only way to decompose K as a connected sum of two knots is as $K \# 0_1$ where 0_1 is the unknotted circle or *unknot*; that is, K is prime if K does not break down as a connected sum of two nontrivial knots. For any knot, to decide whether the knot is prime, we can look at all the ways of intersecting the knot with a circle which meets the knot at exactly two points; this divides the knot into a connected sum of the portion outside the circle (completed by the arc along the circle) and the portion inside the circle (completed with the arc along the circle). If the knot is prime, then every such division will have one side unknotted.



Below is a table of all prime knots with up to eight crossings. Knots are named according to their crossing number with a subscript indicating their position on the table.



Exercises. 1. Take a piece of rope or an extension cord, tie a knot very loosely, and join the ends together. Lay the knot on a flat surface and draw a knot diagram representing the knot. Now, move the knot around to a new position – don't be afraid to add a few twists, just keep the ends joined. Now, draw a diagram of your knot in its new configuration. Repeat a few times; then repeat with a new knot.

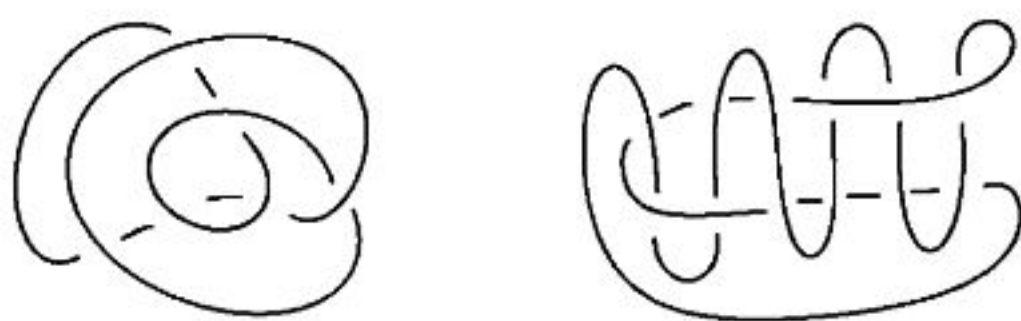
2. Draw all possible knot or link diagrams with exactly two crossings.

3. Draw all possible knot or link diagrams with exactly three crossings.

4. Is it possible, given what you currently know, that the two diagrams below represent the same knot or link?



5. Is it possible, given what you currently know, that the two diagrams below represent the same knot or link?



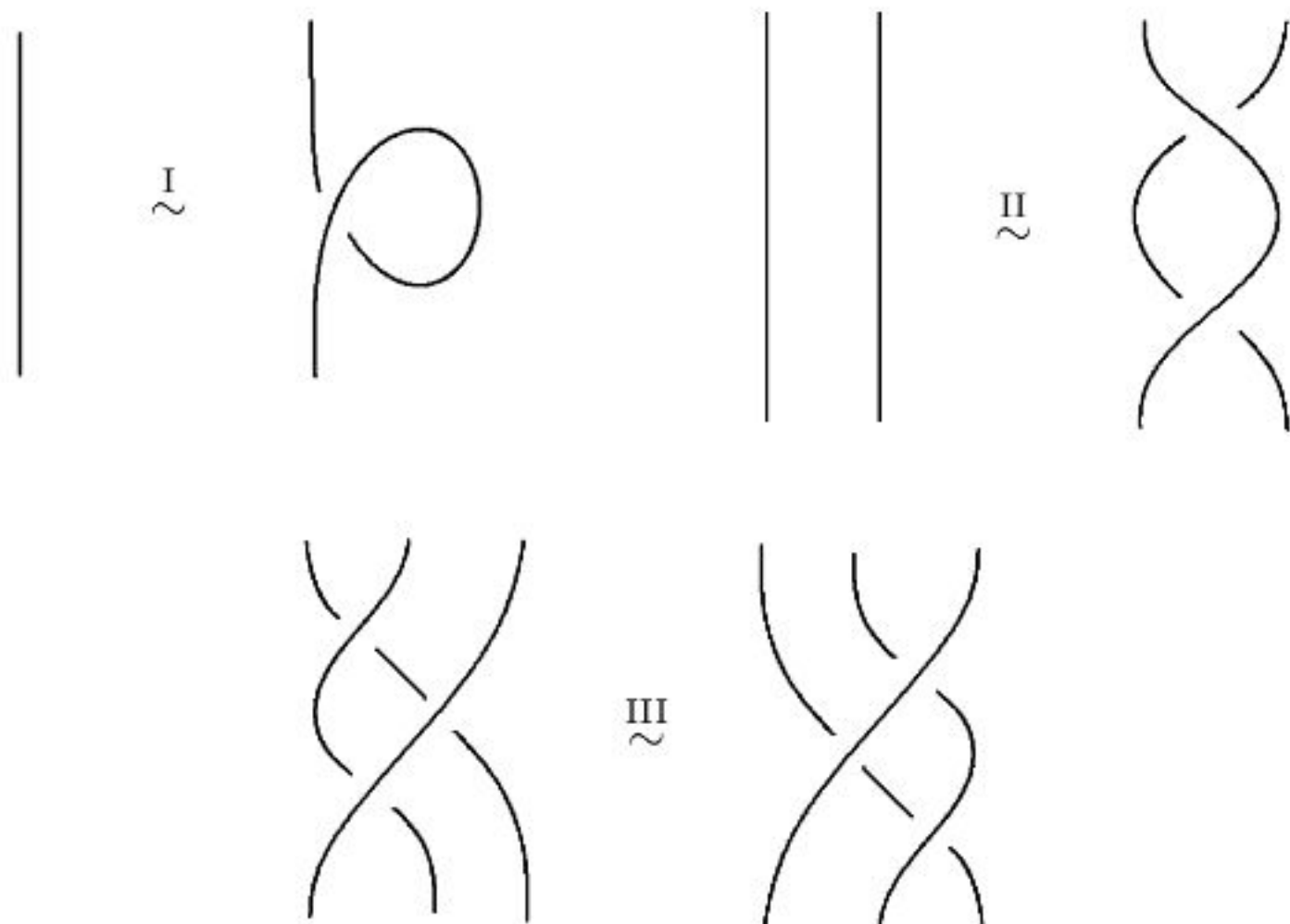
2. Combinatorial Knot Theory

The basic question in knot theory is how to tell when two knot diagrams represent the same knot. This is really two questions: (1) given a knot diagram, what are all possible diagrams which represent the same knot type and (2) how can we prove two diagrams represent different knots. The second question we leave to the next section.

The first question was answered in the 1920s by Kurt Reidemeister. A *local move* on a knot diagram involves replacing one portion of the diagram inside a small disk with something else, while the rest of the diagram outside the disk remains unchanged. A *planar isotopy* is a local move which replaces a strand without crossings with another strand without crossings with the same endpoints.

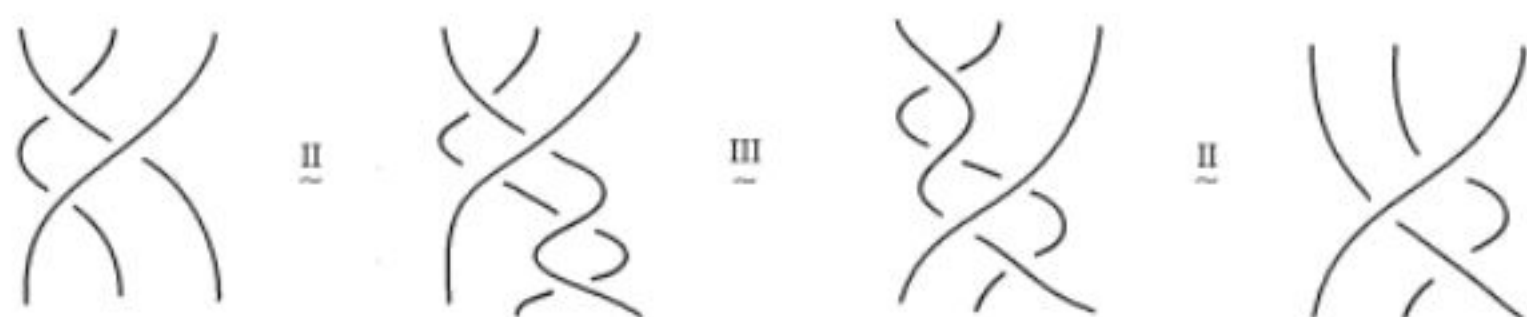


In 1926, Kurt Reidemeister and independently, in 1927, J. W. Alexander and G. B. Briggs proved that two tame knot diagrams, K_0 and K_1 , are ambient isotopic if and only if one can be changed into the other by a finite sequence of planar isotopies and moves of the following three types:



If you look closely, you will find that some other similar moves are implied by the listed moves. For example, move III says you

can move a strand over a crossing where the crossing is the same type (left-over-right) as the other two crossings. We can derive an additional type III move which says you can pass a strand over the other kind of crossing using the given III move and a II move:



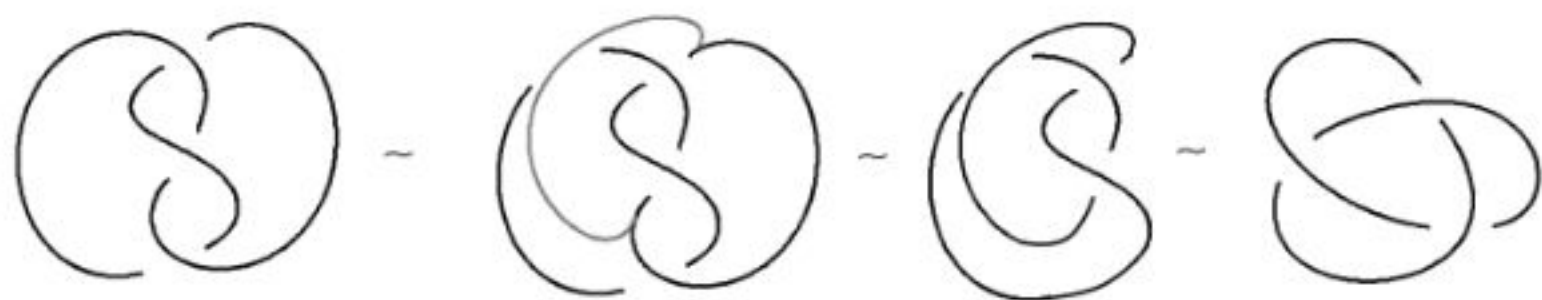
The Reidemeister moves let us translate the topological relationship of ambient isotopy into a combinatorial equivalence relation. That is, we started thinking of knots as geometric objects, simple closed curves in space, but we now have a new way to think of knots: as equivalence classes of knot diagrams under the equivalence relation generated by planar isotopy moves and the Reidemeister moves.

Thus, to prove that two knot diagrams, K_0 and K_1 , represent the same knot type, we can identify an explicit sequence of Reidemeister moves taking K_0 to K_1 . For example, the knot below is secretly the unknot, i.e., an unknotted circle. To prove it, we give a sequence of Reidemeister moves taking it to a circle without crossings.

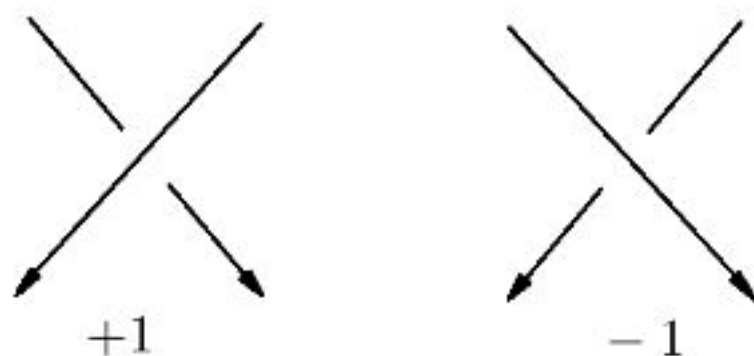


In practice it is often easier to redraw knots using the principle that any portion of a strand with only overcrossings may be replaced with another strand with the same endpoints and all new overcrossings, with the resulting breaks healing. Note that any such “overpass move” can always be broken down into a sequence of Reidemeister

moves and planar isotopies.



Combinatorial Oriented Knots. Introducing an orientation on our knot diagrams gives us two kinds of crossings, which we identify as “positive” and “negative” depending on whether the understrand is directed right-to-left or left-to-right when viewed from the overstrand.



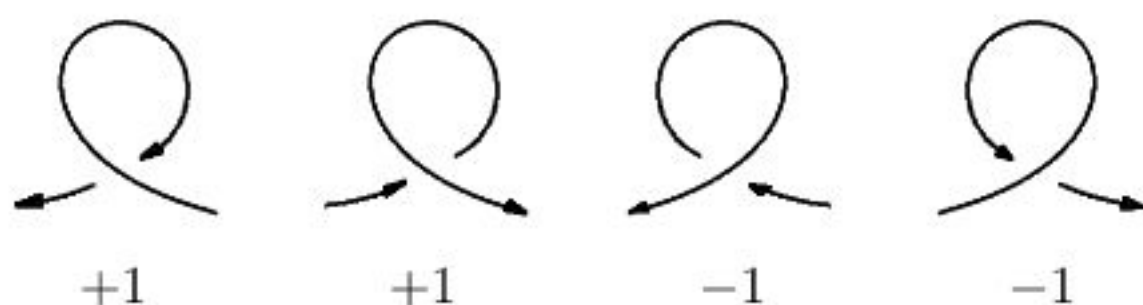
We will denote the sign of a crossing C as $\epsilon(C) = \pm 1$.

Including orientations means we now have more Reidemeister moves than we did before. Instead of two type I moves, we now have four; one type II becomes four – two *direct* moves where the strands are oriented in the same direction, and *reverse* moves where the strands are oriented in opposite directions, and there are eight oriented type III moves.

In practice, many of the moves are implied by the other moves. Indeed, it is an interesting exercise to find a minimal generating set of moves, i.e., a subset containing as few moves as possible from which all of the other oriented moves can be recovered.

The sum of all the crossing signs is a quantity known as the *writhe* of the diagram; writhe is a property of diagrams, not of knots, since starting with a given knot diagram we can adjust the writhe to whatever we want using type I moves. For links, each component has its own writhe determined by counting only crossings where the component crosses itself; multi-component crossings do not contribute to the component writhes.

Note that for any single-component crossing both possible choices of orientation determine the same sign for each crossing since switching the orientation of a component reverses the directions of both strands in the crossing. In particular, writhe is well defined even for unoriented diagrams, and kinks have well-defined signs regardless of orientation choice.



Combinatorial Framed Knots. Given a knot diagram K , there is an easy standard way to choose a framing curve – simply “push off” a copy of K , i.e., draw a framing curve parallel to K . This is traditionally called the *blackboard framing* since it is the easiest framing to draw on a blackboard. More precisely, let F be the knot traced by a normal vector to the knot K . If F is endowed with an orientation parallel to that of K , then the *framing number* of K is the sum of the crossing signs at crossings where F crosses over K . Conversely, if we assign an integer j to a knot K , then we can construct a normal vector to K and a knot F such that j is the framing number. The blackboard framing is then the natural framing with framing number equal to the writhe of K . In particular, every knot or link diagram can be considered as a framed knot or link by using the blackboard framing.

Geometrically, the *framing number* of a framing curve F is the number of times F wraps around the solid torus with K as its core. The framing number of a blackboard-framed knot is equal to its writhe. A little thought reveals that Reidemeister II and III moves do not change the writhe of a diagram, while Reidemeister I moves do. Thus, to preserve the blackboard framing, we must modify the type I move to preserve writhe. In particular, to cancel the $+1$ to writhe from adding a positive kink, we must also add a negative kink. Kinks of both signs come in two versions, clockwise and counterclockwise, also known as kinks of *winding number* -1 and $+1$, respectively.

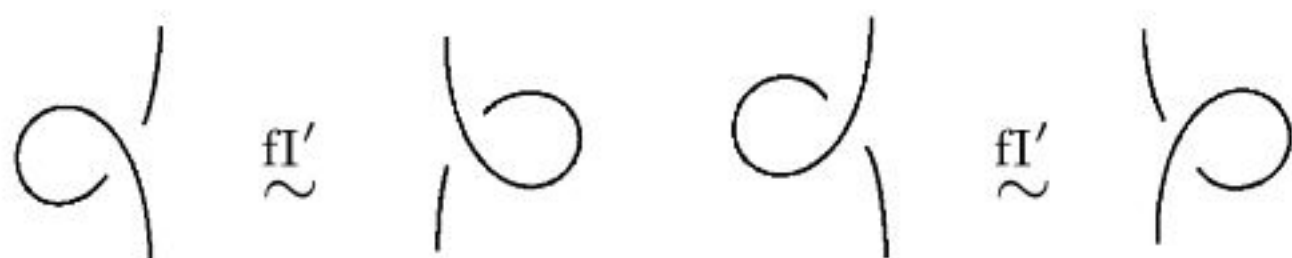
As observed in [FR92], it turns out that if both the winding numbers and crossing signs of the kinks are opposite, we can cancel the kinks using only II and III moves (the following illustration is the simplest “Whitney trick” [Whi44]):



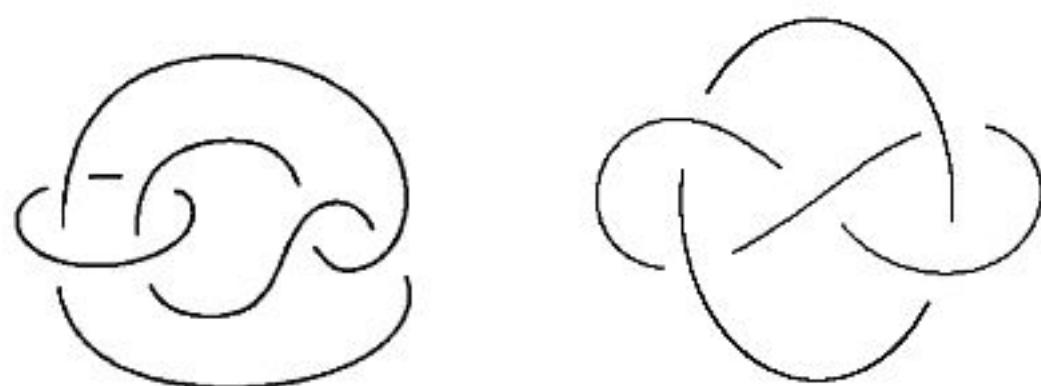
In the case of kinks with equal winding number and opposite writhe, we need an explicit move. These are the *blackboard framed type I moves*:



These moves are equivalent to the alternate framed type I moves:



Exercises. 1. Using Reidemeister moves, show that the diagrams below represent the same link. This link is known as the *Whitehead link*.



2. Using Reidemeister moves, determine whether the knot K below is equivalent to the *trefoil* or the *Figure 8*.



Trefoil

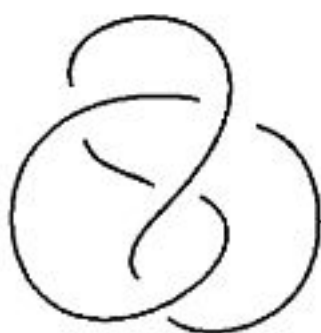
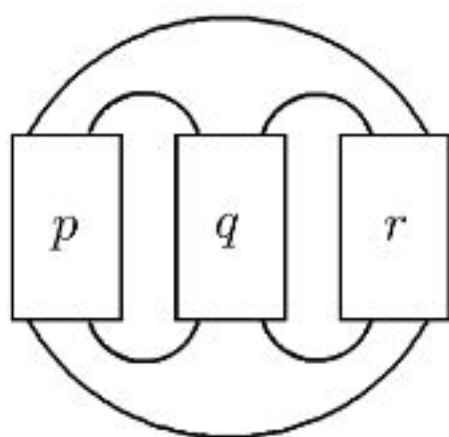
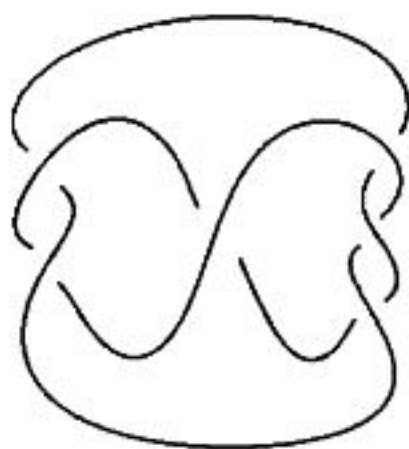
 K 

Figure 8

3. Let p, q, r be three integers. A (p, q, r) -pretzel link is a knot or link of the form



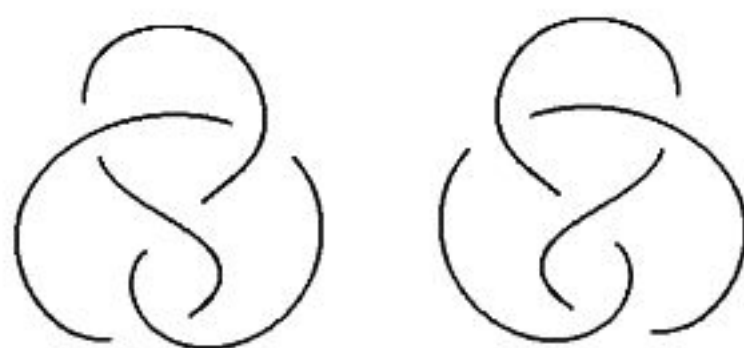
where the boxes are replaced with stacks of p, q and r oriented crossings respectively (a negative value means use negative crossings – also note that the orientation of the crossings in the boxes may not extend to the whole link!) For example, the $(2, 1, -3)$ pretzel link is



How many components are possible in a pretzel link? What conditions on p, q and r ensure that we have a knot? A 2-component link?

4. Show that the figure eight knot 4_1 is ambient isotopic to its mirror image by changing the diagram on the left to the one on the right

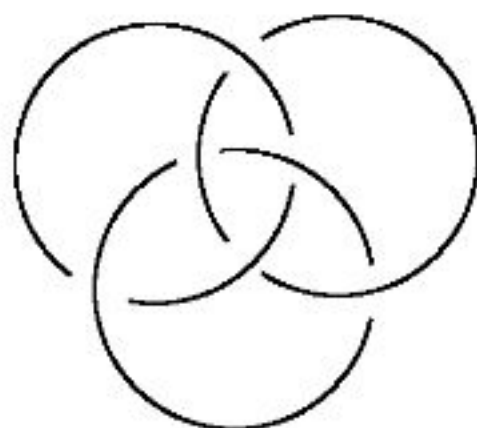
using Reidemeister moves.



5. Using framed Reidemeister moves, show that the knot below is framed isotopic to the unknot with writhe -2 .



6. A link is called *Brunnian* if it is nontrivial, but deleting any component makes the remaining link trivial. Given that the *Borromean rings*



form a nontrivial link, show that the link is Brunnian.

7. Show that the fl and fl' moves are equivalent in the presence of the type II and III moves by deriving the fl' move using only type fl , II, and III moves and then deriving the fl move using only type fl' , II, and III moves.

3. Knot and Link Invariants

Changing K into K' with Reidemeister moves proves that the two diagrams represent the same knot or link. What if we cannot see

a way to change K into K' ? Our inability to change K into K' with Reidemeister moves does not say that K and K' are different; it might be that there is a way, but it's very complicated, perhaps involving hundreds of moves and requiring introducing and removing many crossings. In order to prove that two diagrams represent different knots, we must be more clever.

A *knot invariant* is a function $f : \mathcal{K} \rightarrow X$ from the set of all knot diagrams to a set X such that for each Reidemeister move, we have

$$f(K_1) = f(K_2)$$

where K_1 is the knot diagram before the move and K_2 is the same diagram after the move. If f is a knot invariant, then any two diagrams related by Reidemeister moves must give the same value when we evaluate f .

Knot theory might be described as the search for and the study of knot invariants. Many knot and link invariants have been discovered and studied, mostly in the 20th and 21st centuries. For the remainder of this section we will explore a few well-known knot and link invariants.

Geometric Invariants. One way to define a knot or link invariant is to identify some geometric or topological quantity determined by a knot or link diagram and take the minimum over all diagrams of K . Many examples of this style of invariant have been defined and studied, from basic to more esoteric:

- *Crossing Number* – The minimal number of crossings in any diagram of K .
- *Braid Index* – The smallest number of strands of any braid whose closure is a diagram of K .
- *Bridge Number* – The smallest number of maxima in any diagram of K .
- *Stick Number* – The smallest number of straight line segments needed to form K in \mathbb{R}^3 .
- *Rope Length* – The minimal length of a rope of radius 1 needed to tie K .

- *Genus* – The minimal number of holes in a surface whose edge is K .
- *Unknotting Number* – The minimal number of crossing changes needed to unknot K .

These invariants are easy to define but generally hard to compute. From a particular diagram of K we can compute an upper bound on the actual value of the invariant for K , but finding a diagram of K which realizes the minimal value is not always easy to do.

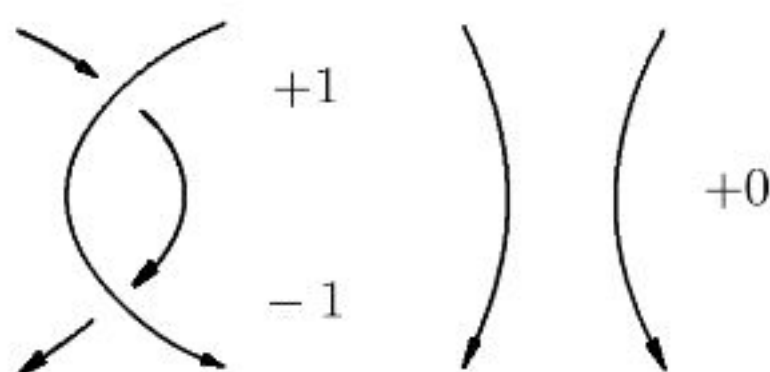
A knot or link invariant f is *computable* if the actual value of $f(K)$ can be determined, not just bounded, using any diagram of K . We will now see three examples of computable knot and link invariants.

Linking Number. Perhaps the easiest example of a computable link invariant is the *linking number*. Let $L = L_1 \cup L_2$ be an oriented link with two components. Let \mathcal{M} be the set of crossings in L with one strand from each component. Then the linking number of L is the sum of the crossing signs of the crossings in \mathcal{M} divided by 2 since this sum is always even:

$$\text{lk}(L) = \frac{1}{2} \sum_{C \in \mathcal{M}} \epsilon(C).$$

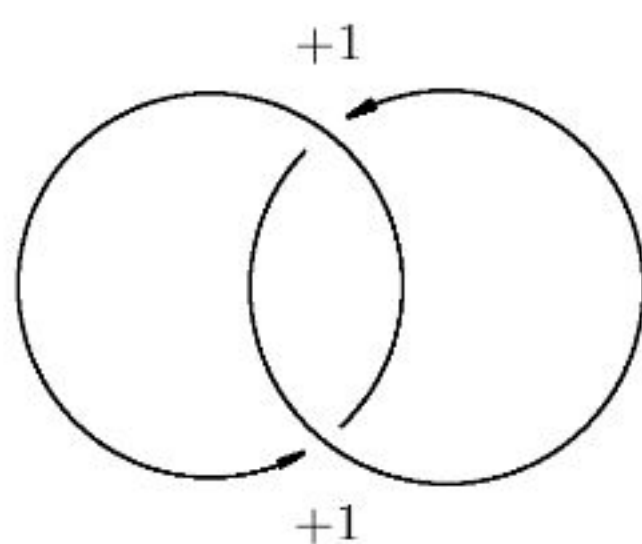
We can verify that the linking number is a link invariant by checking that the contributions match before and after each move. In a type I move, the crossing being introduced or removed is single-component, so it contributes 0 to the linking number, which matches the contribution from the straight strand.

In a type II move, there are two possibilities: either both crossings are multicomponent or both are not. As before, if both crossings are single-component, the contribution of zero matches the zero contribution of the two uncrossed strands. In the multicomponent case, there is always one positive crossing and one negative crossing, so the contribution is $+1 - 1 = 0$.

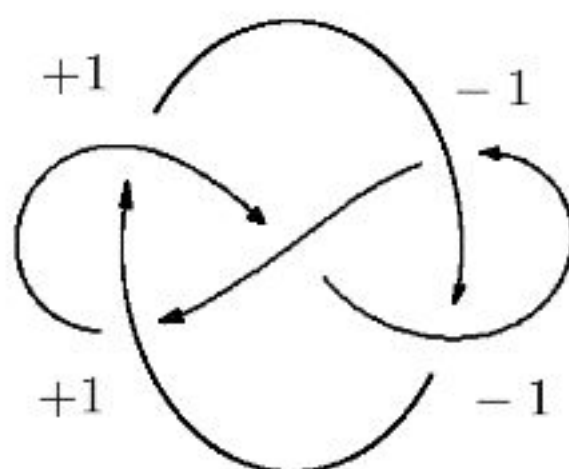


Verifying that $f(L_B) = f(L_A)$ for type III moves is left to the reader as an exercise; see problem 1.

The fact that the linking number is a link invariant lets us distinguish some links from others. For example, the Hopf link below has linking number 1 while the Whitehead link has linking number 0.



$lk = 1$



$lk = 0$

The Jones Polynomial. In 1984 knot theory was reinvigorated by the discovery by Vaughan Jones [Jon85] of a powerful knot and link invariant now known as the *Jones polynomial*. The simplest way to define the Jones polynomial is a recursive definition due to Louis Kauffman using the *Kauffman bracket skein relation*. There are several versions of this invariant related by variable substitution; the version we'll use is from [BN02].

Let K be a knot or link diagram. The skein relation can be understood as a way of interpreting a crossing as a linear combination of *smoothings*:

$$\left\langle \begin{array}{c} \diagup \quad \diagdown \\ \diagdown \quad \diagup \end{array} \right\rangle = \left\langle \begin{array}{c} \frown \\ \smile \end{array} \right\rangle - q \left\langle \begin{array}{c} \frown \\ \frown \end{array} \right\rangle - q^{-1} \left\langle \begin{array}{c} \smile \\ \smile \end{array} \right\rangle.$$

Recursively applying this relation to each crossing lets us replace a knot or link diagram with n crossings with a sum of polynomials in q times diagrams without crossings. We need a rule for evaluating the bracket of a diagram without crossings. Thus, for a disjoint union of n copies of the diagram of the unknot with no crossings, we define

$$\langle \bigcirc \bigcirc \dots \bigcirc \rangle = (q + q^{-1})^{n-1}.$$

In particular, we can erase a closed curve without crossings at the cost of multiplying by $(q + q^{-1})$.

The bracket function is unchanged by Reidemeister III moves: The reader is encouraged to verify that both $\langle \text{diagram 1} \rangle$ and $\langle \text{diagram 2} \rangle$ are equal to

$$-q \left(\langle \text{diagram 3} \rangle + \langle \text{diagram 4} \rangle \right) + q^2 \left(\langle \text{diagram 5} \rangle + \langle \text{diagram 6} \rangle \right) - q^3 \langle \text{diagram 7} \rangle.$$

However, Reidemeister I and II moves do change the value of $\langle K \rangle$, but in a predictable way. More precisely, removing a positive crossing multiplies $\langle K \rangle$ by q^{-1} and removing a negative crossing multiplies $\langle K \rangle$ by $-q^2$:

$$\begin{aligned} \langle \text{diagram 8} \rangle &= \langle \text{diagram 9} \rangle - q \langle \text{diagram 10} \rangle \\ &= (q + q^{-1} - q) \langle \text{diagram 11} \rangle = q^{-1} \langle \text{diagram 12} \rangle \end{aligned}$$

and

$$\begin{aligned} \langle \text{diagram 13} \rangle &= \langle \text{diagram 14} \rangle - q \langle \text{diagram 15} \rangle \\ &= (1 - q(q + q^{-1})) \langle \text{diagram 16} \rangle = -q^2 \langle \text{diagram 17} \rangle. \end{aligned}$$

Likewise, a crossing-removing type II move multiplies $\langle K \rangle$ by a factor of $(-q^2)q^{-1} = -q$:

$$\begin{aligned}
 \left\langle \begin{array}{c} \diagup \diagdown \\ \diagdown \diagup \end{array} \right\rangle &= \left\langle \begin{array}{c} \frown \\ \smile \end{array} \right\rangle - q \left\langle \begin{array}{c} \bigcirc \\ \bigcirc \end{array} \right\rangle - q \left\langle \begin{array}{c} \diagup \\ \diagdown \end{array} \right\rangle \left\langle \begin{array}{c} \diagdown \\ \diagup \end{array} \right\rangle + q^2 \left\langle \begin{array}{c} \cup \\ \cup \end{array} \right\rangle \\
 &= (1 - q(q + q^{-1}) + q^2) \left\langle \begin{array}{c} \frown \\ \smile \end{array} \right\rangle - q \left\langle \begin{array}{c} \diagup \\ \diagdown \end{array} \right\rangle \left\langle \begin{array}{c} \diagdown \\ \diagup \end{array} \right\rangle \\
 &= -q \left\langle \begin{array}{c} \diagup \\ \diagdown \end{array} \right\rangle \left\langle \begin{array}{c} \diagdown \\ \diagup \end{array} \right\rangle.
 \end{aligned}$$

Thus, to cancel the effects of type I and II moves, we need to multiply by $(-1)^n q^{p-2n}$ where p is the number of positive crossings and n is the number of negative crossings. The *Jones polynomial* of a link L is

$$J(L) = (-1)^n q^{p-2n} \langle L \rangle.$$

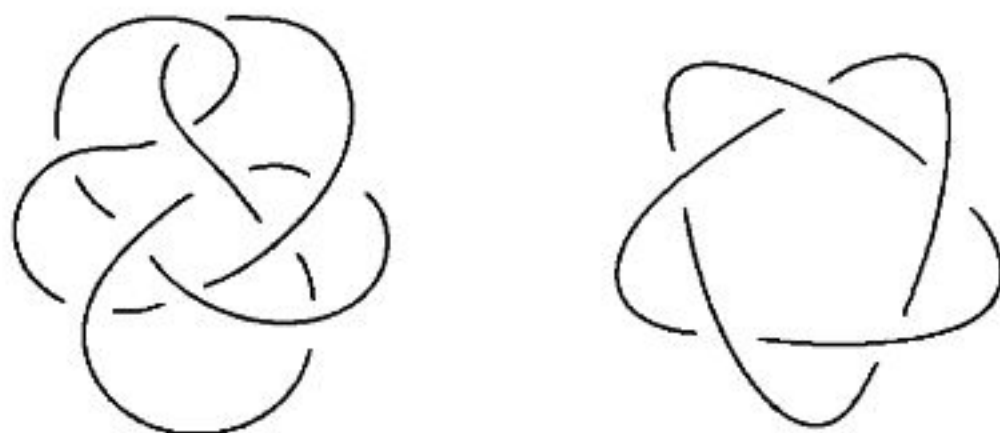
Example 1. Let us compute the Jones polynomial of the Hopf link:

$$\begin{aligned}
 \left\langle \begin{array}{c} \bigcirc \\ \bigcirc \end{array} \right\rangle &= \left\langle \begin{array}{c} \bigcirc \\ \bigcirc \end{array} \right\rangle - q \left\langle \begin{array}{c} \cup \\ \cup \end{array} \right\rangle \\
 &= \left\langle \begin{array}{c} \bigcirc \\ \bigcirc \end{array} \right\rangle - q \left\langle \begin{array}{c} \cup \\ \cup \end{array} \right\rangle \\
 &= -q \left\langle \begin{array}{c} \cup \\ \cup \end{array} \right\rangle + q^2 \left\langle \begin{array}{c} \bigcirc \\ \bigcirc \end{array} \right\rangle \\
 &= q + q^{-1} - q - q + q^2(q + q^{-1}) = q^{-1} + q^3.
 \end{aligned}$$

If we orient the components so that both crossings are positive, we then have $q^2(q^{-1} + q^3) = q + q^5$; if we reverse the orientation of one component while fixing the other, we have two negative crossings and the Jones polynomial becomes $q^{-4}(q^{-1} + q^3) = q^{-1} + q^{-5}$. Thus, the two possible oriented Hopf links have different Jones polynomials and cannot be ambient isotopic to each other.

It turns out that the Jones polynomial of the mirror image of a knot or link K can be obtained from the Jones polynomial of K by

replacing q with q^{-1} . The Jones polynomial is a very powerful invariant, but it is not a complete invariant – there are known examples of pairs of different knots which have the same Jones polynomial, such as the knots below:

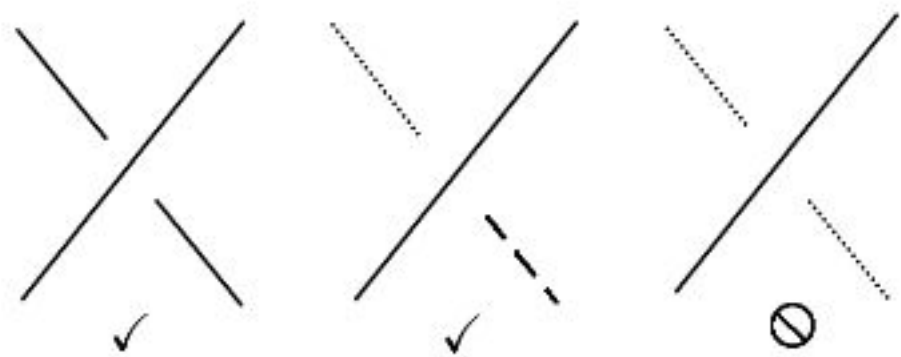


Indeed, one of the more famous unsolved (as of this writing) problems in knot theory is whether the Jones polynomial detects the unknot; that is, is there a nontrivial knot K with Jones polynomial $J(K) = 1$? For links with multiple components, the answer is yes, there are nontrivial links with trivial Jones polynomial. For knots, though, the problem is currently unsolved. Direct computations have shown that no nontrivial knot with fewer than 16 crossings has trivial Jones polynomial.

The Jones polynomial is a powerful knot invariant, but computationally it is very intense. The recursive algorithm described above is an *exponential time* algorithm, meaning each additional crossing doubles the number of computations needed to compute $J(K)$.

Tricoloring. For our final example of a computable knot invariant, we will define *Fox tricoloring*, introduced by Ralph Fox in the 1950s. A *tricoloring* of a knot or link diagram is a choice of color for each arc in the diagram from a set of three colors – we’ll use solid, dotted and dashed, but you can use whatever colors you like. A tricoloring is *valid* if at every crossing we either have all three colors the same or all three colors different. A valid tricoloring is *nontrivial* if it uses

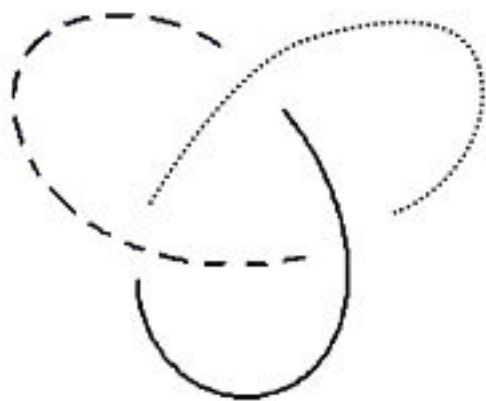
all three colors.



To use tricoloring as a knot invariant, we notice that if we start with a valid tricoloring of a diagram K before doing a Reidemeister move, there is a unique valid tricoloring of the diagram after the move which agrees with the original coloring outside the move area. For example, all strands in a type I move must be the same color. There are two cases for type II moves: both strands the same color before crossing and two different colors before crossing:

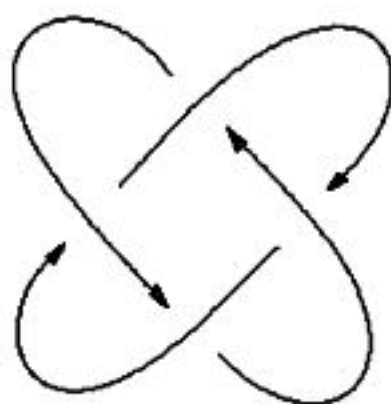


For the type III moves, there are various cases which the reader is encouraged to check. Moreover, if a tricolored diagram is monochrome before a move, the corresponding diagram after the move is also monochrome. Hence, the existence of a nontrivial tricoloring of a knot or link diagram is an invariant of knots and links. For example, the only valid tricolorings of an unknotted diagram are monochrome colorings, while the trefoil has a nontrivial tricoloring; thus there can be no sequence of Reidemeister moves taking the trefoil to the unknot.



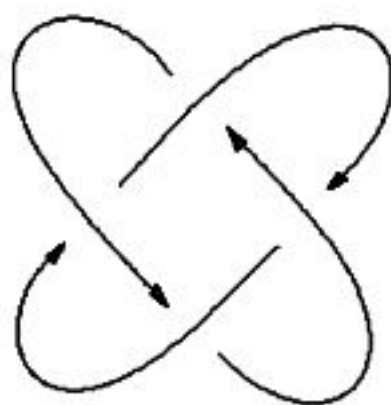
Exercises. 1. Verify that Reidemeister III moves do not change linking number. (Hint: Choose one oriented type III move and consider all cases depending on which strands are from the same component).

2. A link is *split* if it is possible to separate the components so that one component lies entirely on one side of a line and the other component lies entirely on the other side of the line. Prove that the $(4, 2)$ -torus link below is not split.

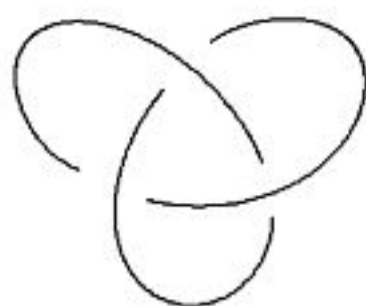
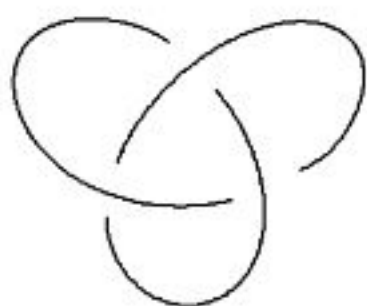


3. Prove that the framing number of a blackboard framed oriented knot is the linking number of the framed knot considered as a 2-component link.

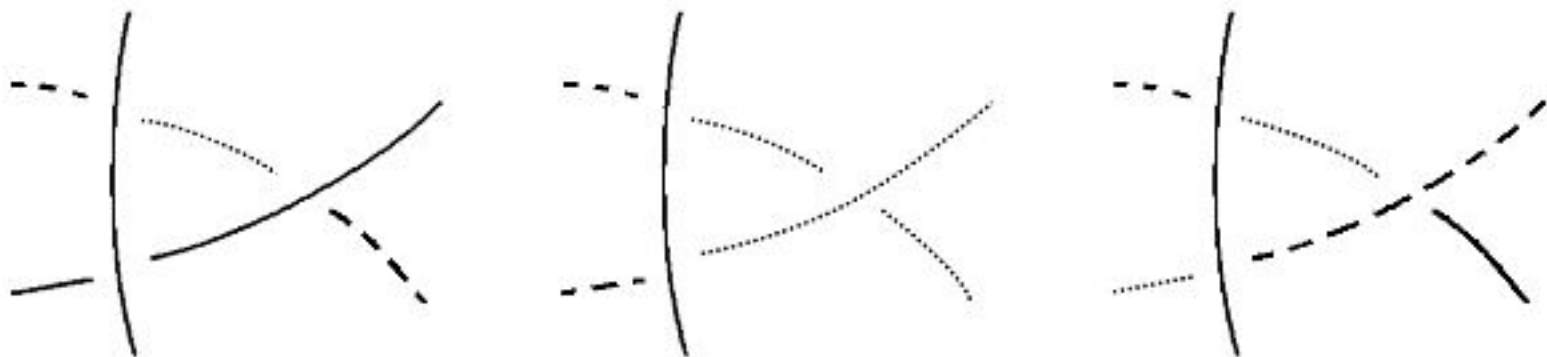
4. Compute the Jones polynomial of the $(4, 2)$ -torus link below.



5. Use the Jones polynomial to prove that the right-handed and left-handed trefoils below are not equivalent.



6. For each of the tricolored tangle diagrams below, find the unique corresponding tricolored tangle diagram after doing a type III move:

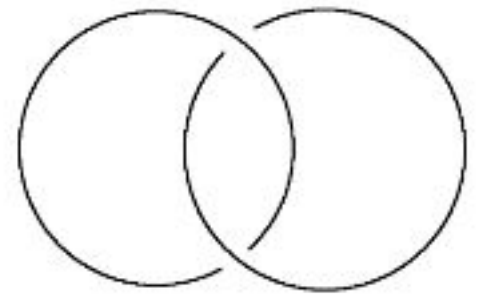


7. Show that there is no nontrivial tricoloring of the figure 8 knot below.



8. How many valid tricolorings of the trefoil knot are possible? How many are nontrivial?

Chapter 2



Algebraic Structures

1. Operation Tables and Isomorphisms

In order to get good at telling knots apart, we need to develop some of the ideas of Algebra. Not high school algebra, like factoring and quadratic formulas; rather, we're talking about *abstract algebra*, also sometimes called *modern algebra*. Abstract algebra is all about *algebraic structures*, i.e., sets with one or more *operations* and the properties they satisfy.

Definition 1. Let X be a set. An *operation* on X is a rule for combining two elements of X to get another element of X . That is, an operation on X is a function from the set of ordered pairs of elements of X to X .

In particular, if $*$ is an operation on X , then X must be *closed under $*$* , meaning if $x, y \in X$ ¹ then $x * y \in X$. For example, addition is an operation on the set of natural numbers $\mathbb{N} = \{0, 1, 2, \dots\}$, but subtraction is not since $2 - 3 \notin \mathbb{N}$.

Example 2. Many examples of operations on sets of numbers are familiar:

- The integers $\mathbb{Z} = \{\dots, -2, -1, 0, 1, 2, \dots\}$,

¹The symbol " \in " is mathematical shorthand for "in", so $x, y \in X$ means x and y are elements of the set X .

- The rational numbers $\mathbb{Q} = \{\frac{a}{b} \mid a, b \in \mathbb{Z}, b \neq 0\}$,
- The real numbers $\mathbb{R} = \{n.d_1d_2\dots, n \in \mathbb{Z}, 0 \leq d_k \leq 9\}$ and
- The complex numbers $\mathbb{C} = \{a + ib \mid a, b \in \mathbb{R}, i^2 = -1\}$

have operations of addition, subtraction, multiplication and division (by everything except 0).

Example 3. Recall from linear algebra that a vector space V over a field \mathbb{F} has an operation of *vector addition* in which $\vec{u}, \vec{v} \in V$ combine to form $\vec{u} + \vec{v} \in V$. Similarly, \mathbb{R}^3 has an operation called *cross product* in which $\vec{u}, \vec{v} \in \mathbb{R}^3$ combine to form $\vec{u} \times \vec{v} \in \mathbb{R}^3$.

These operations satisfy a variety of familiar properties (and some perhaps not so familiar), such as:

- *Commutativity*: $a + b = b + a$, $ab = ba$,
- *Associativity*: $(a + b) + c = a + (b + c)$, $a(bc) = (ab)c$,
- *Distributivity*: $(a + b)c = ac + bc$, $a(b + c) = ab + ac$,
- *Anti-commutativity*: $\vec{x} \times \vec{y} = -\vec{y} \times \vec{x}$,
- *Jacobi Identity*: $\vec{x} \times (\vec{y} \times \vec{z}) + \vec{y} \times (\vec{z} \times \vec{x}) + \vec{z} \times (\vec{x} \times \vec{y}) = \vec{0}$,

and more.

Let us consider a very easy toy example of an algebraic structure: let X be a set and define an operation $*$ on X which simply ignores the second variable and returns the first variable, i.e., $x * y = x$. If X is a finite set, we can express any operation with an *operation table*, that is, a square whose entry in row j and column k is $j * k$. Then the operation above for a set of four elements $X = \{1, 2, 3, 4\}$ is

$*$	1	2	3	4
1	1	1	1	1
2	2	2	2	2
3	3	3	3	3
4	4	4	4	4

This operation satisfies some of the above listed properties: for example, it is associative since

$$(x * y) * z = x * y = x = x * y = x * (y * z).$$

However, it is not commutative:

$$x * y = x \neq y = y * x.$$

Now, suppose that instead of $X = \{1, 2, 3, 4\}$, we have $Y = \{a, b, c, d\}$ and operation table

$*$	a	b	c	d
a	a	a	a	a
b	b	b	b	b
c	c	c	c	c
d	d	d	d	d

At first glance, it might seem like we have a new algebraic structure, but the new one is really just the old one with the names of the elements in X changed. More precisely, we have a one-to-one correspondence (also known as a *bijection*, i.e., a function which is both one-to-one and onto) $f : X \rightarrow Y$ given by $f(1) = a$, $f(2) = b$, $f(3) = c$ and $f(4) = d$, and this correspondence preserves the algebraic structure in the sense that $f(x * x') = f(x) * f(x')$.

If X and Y are algebraic structures and $f : X \rightarrow Y$ is a function which preserves all of the operations, i.e., $f(x *_X x') = f(x) *_Y f(x')$ for each of the operations $*_X$ of X and corresponding operations $*_Y$ of Y , then f is called a *homomorphism*. A bijective homomorphism $f : X \rightarrow Y$, that is, a homomorphism f with an inverse homomorphism $f^{-1} : Y \rightarrow X$, is called an *isomorphism*. An isomorphism $f : X \rightarrow X$ from X to itself is called an *automorphism*. If there is an isomorphism from X to Y , we say that X and Y are *isomorphic*.

Example 4. The exponential function e^x is a homomorphism (indeed, an isomorphism) from the set of all real numbers with the operation of addition to the set of positive real numbers with the operation of multiplication, since

$$e^{x+y} = e^x e^y.$$

The inverse isomorphism is the natural logarithm function $\ln : (\mathbb{R}_+, \cdot) \rightarrow (\mathbb{R}, +)$, which satisfies

$$\ln(xy) = \ln x + \ln y.$$

Example 5. Consider the sets $X = \{1, 2, 3, 4\}$ and $Y = \{a, b\}$ with operations given by the operation tables

$*_X$	1	2	3	4		$*_Y$	a	b
1	1	2	3	4		a	a	b
2	2	1	4	3	and	b	b	a
3	3	4	1	2				
4	4	3	2	1				

Then we can check that the function $f : X \rightarrow Y$ defined by $f(1) = a$, $f(2) = a$, $f(3) = b$, $f(4) = b$ defines a homomorphism by checking that $f(x *_X x') = f(x) *_Y f(x')$ for each pair $(x, x') \in X \times X$. Alternatively, we could observe that replacing each x with $f(x)$ in the operation table of $*_X$ results in a table which collapses onto that of $*_Y$:

$*_X$	1	2	3	4			a	a	b	b		$*_Y$	a	b
1	1	2	3	4		a	a	a	b	b		a	a	b
2	2	1	4	3	→	a	a	a	b	b	→	a	a	b
3	3	4	1	2		b	b	b	a	a		b	b	a
4	4	3	2	1		b	b	b	a	a				

Example 6. The algebraic structure of a vector space V is determined by the operations of scalar multiplication and vector addition. Thus, a homomorphism of vector spaces is a function $f : V \rightarrow W$ which preserves the operations, i.e., a function f such that

$$f(\alpha \vec{u}) = \alpha f(\vec{u}) \quad \text{and} \quad f(\vec{u} + \vec{v}) = f(\vec{u}) + f(\vec{v}).$$

That is, a linear transformation is a vector space homomorphism.

A bijection $\sigma : X \rightarrow X$ is called a *permutation on X* . We can represent a permutation σ with a two-row matrix where the top row is the original ordering of the subscripts $1, 2, \dots, n$ and below each number k is $\sigma(k)$; then the permutation σ on the set $\{1, 2, 3, 4, 5\}$ defined by $\sigma(1) = 3, \sigma(2) = 4, \sigma(3) = 1, \sigma(4) = 5$ and $\sigma(5) = 2$ is expressed by the matrix

$$\begin{bmatrix} 1 & 2 & 3 & 4 & 5 \\ 3 & 4 & 1 & 5 & 2 \end{bmatrix}.$$

Alternatively, if $X = \{1, 2, 3, \dots, n\}$, we can just list the bottom row vector since the top will be understood to be the standard ordering. Then for instance, we can conveniently express σ above as $[3, 4, 1, 5, 2]$.

Now consider our algebraic structure X in Example 5. Suppose we have another operation table with the same set of elements, like

$*$	2	4	3	1
2	1	3	4	2
4	3	1	2	4
3	4	2	1	3
1	2	4	3	1

This is secretly really the same operation table, we just have the rows and columns listed in a different order. We can think of the operation table as a square matrix (not including the row and column labels); then given a permutation $\sigma : X \rightarrow X$, we can obtain a *permutation matrix* P_σ from the identity matrix

$$I = \begin{bmatrix} \vec{e}_1 \\ \vec{e}_2 \\ \vdots \\ \vec{e}_n \end{bmatrix} = \begin{bmatrix} 1 & 0 & \dots & 0 \\ 0 & 1 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 1 \end{bmatrix}$$

by reordering the rows of I according to the permutation σ , to get

$$P_\sigma = \begin{bmatrix} \vec{e}_{\sigma(1)} \\ \vec{e}_{\sigma(2)} \\ \vdots \\ \vec{e}_{\sigma(n)} \end{bmatrix}$$

where \vec{e}_k is the k th standard basis vector, i.e., the ordered n -tuple with a 1 in the k th position and 0s elsewhere. Then left multiplication by P_σ reorders the rows of a matrix by σ and right multiplication by $P_\sigma^{-1} = P_\sigma^T$ reorders the columns of our matrix. In the operation table above, we have $\sigma(1) = 2, \sigma(2) = 4, \sigma(3) = 3$ and $\sigma(4) = 1$, so the permutation σ is $[2, 4, 3, 1]$. Then we have

$$P_\sigma = \begin{bmatrix} \vec{e}_2 \\ \vec{e}_4 \\ \vec{e}_3 \\ \vec{e}_1 \end{bmatrix} = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 \end{bmatrix}$$

and

$$P_{\sigma}^{-1} = \begin{bmatrix} 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix}.$$

Then we can unscramble our scrambled operation table matrix M by multiplying on the right by P_{σ} and on the left by P_{σ}^{-1} :

$$\begin{aligned} P_{\sigma}^{-1}MP_{\sigma} &= \begin{bmatrix} 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix} \begin{bmatrix} 1 & 3 & 4 & 2 \\ 3 & 1 & 2 & 4 \\ 4 & 2 & 1 & 3 \\ 2 & 4 & 3 & 1 \end{bmatrix} \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 \end{bmatrix} \\ &= \begin{bmatrix} 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix} \begin{bmatrix} 2 & 1 & 4 & 3 \\ 4 & 3 & 2 & 1 \\ 3 & 4 & 1 & 2 \\ 1 & 2 & 3 & 4 \end{bmatrix} \\ &= \begin{bmatrix} 1 & 2 & 3 & 4 \\ 2 & 1 & 4 & 3 \\ 3 & 4 & 1 & 2 \\ 4 & 3 & 2 & 1 \end{bmatrix}. \end{aligned}$$

Given any finite set $X = \{x_1, \dots, x_n\}$, we can define an operation on X with an $n \times n$ matrix M with entries in $\{1, 2, \dots, n\}$. Then any permutation σ gives us an isomorphic algebraic structure on X with operation table matrix M' given by

$$M' = P_{\sigma}^{-1}\sigma(M)P_{\sigma}$$

where $\sigma(M)$ is the matrix obtained from M by replacing each entry m_{ij} with $\sigma(m_{ij})$.

Example 7. Let $X = \{1, 2, 3, 4\}$ and define an operation $*$ on X by the operation table

$*$	1	2	3	4
1	1	1	2	2
2	2	2	1	1
3	3	3	4	4
4	4	4	3	3

as above. Then the permutation $\sigma = [1, 3, 4, 2]$ gives us an isomorphic algebraic structure on X with operation matrix

$$\begin{aligned}
 P_{\sigma}^{-1}\sigma(M)P_{\sigma} &= \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} 1 & 1 & 3 & 3 \\ 3 & 3 & 1 & 1 \\ 4 & 4 & 2 & 2 \\ 2 & 2 & 4 & 4 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 \end{bmatrix} \\
 &= \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} 1 & 3 & 1 & 3 \\ 3 & 1 & 3 & 1 \\ 4 & 2 & 4 & 2 \\ 2 & 4 & 2 & 4 \end{bmatrix} \\
 &= \begin{bmatrix} 1 & 3 & 1 & 3 \\ 2 & 4 & 2 & 4 \\ 3 & 1 & 3 & 1 \\ 4 & 2 & 4 & 2 \end{bmatrix}.
 \end{aligned}$$

Thus, the algebraic structures on $X = \{1, 2, 3, 4\}$ given by the operation tables

*	1	2	3	4
1	1	1	2	2
2	2	2	1	1
3	3	3	4	4
4	4	4	3	3

and

*	1	2	3	4
1	1	3	1	3
2	2	4	2	4
3	3	1	3	1
4	4	2	4	2

are isomorphic.

An isomorphism σ such that the new operation matrix $M' = P_{\sigma}^{-1}\sigma(M)P_{\sigma}$ is just the original matrix M is an automorphism. For any algebraic structure, the set of automorphisms, denoted $\text{Aut}(X)$, is an invariant: if X is isomorphic to Y , then there is a one-to-one correspondence between elements of $\text{Aut}(X)$ and elements of $\text{Aut}(Y)$. In fact, the set of automorphisms itself has an algebraic structure known as a *group*, which we will see more about shortly.

Example 8. Consider the set $X = \{1, 2, 3, 4\}$ with operation table

$*$	1	2	3	4
1	1	2	3	4
2	2	3	4	1
3	3	4	1	2
4	4	1	2	3

and the permutation $\sigma = [1, 4, 3, 2]$ defined by $\sigma(1) = 1, \sigma(2) = 4, \sigma(3) = 3$ and $\sigma(4) = 2$. Then we have

$$\begin{aligned}
 P_{\sigma}^{-1} \sigma(M) P_{\sigma} &= \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix} \begin{bmatrix} 1 & 4 & 3 & 2 \\ 4 & 3 & 2 & 1 \\ 3 & 2 & 1 & 4 \\ 2 & 1 & 4 & 3 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix} \\
 &= \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix} \begin{bmatrix} 1 & 2 & 3 & 4 \\ 4 & 1 & 2 & 3 \\ 3 & 4 & 1 & 2 \\ 2 & 3 & 4 & 1 \end{bmatrix} \\
 &= \begin{bmatrix} 1 & 2 & 3 & 4 \\ 2 & 3 & 4 & 1 \\ 3 & 4 & 1 & 2 \\ 4 & 1 & 2 & 3 \end{bmatrix},
 \end{aligned}$$

which is the original operation table. Thus, this σ is an automorphism.

Some algebraic structures have more than one operation; for instance, the integers have both addition and multiplication (and others like subtraction which are really just addition of negatives). For these structures, an automorphism must fix not just one but *all* of the operation tables which define the structure.

Example 9. Consider the set $X = \{1, 2, 3, 4\}$ with operations

$*$	1	2	3	4		\circ	1	2	3	4
1	1	2	3	4		1	1	1	1	1
2	2	3	4	1	and	2	1	2	3	4
3	3	4	1	2		3	1	3	1	3
4	4	1	2	3		4	1	4	3	2

and the permutation $\sigma = [1, 4, 3, 2]$. We have already seen that σ is an automorphism for $*$; what about \circ ? Well,

$$\begin{aligned}
 P_{\sigma}^{-1}\sigma(M)P_{\sigma} &= \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix} \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & 4 & 3 & 2 \\ 1 & 3 & 1 & 3 \\ 1 & 2 & 3 & 4 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix} \\
 &= \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix} \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & 2 & 3 & 4 \\ 1 & 3 & 1 & 3 \\ 1 & 4 & 3 & 2 \end{bmatrix} \\
 &= \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & 4 & 3 & 2 \\ 1 & 3 & 1 & 3 \\ 1 & 2 & 3 & 4 \end{bmatrix}
 \end{aligned}$$

which is *not* the original operation table. Thus, this σ is not an automorphism for the full algebraic structure defined by $*$ and \circ .

For nonfinite algebraic structures (or really, even for finite but large ones) the operation table approach is not as useful since the table needs infinitely many rows and columns. In this situation, we can instead rely on algebraic formulas to describe isomorphisms. For example, the function $f : \mathbb{Z} \rightarrow \mathbb{Z}$ defined by $f(x) = -x$ is an automorphism of the addition operation since

$$f(x + y) = -(x + y) = -x + -y = f(x) + f(y)$$

and $f^{-1}(x) = f(x)$; however, f is not an automorphism of the multiplication operation since $f(xy) = -xy$ but $f(x)f(y) = (-x)(-y) = xy \neq -xy$.

Example 10. An automorphism of a vector space V is an invertible linear transformation $f : V \rightarrow V$. Given a choice of basis for V , each such automorphism is represented by a matrix with nonzero determinant. The set of automorphisms of V is known as the *general linear group* $GL(V)$, or if V is the set \mathbb{F}^n of ordered n -tuples of elements of a field \mathbb{F} , we generally write $GL_n(\mathbb{F})$.

Exercises. 1. Find the operation tables for all 16 operations on the set $X = \{1, 2\}$.

2. Determine whether the permutation $\sigma = [2, 3, 1]$ is an automorphism of the algebraic structure defined by the operation

$*$	1	2	3
1	1	3	1
2	2	2	2
3	3	1	3

3. Find all automorphisms of the algebraic structure on $X = \{1, 2, 3\}$ given by

$*$	1	2	3
1	1	1	1
2	2	2	2
3	3	3	3

4. Find all automorphisms of the algebraic structure on $X = \{1, 2, 3\}$ given by

$*$	1	2	3
1	1	2	2
2	2	1	1
3	3	3	3

5. Let X be the set of permutations of length three, i.e., bijective maps $\sigma : \{1, 2, 3\} \rightarrow \{1, 2, 3\}$. Define an operation $*$ on X by function composition, so for instance, if $f = [2, 3, 1]$ and $g = [2, 1, 3]$, then $gf = [2, 1, 3] * [2, 3, 1] = [1, 3, 2]$. Find the operation table for this algebraic structure.

6. Find an algebraic structure on a set of three elements that has the identity function as its only automorphism.

2. Quotient Sets and Equivalence Relations

Let X be a set. A *relation* on X is a way of comparing pairs of elements of X . More formally, recall that the *Cartesian product* of X

with itself is the set

$$X \times X = \{(x, x') \mid x, x' \in X\}$$

of ordered pairs of elements of X . Then a *relation* on X is a function $R : X \times X \rightarrow \{\text{True}, \text{False}\}$ which compares pairs of elements of X . We often write xRy in place of the phrase “ $R(x, y) = \text{True}$ ”.

Example 11. Any subset $S \subset X \times X$ of the Cartesian product defines a relation by saying xRy iff $(x, y) \in S$. The complement $X \times X \setminus S$ defines a relation \bar{R} . Thus, if $X = \{1, 2, 3\}$, the Cartesian product $X \times X$ is

$$\begin{aligned} X \times X = \{ & (1, 1), \quad (1, 2), \quad (1, 3), \\ & (2, 1), \quad (2, 2), \quad (2, 3), \\ & (3, 1), \quad (3, 2), \quad (3, 3) \}. \end{aligned}$$

The subset $\Delta = \{(1, 1), (2, 2), (3, 3)\}$, called the *diagonal* of $X \times X$, corresponds to the relation “ $=$ ”. The subset $\{(1, 2), (1, 3), (2, 3)\}$ corresponds to the relation “ $<$ ”.

Definition 2. A relation R on X is an *equivalence relation* if it satisfies:

- (i) For all $x \in X$, we have xRx (R is *reflective*).
- (ii) For all $x, y \in X$, xRy implies yRx (R is *symmetric*).
- (iii) For all $x, y, z \in X$, xRy and yRz implies xRz (R is *transitive*).

We often use the symbol \sim for equivalence relations in place of R .

Example 12. A *formal fraction* is a pair of two integers n (for “numerator”) and $d \neq 0$ (for “denominator”), written as

$$\frac{n}{d}.$$

Then we have an equivalence relation on the set of formal fractions defined by

$$\frac{n_1}{d_1} \sim \frac{n_2}{d_2}$$

if $n_1 d_2 = n_2 d_1$. Let us verify that this defines an equivalence relation:

- (i) For any formal fraction $\frac{n}{d}$, we have $nd = nd$, so $\frac{n}{d} \sim \frac{n}{d}$.

(ii) For any two formal fractions $\frac{n_1}{d_1} \cdot \frac{n_2}{d_2}$, if $\frac{n_1}{d_1} \sim \frac{n_2}{d_2}$, then

$$n_1 d_2 = n_2 d_1 \quad \text{implies} \quad n_2 d_1 = n_1 d_2$$

and $\frac{n_2}{d_2} \sim \frac{n_1}{d_1}$.

(iii) If $\frac{n_1}{d_1} \sim \frac{n_2}{d_2}$ and $\frac{n_2}{d_2} \sim \frac{n_3}{d_3}$, then we have $n_1 d_2 = n_2 d_1$ and $n_2 d_3 = n_3 d_2$. Then consider $n_1 d_3$; we have

$$d_2 n_1 d_3 = d_1 n_2 d_3 = d_1 n_3 d_2$$

and hence $d_2(n_1 d_3) = d_2(n_3 d_1)$; then

$$d_2(n_1 d_3 - n_3 d_1) = 0,$$

and since $d_2 \neq 0$, it follows that $n_1 d_3 = d_1 n_3$ and hence $\frac{n_1}{d_1} \sim \frac{n_3}{d_3}$.

Example 13. Let $X = \mathbb{Z}$ be the set of integers and say $n \sim m$ if and only if $n - m$ is even. Then \sim is an equivalence relation since

- (i) $n - n = 0$ is even,
- (ii) $m - n = -(n - m)$ so $m - n$ is even iff $n - m$ is, and
- (iii) $n - m$ even and $m - p$ even imply

$$n - p = n - m + m - p$$

is a sum of even numbers, which is even.

Definition 3. Let X be a set and let \sim be an equivalence relation on X . For each $x \in X$, the set

$$[x] = \{y \in X \mid y \sim x\}$$

of all elements of X equivalent to x is called the *equivalence class* of x . It turns out (see Exercise 2) the equivalence classes do not intersect, i.e.,

$$[x] \cap [y] \neq \emptyset \iff [x] = [y].$$

The division of X into equivalence classes forms a *partition* of X , i.e., a separation of X into disjoint (nonoverlapping) subsets whose union is X .

Example 14. Let $X = \mathbb{Z}$ and consider the equivalence relation \sim in Example 13 above. The class

$$[0] = \{0 + 2n \mid n \in \mathbb{Z}\} = \{0, \pm 2, \pm 4, \dots\}$$

includes all integers equivalent to 0, i.e., all integers which differ from 0 by an even number – that is, all even integers. The class

$$[1] = \{1 + 2n \mid n \in \mathbb{Z}\} = \{\pm 1, \pm 3, \pm 5, \dots\}$$

includes all integers which differ from 1 by an even integer; these are all of the odd integers. Note that $[0] \cap [1] = \emptyset$ and $[0] \cup [1] = \mathbb{Z}$. In particular, as a set X/\sim can be identified with any set which contains exactly one representative of each equivalence class.

The set of equivalence classes of X with equivalence relation \sim is called a *quotient set*, written X/\sim . Thus in our previous example we have $\mathbb{Z}/\sim = \{[0], [1]\}$. The quotient set may be pictured as the result of “collapsing” or “crushing” the equivalence classes $[x]$ into single points. Each element of an equivalence class is a *representative* of its class. Sometimes the equivalence classes have a natural choice for *canonical representative*, i.e., a notion of “best” representative for each class.

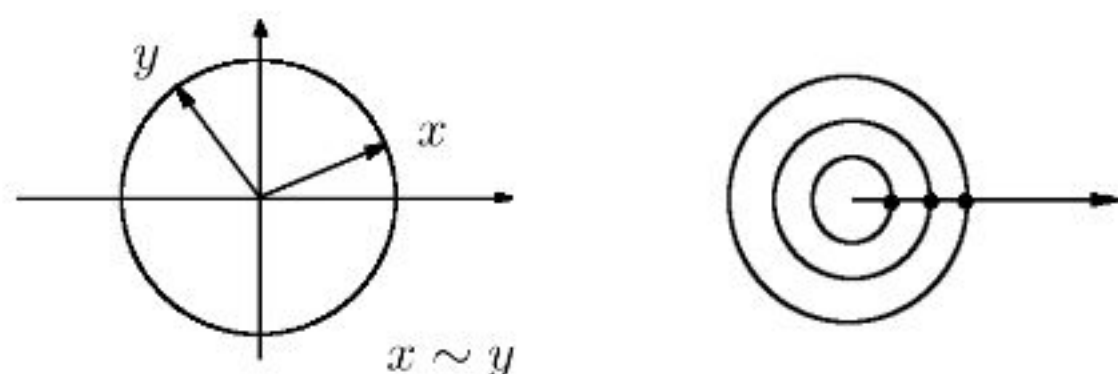
Example 15. The set of equivalence classes of formal fractions under the equivalence relation in Example 12 is the set of rational numbers \mathbb{Q} ; each fraction represents a ratio, differing from other members of its equivalence class by a cancelable factor in the numerator and denominator. For instance, we have

$$\left[\frac{1}{2}\right] = \left[\frac{2}{4}\right] = \left[\frac{-3}{-6}\right] = \dots$$

Each equivalence class of formal fractions under the equivalence relation in Example 12 has a unique canonical representative, namely the fraction written in *least terms* $\frac{a}{b}$ where the greatest common divisor of a and b is 1.

For general sets X and equivalence relations \sim there may not always be a natural choice of canonical representative. Indeed, in general, it can be quite hard to determine in practice or even provably impossible to determine whether two elements of a set are equivalent.

Example 16. For a geometric example, let two points in \mathbb{R}^2 be equivalent if the points are the same distance from the origin. Then the equivalence classes are circles centered at the origin, and the quotient set \mathbb{R}^2/\sim is the set of circles centered at the origin. We can choose canonical representatives to be points on the positive part of the x -axis, so the quotient set can be understood as the set of nonnegative real numbers $[0, \infty)$.



Equivalence relations determine partitions, and the converse is also true: every partition of a set X defines an equivalence relation by setting $x \sim y$ if and only if x and y are in the same subset of the chosen partition.

Example 17. There are eight partitions of the set $\{1, 2, 3, 4\}$ into two disjoint subsets, as listed below:

A	B	A	B
\emptyset	$\{1, 2, 3, 4\}$	$\{4\}$	$\{1, 2, 3\}$
$\{1\}$	$\{2, 3, 4\}$	$\{1, 2\}$	$\{3, 4\}$
$\{2\}$	$\{1, 3, 4\}$	$\{1, 3\}$	$\{2, 4\}$
$\{3\}$	$\{1, 2, 4\}$	$\{1, 4\}$	$\{2, 3\}$

Thus, there are eight equivalence relations on $\{1, 2, 3, 4\}$ which have two equivalence classes.

Similarly, if we have a set X and we would like to make certain equations

$$x_1 = y_1, x_2 = y_2, \dots, x_n = y_n$$

true, then we can ask for the the equivalence relation *generated by* these equations, that is, the smallest equivalence relation classes satisfying the given equations. The idea here is that some equations which are not explicitly listed may nonetheless be implied by the listed equations, and these must be taken into account.

Example 18. Let $X = \{a, b, c, d, e, f, g\}$. Then the equivalence relation generated by $a \sim c$, $d \sim a$ and $f \sim g$ includes an additional relation $c \sim d$ and has quotient set

$$X/\sim = \{[a, c, d], [b], [e], [f, g]\}.$$

As we have seen, the rational number system is really a quotient set. The same is true for some other number systems as well.

Example 19. The real numbers \mathbb{R} can be understood as equivalence classes of Cauchy sequences of rational numbers, i.e. sequences a_n where the terms get closer together as $n \rightarrow \infty$, under the equivalence relation given by $a_n \sim b_n$ if $\lim_{n \rightarrow \infty} (a_n - b_n) = 0$. A decimal expansion for a real number is a Cauchy sequence:

$$n.d_1d_2d_3 \cdots \leftrightarrow \left\{ n, n + \frac{d_1}{10}, n + \frac{d_1}{10} + \frac{d_2}{100}, \cdots \right\}.$$

For nonterminating decimals there is a canonical representative for each class given by the decimal expansion, but real numbers represented by terminating decimals have two decimal expansions, e.g., $1.000 \cdots = 0.999 \cdots$ or $0.5000 \cdots = 0.4999 \cdots$. It might seem strange at first to realize that some numbers have more than one decimal expansion, but really these are just two of infinitely many equivalent Cauchy sequences converging to the same limit.

Example 20. The complex numbers

$$\mathbb{C} = \{a + ib \mid a, b \in \mathbb{R}, i^2 = -1\}$$

are equivalence classes of polynomials in i with real number coefficients

$$a_0 + a_1i + a_2i^2 + \cdots + a_ni^n$$

where the equivalence relation is generated by $i^2 = -1$. That is, complex arithmetic is polynomial arithmetic with the extra rule that we can replace each i^2 with the equivalent expression -1 .

Congruences and Modular Arithmetic. Suppose X is a set with an operation $* : X \times X \rightarrow X$, e.g., the integers \mathbb{Z} with addition. An equivalence relation on X is called a *congruence* with respect to $*$ if it is compatible with $*$ in the sense that

$$x \sim x' \quad \text{and} \quad y \sim y' \quad \Rightarrow \quad x * y \sim x' * y'.$$

If \sim is a congruence with respect to $*$, then $*$ defines an operation on the set of equivalence classes by setting

$$[x] * [y] = [x * y].$$

Example 21. Let $X = \mathbb{Z}$, the set of integers. One very important example of an equivalence relation is *equivalence modulo n* where $n \in \mathbb{Z}$ is a fixed integer. Say that two integers x and y are *equivalent modulo n* , denoted $x \sim_n y$ or $x \equiv y \pmod{n}$, if $x - y = nz$ for some integer z . That is, two integers are equivalent mod n if they differ by a multiple of n . Let's verify that this definition gives us an equivalence relation:

(i) For any $x \in \mathbb{Z}$ we have

$$x - x = 0 = 0n \quad \Rightarrow \quad x \sim_n x.$$

(ii) For any $x, y \in \mathbb{Z}$, suppose $x \sim_n y$. Then $x - y = nz$, and

$$y - x = -(x - y) = -nz = n(-z) \quad \Rightarrow \quad y \sim_n x.$$

(iii) For any $x, y, z \in \mathbb{Z}$ suppose $x \sim_n y$ and $y \sim_n z$. Then $x - y = nu$ and $y - z = nv$ and we have

$$x - z = x - y + y - z = nu + nv = n(u + v) \quad \Rightarrow \quad x \sim_n z.$$

The equivalence relation \sim_n on \mathbb{Z} is a congruence with respect to addition since if $x \sim_n x'$ and $y \sim_n y'$ we have $x' = x + nz$ and $y' = y + nw$; then

$$x' + y' = (x + nz) + (y + nw) = (x + y) + n(z + w)$$

and $x' + y' \sim x + y$. Thus, the set of equivalence classes has a well-defined addition; this structure is known as the *integers modulo n* , denoted $\mathbb{Z}/n\mathbb{Z}$ or just \mathbb{Z}_n .

Indeed, \mathbb{Z}_n is a very important structure and will be used frequently throughout this book. It turns out (see exercise 5) that \sim_n is a congruence with respect to multiplication as well, so \mathbb{Z}_n has many of the useful features of \mathbb{Z} and even \mathbb{Q} . Modular arithmetic may seem strange at first, but we use it all the time to tell time; clock arithmetic is mod 12 (or mod 24 in some cases). Indeed, mod n arithmetic is just clock arithmetic with n hours instead of 12.

When doing arithmetic in a quotient set, you are free to choose any representative for each equivalence class. Thus, in \mathbb{Z}_5 , we have

$$[176] + [-422] = [1] + [-2] = [1] + [3] = [4].$$

It is common in algebraic structures defined by congruences to drop the square brackets and just write “ x ” for the equivalence class of x , using the rule that elements can be replaced by equivalent elements at any point. Thus, as long as we know we are working in \mathbb{Z}_5 , the above equation can be written more simply as

$$176 - 422 = 1 + 3 = 4.$$

The integers mod n form an algebraic structure with two operations, addition and multiplication. For example, here we have the addition and multiplication tables for \mathbb{Z}_5 :

+	0	1	2	3	4	×	0	1	2	3	4
0	0	1	2	3	4	0	0	0	0	0	0
1	1	2	3	4	0	1	0	1	2	3	4
2	2	3	4	0	1	2	0	2	4	1	3
3	3	4	0	1	2	3	0	3	1	4	2
4	4	0	1	2	3	4	0	4	3	2	1

It can be useful to consider vectors where the different components have different rules of arithmetic.

Definition 4. Suppose we have two number systems A and B , e.g., $A = \mathbb{Z}_n$ and $B = \mathbb{Z}_m$ or $A = \mathbb{Z}$ and $B = \mathbb{Z}_n$. The *direct sum* $A \oplus B$ is the set of ordered pairs

$$A \oplus B = \{(x, y) \mid x \in A, y \in B\}$$

where we use A arithmetic rules for the first component and B arithmetic rules for the second component.

Example 22. In $\mathbb{Z}_3 \oplus \mathbb{Z}_2$, we have

$$(2, 1) + (1, 1) = (3, 2) = (0, 0).$$

Example 23. We can form direct sums with any number of components. For instance, the set of ordered triples with integer first

component, second components in \mathbb{Z}_4 and third in \mathbb{Z}_6 is the direct sum

$$\mathbb{Z} \oplus \mathbb{Z}_4 \oplus \mathbb{Z}_6.$$

The infinite part \mathbb{Z} is called the *free part* while the finite part $\mathbb{Z}_4 \oplus \mathbb{Z}_6$ is called the *torsion part*. Then, for instance, in $\mathbb{Z} \oplus \mathbb{Z}_4 \oplus \mathbb{Z}_6$ we have

$$2(3, 3, 3) = (6, 2, 0).$$

Quotient Vector Spaces. Another important example we will need is the idea of a *quotient vector space*. Let V be a vector space with field of scalars \mathbb{F} , and let $S \subset V$ be a subspace of V . If $\{\vec{b}_1, \dots, \vec{b}_n\}$ is a spanning set for V , that is, if V is the set of linear combinations of $\vec{b}_1, \dots, \vec{b}_n$ with \mathbb{F} coefficients, then we will write

$$V = \mathbb{F}[\vec{b}_1, \dots, \vec{b}_n].$$

Then the relation \sim on V defined by $\vec{x} \sim \vec{y}$ iff $\vec{x} - \vec{y} \in S$ is an equivalence relation (see problem 9), and indeed a congruence with respect to both vector addition and scalar multiplication. The equivalence classes are *affine subspaces*

$$[\vec{x}] = \vec{x} + S,$$

i.e., copies of S shifted away from the origin by a vector \vec{x} , and we have

$$[\vec{x}] + [\vec{y}] = \vec{x} + S + \vec{y} + S = \vec{x} + \vec{y} + S = [\vec{x} + \vec{y}]$$

and

$$\lambda[\vec{x}] = \lambda\vec{x} + \lambda S = \lambda\vec{x} + S = [\lambda\vec{x}].$$

The set of equivalence classes V/\sim is itself a vector space called the *quotient vector space* V modulo S , denoted V/S .

If $\{\vec{b}_1, \dots, \vec{b}_k, \vec{b}_{k+1}, \dots, \vec{b}_n\}$ is a basis for V with $\{\vec{b}_1, \dots, \vec{b}_k\}$ a basis for S , then $\{\vec{b}_{k+1} + S, \dots, \vec{b}_n + S\}$ is a basis for V/S . It is common to drop the S and identify V/S with the span of $\{\vec{b}_{k+1}, \dots, \vec{b}_n\}$. In particular, we have

Theorem 1. *Let V be a vector space and $S \subset V$ a subspace. Then*

$$\dim(S) + \dim(V/S) = \dim(V).$$

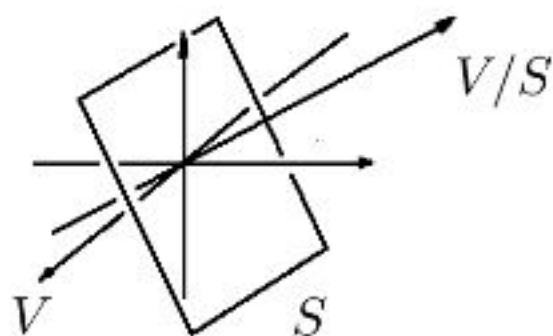
There is a natural choice of canonical representative for elements of the quotient space, given by the elements of S^\perp , the subspace of V of vectors \vec{w} which have dot product 0 with all elements of S . Recall that given a spanning set $\{\vec{s}_1, \dots, \vec{s}_k\}$ for S , we can find a basis for V/S by writing the vectors $\vec{s}_1, \dots, \vec{s}_k$ as row vectors in a matrix and row-reducing to reduced echelon form. The resulting basis for the null space of the matrix gives us a basis for the quotient space.

Example 24. Let $V = \mathbb{R}^4$ and $S = \mathbb{R}[(1, 2, 0, 1), (2, 2, -1, 1)]$. Then we have

$$\begin{bmatrix} 1 & 2 & -1 & 3 \\ 2 & 4 & -1 & 1 \end{bmatrix} \sim \begin{bmatrix} 1 & 2 & 0 & -2 \\ 0 & 0 & 1 & 5 \end{bmatrix}.$$

Interpreting this as the coefficient matrix for a homogeneous system of linear equations, we have basis $\{(-2, 1, 0, 0) + S, (2, 0, -5, 1) + S\}$ for V/S .

Geometrically, we can think of the vector space V as a stack of copies of S parametrized by elements of V/S ; then the quotient space V/S is the result of collapsing the copies of the subspace S down to single points:



Universal Algebra. Another example of quotient sets we will find very useful is in *universal algebra*. In a universal algebraic object (sometimes just called a “universal algebra”), we have a set of letters we call *generators* and a set of symbols usually including operator symbols like \cdot , $+$ or others as well as parentheses. These letters and symbols are then put together to form *well-formed words*; usually the rules for what constitutes as well-formed word are fairly obvious, like $(a * b) * c$ is a well-formed word where $((a * ($ is not, but to be clear the rules are generally spelled out explicitly. Think of the generators as basis vectors and well-formed words as linear combinations of the basis vectors. We then have a set of equations setting one word equivalent to another, (perhaps confusingly) called *relations*; these

equivalences then generate an equivalence relation on the set of well-formed words, and the universal algebraic object is a quotient set, i.e., the set of equivalence classes of well-formed words under the equivalence relation generated by the given relation. Such a description of an algebraic structure is called a *presentation by generators and relations*.

Example 25. We can describe the natural numbers

$$\mathbb{N} = \{0, 1, 2, 3, \dots\}$$

via a presentation with generator 1 and operation $+$, with well-formed words defined by the rules that

- $1 \in W$ and
- if $w, w' \in W$, then $(w + w') \in W$.

Then the presentation with relations

- $(w + w') + w'' \sim w + (w' + w'')$,
- $w + w' \sim w' + w$ and

for all well-formed words w describes \mathbb{N} as a quotient set of W . We recognize 0 as the empty word and any word with n copies of 1 as a representative of the natural number n .

Example 26. Adding a relation $n \times 1 \sim 0$ where $n \times 1$ is an abbreviation for a word with n copies of 1 in Example 25 gives us a presentation of \mathbb{Z}_n .

We will use presentations of algebraic structures by generators and relations later on when we look at groups and again when we study quandles and their various related objects.

Exercises. 1. Find all partitions on the set $\{1, 2, 3\}$.

2. Let \sim be an equivalence relation on a set X . Show that the equivalence classes are disjoint.

3. Let $X = \{a, b, c, d, e, f, h, i, j\}$. Find the equivalence relation generated by $a = c$, $f = c$, $d = b$, $c = d$, $i = j$ and $h = i$.

4. Let $f : X \rightarrow Y$ be a surjective (onto) function. Say that $x \sim x'$ iff $f(x) = f(x')$. Show that \sim is an equivalence relation. What are the equivalence classes? What is X/\sim ?
5. Show that \sim_n is a congruence with respect to multiplication of integers.
6. Let P be the set of polynomial functions with \mathbb{Z}_2 coefficients and let \sim be the congruence on P generated by $x^2 = 1 + x$. Identify the elements of P/\sim .
7. Construct the operation tables for $+$ and \times for \mathbb{Z}_3 , \mathbb{Z}_4 and \mathbb{Z}_5 .
8. Let \sim be the congruence on \mathbb{R}^2 generated by $(x, y) \sim (x + n, y + n)$ for $n \in \mathbb{Z}$. Identify the quotient set \mathbb{R}^2/\sim .
9. Show that the relation \sim on a vector space V defined by $\vec{x} \sim \vec{y}$ iff $\vec{x} - \vec{y} \in S$ for a subspace $S \subset V$ is an equivalence relation.
10. Let $S = \mathbb{R}[(2, 1, 1, -1, 1, 0), (1, 0, 2, -2, 0, 0), (1, 1, -1, 1, 1, 0)] \subset \mathbb{R}^6$. Find a basis for \mathbb{R}^6/S .

3. Modules

Recall from linear algebra that a vector space V has operations of *vector addition* $+: V \times V \rightarrow V$, i.e.,

$$\vec{u}, \vec{v} \in V \Rightarrow \vec{u} + \vec{v} \in V$$

and *scalar multiplication*

$$\alpha \in \mathbb{F}, \vec{v} \in V \Rightarrow \alpha \vec{v} \in V.$$

In linear algebra the scalars come from a *field*, i.e., a set \mathbb{F} with operations of addition and multiplication which are both

(i) *Associative*:

$$\begin{aligned} (\alpha + \beta) + \gamma &= \alpha + (\beta + \gamma), \\ (\alpha\beta)\gamma &= \alpha(\beta\gamma). \end{aligned}$$

(ii) *Commutative*:

$$\begin{aligned} \alpha + \beta &= \beta + \alpha, \\ \alpha\beta &= \beta\alpha. \end{aligned}$$

(iii) Multiplication *distributes* over addition:

$$\begin{aligned}(\alpha + \beta)\gamma &= \alpha\gamma + \beta\gamma, \\ \alpha(\beta + \gamma) &= \alpha\beta + \alpha\gamma.\end{aligned}$$

(iv) \mathbb{F} contains *additive and multiplicative identities* $0, 1 \in \mathbb{F}$ such that

$$0 + \alpha = \alpha \quad \text{and} \quad 1\alpha = \alpha.$$

(v) With *additive and multiplicative inverses* $-\alpha, \alpha^{-1}$ for every element α

$$-\alpha + \alpha = 0 \quad \text{and} \quad \alpha^{-1}\alpha = 1$$

with the exception that the additive identity 0 does not have a multiplicative inverse.

Examples of fields include $\mathbb{Q}, \mathbb{R}, \mathbb{C}$ and \mathbb{Z}_p for p prime (and many more). There are many sets which are almost but not quite fields, in the sense that they have addition and multiplication operations which satisfy some but not all of the conditions (i)–(v) above. For example, \mathbb{Z} is a set with associative and commutative addition and multiplication satisfying the distributive laws with identities 0, 1 and additive inverses for every element, but no multiplicative inverses for any elements except 1 and -1 . For instance, the integer 2 has no multiplicative inverse in the set of integers – it is true that $\frac{1}{2}$ is a multiplicative inverse for 2, but $\frac{1}{2}$ is not an integer – to get multiplicative inverses for every nonzero integer, we have to go beyond the set of integers, and thus the set of integers is not a field.

An “almost field” which satisfies the conditions (i)–(v) above except the multiplicative inverses requirement is called a *ring*; technically, this is just one type of ring, a *commutative ring with identity*, but we will not encounter other types of rings in this book. We can still do linear algebra with scalars from \mathbb{Z} and other rings provided the basic properties of scalar multiplication are satisfied. More precisely:

Definition 5. Let R be a ring. An R -module is a set M with an associative commutative vector addition operation with inverses and

an identity vector $\vec{0} \in M$ and a scalar multiplication satisfying

$$\begin{aligned}(\alpha + \beta)\vec{v} &= \alpha\vec{v} + \beta\vec{v}, \\ \alpha(\vec{v} + \vec{u}) &= \alpha\vec{v} + \alpha\vec{u}, \\ (\alpha\beta)\vec{v} &= \alpha(\beta\vec{v}), \text{ and} \\ 1\vec{v} &= \vec{v},\end{aligned}$$

for all $\vec{u}, \vec{v} \in M$ and $\alpha, \beta \in R$.

An R -module is essentially a vector space in which the scalars come from R ; that is, we have vector addition and scalar multiplication operations as usual, it's just that there are now some scalars we can't divide by. In particular, a field is a particular kind of ring, and linear algebra is a special case of module theory.

Let M and N be R -modules. A function $f : M \rightarrow N$ is a *linear transformation* if for all $\alpha \in R$ and $\vec{u}, \vec{v} \in M$ we have

$$f(\alpha\vec{u}) = \alpha f(\vec{u}) \quad \text{and} \quad f(\vec{u} + \vec{v}) = f(\vec{u}) + f(\vec{v}).$$

Some functions are linear transformations and some are not.

A set of vectors $B = \{\vec{b}_1, \dots, \vec{b}_m\}$ is a *basis* for a module M if every element of M can be written in a unique way as a linear combination of the vectors in B . If R is a field, this is equivalent to B being a linearly independent spanning set, but for modules the situation can be more complicated. A module is called *free* if it has a basis. If $B = \{\vec{b}_1, \dots, \vec{b}_m\}$ is basis for a module M , then every vector $\vec{u} \in M$ can be written in a unique way as

$$\vec{u} = \alpha_1\vec{b}_1 + \alpha_2\vec{b}_2 + \dots + \alpha_m\vec{b}_m.$$

We call the ordered m -tuple $(\alpha_1, \alpha_2, \dots, \alpha_m)$ the *coordinate m -tuple* for \vec{u} in the B basis and write

$$\vec{u}_B = (\alpha_1, \alpha_2, \dots, \alpha_m).$$

We write $M = R[\vec{b}_1, \dots, \vec{b}_n]$ if B is a basis for M .

If $B = \{\vec{b}_1, \dots, \vec{b}_m\}$ is a basis for M and $C = \{\vec{c}_1, \dots, \vec{c}_n\}$ is a basis for N , then if $f : M \rightarrow N$ is a linear transformation and $\vec{u} \in M$, we have

$$\begin{aligned}f(\vec{u})_C &= f(\alpha_1\vec{b}_1 + \dots + \alpha_m\vec{b}_m) \\ &= \alpha_1 f(\vec{b}_1)_C + \dots + \alpha_m f(\vec{b}_m)_C.\end{aligned}$$

The $n \times m$ matrix A whose columns are $f(\vec{b}_1)_C, \dots, f(\vec{b}_m)_C$ satisfies $f(\vec{u})_C = A\vec{u}_B$ where $A\vec{u}_B$ is the matrix product of A times the column vector \vec{u}_B . In particular, choosing bases for M and N determines a matrix for each linear transformation $f : M \rightarrow N$.

A subset S of an R -module is a *submodule* (or a *subspace* if R is a field) if S is closed under vector addition and scalar multiplication, i.e., if for all $\vec{u}, \vec{v} \in S$ and $\alpha \in R$ we have

$$\vec{u} + \vec{v} \in S \quad \text{and} \quad \alpha\vec{u} \in S.$$

In linear algebra, many problems can be solved by row-reduction of matrices to *reduced echelon form*, i.e., every row has a leading 1, the leading 1 in each row is to the right of the leading 1 in the row above, and each leading 1 is the only nonzero entry in its column. During row-reduction, we sometimes divide a row by a leading entry to get a leading 1, often resulting in a matrix full of fractions. We may even find ourselves doing extra row operations to try to avoid the dreaded fractions. Module theory is a bit like that – since the entries in the matrix are required to stay in the specified ring of scalars, we may not be able to divide to get a leading 1 in every row. Indeed, for \mathbb{Z} -modules, no fractions of any sort are allowed, and hence it is not always possible to get leading 1s; sometimes we have to settle for leading 2s or 7s. For example, the matrix

$$\begin{bmatrix} 2 & 0 & 1 \\ 0 & 4 & -1 \\ 0 & 0 & 0 \end{bmatrix}$$

is fully row-reduced over \mathbb{Z} .

Row moves on a matrix reflect changes to the output basis of the matrix; similarly, if we do column moves on a matrix, we change the input basis. We can often simplify a matrix further using column moves in addition to row moves, keeping in mind that the new matrix represents the original linear transformation with respect to new input and output bases. For example, the above matrix column-reduces over

\mathbb{Z} to

$$\begin{bmatrix} 2 & 0 & 1 \\ 0 & 4 & -1 \\ 0 & 0 & 0 \end{bmatrix} \sim \begin{bmatrix} 1 & 0 & 1 \\ 1 & 4 & -1 \\ 0 & 0 & 0 \end{bmatrix} \sim \begin{bmatrix} 1 & 0 & 0 \\ 1 & 4 & -2 \\ 0 & 0 & 0 \end{bmatrix} \\ \sim \begin{bmatrix} 1 & 0 & 0 \\ 0 & 4 & -2 \\ 0 & 0 & 0 \end{bmatrix} \sim \begin{bmatrix} 1 & 0 & 0 \\ 0 & 2 & -2 \\ 0 & 0 & 0 \end{bmatrix} \sim \begin{bmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 0 \end{bmatrix}.$$

A matrix is in *Smith normal form* if its only nonzero entries are on the diagonal and each nonzero diagonal entry divides the next. Not every ring allows a Smith normal form, but every matrix with integer entries has a Smith normal form.

Recall that there are two important submodules associated with a linear transformation $f : M \rightarrow N$:

- The *kernel* of f , also called the *null space* or *solution space*, is the set of all vectors in M which f maps to zero:

$$\text{Ker}(f) = \{\vec{x} \in M \mid f(\vec{x}) = \vec{0}\}.$$

- The *image* of f , denoted $\text{Im}(f)$ or $f(M)$, is the set of all elements of N that get hit by f , i.e.,

$$\text{Im}(f) = \{f(\vec{x}) \in N \mid \vec{x} \in M\}.$$

The image is spanned by the columns of the matrix of f and thus is sometimes called the *column space* of f .

For matrices with entries in a field, we can find bases for the image and kernel of a linear transformation by row-reducing the matrix to reduced echelon form; then the columns containing leading 1s form a basis for the image and the kernel has a basis vector for every column *without* a leading 1.

Example 27. The matrix

$$A = \begin{bmatrix} 1 & 0 & 1 & 0 & 3 \\ 0 & 1 & -1 & 0 & 1 \\ 0 & 0 & 0 & 1 & 0 \end{bmatrix}$$

representing the linear transformation $f : \mathbb{Q}^5 \rightarrow \mathbb{Q}^3$ defined by

$$f(x_1, x_2, x_3, x_4, x_5) = (x_1 + x_3 + 3x_5, x_2 - x_3 + x_5, x_4)$$

has image given by

$$\mathbb{Q} \left[\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} \right] \cong \mathbb{Q}^3.$$

The kernel is isomorphic to \mathbb{Q}^2 , which we can see by setting $x_3 = \alpha$ and $x_5 = \beta$; then

$$\begin{aligned} x_1 + \alpha + 3\beta &= 0 & x_1 &= -\alpha - 3\beta, \\ x_2 - \alpha + \beta &= 0 & x_2 &= \alpha - \beta, \\ x_3 &= \alpha \Rightarrow x_3 &= \alpha, \\ x_4 &= 0 & x_4 &= 0, \\ x_5 &= \beta & x_5 &= \beta, \end{aligned}$$

so every kernel element has the form

$$(-\alpha - 3\beta, \alpha - \beta, \alpha, 0, \beta) = \alpha(-1, 1, 1, 0, 0) + \beta(-3, -1, 0, 0, 1).$$

These are linearly independent since setting

$$(-\alpha - 3\beta, \alpha - \beta, \alpha, 0, \beta) = \vec{0}$$

implies that α and β are zero, so the kernel is

$$\mathbb{Q}[(-1, 1, 1, 0, 0), (-3, -1, 0, 0, 1)] \cong \mathbb{Q}^2.$$

If our matrix has entries in a ring, then finding a basis for the image and kernel may not be possible since they may not be free modules. However, we can still identify the image and kernel up to isomorphism from the Smith normal form of the matrix for \mathbb{Z} -modules. The image is still the span of the column vectors, while the kernel is the direct sum of \mathbb{Z}_a for diagonal entries a where we interpret $\mathbb{Z}_1 = \{0\}$ and $\mathbb{Z}_0 = \mathbb{Z}$.

Example 28. The matrix over \mathbb{Z} with Smith normal form

$$\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 2 & 0 & 0 \\ 0 & 0 & 6 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

has solution space given by $\mathbb{Z}_2 \oplus \mathbb{Z}_6 \oplus \mathbb{Z}$ as we can see by assigning free variables $\alpha = x_1$, $\beta = x_2$, $\gamma = x_3$ and $\delta = x_4$; then the system of homogeneous equations says $1\alpha = 0$, so α contributes a zero direct

summand; $2\beta = 0$ so β is a \mathbb{Z}_2 variable, $6\gamma = 0$ so γ is a \mathbb{Z}_6 variable and $0\delta = 0$ so δ has no constraints, i.e., δ contributes a \mathbb{Z} summand.

Exercises. 1. Prove that if R is a ring, then the set $M_n(R)$ of square $n \times n$ matrices is also a ring.

2. Prove that if R is a ring, then the set $R[x]$ of polynomials with coefficient in R is also a ring. Try the same question for $R[x, y]$ the set of polynomials with two variables x and y .

3. Using row operations, find bases for the image and kernel of the matrix over \mathbb{Q} :

$$\begin{bmatrix} 1 & 1 & 2 & 1 & 1 \\ 3 & -1 & -1 & 2 & 1 \\ 0 & 1 & 1 & 0 & -1 \end{bmatrix}.$$

4. Using row and column operations, find the Smith normal form for the matrix

$$\begin{bmatrix} 2 & 1 & 1 & 0 & 1 \\ 1 & 1 & -1 & 2 & 1 \\ 0 & 1 & 2 & 1 & 0 \end{bmatrix}.$$

5. Prove that if R is a ring, then the Cartesian product R^n , where n is a natural number, is an R -module.

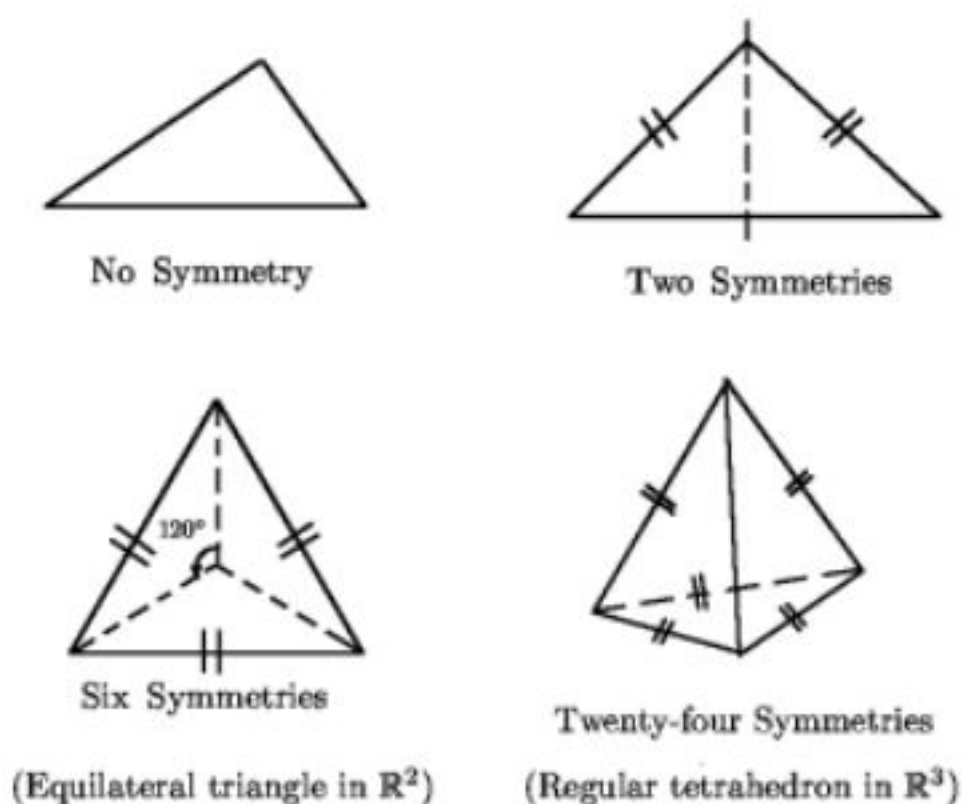
6. A nonzero element u in a ring R is called a *proper zero divisor* if there exists a nonzero element x in R such that $xu = 0$. Show that a proper zero divisor cannot be multiplicatively invertible in R .

7. Prove that the set $\{6, 14, 21\}$ generates \mathbb{Z} but no subset of it generates \mathbb{Z} . Hint: $\gcd(6, 14, 21) = 1$.

4. Groups

We start by asking the question, “given an object, how many symmetries does it have?” For example, a circle has more symmetries than a triangle. An equilateral triangle has more symmetries than an

isosceles triangle. The following figures show the symmetries of three triangles in the plane \mathbb{R}^2 and a regular tetrahedron in the 3-space \mathbb{R}^3 .



The notion of symmetry was at the origin of the notion of a *group*. The famous German mathematician Christian Felix Klein (25 April 1849–22 June 1925) was instrumental in the development of the theory of groups. His 1872 *Erlangen Program*, classifying geometries by their underlying symmetry groups, was a hugely influential synthesis of much of the mathematics of the day. The notion of a group is a central idea in modern mathematics such as Galois theory. The French mathematician Galois associated a group to a given equation such as

$$a_n x^n + a_{n-1} x^{n-1} + \cdots + a_1 x + a_0 = 0$$

in such a way that the properties of the group allow answering the question of whether the solutions can be derived from the coefficients a_0, \dots, a_n only by addition, subtraction, multiplication, division and extraction of roots (this is called solvable by radicals). Groups also appear in the study of combinatorics, crystallography, physics, etc.

Now we state the formal definition of a group.

Definition 6. A *group* is a set $(G, *)$ with a binary operation $(a, b) \mapsto a * b$ such that the following three axioms hold:

- (i) For all $a, b, c \in G$, $(a * b) * c = a * (b * c)$, (associative property).

- (ii) There exists an element $e \in G$ such that for all $a \in G$, $e * a = a * e = a$, (existence of an identity element).
- (iii) For all $a \in G$, there exists an element denoted a^{-1} such that $a * a^{-1} = e = a^{-1} * a$ (existence of inverses).

Many familiar sets have a group structure:

Example 29. The set of integers \mathbb{Z} is a group with addition as the operation. The set of integers modulo n , \mathbb{Z}_n is also a group with addition.

Example 30. The set of bijections (called also permutations) S_X from a set X to itself with composition as the operation is a group. When the set X has n elements, this group is denoted by S_n . We can specify a permutation $\sigma : \{1, 2, \dots, n\} \rightarrow \{1, 2, \dots, n\}$ conveniently in two ways:

- As we have seen, by giving a vector specifying the list

$$\sigma = [\sigma(1), \sigma(2), \dots, \sigma(n)]$$

of the images of the elements of σ in order, e.g.,

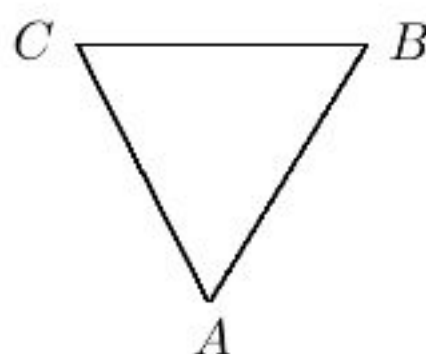
$$\sigma = [1, 3, 2]$$

represents the permutation on $\{1, 2, 3\}$ fixing 1 and switching 2 with 3, or

- Using *cycle notation*, where we write an element followed by its image with parentheses closing cycles, e.g., the permutation with image vector $[2, 3, 1, 5, 4]$ has cycle notation $(123)(45)$. Note that fixed points are usually left out of cycle notation, so the identity is the empty cycle $()$.

Example 31. The set of rotations in the plane around a fixed point is a group with composition.

Example 32. The set of n by n matrices with real coefficients is a group with addition.

Example 33.

Let ABC be an equilateral triangle in the plane ($AB = BC = CA$). Let G be the group of all symmetries of this triangle. Any symmetry permutes the three vertices A , B and C . Geometrically, G is made of reflections $[A, C, B]$, $[C, B, A]$ and $[B, A, C]$ about each vertex, the rotation of 120 degree angle $[B, C, A]$, the rotation of 240 degree angle $[C, A, B]$ and the rotation of 360 degree angle which corresponds to the identity $[A, B, C]$. In cycle notation these are

$$\{(BC), (AC), (AB), (ABC), (ACB), ()\}.$$

This group is usually denoted by D_3 and called the *dihedral group* of order six. More generally, the group of symmetries of a regular n -gon is the dihedral group of order $2n$, denoted D_n .

A group $(G, *)$ is called *abelian* if $x * y = y * x$ for all x and y in G . That is, an abelian group has commutative group operation. For example the set \mathbb{Z} of integers with addition is an abelian group, while the group of symmetries of an equilateral triangle is not abelian.

Example 34. A commutative ring with identity is an abelian group under addition, and the ring without zero is an abelian group under multiplication. Indeed, this is the easiest way to remember the commutative ring with identity axioms: a commutative ring with identity is a set with two abelian group structures (with the exception that 0 has no multiplicative inverse) with one operation distributing over the other.

In a group, the identity element is always unique since if both e and e' are identities, we have $e = ee' = e'$. Similarly, each element of a group has a unique inverse.

Definition 7. Given two groups $(G, *)$ and (K, \circ) , a group *homomorphism* is a function $f : G \rightarrow K$ that satisfies

$$f(x * y) = f(x) \circ f(y) \quad \text{for all } x, y \in G.$$

Example 35. For a positive integer n , define $f : \mathbb{Z} \rightarrow \mathbb{Z}_n$ by for all m in \mathbb{Z} , $f(m) = [m]$ (class of m modulo n). Then f is a homomorphism that is onto but not one-to-one.

If $f : (G, *) \rightarrow (K, \circ)$ is a group homomorphism, then $f(e_G) = e_K$ and $f(x^{-1}) = f(x)^{-1}$, for all x in G .

A group homomorphism that is bijective is called a group *isomorphism*.

Definition 8. Given a group homomorphism $f : (G, *) \rightarrow (K, \circ)$:

- (i) The *kernel* of f is the set of all elements g of G such that $f(g) = e_K$. It is denoted by $\text{Ker}(f)$.
- (ii) The *image* of f is the set of all elements $f(g)$ where $g \in G$. This is a subset of K and it is denoted by $\text{Im}(f)$.

Subgroups, Normal Subgroups and Quotients of Groups. A subset H of a group G is called a *subgroup* of $(G, *)$ if $(H, *)$ is a group with respect to the operation of G . This means that H is closed under the operation $*$ and under taking inverses, so $g, h \in H$ implies $g * h \in H$ and $g^{-1}, h^{-1} \in H$.

Example 36. $(\mathbb{Z}, +)$ is a subgroup of $(\mathbb{Q}, +)$.

Example 37. The symmetry group of the square is a subgroup of the symmetric group S_4 .

If $(H, *)$ is a subgroup of $(G, *)$, then the identity element in H is the same as the identity element in G . Also, the inverse of any element h in H is the same as the inverse of h in G . Now we have the following characterization of subgroups.

Theorem 2. Let $(G, *)$ be a group and H be a subset of G . Then $(H, *)$ is subgroup of $(G, *)$ if and only if

- (i) H is nonempty set, $(H \neq \emptyset)$
- (ii) H is closed under the operation $*$, and

(iii) For all $a \in H$, $a^{-1} \in H$ (H is closed under the operation inverse).

Example 38. The set $3\mathbb{Z} = \{3n, n \in \mathbb{Z}\}$ is a subgroup of $(\mathbb{Z}, +)$.

Example 39. Consider an equilateral triangle and let r be the rotation of 120-degrees around the center of gravity of the triangle. Then the set $\{e, r, r^2 = r \circ r\}$ is a subgroup of D_3 .

Definition 9. A subgroup N of a group G is called a *normal* subgroup if for any $g \in G$ and any $h \in N$, $ghg^{-1} \in N$. In other words, N is closed under conjugation by all elements g of G .

This definition means, in other words, that for all g in G and all h in N , there exists h' in N such that $gh = h'g$.

Example 40. If G is an abelian group, then any subgroup of G is normal since conjugation is trivial ($ghg^{-1} = h$).

Example 41. From Example 33 we conclude that the set of rotations $\{e, r, r^2 = r \circ r\}$, is a *normal* subgroup of S_3 . Furthermore, if y is a reflection about a vertex of the equilateral triangle then we have, for example, $yry^{-1} = r^{-1}$. But the subgroup $\{e, y\}$ is not normal since the element ryr^{-1} is neither the identity nor the reflection y .

Theorem 3. If $f : (G, *) \rightarrow (K, \circ)$ is group homomorphism, then $\text{Im}(f)$ is a subgroup of K , and $\text{Ker}(f)$ is a normal subgroup of G .

If N is a normal subgroup of G , then the set of equivalence classes G/N (where $x \sim y \iff y = nx$ for some $n \in N$) has a group structure given by $[x][y] := [xy]$. It is called the *quotient* group G by N and denoted G/N . We leave this as an exercise. In this context equivalence classes are also called cosets and $[x]$ is denoted sometimes by Nx .

We will see shortly in the next theorem that quotient groups are essentially the same as homomorphic images. For example, the group \mathbb{Z}_3 is constructed in an easy way from the group of integers. The set of all multiples of 3, denoted $3\mathbb{Z}$, form a normal subgroup of \mathbb{Z} since the latter is abelian. The elements of \mathbb{Z}_3 are the cosets of the subgroup $3\mathbb{Z}$.

Theorem 4. *If N is a normal subgroup of a group G , then the natural map $f : G \rightarrow G/N$ defined by $f(x) = [x]$ is a surjective homomorphism and the kernel of f is N .*

Direct Product of Groups. Let G and H be two groups. Then the Cartesian product $G \times H$ becomes a group with respect to the operation

$$(x, a) \times (y, b) = (xy, ab)$$

for all $x, y \in G$ and $a, b \in H$. The group $G \times H$ with this operation is called the *direct product* of G and H . For example, the cyclic group \mathbb{Z}_{15} is isomorphic to the direct product of the cyclic groups \mathbb{Z}_3 and \mathbb{Z}_5 .

If G and H are abelian groups with operation written as addition, then the direct product $G \times H$ is the same as the direct sum $G \oplus H$.

Finite Abelian Groups. Recall that a group G is called *abelian* if the group operation is commutative, that is $ab = ba$ for all $a, b \in G$. Usually the operation in an abelian group is denoted by the sign $+$, the identity element is denoted by 0 and the inverse of an element x is denoted $-x$. We then adopt the notation $(G, +)$ for an abelian group. We define $0x$ to be 0 . For any $x \in G$ and a positive integer n we define nx to be the sum of n copies of x , i.e.,

$$nx = \overbrace{x + x + \cdots + x}^n.$$

Then $(-n)x$ is defined by $(-n)x = n(-x)$. Now we state the following theorem which describes all finite abelian groups (up to isomorphism) in a standard way.

Theorem 5. *Every finite Abelian group is a direct product of cyclic groups of prime-power order. Furthermore, the number of terms in the product and the orders of the cyclic groups are uniquely determined by the finite abelian group.*

This theorem states that if G is a finite abelian group then G is isomorphic to

$$\mathbb{Z}_{p_1^{n_1}} \times \mathbb{Z}_{p_2^{n_2}} \times \cdots \times \mathbb{Z}_{p_l^{n_l}}$$

where the p_j 's are not necessarily distinct primes. The prime powers $p_1^{n_1}, p_2^{n_2}, \dots, p_l^{n_l}$ are completely and uniquely determined by the group G .

Example 42. Let p be a prime number and k a positive integer. Any abelian group of order p^k is isomorphic to a direct product $\mathbb{Z}_{p^{n_1}} \times \mathbb{Z}_{p^{n_2}} \times \dots \times \mathbb{Z}_{p^{n_l}}$ where the n_i 's are positive integers and $k = n_1 + n_2 + \dots + n_l$ (this is called a partition of k). For example any abelian group of order 4 is either isomorphic to \mathbb{Z}_4 or to $\mathbb{Z}_2 \times \mathbb{Z}_2$.

Finitely Generated Abelian Groups. Every abelian group can be understood as a \mathbb{Z} -module, with nx meaning add x to itself n times. It is then natural to think about abelian groups in terms of bases. An Abelian group G is called *finitely generated* if there exist finitely many elements g_1, \dots, g_n of G such that any element x of G can be written as

$$x = k_1g_1 + \dots + k_ng_n$$

where k_i are integers for $i = 1, \dots, n$. The set $\{g_1, \dots, g_n\}$ is called a system of *generators* of G and is denoted by $\langle g_1, \dots, g_n \rangle$.

Example 43. We have the following example and nonexample of finitely generated abelian groups:

- The groups $(\mathbb{Z}, +)$ and $(\mathbb{Z}_n, +)$ are finitely generated abelian groups (for example 1 and $[1]$ are respectively systems of generators for these groups).
- The group $(\mathbb{Q}, +)$ of rational numbers with addition is *not* a finitely generated abelian group. We leave this as an exercise.

A system of generators (g_1, \dots, g_n) of an abelian group G is called *free* if for all $k_1, \dots, k_n \in \mathbb{Z}$ the relation

$$k_1g_1 + \dots + k_ng_n = 0$$

implies $k_1 = \dots = k_n = 0$. This gives us the fact that any element x of G is written *uniquely* as

$$x = k_1g_1 + \dots + k_ng_n.$$

In this case the system is called a *basis* of the *free abelian group* G (this is similar to the case of vector spaces over fields). Any two bases of a free abelian group have the same number of elements called the *rank* of the group.

Example 44. The set of two by two matrices over the integers with addition is a free abelian group of rank 4. This is because any matrix $\begin{bmatrix} a & b \\ c & d \end{bmatrix}$ can be written uniquely as

$$a \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} + b \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} + c \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix} + d \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}.$$

Presentations of Groups by Generators and Relations. Sometimes it is convenient to define a group with certain prescribed properties. Simply stated, we start with a set of elements that we want to generate the group, and a set of equations (called relations) that the generators have to satisfy. Let us start with a concrete example: consider D_5 , the group of symmetries of a regular pentagon. Let R be the rotation of angle $\frac{2\pi}{5}$, and S the reflection around a line passing through a vertex of the pentagon and the midpoint of its opposite side. Recall that R and S generate the group D_5 . Notice that R and S are related by the equation

$$(1) \quad R^5 = S^2 = (RS)^2 = 1.$$

There are obviously other equations between R and S , such as $SR = R^4S$ and $RSR = S$, but they can be derived from those given in equation (1). In fact, any relation between R and S can be obtained from those given in equation (1). Then, the group D_5 is generated by a pair of elements x and y subject to the relations

$$x^5 = y^2 = (xy)^2 = 1.$$

It is then natural to ask if this description of the group D_5 applies to some other group as well. The answer is NO! This means that any group generated by two elements u and v such that $u^5 = v^2 = (uv)^2 = 1$ is isomorphic to the group D_5 . We traditionally write

$$D_5 = \langle x, y \mid x^5 = y^2 = (xy)^2 = 1 \rangle.$$

As another example, the cyclic group of order n can be presented by one generator denoted x and the relation $x^n = 1$. One may think of x as the rotation of angle $\frac{2\pi}{n}$. Since composition of rotations corresponds to adding their angles, we see that the rotation of the angle $\frac{2\pi}{n}$ composed n -times gives the identity transformation.

In fact, this cyclic group is the same as the integers mod n considered as an abelian group under addition. We can then write

$$\mathbb{Z}_n = \langle x \mid x^n = 1 \rangle.$$

The advantages of defining groups this way include that it is a very compact notation (much smaller than giving the operation table, for instance) and also because many groups in algebraic topology arise naturally this way. For example we will see that the *fundamental group* of the trefoil knot 3_1 is given by the presentation

$$\langle a, b \mid aba = bab \rangle.$$

Another presentation of this group is

$$\langle x, y \mid x^3 = y^2 \rangle.$$

We will leave this as an exercise to check that the two presentations give the same group. We think of isomorphic groups as being the same.

In order to give the general definition of a group in term of generators and relations, we need some notation. For any set $X = \{x_1, \dots, x_n\}$ of distinct elements (called *symbols* in this context), we form a new set denoted $X^{-1} = \{x_1^{-1}, \dots, x_n^{-1}\}$ (again at this level these elements are just symbols). We define the set $W(X)$ as the collection of all formal finite strings of the form $a_1 a_2 \dots a_m$, where a_j is in the union $X \cup X^{-1}$. The set $W(X)$ is called the set of *words* on X . We allow the string with no elements to be in $W(X)$, call it the empty word and denote it symbolically by 1. There is a natural “multiplication” in $W(X)$ which is juxtaposition of words; that is, the multiplication of $a_1 a_2 \dots a_l$ and $b_1 b_2 \dots b_m$ is the word $a_1 a_2 \dots a_l b_1 b_2 \dots b_m$. It is clear that this binary operation is associative and its identity is the empty word. Be aware that at this level the word xx^{-1} is not the empty word, and this is because we are interpreting the elements purely as symbols with no meaning yet. So

far we have everything (associativity and identity) to make a group except the notion of inverse. Recall from linear algebra that if A and B are two invertible matrices, then the inverse of AB is $B^{-1}A^{-1}$ (a fact sometimes known as the “shoe-sock theorem”, since taking off socks and shoes must be done in the opposite order from which they were put on). We expect the inverse of the word xy to be $y^{-1}x^{-1}$, but again $xyy^{-1}x^{-1}$ is not the empty word, as we explained earlier that elements are thought of as just symbols with no meaning. To remedy this problem we need to define an equivalence relation on $W(X)$.

Definition 10. Let X be a finite set of symbols and $W(X)$ be the set of words on X . Given any pair of elements w and w' in $W(X)$, we say that $w \sim w'$ if and only if w' can be obtained by a finite sequence of insertions or deletions of words of the form $x^{-1}x$ or xx^{-1} , where x is in X .

Example 45. Consider the set $X = \{x, y, z, u\}$. Then the word xyz is equivalent to $xyuu^{-1}z$, the word $xzyy^{-1}yyuxz$ is equivalent to $xzyyuxz$, and $xyzz^{-1}y^{-1}x^{-1}$ is equivalent to the empty word 1. However, the word $uxzx^{-1}$ is not equivalent to uz .

The relation defined in the previous definition is an equivalence relation.

Definition 11. Let X be a finite set of symbols and $W(X)$ be the set of words on X . For any word w in $W(X)$, let $[w]$ denote the equivalence class of w . The set of all equivalence classes of elements of $W(X)$ is a group under the binary operation $[w] \cdot [w'] = [ww']$. This group is called the *free group* on X .

It turns out that every group is a homomorphic image of a *free group*. The proof can be found in some of the classical text books on group theory.

Now we have the foundation of defining a group by generators and relations. Before we give the precise definition, we revisit the example of D_5 , the symmetry group of a regular pentagon we discussed earlier in the motivation.

Example 46. Let F be the free group on the set $X = \{x, y\}$ and let N be the smallest normal subgroup of F containing the set

$\{x^5, y^2, (xy)^2\}$. We claim that the quotient group F/N is isomorphic to D_5 . To see this, consider the group homomorphism $h : F \rightarrow D_5$ such that $h(x) = R$ and $h(y) = S$. Notice that h is a surjective homomorphism whose kernel contains N . We leave it as an exercise to check that in fact the kernel of h is exactly N .

Definition 12. Let G be a group generated by some set of symbols $X = \{x_1, \dots, x_m\}$ and let $F(X)$ be the free group on X . Let $W = \{r_1, \dots, r_l\}$ be a subset of $F(X)$ and N be the smallest normal subgroup of $F(X)$ containing W . We then say that the group G is given by generators x_1, \dots, x_m and relations $r_1 = 1, \dots, r_l = 1$ if there is an isomorphism from F/N to G which sends the class $[x_j]$ modulo N to x_j . We write $G = \langle x_1, \dots, x_m \mid r_1 = \dots = r_l = 1 \rangle$.

Example 47. The set of integers as a group can be given by the presentation $\mathbb{Z} = \langle x \rangle$ with only one generator and no relations. The cyclic group \mathbb{Z}_n can be given by the presentation $\langle x : x^n = 1 \rangle$ and the dihedral group D_n (the group of symmetry of a regular n -gon) can be given by $D_n = \langle x, y \mid x^n = y^2 = (xy)^2 = 1 \rangle$.

Exercises. 1. Make the operation tables for \mathbb{Z}_4 and $\mathbb{Z}_2 \times \mathbb{Z}_2$ and show that these groups are not isomorphic.

2. Does the binary operation $x * y = x\sqrt{1+y^2} + y\sqrt{1+x^2}$ define a group structure on the real line \mathbb{R} ?

3. Let N be a normal subgroup of G . Prove that the following relation on G is an equivalence relation,

$$\forall x, y \in G, x \sim y \iff y = nx, \quad \text{for some } n \in N.$$

Define the binary operation on the quotient set G/N by $[x][y] = [xy]$. First check that this operation is well defined, and then prove that G/N with this operation is a group.

4. Prove that the group $(\mathbb{Q}, +)$ of rational numbers with addition is not a finitely generated abelian group. Hint: Assume that the following n rational numbers $\frac{p_1}{q_1}, \dots, \frac{p_n}{q_n}$ generate all rationals and find a contradiction.

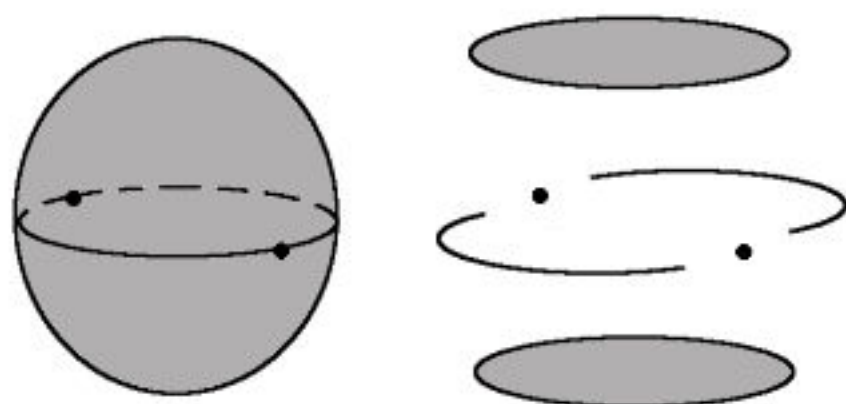
5. Let m and n be two positive integers. Find all group homomorphisms from $(\mathbb{Z}_m, +)$ to $(\mathbb{Z}_n, +)$, and all group automorphisms of $(\mathbb{Z}_m, +)$.
6. Prove that the set of transpositions switching 1 with i where $2 \leq i \leq n$ generates the symmetric group S_n .
7. Use the figure of a regular tetrahedron, mentioned in the beginning of this section, to compute the group G of its symmetries. Conclude that G is isomorphic to the group S_4 of permutations of four letters. Prove that the set of rotations H form a subgroup of G . The group H viewed as a subgroup of S_4 is called the alternating group and denoted A_4 .
8. Let n be an integer greater than or equal to 2 and m be a positive odd integer. Prove that the only group homomorphism from S_n to \mathbb{Z}_m is the zero map (sending any element to zero).
9. Let n be a nonnegative integer. Prove that the linear groups $GL_n(\mathbb{R})$ and $GL_n(\mathbb{C})$ cannot be isomorphic. (Hint: you may use the facts that the *centers*, i.e., sets of elements which commute with all other elements, of $GL_n(\mathbb{R})$ and $GL_n(\mathbb{C})$ are, respectively, the multiplicative groups $\mathbb{R} - \{0\}$ and $\mathbb{C} - \{0\}$).
10. Let G be a finite subgroup of the group of affine bijections, that is, bijections of the form $f(\vec{x}) = A\vec{x} + \vec{y}$ for an invertible matrix A , of a real vector space V .
 - (i) Prove that there exists a point that is invariant by all elements of G .
 - (ii) Determine all finite subgroups of the group of nonzero complex numbers $(\mathbb{C} - \{0\}, \times)$ with multiplication.
 - (iii) First, recall that a *similarity* is a plane transformation that preserve the ratio of the distances. It is well known that a plane transformation is a similarity if and only if it multiplies the distances by a positive real number k (called the ratio of the similarity). If $k = 1$, the similarity is called an *isometry*.

Determine all finite subgroups of the group of plane similarities which preserve orientation.

5. Cohomology

Cohomology has been called one of the most important contributions to mathematics made during the 20th century. It is a way of translating geometric or topological questions into algebraic questions, and shows up in many places in modern mathematics. We will not need to develop cohomology theory in great depth (which is fortunate as it would take another full book to do it justice) but will content ourselves with a brief introduction using linear algebra.

Cohomology has its origins in geometry. A *cell decomposition* of a subset $X \subset \mathbb{R}^n$ divides X into *cells* of various dimensions; each cell has a *boundary* consisting of cells of lower dimensions.



By thinking of the boundary of a cell as a linear combination of cells one dimension down, we can describe the overall set X with a set of vector spaces generated by various cells related to each other by linear transformations encoding the boundary maps. The key observation is that the boundary of a boundary is empty; in terms of linear maps, this means that the composition of two boundary transformations must be the zero map.

Cohomology also appears when we generalize the fundamental theorem of calculus to higher dimensions using *differential forms*, which provide an elegant way of unifying Green's and Stokes' theorems. Simply stated, a 0-form on a region in the xy plane is a scalar-valued function $f(x, y)$, an expression of the type $f(x, y)dx + g(x, y)dy$ is a 1-form and an expression of the form $F(x, y)dx \, dy$ is a 2-form. That is, a k -form is a product of k differentials with scalar function coefficients. The “ d ” is a differential operator with the property that

$d(d\omega)$ is always zero, a fact which gives as cohomology the theory known as *de Rham cohomology*. See the exercises for more.

Let \mathbb{F} be a field and let $C^0, C^1, C^2, \dots, C^n, \dots$ be \mathbb{F} -vector spaces. For each $k = 1, 2, \dots$ let $d^k : C^{k-1} \rightarrow C^k$ be a linear transformation represented by a matrix A_k . If for all $k = 1, 2, \dots$ the matrix product $A_{k+1}A_k$ is the zero matrix, i.e., if the composite maps $d^{k+1} \circ d^k$ are all equal to the zero map so we have $d^{k+1}(d^k(\vec{v})) = \vec{0}$ for all $\vec{v} \in C^{k-1}$, then the sequence of vector spaces and linear transformations

$$\dots \xleftarrow{d^{n+1}} C^n \xleftarrow{d^n} C^{n-1} \xleftarrow{d^{n-1}} \dots \xleftarrow{d^3} C^2 \xleftarrow{d^2} C^1 \xleftarrow{d^1} C^0$$

is called a *cochain complex*. Note that we have written the maps going from right to left so the order of the maps agrees with the usual convention for matrix multiplication. In particular, the cochain complex condition says that the column space of A_k is always a subspace of the null space or kernel of A_{k+1} . In terms of linear transformations, this says $\text{Im}(d^k) \subset \text{Ker}(d^{k+1})$.

Vectors in C^k are called *k-cochains* and the linear transformations d^k are called *coboundary maps* or *differentials*. The column space of A_k is usually denoted B^k and vectors in B^k are called *k-coboundaries*; the null space of A_{k+1} is denoted Z^k and its elements are called *k-cocycles*. Thus, in a cochain complex we always have $B^k \subset Z^k$. The quotient vector space

$$H^k = Z^k / B^k = \text{Ker}(d^{k+1}) / \text{Im}(d^k)$$

is called the *k-th cohomology space* of the chain complex.

Note that if the indices are going down rather than up with application of d , we have a *chain complex* with *homology spaces* rather than cohomology. For example, if the coboundary maps are expressed as matrices, taking the transpose of each map reverses the direction and switches from cohomology to homology. We will primarily need cohomology in this book.

Many cochain complexes are effectively finite, in that there is a largest n beyond which all the cochain spaces are the zero vector space $0 = \{\vec{0}\}$ and all the differentials are the zero map $0(\vec{v}) = \vec{0}$. Indeed, unspecified spaces and maps will be assumed to be zero. We

will usually write such a cochain complex as a finite sequence:

$$C^n \xleftarrow{d^n} C^{n-1} \dots \xleftarrow{d^3} C^2 \xleftarrow{d^2} C^1 \xleftarrow{d^1} C^0.$$

In particular, since C^{-1} and $d^0 : C^{-1} \rightarrow C^0$ are not listed, we have $\text{Im}(d^0) = 0$ and $B^0 = 0$; then we have $H^0 = Z^0/0 = Z^0$; similarly, $C^{n+1} = 0$ and $d^{n+1} : C^n \rightarrow 0$ is the zero map, so $Z^n = \ker(d^{n+1}) = C^n$ and we have $H^n = C^n/B^n$.

Example 48. For a first example, suppose we have a cochain complex in which all of the differentials d^k are the zero map. Then for every k , we have $Z^k = C^k$ and $B^k = 0$, so $H^k = Z^k/B^k = C^k/0 = C^k$.

Example 49. Suppose we have a sequence of vector spaces and linear maps

$$C^n \xleftarrow{d^n} C^{n-1} \dots \xleftarrow{d^3} C^2 \xleftarrow{d^2} C^1 \xleftarrow{d^1} C^0$$

such that $\text{Im}(d^k) = \ker(d^{k+1})$. Then we have $Z^k = B^k$ for all k and $H^k = Z^k/Z^k = 0$ for all k . Such a sequence is called *exact*. Indeed, cohomology can be understood as measuring the failure of a sequence to be exact.

Example 50. For a nontrivial example of cohomology, consider the sequence

$$0 \xleftarrow{0} \mathbb{Q}^2 \xleftarrow{\begin{bmatrix} 1 & 1 \\ 2 & 2 \end{bmatrix}} \mathbb{Q}^3 \xleftarrow{\begin{bmatrix} 1 & -1 & 1 \\ -1 & 1 & -1 \end{bmatrix}} \mathbb{Q}^2 \xleftarrow{\begin{bmatrix} 0 & 0 \\ 1 & 1 \\ 1 & 1 \end{bmatrix}} \mathbb{Q}^2 \xleftarrow{0} 0.$$

Here we have $C^0 = 0$, $C^1 = \mathbb{Q}^2$, $C^2 = \mathbb{Q}^3$, $C^3 = \mathbb{Q}^2$, $C^4 = \mathbb{Q}^2$ and $C^5 = 0$. We have $d^5 = d^1 = 0$ and d^4, d^3 and d^2 are multiplication by

$$A_4 = \begin{bmatrix} 1 & 1 \\ 2 & 2 \end{bmatrix}, \quad A_3 = \begin{bmatrix} 1 & -1 & 1 \\ -1 & 1 & -1 \end{bmatrix}, \quad \text{and} \quad A_2 = \begin{bmatrix} 0 & 0 \\ 1 & 1 \\ 1 & 1 \end{bmatrix}.$$

We clearly have $d^5 \circ d^4 = 0$ and $d^2 \circ d^1 = 0$; let us check that $d^4 \circ d^3 = 0$ and $d^3 \circ d^2 = 0$:

$$\begin{aligned} \begin{bmatrix} 1 & 1 \\ 2 & 2 \end{bmatrix} \begin{bmatrix} 1 & -1 & 1 \\ -1 & 1 & -1 \end{bmatrix} &= \begin{bmatrix} 1-1 & -1+1 & 1-1 \\ -2+2 & 2-2 & -2+2 \end{bmatrix} \\ &= \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \end{aligned}$$

and

$$\begin{bmatrix} 1 & -1 & 1 \\ -1 & 1 & -1 \end{bmatrix} \begin{bmatrix} 0 & 0 \\ 1 & 1 \\ 1 & 1 \end{bmatrix} = \begin{bmatrix} 0-1+1 & 0-1+1 \\ 0+1-1 & 0+1-1 \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$$

hence we have a cochain complex. Let us now compute the cocycles and coboundaries. Starting with cocycles, we can immediately observe that $Z^4 = \ker(d^5) = C^4 = \mathbb{Q}^2$ and $Z^0 = \ker(d^1) = 0$. To find Z^3 , Z^2 and Z^1 we find the null spaces of the matrices by row-reduction to reduced echelon form:

$$\begin{bmatrix} 1 & 1 \\ 2 & 2 \end{bmatrix} \sim \begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix}$$

so $Z^3 = \mathbb{Q}[(-1, 1)] \cong \mathbb{Q}^1$,

$$\begin{bmatrix} 1 & -1 & 1 \\ -1 & 1 & -1 \end{bmatrix} \sim \begin{bmatrix} 1 & -1 & 1 \\ 0 & 0 & 0 \end{bmatrix}$$

so $Z^2 = \mathbb{Q}[(1, 1, 0), (0, 1, 1)] \cong \mathbb{Q}^2$, and

$$\begin{bmatrix} 0 & 0 \\ 1 & 1 \\ 1 & 1 \end{bmatrix} \sim \begin{bmatrix} 1 & 1 \\ 0 & 0 \\ 0 & 0 \end{bmatrix}$$

and $Z^1 = \mathbb{Q}[(-1, 1)] = \mathbb{Q}^1$. For the boundaries, we need to find the image of each linear transformation, or in terms of matrices, the column spaces of the matrices. The zero map has image $\text{Im}(0) = \{\vec{0}\}$, so $B^5 = B^1 = \{\vec{0}\}$. As we can see from the echelon forms of the matrices, each of the differential maps d^4 , d^3 and d^2 has one-dimensional column space $B^4 \cong B^3 \cong B^2 = \mathbb{Q}^1$. Finally, to identify H^k up to isomorphism, we can use the fact that

$$\dim(H^k) = \dim(Z^k/B^k) = \dim(Z^k) - \dim(B^k).$$

Thus, we have:

k	Z^k	B^k	H^k
4	\mathbb{Q}^2	\mathbb{Q}	\mathbb{Q}
3	\mathbb{Q}^1	\mathbb{Q}	0
2	\mathbb{Q}^2	\mathbb{Q}	\mathbb{Q}
1	\mathbb{Q}^1	0	\mathbb{Q}

Example 51. Not every sequence of vector spaces and linear maps defines a cochain complex; we need $d^{k+1}d^k$ to equal zero for every k . Consider the sequence

$$0 \xleftarrow{0} \mathbb{Q}^2 \xleftarrow{\begin{bmatrix} 1 & 2 \\ 0 & 1 \end{bmatrix}} \mathbb{Q}^2 \xleftarrow{\begin{bmatrix} 2 & 1 \\ -1 & 0 \end{bmatrix}} \mathbb{Q}^2 \xleftarrow{0} 0.$$

Here we have $C^4 = C^0 = 0$ and $C^3 = C^2 = C^1 = \mathbb{Q}^2$, and we have $d^2 \circ d^1 = 0$ and $d^4 \circ d^3 = 0$. However, $d^3 \circ d^2 \neq 0$ since

$$\begin{bmatrix} 1 & 2 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 2 & 1 \\ -1 & 0 \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} \neq \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}.$$

This means that $\text{Im}(d^2) \not\subset \text{Ker}(d^3)$, and we cannot talk about cohomology in this situation.

Now if we replace the field of rational numbers \mathbb{Q} by the ring of integers \mathbb{Z} we can get similar examples to the previous ones with modules in place of vector spaces.

Example 52. Consider the sequence

$$0 \xleftarrow{0} \mathbb{Z} \xleftarrow{\begin{bmatrix} 5 & 5 \end{bmatrix}} \mathbb{Z}^2 \xleftarrow{0} 0.$$

Here we have $C^0 = 0$, $C^1 = \mathbb{Z}^2$, $C^2 = \mathbb{Z}$ and $C^3 = 0$. We have $d^2(x, y) = 5x + 5y$ which says

$$\begin{aligned} \text{Ker}(d^2) &= \{(x, -x) \mid x \in \mathbb{Z}\} \cong \mathbb{Z} \quad \text{and} \\ \text{Im}(d^2) &= \{5x + 5y \mid x, y \in \mathbb{Z}\} \cong 5\mathbb{Z}. \end{aligned}$$

Then $H^1 = \mathbb{Z}$ and $H^2 = \mathbb{Z}/5\mathbb{Z} \cong \mathbb{Z}_5$.

Example 53. Let us see one more nontrivial example of cohomology. Consider the sequence

$$0 \xleftarrow{0} \mathbb{Z} \xleftarrow{0} \mathbb{Z}^2 \xleftarrow{\begin{bmatrix} 1 & 2 \\ 1 & 4 \end{bmatrix}} \mathbb{Z}^2 \xleftarrow{0} 0.$$

Here we have $C^0 = 0$, $C^1 = \mathbb{Z}^2$, $C^2 = \mathbb{Z}^2$, $C^3 = \mathbb{Z}$ and $C^4 = 0$. We have $d^1 = 0$, d^2 is left multiplication by $A = \begin{bmatrix} 1 & 2 \\ 1 & 4 \end{bmatrix}$, $d^3 = 0$ and $d^4 = 0$. We have $d^1 \circ d^2 = 0$, $d^2 \circ d^3 = 0$ and $d^3 \circ d^4 = 0$, so we have a cochain complex. Let us now determine the cocycles and

coboundaries. Starting with cocycles, we can immediately observe that $Z^3 = \mathbb{Z}$, $Z^2 = \mathbb{Z}^2$, and to find Z^1 we need to find the null space of the matrix A by row and column reduction, keeping in mind that we are working over \mathbb{Z} . Then we have

$$\begin{bmatrix} 1 & 2 \\ 1 & 4 \end{bmatrix} \sim \begin{bmatrix} 1 & 2 \\ 0 & 2 \end{bmatrix} \sim \begin{bmatrix} 1 & 0 \\ 0 & 2 \end{bmatrix}$$

which is in Smith normal form. Then the kernel Z^1 is isomorphic to $\mathbb{Z}_1 \oplus \mathbb{Z}_2 = \mathbb{Z}_2$. Turning to coboundaries, we have $B^1 = 0$, $B^2 = \mathbb{Z}^2$ and $B^3 = 0$, so we have cohomology groups $H^1 = Z^1/B^1 = \mathbb{Z}_2$, $H^2 = \mathbb{Z}^2/\mathbb{Z}^2 = 0$, and $H^3 = \mathbb{Z}/0 = \mathbb{Z}$.

Exercises. 1. Compute the cohomology spaces of the cochain complex

$$0 \xleftarrow{0} \mathbb{R}^2 \xleftarrow{\begin{bmatrix} 1 & 1 & 0 \\ -1 & -1 & 0 \end{bmatrix}} \mathbb{R}^3 \xleftarrow{\begin{bmatrix} 1 & 1 \\ -1 & -1 \\ -1 & 1 \end{bmatrix}} \mathbb{R}^2 \xleftarrow{0} 0.$$

2. Compute the cohomology spaces of the cochain complex

$$0 \xleftarrow{0} \mathbb{Q}^3 \xleftarrow{\begin{bmatrix} 1 & 2 & 1 \\ -1 & -1 & 1 \\ 1 & 0 & 2 \end{bmatrix}} \mathbb{Q}^3 \xleftarrow{0} \mathbb{Q}^3 \xleftarrow{\begin{bmatrix} 1 & 1 \\ 2 & 1 \\ 1 & 0 \end{bmatrix}} \mathbb{Q}^2 \xleftarrow{0} 0.$$

3. Consider the sequence of vector spaces and maps

$$0 \xleftarrow{0} \mathbb{Q}^2 \xleftarrow{\begin{bmatrix} a & b \\ -1 & -1 \end{bmatrix}} \mathbb{Q}^2 \xleftarrow{\begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}} \mathbb{Q}^2 \xleftarrow{0} 0.$$

What values of a and b will make this a cochain complex?

4. Consider a sequence

$$0 \xleftarrow{0} \mathbb{Q}^2 \xleftarrow{f} \mathbb{Q}^2 \xleftarrow{0} 0$$

What cohomology spaces H^1 and H^2 are possible? Give an example of a matrix f realizing each case.

5. Let C^0 be the set of 3-times differentiable functions on \mathbb{R}^3 . A *differential 1-form* is an expression of the form $f(\vec{x})dx_j$ where $j = 1, 2$ or 3 and $\vec{x} \in \mathbb{R}^3$. We define a product \wedge on 1-forms (called the *wedge product*) satisfying the rules that

- $f \wedge (gdx_j) = fgdx_j$,
- $f dx_j \wedge (gdx_k + hdx_l) = f dx_j \wedge gdx_k + f dx_j \wedge hdx_l$,
- \wedge is *anticommutative*, i.e.,

$$dx_j \wedge dx_k = -dx_k \wedge dx_j,$$

and

- \wedge is *square-free*, i.e.,

$$dx_j \wedge dx_j = 0.$$

Let

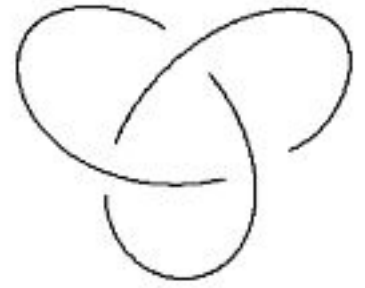
$$\begin{aligned} C^1 &= \{f dx_1 + g dx_2 + h dx_3\}, \\ C^2 &= \{f dx_1 \wedge dx_2 + g dx_1 \wedge dx_3 + h dx_2 \wedge dx_3\}, \\ C^3 &= \{f dx_1 \wedge dx_2 \wedge dx_3\}. \end{aligned}$$

Then define a map $d : C^k \rightarrow C^{k+1}$ by

$$d(f\omega) = \left(\frac{\partial f}{\partial x_1} dx_1 + \frac{\partial f}{\partial x_2} dx_2 + \frac{\partial f}{\partial x_3} dx_3 \right) \wedge \omega.$$

Show that d is a differential. (This d is called *exterior differentiation*, and the resulting cohomology spaces are called *de Rham cohomology*.)

Chapter 3



Quandles

1. Kei

A very natural question is how we can generalize the tricoloring idea from Chapter 1 to get stronger invariants with more colors. It turns out that hidden in the simplicity of the tricoloring rules is a new kind of algebra, a powerful algebraic structure which ultimately gives us a complete invariant of knots. Let us start at the beginning.

Let X be a set. An operation \triangleright which takes two elements $x, y \in X$ and gives us back an element $x \triangleright y \in X$ is a *Kei operation* if it satisfies the following three axioms:

- (i) For all $x \in X$, $x \triangleright x = x$.
- (ii) For all $x, y \in X$, $(x \triangleright y) \triangleright y = x$.
- (iii) For all $x, y, z \in X$, $(x \triangleright y) \triangleright z = (x \triangleright z) \triangleright (y \triangleright z)$.

These axioms are quite unlike the usual rules obeyed by more familiar operations like addition and multiplication. Let's unravel them one by one.

The first kei axiom says $x \triangleright x = x$ for every $x \in X$. The property is known as *idempotency* in standard mathematical jargon; for example, a matrix A is idempotent if $A^2 = A$, e.g. a projection map onto a

coordinate axis. If \triangleright were like addition, this axiom would mean that every element acted like 0.

The second axiom says $(x \triangleright y) \triangleright y = x$ for all $x, y \in X$. This says that if we triangle x with y twice, we get x back. Thus, where the first kei axiom says elements act trivially on themselves, the second axiom says that elements act on other elements by *involutions*, i.e. the function $\beta_y : X \rightarrow X$ defined by $\beta_y(x) = x \triangleright y$ is its own inverse. If addition were involutory, addition and subtraction would be the same.

The third kei axiom is perhaps the strangest of all. The requirement that $(x \triangleright y) \triangleright z = (x \triangleright z) \triangleright (y \triangleright z)$ says that the \triangleright operation is *self-distributive*, i.e. \triangleright distributes over \triangleright on the right the same way that multiplication distributes over addition. In particular, the kei operation is in general *nonassociative*, i.e.,

$$(x \triangleright y) \triangleright z \neq x \triangleright (y \triangleright z).$$

Thus, it is very important to keep track of the order of the elements as well as the parentheses when doing kei computation.

Where do these bizarre axioms come from, and what is their connection to knots and links? You might have noticed that there are three axioms, one making a statement about a single element, one making a statement about two elements, and one making a statement about three elements. You might also recall that there are three Reidemeister moves, one involving a single strand, one involving two strands, and one involving three strands. If you noticed both of these things, you probably suspect that the similarities are not a coincidence. If so, you're correct!

The idea is that each “color” or element of X corresponds to an arc in a diagram and the $x \triangleright y$ operation corresponds to one arc x passing under another arc y to become $x \triangleright y$. Notice that unlike in addition or multiplication, the two operands here are playing different roles – when x crosses under y , y is unchanged but $x \triangleright y$ is a new arc; y is doing something to x , not the other way around. The kei operation can be understood as an *action* of the set X on itself. Thus, we don't expect \triangleright to be commutative, and in general it's not.

Kei axiom (i) follows from the type I Reidemeister move:

Kei axiom (ii) follows from the type II Reidemeister move:

Kei axiom (iii) follows from the type III Reidemeister move:

The term “kei” was chosen by Mituhisa Takasaki [Tak42].

Example 54. Perhaps the simplest nontrivial example of a kei operation is known as a *Takasaki kei*, also sometimes called a *cyclic kei* or *dihedral quandle*. Let $X = \mathbb{Z}$ or \mathbb{Z}_n and define

$$x \triangleright y = 2y - x.$$

To see that this \triangleright is a kei operation, we just need to verify that all three kei axioms are satisfied:

(i)

$$x \triangleright x = 2x - x = x \quad \checkmark$$

(ii)

$$(x \triangleright y) \triangleright y = 2y - (x \triangleright y) = 2y - (2y - x) = 2y - 2y + x = x \quad \checkmark$$

(iii)

$$(x \triangleright y) \triangleright z = 2z - (x \triangleright y) = 2z - (2y - x) = 2z - 2y + x$$

while

$$\begin{aligned} (x \triangleright z) \triangleright (y \triangleright z) &= 2(y \triangleright z) - (x \triangleright z) = 2(2z - y) - (2z - x) \\ &= 4z - 2y - 2z + x = 2z - 2y + x. \quad \checkmark \end{aligned}$$

Note that this example can be generalized by replacing \mathbb{Z} or \mathbb{Z}_n by any abelian group A .

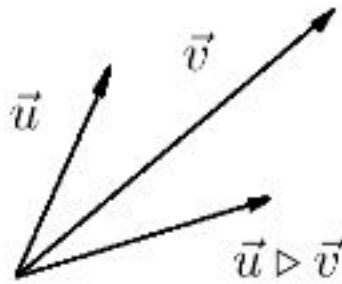
Example 55. Let V be an \mathbb{F} -vector space and $\langle \cdot, \cdot \rangle : V \times V \rightarrow \mathbb{F}$ a *symmetric bilinear form*, i.e.,

- $\langle \vec{u} + \vec{v}, \vec{w} \rangle = \langle \vec{u}, \vec{w} \rangle + \langle \vec{v}, \vec{w} \rangle$,
- $\langle \alpha \vec{u}, \vec{v} \rangle = \alpha \langle \vec{u}, \vec{v} \rangle$, and
- $\langle \vec{u}, \vec{v} \rangle = \langle \vec{v}, \vec{u} \rangle$.

Let X be the subset of V consisting of vectors \vec{u} such that $\langle \vec{u}, \vec{u} \rangle \neq 0$. Then the operation

$$\vec{u} \triangleright \vec{v} = \frac{2\langle \vec{u}, \vec{v} \rangle}{\langle \vec{u}, \vec{u} \rangle} \vec{v} - \vec{u}$$

defines a kei structure on X . Geometrically, $\vec{u} \triangleright \vec{v}$ is the result of reflecting \vec{u} across \vec{v} .

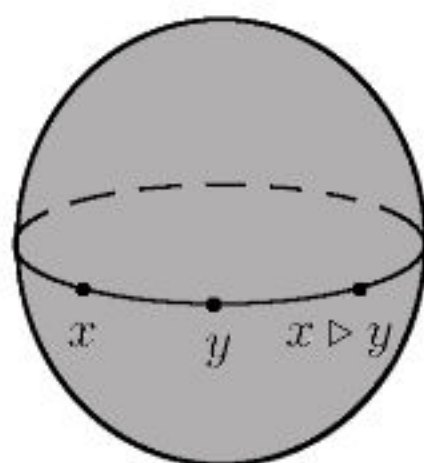


This type of kei is called a *Coxeter kei*.

Example 56. A related example is a *symmetric space*, a geometric space in which every point has an involutory *point symmetry*. For example, consider the 2-sphere

$$S^2 = \{\vec{x} \in \mathbb{R}^3 \mid \|\vec{x}\| = 1\}.$$

Connecting any two points on S^2 is a *geodesic* or path of least distance, given by an arc of the great circle containing those two points; this geodesic is unique unless the two points are *antipodes* x and $-x$, in which case there are infinitely many geodesics connecting x and $-x$. Then for any two points x and y on S^2 , define $x \triangleright y$ as the result of finding the geodesic connecting x to y and then going from x to y and then past y along the geodesic by the same distance.



If we think of the geodesic as the “straight line” connecting x to y , then $x \triangleright y$ is the point on the other side of y the same distance from y as x . In terms of unit vectors, we have

$$\vec{x} \triangleright \vec{y} = 2(\vec{x} \cdot \vec{y})\vec{y} - \vec{x}.$$

To understand a kei operation, it is helpful to look at the operation table. To find $x \triangleright y$ in the operation table, look in the row labeled with x and the column labeled with y ; since kei operations are generally noncommutative, this is usually a different element than the entry in row y column x , so it is important to pay attention to the order.

Example 57. For example, if we take $X = \mathbb{Z}_4$ with $x \triangleright y = 2y - x$, we get the operation table

\triangleright	0	1	2	3
0	0	2	0	2
1	3	1	3	1
2	2	0	2	0
3	1	3	1	3

For computational purposes, we can represent a kei operation on a set $X = \{x_1, \dots, x_n\}$ with n elements with an $n \times n$ matrix M_X which encodes the operation table by dropping the “ x ”s:

Example 58. Dropping the x s from the operation table yields the following matrix:

$$\begin{array}{c|ccc} \triangleright & x_1 & x_2 & x_3 \\ \hline x_1 & x_1 & x_3 & x_2 \\ x_2 & x_3 & x_2 & x_1 \\ x_3 & x_2 & x_1 & x_3 \end{array} \rightarrow \begin{bmatrix} 1 & 3 & 2 \\ 3 & 2 & 1 \\ 2 & 1 & 3 \end{bmatrix} = M_X.$$

This matrix notation allows us to compute colorings with a kei for which we may not have nice algebraic formulas.

The Fundamental Kei of a Knot. For every knot, link, or tangle K there is an associated kei (i.e., a set X with a kei operation \triangleright) called the *fundamental kei* of the knot, $\mathcal{K}(K)$, which we can define from a diagram of K via universal algebra.

To start, let $X = \{x_1, \dots, x_n\}$ be a set. The elements of X will be called *generators*. The set $W_{\mathcal{K}}(K)$ of *kei words* in X is defined recursively by the rules that

- (i) $x \in X$ implies $x \in W_{\mathcal{K}}(X)$ and
- (ii) $x, y \in W_{\mathcal{K}}X$ implies $x \triangleright y \in W_{\mathcal{K}}(X)$.

Thus, a kei word in X is a finitely long string of elements of X and the symbol \triangleright and parentheses which makes sense as a kei product. For example, if $X = \{x, y, z\}$, then $W_{\mathcal{K}}(X)$ includes such expressions as

$$x \triangleright y, \quad z \triangleright ((x \triangleright x) \triangleright y), \quad ((x \triangleright y) \triangleright (y \triangleright x)) \triangleright z,$$

etc.

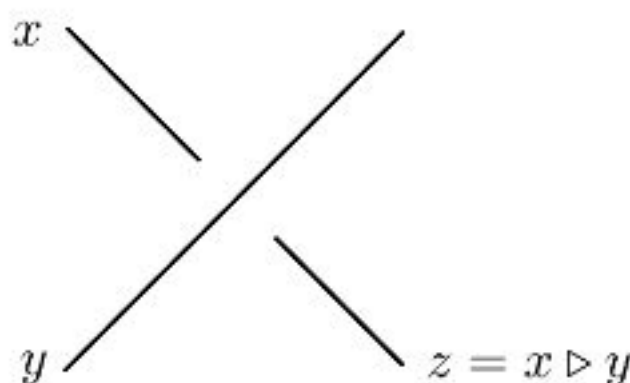
The *free kei* on X is then the set of equivalence classes of kei words in X modulo the equivalence relation generated by

$$x \triangleright x \sim x, \quad (x \triangleright y) \triangleright y \sim x, \quad (x \triangleright y) \triangleright z \sim (x \triangleright z) \triangleright (y \triangleright z)$$

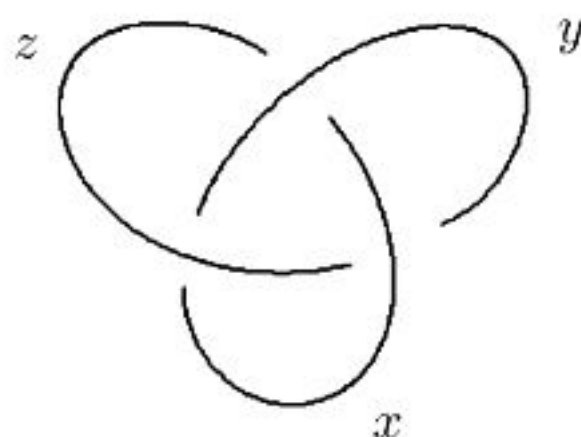
for all $x, y, z \in W_{\mathcal{K}}(X)$.

Now, let $X = \{x_1, \dots, x_n\}$ be a set with one element for each arc in a diagram of a knot, link or tangle K . Each crossing in our diagram K gives us an equation, called a *crossing relation*, of the

form $x \triangleright y = z$.



The *fundamental kei* of the knot K is then the set of equivalence classes of elements of the free kei on X modulo the equivalence relation generated by the crossing relations. We usually express this with a *kei presentation* listing the elements of X , known as *generators* and the crossing relations. For example, the trefoil knot has fundamental kei presentation:



$$\mathcal{K}(K) = \langle x, y, z \mid x \triangleright y = z, y \triangleright z = x, z \triangleright x = y \rangle.$$

Note that, in general, most elements of the fundamental kei of a knot K do not correspond to arcs in any given diagram of K . It is also important to notice that different diagrams of K will give us different-looking presentations of $\mathcal{K}(K)$. Do not be fooled; these different presentations nevertheless describe the same set of equivalence classes. We can change one presentation into another by a sequence of *Tietze moves*:

- (i) Add or delete a generator x and a relation of the form $x = W$ where W is a word not involving x .
- (ii) Add or delete a relation which is a consequence of the other relations and the kei axioms.

For example, in the presentation

$$\mathcal{K}(K) = \langle x, y, z \mid x \triangleright y = z, y \triangleright z = x, z \triangleright x = y \rangle$$

we could interpret the relation $x \triangleright y = z$ as saying that z is an abbreviation for $x \triangleright y$; we can thus replace every instance of z with $x \triangleright y$ and obtain a simpler presentation

$$\mathcal{K}(K) = \langle x, y \mid y \triangleright (x \triangleright y) = x, (x \triangleright y) \triangleright x = y \rangle.$$

Indeed, each Reidemeister move determines a Tietze move (or a set of Tietze moves) on the fundamental kei. Unfortunately, the converse is not true in general; most Tietze moves cannot be interpreted as Reidemeister moves on diagrams.

The generators in a kei presentation can be understood as analogous to basis vectors in a vector space – every vector can be expressed as a linear combination of the basis vectors, and every element of a kei can be expressed as a “kei combination” of the generators. The kei axiom relations and crossing relations make the situation more closely analogous to quotient vector spaces, in that elements are equivalence classes.

Indeed, we can represent the fundamental kei of a knot with a *presentation matrix*, a kind of partially-filled-in operation table with a row and column for each generator. When we have a relation $x \triangleright y = z$, we put a z in row x column y ; otherwise, the entries are blank, represented by 0. Then the presentation

$$\mathcal{K}(K) = \langle x, y, z \mid x \triangleright y = z, y \triangleright z = x, z \triangleright x = y \rangle$$

can be expressed with the table (or corresponding matrix with $x = x_1, y = x_2, z = x_3$)

$$\begin{array}{c|ccc} \triangleright & x & y & z \\ \hline x & 0 & z & 0 \\ y & 0 & 0 & x \\ z & y & 0 & 0 \end{array} \quad \begin{bmatrix} 0 & 3 & 0 \\ 0 & 0 & 1 \\ 2 & 0 & 0 \end{bmatrix}.$$

Note that in this case we can fill in the zeroes using the kei axioms: since $(x \triangleright y) \triangleright y = x$, we also have $z \triangleright y = (x \triangleright y) \triangleright y = x$, $x \triangleright z = (y \triangleright z) \triangleright z = y$ and $y \triangleright x = (z \triangleright x) \triangleright x = y$, and we also have $x \triangleright x = x$, $y \triangleright y = y$ and $z \triangleright z = z$. Each of these new relations is a consequence

of the kei axioms and the relations from K . Then we have

$$\begin{array}{c|ccc} \triangleright & x & y & z \\ \hline x & x & z & y \\ y & z & y & x \\ z & y & x & z \end{array} \quad \begin{bmatrix} 1 & 3 & 2 \\ 3 & 2 & 1 \\ 2 & 1 & 3 \end{bmatrix}.$$

In particular, this shows that the fundamental kei of the trefoil is none other than the 3-element Takasaki kei, also known as the dihedral quandle of three elements or the Fox tricoloring quandle.

Not all knots have finite fundamental kei, but the set of knots with finite fundamental kei is infinite. See [Win84] for more.

Homomorphisms and Colorings. Let X and Y be keis. A *kei homomorphism* is a function $f : X \rightarrow Y$ satisfying

$$f(x \triangleright y) = f(x) \triangleright f(y)$$

for all $x, y \in X$. Notice that the \triangleright in $x \triangleright y$ is the kei operation in X , while the \triangleright in $f(x) \triangleright f(y)$ is the kei operation in Y . Thus, a kei homomorphism is a function between keis which preserves or respects the kei structure. Kei homomorphisms are analogous to linear transformations, which are functions between vector spaces which preserve the vector space structure.

Example 59. For example, let $X = \mathbb{Z}$ with $x \triangleright y = 2y - x$. Then $f : X \rightarrow X$, defined by $f(x) = lx$ where $l \in \mathbb{Z}$, is a kei homomorphism since

$$f(x \triangleright y) = f(2x - y) = l(2x - y) = 2(lx) - ly = f(x) \triangleright f(y).$$

Similarly, $f(x) = x + l$ is a homomorphism:

$$f(x \triangleright y) = 2y - x + l = 2(y + l) - (x + l) = f(x) \triangleright f(y).$$

However, $f(x) = x^2$ is not a kei homomorphism:

$$f(x \triangleright y) = (2y - x)^2 = 4y^2 - 4xy + x^2$$

while

$$f(x) \triangleright f(y) = 2y^2 - x^2.$$

Definition 13. Let (X, \triangleright) be a kei. A subset $S \subset X$ is a *subkei* of X if (S, \triangleright) is itself a kei. In particular, to be a subkei, S must be *closed under* \triangleright : if $x, y \in S$, then we need $x \triangleright y \in S$.

Note that since the kei axioms are satisfied in X , they are automatically satisfied in S , so closure under \triangleright is necessary and sufficient for $S \subset X$ to be a subkei.

If $f : X \rightarrow Y$ is a kei homomorphism, then the subset

$$\text{Im}(f) = \{y \in Y \mid y = f(x) \text{ for some } x \in X\}$$

is a subkei, called the *image subkei* of f . To see that $\text{Im}(f)$ is closed under \triangleright , note that if $y, y' \in \text{Im}(f)$ then there exist $x, x' \in X$ such that $y = f(x)$ and $y' = f(x')$; then

$$y \triangleright y' = f(x) \triangleright f(x') = f(x \triangleright x')$$

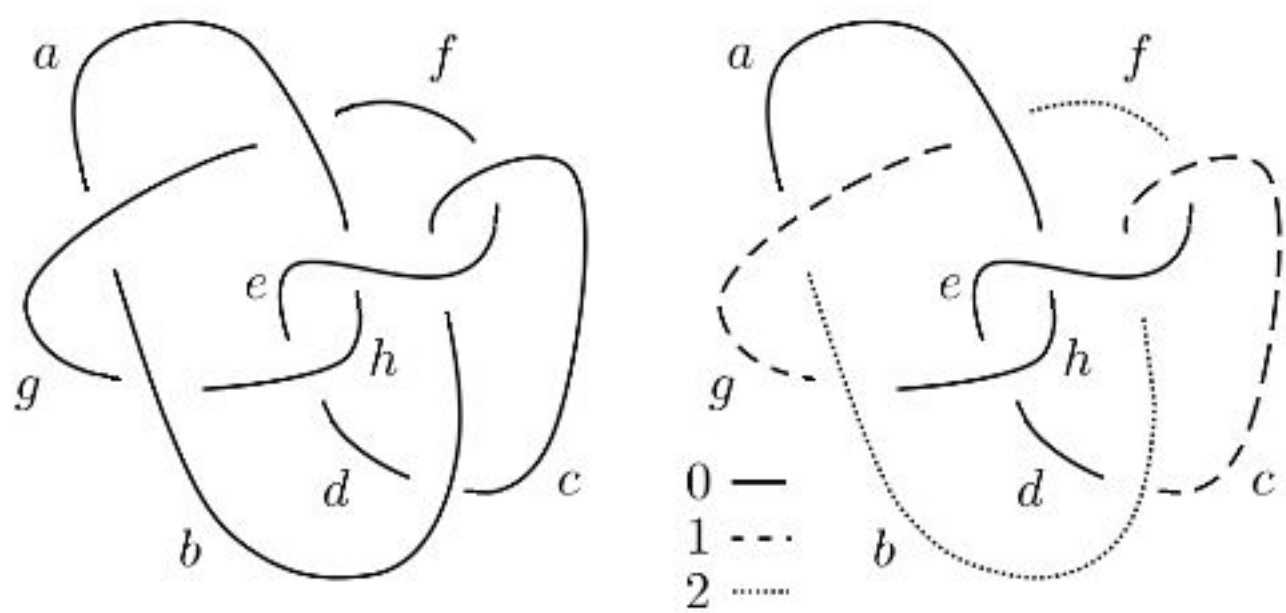
and we have $y \triangleright y' \in \text{Im}(f)$.

Now, let K be a knot, link or tangle, (X, \triangleright) be a kei and consider a map $f : \mathcal{K}(K) \rightarrow X$. Since $\mathcal{K}(K)$ is generated by arcs in a diagram of K , if we choose an image $f(x_k)$ for every arc x_k , then the homomorphism condition determines the function f for the entire fundamental kei; that is, once we know what $f(x_k)$ and $f(x_j)$ are, we simply define $f(x_k \triangleright x_j)$ to be $f(x_k) \triangleright f(x_j)$ and so forth. The only potential problem arises at the crossings – if the fundamental kei has a relation $x \triangleright y = z$, then we must be careful to choose $f(x)$, $f(y)$ and $f(z)$ so that $f(x) \triangleright f(y)$ equals $f(z)$. Provided our choice of images $f(x_k)$ respect the crossing relations at every crossing, then we have a homomorphism $f : \mathcal{K}(K) \rightarrow X$.

Tricolorings of a diagram K are actually homomorphisms from the fundamental kei $\mathcal{K}(K)$ to the Takasaki kei \mathbb{Z}_3 . Thus, Takasaki keis give us a generalization of tricoloring – say a valid n -coloring of a knot diagram is a homomorphism from $\mathcal{K}(K)$ to the Takasaki kei \mathbb{Z}_n . Then as with tricoloring, any n -colored diagram before a move corresponds to a unique n -colored diagram after the move. We might say a diagram is n -colorable if it has a coloring which used all n colors; in terms of homomorphisms, such a coloring is a *surjective* or *onto* function.

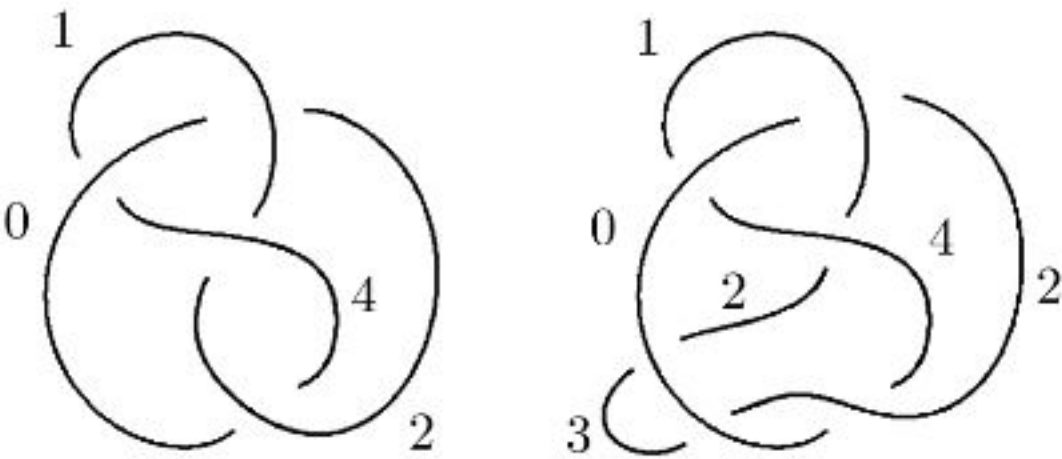
Example 60. Let us illustrate how Fox tricolorability is about the existence of surjective colorings by the Takasaki kei \mathbb{Z}_3 . More precisely, a Fox tricoloring of the knot 8_{19} gives a nontrivial kei homomorphism from the fundamental quandle of 8_{19} to the kei \mathbb{Z}_3 with operation

$x \triangleright y = 2y - x.$



A surjective coloring may not be obviously surjective in that it may not use every color in every diagram of K if $n > 3$. To determine whether a coloring is surjective, we must check whether $\text{Im}(f)$ equals X or is a proper subkei.

The image subkei of f is the smallest subkei of X which contains $f(x_k)$ for all generators x_k . For example, the figure 8 knot diagram below depicts a surjective coloring by the Takasaki kei \mathbb{Z}_5 , even though only four colors actually appear in the diagram. Notice that the fourth color appears when we change the diagram by a type II move.



The Counting Invariant. The existence or nonexistence of a surjective n -coloring is a computable invariant, but a rather coarse one in that it only has two possible values, “colorable” or “not colorable”. We would like to find a stronger, more sensitive invariant of knots and links that can be computed using kei colorings. One solution is to count the number of colorings of any diagram of K by a kei X ; since the fundamental kei $\mathcal{K}(K)$ does not depend on our choice of

diagram of K , the set of all kei homomorphisms

$$\text{Hom}(\mathcal{K}(K), X) = \{f : \mathcal{K}(K) \rightarrow X \mid f(x \triangleright y) = f(x) \triangleright f(y)\}$$

is an invariant of the knot K . In particular, for any choice of diagram of K , there is one coloring for each homomorphism: on the one hand, a homomorphism f assigns only one image $f(x)$ to each generator x ; on the other hand, the fact that f is a homomorphism says that if we know the images of the generators $f(x_1), \dots, f(x_n)$ then we know the images of every element of the fundamental kei, e.g.

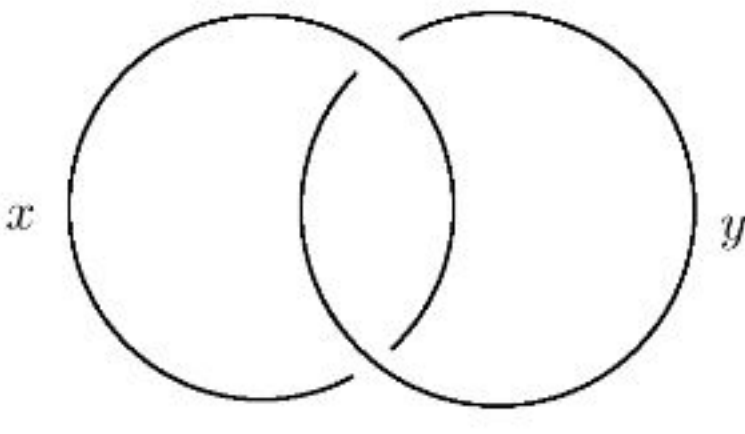
$$f(x_1 \triangleright (x_2 \triangleright x_4)) = f(x_1) \triangleright (f(x_2) \triangleright f(x_4)),$$

etc.

Moreover, if X is a finite set, then there are only finitely many possible colorings of a given diagram of K – if K has n crossings and thus n arcs and if X has m elements, then there are at most m^n possible kei colorings of K by X . Thus, to compute the number of colorings by brute force, we can simply list all assignments of kei elements to each arc of K and check which ones satisfy all of the crossing conditions.

Thus, the cardinality of the set $\text{Hom}(\mathcal{K}(K), X)$ is a computable link invariant known as the *kei counting invariant*.

Example 61. Let us compute the kei counting invariant for the Hopf link with respect to the four element Takasaki kei \mathbb{Z}_4 . The crossing relations R_1 and R_2 are $x \triangleright y = x$ and $y \triangleright x = y$. Then we have



\triangleright	0	1	2	3
0	0	2	0	2
1	3	1	3	1
2	2	0	2	0
3	1	3	1	3

$f(x)$	$f(y)$	$R_1?$	$R_2?$	$f(x)$	$f(y)$	$R_1?$	$R_2?$
0	0	✓	✓	2	0	✓	✓
0	1			2	1		
0	2	✓	✓	2	2	✓	✓
0	3			2	3		
1	0			3	0		
1	1	✓	✓	3	1	✓	✓
1	2			3	2		
1	3	✓	✓	3	3	✓	✓

Thus, we have $|\text{Hom}(\mathcal{K}(K), \mathbb{Z}_4)| = 8$.

Example 62. The kei counting invariant for two unlinked circles is $|\text{Hom}(\mathcal{K}(K), \mathbb{Z}_4)| = 16$ since there are no crossing relations and hence no conditions imposed on $f(x)$ and $f(y)$.

An improved method for computing a kei counting invariant is to first reduce the presentation of $\mathcal{K}(K)$ to make the table as small as possible. Computing the colorings of the figure eight knot 4_1 by \mathbb{Z}_4 with the presentation coming directly from the diagram requires a table with $4^4 = 256$ lines, with four crossing relations to check at each line; reducing beforehand gives us a two-generator presentation with two relations, resulting in a table with only $4^2 = 16$ lines and only two (admittedly longer) relations to check on each line.



$$\begin{aligned}
\mathcal{K}(K) &= \langle x, y, z, w \mid x \triangleright y = z, y \triangleright w = z, y \triangleright x = w, x \triangleright z = w \rangle \\
&= \langle x, y, z \mid x \triangleright y = z, y \triangleright (y \triangleright x) = z, x \triangleright z = y \triangleright x \rangle \\
&= \langle x, y \mid y \triangleright (y \triangleright x) = x \triangleright y, x \triangleright (x \triangleright y) = y \triangleright x \rangle
\end{aligned}$$

Let R_1 be the equation $f(y) \triangleright (f(y) \triangleright f(x)) = f(x) \triangleright f(y)$ and R_2 the equation $f(x) \triangleright (f(x) \triangleright f(y)) = f(y) \triangleright f(x)$. Then we have the table

$f(x)$	$f(y)$	$R_1?$	$R_2?$	$f(x)$	$f(y)$	$R_1?$	$R_2?$
0	0	✓	✓	2	0		
0	1			2	1		
0	2			2	2	✓	✓
0	3			2	3		
1	0			3	0		
1	1	✓	✓	3	1		
1	2			3	2		
1	3			3	3	✓	✓

Exercises. 1. Show that if \triangleright is a kei operation which is associative, i.e. if \triangleright also satisfies

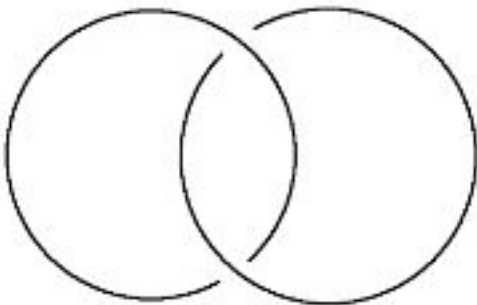
$$(x \triangleright y) \triangleright z = x \triangleright (y \triangleright z),$$

then the operation is *trivial*, i.e. $x \triangleright y = x$ for all x .

2. Compute the operation tables for the Takasaki kei operations on \mathbb{Z}_5 , \mathbb{Z}_6 , and \mathbb{Z}_7 .
3. Find a presentation for the fundamental kei of the knot below with as few generators as possible.



4. Prove that the Hopf link below has fundamental kei isomorphic to the trivial kei of two elements.



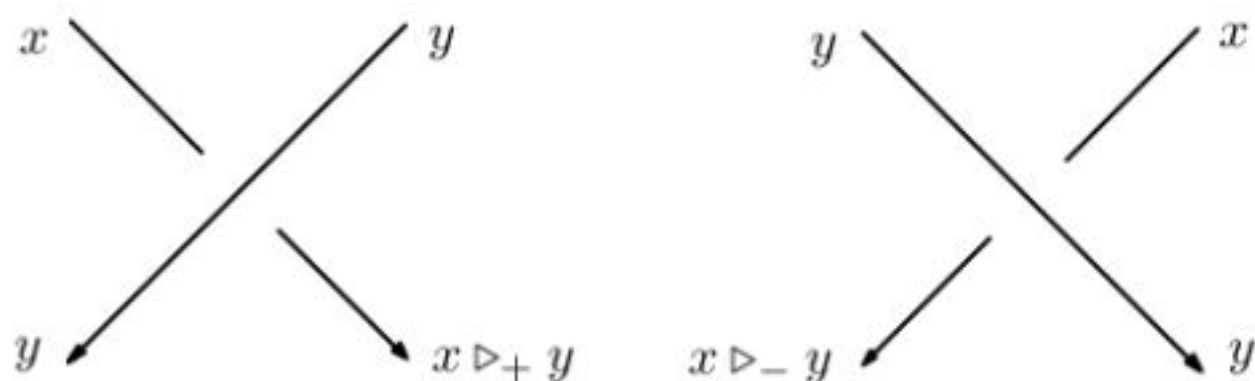
5. Show that every valid tricoloring determines a unique valid homomorphism from $\mathcal{K}(K)$ to the Takasaki kei \mathbb{Z}_3 .
6. Compute the kei counting invariant for the knot in problem 3 with respect to the Takasaki keis \mathbb{Z}_3 and \mathbb{Z}_4 . For each coloring, identify the image subkei.
7. Verify that the Coxeter kei definition in example 55 satisfies the kei axioms.
8. From the diagram of 8_{19} in example 60, find a presentation of fundamental quandle using generators a, b, c, d, e, f, g and h and eight relations at the crossings. Show that the map $\phi : \mathcal{K}(8_{19}) \rightarrow \mathbb{Z}_3$ with $\phi(a) = \phi(d) = \phi(e) = \phi(h) = 0$, $\phi(c) = \phi(g) = 1$ and $\phi(b) = \phi(f) = 2$ is a kei homomorphism by checking that the relations at crossings still hold after mapping by ϕ .

2. Quandles

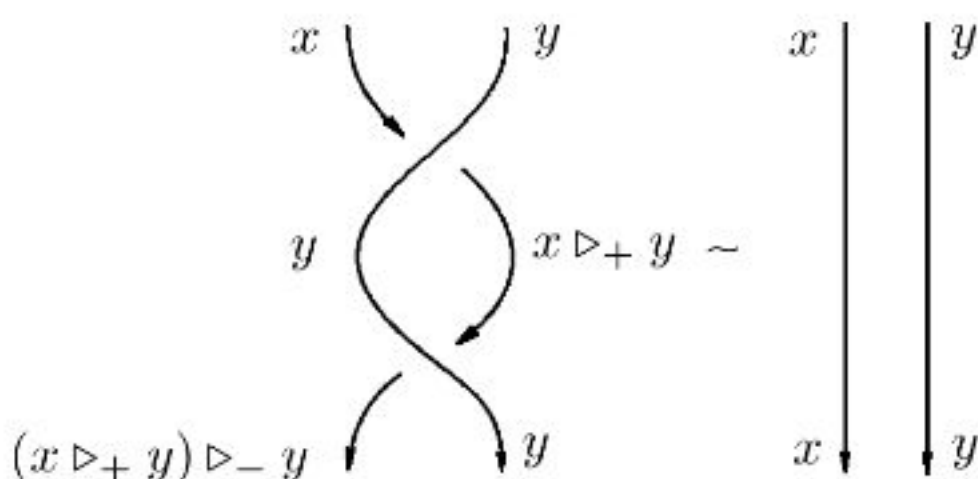
In the last section we introduced *kei*, an algebraic structure whose laws or axioms encode the Reidemeister moves for *unoriented* knots, i.e., for simple closed curves in \mathbb{R}^3 . As we recall from multivariable calculus, one common way to describe a curve in \mathbb{R}^3 is by giving a parametrization $P(t) = (x(t), y(t), z(t))$ where we can think of t as a time variable. In particular, a parametrized curve comes with a choice of direction or *orientation* corresponding to the forward direction of time. Let us think about how including a choice of orientation for our knots can change the algebraic structure.

As we have seen, there are two types of oriented crossing, usually called positive (+1) and negative (−1). Thus, instead of one “crossing under” operation, we have two crossing under operations; we can temporarily set $x \triangleright_+ y$ to be the result of x crossing under y at a positive crossing and $x \triangleright_- y$ to be the result of x crossing under y at

a negative crossing.



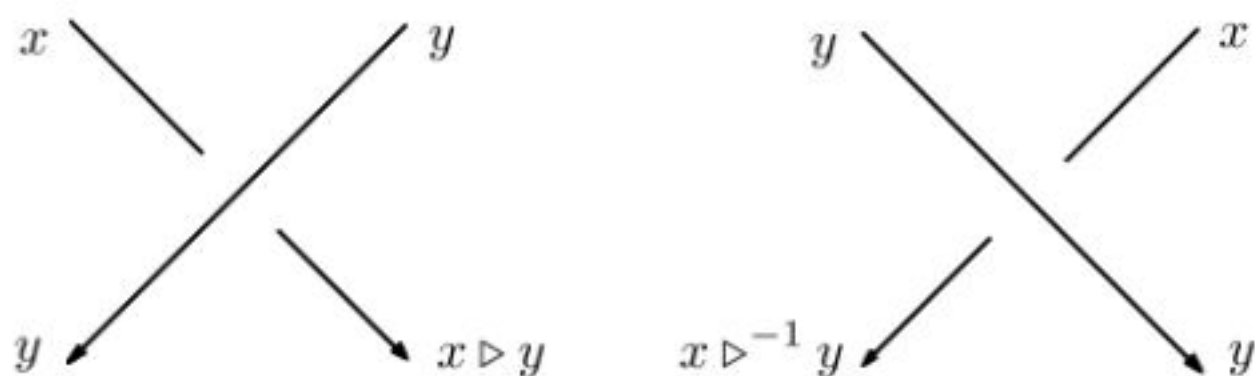
The second Reidemeister move then says that \triangleright_+ and \triangleright_- are inverse operations, like addition and subtraction or multiplication and division.



More precisely, we have

$$(x \triangleright_+ y) \triangleright_- y = x = (x \triangleright_- y) \triangleright_+ y.$$

We usually drop the $+$ and write $x \triangleright y$ for $x \triangleright_+ y$ and write $x \triangleright^{-1} y$ for $x \triangleright_- y$.



We can also think about this from an algebraic point of view. The second kei axiom says

$$(x \triangleright y) \triangleright y = x.$$

For each fixed element y in X , let us define a function $\beta_y : X \rightarrow X$ by setting $\beta_y(x) = x \triangleright y$. Then the second kei axiom says $\beta_y(\beta_y(x)) = x$. That is, the function β_y is its own inverse function, $\beta_y^{-1} = \beta_y$. A

function which is its own inverse (like flipping a light switch, or like $f(x) = x + 2$ in \mathbb{Z}_4) is called an *involution*.

A function need not be an involution to be invertible; for instance, the function $f(x) = x + 2$ has inverse function $f^{-1}(x) = x - 2 \neq f(x)$, and similarly the function $f(x) = \sqrt{x - 3}$ has inverse function $f^{-1}(x) = x^2 + 3 \neq f(x)$. Thus, in generalizing from unoriented knots to oriented, we are replacing the involutions β_y with merely invertible functions β_y with $\beta_y \neq \beta_y^{-1}$.

It would seem natural to call the new structure *oriented kei*, but for historical reasons the standard name is *quandle*. Thus, we can state our new definition:

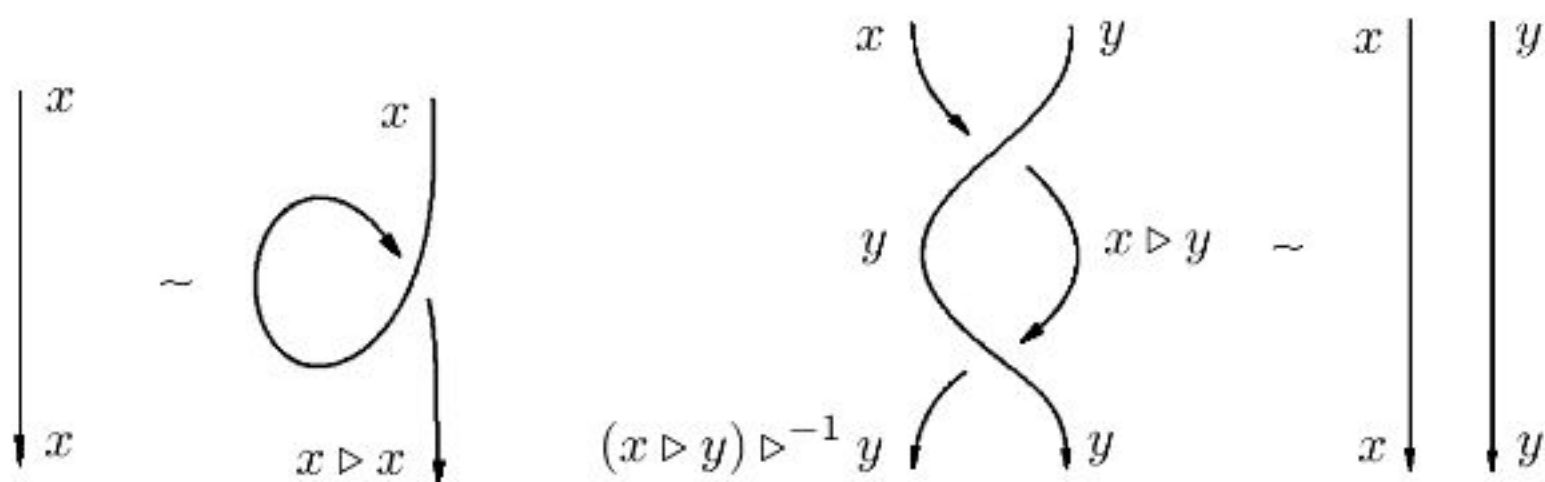
Definition 14. A *quandle* is a set X with a binary operation $\triangleright : X \times X \rightarrow X$ satisfying:

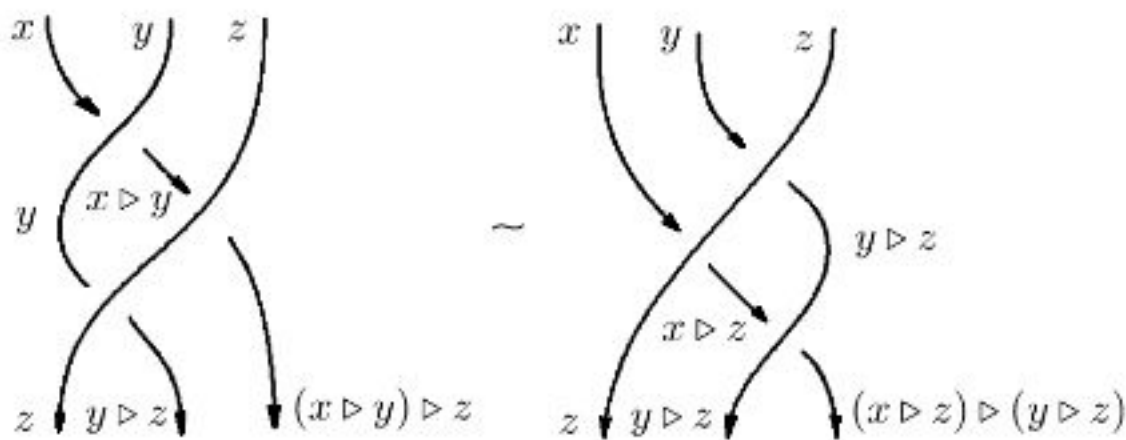
- (i) For all $x \in X$, $x \triangleright x = x$.
- (ii) For all $y \in X$, the map $\beta_y : X \rightarrow X$ defined by $\beta_y(x) = x \triangleright y$ is invertible.
- (iii) For all $x, y, z \in X$, $(x \triangleright y) \triangleright z = (x \triangleright z) \triangleright (y \triangleright z)$.

We write $x \triangleright^{-1} y$ for $\beta_y^{-1}(x)$.

With this definition, we can see that kei are a type of quandle, namely quandles for which the maps β_y are involutions. For this reason, kei are often called *involutory quandles*.

As with kei, we can understand the quandle axioms in terms of knot diagrams:





Example 63. Any set X with the operation $x \triangleright y = x$ for all $x, y \in X$ is a quandle, called a *trivial* quandle. We will use the notation T_n to denote a trivial quandle with n elements.

Example 64. Let \mathbb{F} be a field, e.g. $\mathbb{Q}, \mathbb{R}, \mathbb{C}$ or \mathbb{Z}_p for p prime. Recall that the set of invertible $n \times n$ matrices with entries in \mathbb{F} is denoted $GL_n(\mathbb{F})$. Then we can check that $GL_n(\mathbb{F})$ is a quandle with quandle operation

$$A \triangleright B = B^{-1}AB.$$

For instance, we can easily check that

$$A \triangleright A = A^{-1}AA = A$$

so the first quandle axiom is satisfied. To verify the second quandle axiom, we need to show that we can solve the equation $A \triangleright B = C$ for A . In this case, we have

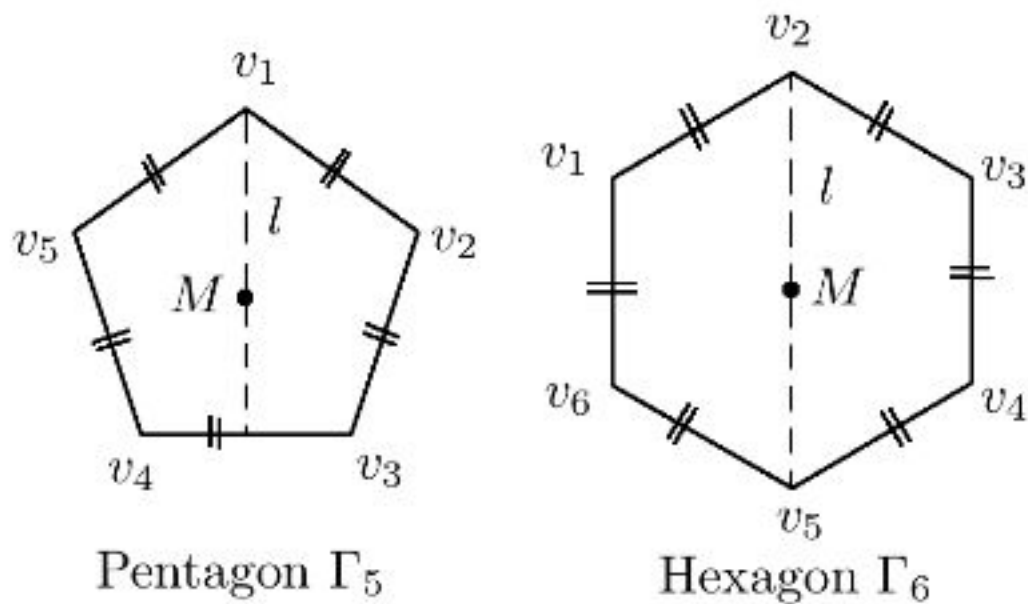
$$A \triangleright B = B^{-1}AB \iff B(A \triangleright B) = AB \iff B(A \triangleright B)B^{-1} = A$$

so we have $A \triangleright^{-1} B = BAB^{-1}$. Verification of axiom (iii) is exercise 8. The operation $A \triangleright B = B^{-1}AB$ is called *conjugation*, and a quandle in which the quandle operation is conjugation is called a *conjugation quandle*. Note that when multiplication is commutative, the conjugation operation is trivial.

Example 65. More generally, let G be any group. Then G is a quandle under the operation of conjugation, i.e.

$$x \triangleright y = y^{-1}xy.$$

Example 66. Let n be an integer ≥ 3 and consider a regular plane n -gon Γ_n , say with center M and vertices v_1, \dots, v_n as depicted.



Let l be the line segment which goes from v_1 through M and continues until it hits Γ_n again. Recall that this regular n -gon has $2n$ symmetries which form a group called the *dihedral group*. To describe these symmetries, let u be the rotation of Γ_n about M through an angle of $\frac{2\pi}{n}$ and v be a reflection (flip) about the line l . Then u^n and v^2 are the identity symmetry, and we have $vu v^{-1} = vu v = u^{-1}$. In fact, using these rules, any element of the group of symmetry of Γ_n can be written as $u^i v^j$, where $0 \leq i \leq n-1$ and $0 \leq j \leq 1$. In other words, there are n rotations (u^i , $0 \leq i \leq n-1$) corresponding to $j = 0$, and n flips ($u^i v$, $0 \leq i \leq n-1$) corresponding to $j = 1$. The conjugation $x \triangleright y = yxy^{-1}$ on the set of reflections of Γ_n is given by

$$(u^a v) \triangleright (u^b v) = u^b v u^a v (u^b v)^{-1} = u^b u^{-a} v u^{-b} = u^{2b-a} v.$$

By considering the one to one correspondence $u^a v \leftrightarrow a$ between the set of reflections of Γ_n and \mathbb{Z}_n we can transfer the quandle operation from the set of reflections of Γ_n to \mathbb{Z}_n by defining $a \triangleright b \equiv 2b - a \pmod{n}$ for $a, b \in \mathbb{Z}_n$ (integers modulo n). The set \mathbb{Z}_n with this quandle structure called the *dihedral quandle*, denoted by R_n .

Example 67. For any vector space V and an invertible linear transformation $t : V \rightarrow V$ of V , define a quandle structure on V by

$$\vec{u} \triangleright \vec{v} = t(\vec{u} - \vec{v}) + \vec{v}.$$

Such a quandle is called an *Alexander quandle*; we will look at Alexander quandles in more detail in the next section.

Quandle Colorings. As with kei , given a finite quandle X we have a counting invariant $\Phi_X^{\mathbb{Z}}$ defined by counting the number of quandle colorings of our oriented knot or link diagram. For much of the remainder of this book, we will study the quandle counting invariant and certain invariants known as *enhancements*.

Example 68. As with the fundamental kei of a knot, there is a *fundamental quandle*, sometimes called the *knot quandle*, associated to an oriented knot or link given by a presentation with generators corresponding to arcs and quandle relations at crossings. We will study the fundamental quandle, also called the *knot quandle* in more detail in Chapter 4.

As with kei , a function $\phi : (X_1, \triangleright_1) \rightarrow (X_2, \triangleright_2)$ is a *quandle homomorphism* if

$$\phi(a \triangleright_1 b) = \phi(a) \triangleright_2 \phi(b)$$

for all $a, b \in X_1$. Axiom (iii) of the quandle definition states that for each $u \in X$, the map $\beta_b : X \rightarrow X$ defined by $\beta_b(a) = a \triangleright b$ is a quandle homomorphism. As expected, a subquandle of a quandle is a subset closed under \triangleright , and every quandle homomorphism has an image subquandle contained in the codomain quandle. Quandle colorings of a knot K by a quandle X are really quandle homomorphisms from the fundamental quandle of K to X .

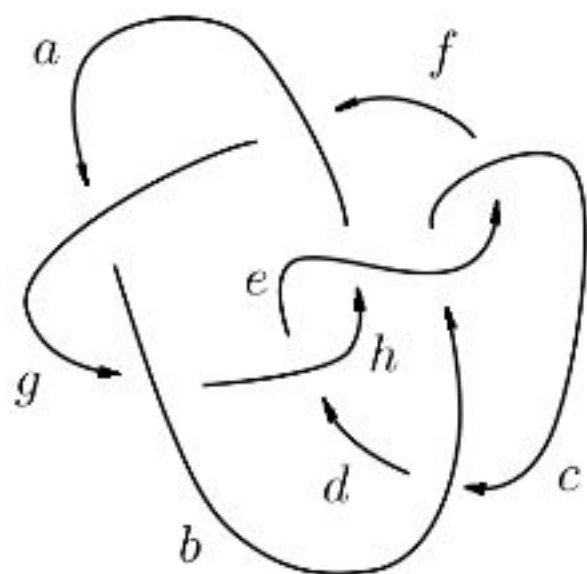
Example 69. Consider the knot 8_{19} from example 60 and the Alexander quandle $X = \Lambda_7/(t - 3)$, i.e. \mathbb{Z}_7 with quandle operation

$$x \triangleright y = 3x + 5y.$$

X has the operation table

\triangleright	0	1	2	3	4	5	6
0	0	5	3	1	6	4	2
1	3	1	6	4	2	0	5
2	6	4	2	0	5	3	1
3	2	0	5	3	1	6	4
4	5	3	1	6	4	2	0
5	1	6	4	2	0	5	3
6	4	2	0	5	3	1	6

Since this quandle operation is a linear function, we can compute the set of all quandle colorings of 8_{19} by X using row-reduction over \mathbb{Z}_7 . Specifically, if we give 8_{19} the pictured orientation



then the system of linear equations determined by the crossing relations has the coefficient matrix

$$\begin{bmatrix} 3 & 6 & 0 & 0 & 0 & 0 & 5 & 0 \\ 0 & 3 & 6 & 0 & 5 & 0 & 0 & 0 \\ 0 & 5 & 3 & 6 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 3 & 6 & 0 & 0 & 5 \\ 0 & 0 & 5 & 0 & 3 & 6 & 0 & 0 \\ 5 & 0 & 0 & 0 & 0 & 3 & 6 & 0 \\ 0 & 5 & 0 & 0 & 0 & 0 & 3 & 6 \\ 6 & 0 & 0 & 0 & 5 & 0 & 0 & 3 \end{bmatrix}.$$

After row-reduction over \mathbb{Z}_7 , we find there is a 2-dimensional space of quandle homomorphisms $\text{Hom}(\mathcal{Q}(8_{19}), X)$ (see exercises), and hence the counting invariant is

$$|\text{Hom}(\mathcal{Q}(8_{19}), X)| = 7^2 = 49.$$

Automorphism Groups. Let $\text{Aut}(X)$ denote the group of all automorphisms of X . The subgroup of $\text{Aut}(X)$, generated by the permutations β_x , is called the *inner* automorphism group of X and is denoted by $\text{Inn}(X)$. Quandle axiom (iii) implies that the map $\beta : X \rightarrow \text{Inn}(X)$, sending u to β_u satisfies the equation

$$\beta_z \beta_y = \beta_{y \triangleright z} \beta_z$$

for all $y, z \in X$, which can be rewritten as

$$\beta_z \beta_y \beta_z^{-1} = \beta_{y \triangleright z}.$$

Thus, if the group $\text{Inn}(X)$ is considered as a quandle with conjugation then the map β becomes a quandle homomorphism.

The subgroup of $\text{Aut}(X)$ generated by $\beta_x \beta_y^{-1}$ for all $x, y \in X$ is called the *transvection* group of X , denoted by $\text{Transv}(X)$. It turns out that the transvection group is a normal subgroup of the inner automorphism group and that the inner automorphism group is a normal subgroup of the automorphism group of X since $f \beta_x f^{-1} = \beta_{f(x)}$ for all $x \in X$ and for all $f \in \text{Aut}(X)$. The quotient group $\text{Inn}(X)/\text{Transv}(X)$ is a cyclic group since any two generators β_x and β_y are equivalent modulo $\text{Transv}(X)$.

The *orbit* of an element x in X , denoted by $\text{Orb}(x)$, is the subset of elements y in X such that there exists some element $f \in \text{Inn}(X)$ that maps x to y . That is, the orbit of $x \in X$ is the set of elements one can get to from x by quandle operations. For example, the dihedral quandle R_4 has two orbits: $\text{orb}(0) = \text{orb}(2) = \{0, 2\}$ and $\text{orb}(1) = \text{orb}(3) = \{1, 3\}$. In the operation table of a quandle, the orbit of an element x_k includes all the elements in the row of x_k together with all the elements in the rows of those elements, etc.

Quandles can have various extra properties; we list some of the more common types.

- A quandle X is *connected* if it has a single orbit. That is, X is connected if for all x, y in X , there exists an element f in $\text{Inn}(X)$ that maps x to y .
- A quandle is *Latin* if for each $a \in X$, the map $\lambda_a : X \rightarrow X$ defined by $\lambda_a(b) = a \triangleright b$ is a bijection. That is, X is Latin if the multiplication table of the quandle is a *Latin square*, i.e. a square with no repeated elements in any row or column.
- A quandle X is *medial* if for all $a, b, c, d \in X$ we have

$$(a \triangleright b) \triangleright (c \triangleright d) = (a \triangleright c) \triangleright (b \triangleright d).$$

It turns out that a quandle is medial if and only if its transvection group is abelian; thus, medial quandles are also called *abelian*. For example, Alexander quandles are medial.

- A quandle is *faithful* if the mapping $a \mapsto \beta_a$ is an injection from X to $\text{Inn}(X)$. That is, a quandle is faithful if no two elements have the same β_x map, or equivalently, if no two columns in the quandle operation table are the same.
- A quandle X is called *simple* if the only surjective quandle homomorphisms $f : X \rightarrow Y$ have trivial images or are bijective.

Exercises. 1. Let (X, \triangleright) be a quandle and recall that the inverse map of β_y is the map $x \mapsto x \triangleright^{-1} y$. Prove that (X, \triangleright^{-1}) is a quandle (called the *dual quandle* of (X, \triangleright)).

2. Using the quandle axioms, prove that the quandle operation \triangleright and the inverse quandle operation \triangleright^{-1} are mutually right-distributive; that is,

$$\begin{aligned} (x \triangleright^{-1} y) \triangleright z &= (x \triangleright z) \triangleright^{-1} (y \triangleright z) \quad \text{and} \\ (x \triangleright y) \triangleright^{-1} z &= (x \triangleright^{-1} z) \triangleright (y \triangleright^{-1} z). \end{aligned}$$

3. Find all quandle structures with three elements: first, note that any such quandle has a 3×3 operation matrix with diagonal entries 1, 2, 3 by quandle axiom (i) and columns which are permutations by quandle axiom (ii). Which of the ways of completing such a table satisfy the self-distributive property?

4. Of the quandles you identified in problem 3, which are isomorphic to which?

5. Using operation tables like in problems 3 and 4, identify all four-element Latin quandles up to isomorphism.

6. Prove that a quandle can be decomposed as a disjoint union of its orbits and that each orbit set forms a subquandle.

7. Using row-reduction over \mathbb{Z}_7 , find a basis for the space of quandle homomorphisms in example 69.

8. Verify that conjugation in a group is self-distributive.

3. Alexander Quandles and the Alexander Polynomial

For most of the 20th century, the quest for knot invariants seemed to come back to one of the first strong knot invariants to be discovered, the Alexander polynomial $\Delta(K)$. It turns out that perhaps the simplest way to understand the Alexander polynomial is in terms of a type of quandles known as *Alexander quandles*. These quandles are modules over $\Lambda = \mathbb{Z}[t^{\pm 1}]$, the set of Laurent polynomials in one variable with integer coefficients. This set Λ is not a field, but it does include negative powers of t ; in fact, the only division we can do in Λ is division by (plus or minus) powers of t . We will denote by Λ_n the ring $\mathbb{Z}_n[t^{\pm 1}]$ of Laurent polynomials with \mathbb{Z}_n coefficients.

Definition 15. Let A be a module over $\Lambda = \mathbb{Z}[t^{\pm 1}]$. Then A is a quandle under the operation

$$\vec{x} \triangleright \vec{y} = t\vec{x} + (1 - t)\vec{y}$$

known as an *Alexander quandle*.

To verify that we actually have a quandle, we must check that the quandle axioms are satisfied. So, suppose A is a Λ -module and define $\vec{x} \triangleright \vec{y}$ as above. Then for the first axiom, we have

$$\vec{x} \triangleright \vec{x} = t\vec{x} + (1 - t)\vec{x} = (t + 1 - t)\vec{x} = \vec{x}$$

as required. For the second axiom, we need to identify $\vec{x} \triangleright^{-1} \vec{y}$. Let us write $\vec{x} = \vec{z} \triangleright \vec{y}$ and solve for \vec{z} :

$$\begin{aligned} \vec{x} &= t\vec{z} + (1 - t)\vec{y}, \\ \vec{x} - (1 - t)\vec{y} &= t\vec{z}, \\ t^{-1}\vec{x} - t^{-1}(1 - t)\vec{y} &= \vec{z}, \\ t^{-1}\vec{x} - (t^{-1} - 1)\vec{y} &= \vec{z}, \\ t^{-1}\vec{x} + (1 - t^{-1})\vec{y} &= \vec{z}, \end{aligned}$$

so we have $\vec{x} \triangleright^{-1} \vec{y} = t^{-1}\vec{x} + (1 - t^{-1})\vec{y}$.

Finally, let us check self-distributivity:

$$\begin{aligned} (\vec{x} \triangleright \vec{y}) \triangleright \vec{z} &= t(\vec{x} \triangleright \vec{y}) + (1-t)\vec{z} \\ &= t(t\vec{x} + (1-t)\vec{y}) + (1-t)\vec{z} \\ &= t^2\vec{x} + t(1-t)\vec{y} + (1-t)\vec{z}, \end{aligned}$$

while

$$\begin{aligned} (\vec{x} \triangleright \vec{z}) \triangleright (\vec{y} \triangleright \vec{z}) &= t(\vec{x} \triangleright \vec{z}) + (1-t)(\vec{y} \triangleright \vec{z}) \\ &= t(t\vec{x} + (1-t)\vec{z}) + (1-t)(t\vec{y} + (1-t)\vec{z}) \\ &= t^2\vec{x} + t(1-t)\vec{y} + [t(1-t) + (1-t)^2]\vec{z} \\ &= t^2\vec{x} + t(1-t)\vec{y} + [t - t^2 + 1 - 2t + t^2]\vec{z} \\ &= t^2\vec{x} + t(1-t)\vec{y} + (1-t)\vec{z} \end{aligned}$$

as required.

Example 70. Any vector space V becomes an Alexander quandle when we select an invertible linear transformation $t : V \rightarrow V$ and define

$$\vec{x} \triangleright \vec{y} = t\vec{x} + (I - t)\vec{y}$$

where I is the identity matrix. For example, consider $V = \mathbb{R}^2$ and choose $t = \begin{bmatrix} 1 & 2 \\ 1 & 3 \end{bmatrix}$. Then t is invertible with

$$t^{-1} = \begin{bmatrix} 3 & -2 \\ -1 & 1 \end{bmatrix} \quad \text{and} \quad I - t = \begin{bmatrix} 0 & -2 \\ -1 & -2 \end{bmatrix};$$

then we have quandle operation

$$\begin{aligned} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \triangleright \begin{bmatrix} y_1 \\ y_2 \end{bmatrix} &= \begin{bmatrix} 1 & 2 \\ 1 & 3 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} + \begin{bmatrix} 0 & -2 \\ -1 & -2 \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \end{bmatrix} \\ &= \begin{bmatrix} x_1 + 2x_2 \\ x_1 + 3x_2 \end{bmatrix} + \begin{bmatrix} -2y_2 \\ -y_1 - 2y_2 \end{bmatrix} \\ &= \begin{bmatrix} x_1 + 2x_2 - y_2 \\ x_1 + 3x_2 - y_1 - 2y_2 \end{bmatrix}. \end{aligned}$$

Example 71. The integers mod n , \mathbb{Z}_n , form an Alexander quandle with the choice of any invertible element $t \in \mathbb{Z}_n$, i.e., any t whose greatest common divisor with n is 1. Then we have Alexander quandle operation

$$x \triangleright y = tx + (1-t)y.$$

For example, in \mathbb{Z}_3 we can choose $t = 1$ or $t = 2$; then we get Alexander quandles with operations as listed:

$x \triangleright y = x$				$x \triangleright y = 2x + 2y$			
\triangleright	0	1	2	\triangleright	0	1	2
0	0	0	0	0	0	2	1
1	1	1	1	1	2	1	0
2	2	2	2	2	1	0	2

We can get more examples of finite Alexander quandles by taking quotients of $\Lambda_n = \mathbb{Z}_n[t^{\pm 1}]$ by monic Laurent polynomials $P \in \Lambda_n$, i.e., polynomials with top degree term t^{k+1} for some integer k . In fact, we can without loss of generality assume P is a genuine polynomial by multiplying P by t^n to get a polynomial with nonzero constant term. Then as a set, our finite quandle consists of \mathbb{Z}_n -linear combinations of $1, t, t^2, \dots, t^k$ where $\deg(P) = k + 1$ with the rule that t^{k+1} gets replaced by $t^{k+1} - P$ in our computations.

Example 72. In the Alexander quandle $A = \Lambda_3/(2 + t + t^2)$, we have $2 + t + t^2 = 0$ which implies $t^2 = -2 - t = 1 + 2t$ (since we have \mathbb{Z}_3 coefficients). Then the elements of A are $\{0, 1, 2, t, 1 + t, 2 + t, 2t, 1 + 2t, 2 + 2t\}$. Then for instance we have

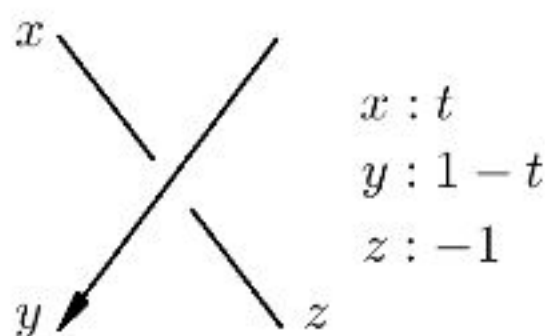
$$\begin{aligned}
 (1 + t) \triangleright 2t &= t(1 + t) + (1 - t)(2t) \\
 &= t + t^2 + 2t - 2t^2 \\
 &= 3t - t^2 \\
 &= 2t^2 \\
 &= 2(1 + 2t) \\
 &= 2 + 4t \\
 &= 2 + t.
 \end{aligned}$$

Example 73. In the Alexander quandle $A = \Lambda_2/(1 + t^2)$, we have $1 + t^2 = 0$ which implies $t^2 = -1 = 1$ (since we have \mathbb{Z}_2 coefficients). Then the elements of A are $\{0, 1, t, 1 + t\}$, and we have, for instance, $(1 + t)^2 = 1 + 2t + t^2 = 1 + 0 + t^2 = 1 + t^2 = 1 + 1 = 0$ and $t(1 + t) = t^2 + t = 1 + t$. We can then find the complete operation

table of A :

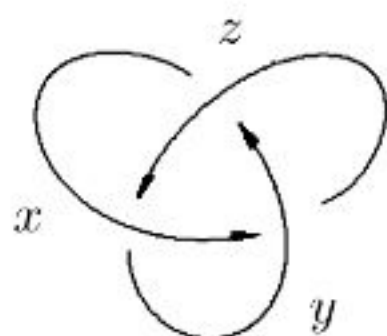
\triangleright	0	1	t	$1+t$
0	0	$1+t$	$1+t$	0
1	t	1	1	t
t	1	t	t	1
$1+t$	$1+t$	0	0	$1+t$

The Alexander Module. What if we apply the Alexander quandle idea to the fundamental quandle of a knot or link? Specifically, suppose we take an oriented knot diagram and label each of the arcs x_1, x_2, \dots, x_n . At each crossing, we get a quandle relation $x_i \triangleright x_j = x_k$; let us interpret this as an Alexander quandle relation, so we have equation $tx_i + (1-t)x_j = x_k$ or $tx_i + (1-t)x_j - x_k = 0$. Thus, we have a homogeneous system of linear equations, which we can express as a matrix equation $A\vec{x} = \vec{0}$. In particular, the matrix A has a row for each crossing with entries $t, 1-t$, or -1 for the arcs involved in the crossing and 0 otherwise.



As we saw in Chapter 2, such an equation has several associated vector spaces; if A has m rows and n columns, then A represents a linear transformation $f : \Lambda^n \rightarrow \Lambda^m$; the solution space to the system $A\vec{x} = \vec{0}$ is called the *kernel* of f , and we can form the quotient module $\Lambda^n / \text{Ker}(f)$. In this case, the quotient module of the free Λ -module generated by the arcs of K modulo the kernel of A is called the *Alexander module* of K ; it can be understood as the fundamental quandle of the knot K interpreted as an Alexander quandle. The matrix A is known as a *presentation matrix* for the Alexander module.

Example 74. The trefoil knot below has Alexander module with listed presentation matrix.



$$A = \begin{bmatrix} 1-t & -1 & t \\ t & 1-t & -1 \\ -1 & t & 1-t \end{bmatrix}.$$

The first really powerful knot invariant, discovered back in the 1920s by Alexander, was the *Alexander polynomial*. Like the Jones polynomial, it is actually a Laurent polynomial, meaning it can have negative as well as positive powers of its variable t , though it turns out we can always “normalize” it to get a genuine polynomial.

Here’s how it works: a subset S of Λ is called an *ideal* if it satisfies the properties:

- If $x, y \in S$ then $x + y \in S$.
- If $\lambda \in \Lambda$ and $x \in S$, then $\lambda x \in S$.

Ideals are very similar to subspaces in linear algebra: both are subsets which are closed under addition and a kind of multiplication, but where a subspace must be closed under scalar multiplication, an ideal must be closed under multiplication by everything in Λ . A *generating set* for an ideal S is a set $G \subset S$ such that everything in S can be written as sums of multiples of elements of G , analogous to a basis but with multiplication in Λ instead of scalar multiplication. An ideal is called *principal* if it has a generating set consisting of a single element. In particular, if $x = yz$ where y is invertible, then the principal ideals generated by x and z are the same since every multiple λz of z is a multiple $(\lambda y^{-1})x$ of x . On the other hand, if two elements x and z generate the same principal ideal, then $x = yz$ for some invertible element y of Λ .

Now, consider an $n \times n$ matrix with entries in Λ . For any non-negative integer k , let I_k be the ideal with generating set given by all of the $(n - k)$ minors of A , i.e. the determinants of the matrices obtained from A by eliminating k rows and columns. I_k is called the *kth elementary ideal* of A . Next, let P_k be the smallest principal

ideal containing I_k – since Λ itself is the principal ideal generated by 1, there is always a P_k for every I_k – and let Δ_k be a generator for P_k . Indeed, it turns out that the greatest common divisor of a set of generators for I_k is always a generator for P_k . Then if our matrix is a presentation matrix A for the Alexander module of a knot K , Δ_k is the k th Alexander polynomial of K .

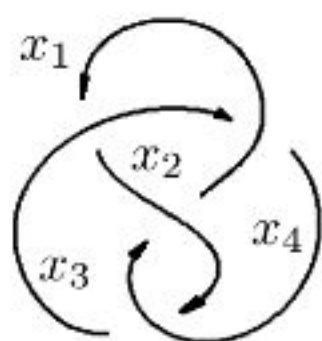
Alexander proved that every knot has $\Delta_0 = 0$ since the matrix A always ends up being singular, and that the first elementary ideal I_1 of a presentation matrix A for the Alexander module of a knot is always principal. Then in particular, we can compute the Alexander polynomial of a knot by writing down the matrix A and taking any $(n - 1)$ minor. Since Δ_k is only defined up multiplication by $\pm t^n$, to get a canonical value for the invariant we can multiply by an appropriate $\pm t^n$ to get a positive constant term. For instance, if we get

$$-t^{-2} + 1 + t,$$

we can multiply through by $(-t^2)$ to get normalized polynomial

$$1 - t^2 - t^3.$$

Example 75. Let us compute the Alexander polynomial of the figure eight knot 4_1 . First, we need a presentation matrix for the Alexander module, which we can obtain by labeling the arcs in a diagram of 4_1 and interpreting the crossing relations as Alexander quandle operations:



$$A = \begin{bmatrix} -1 & t & 1-t & 0 \\ 1-t & 0 & -1 & t \\ -1 & 1-t & 0 & t \\ 0 & t & -1 & 1-t \end{bmatrix}.$$

Next, we can choose any row and column to eliminate, then take the determinant. For example, suppose we eliminate row 1 and column 3:

$$\begin{aligned}
 \begin{vmatrix} 1-t & 0 & t \\ -1 & 1-t & t \\ 0 & t & 1-t \end{vmatrix} &= (1-t) \begin{vmatrix} 1-t & t \\ t & 1-t \end{vmatrix} + 0 \\
 &\quad + t \begin{vmatrix} -1 & 1-t \\ 0 & t \end{vmatrix} \\
 &= (1-t)[(1-t)^2 - t^2] + t(-t) \\
 &= (1-t)[1 - 2t + t^2 - t^2] - t^2 \\
 &= (1-t)[1 - 2t] - t^2 \\
 &= 1 - t - 2t + 2t^2 - t^2 \\
 &= 1 - 3t + t^2
 \end{aligned}$$

and in this case, we don't need to normalize since we already have a positive constant term. Alternatively, we could instead eliminate row 2 and column 4:

$$\begin{aligned}
 \begin{vmatrix} -1 & t & 1-t \\ -1 & 1-t & 0 \\ 0 & t & -1 \end{vmatrix} &= - \begin{vmatrix} 1-t & 0 \\ t & -1 \end{vmatrix} - t \begin{vmatrix} -1 & 0 \\ 0 & -1 \end{vmatrix} \\
 &\quad + (1-t) \begin{vmatrix} -1 & 1-t \\ 0 & t \end{vmatrix} \\
 &= -(-(1-t)) - t(1) + (1-t)(-t) \\
 &= 1 - t - t - t + t^2 \\
 &= 1 - 3t + t^2
 \end{aligned}$$

or row 3 and column 1:

$$\begin{aligned}
 \begin{vmatrix} t & 1-t & 0 \\ 0 & -1 & t \\ t & -1 & 1-t \end{vmatrix} &= t \begin{vmatrix} -1 & t \\ -1 & 1-t \end{vmatrix} + 0 + t \begin{vmatrix} 1-t & 0 \\ -1 & t \end{vmatrix} \\
 &= t(-1(1-t) - (-t)) + t((1-t)t) \\
 &= t(-1 + 2t^2) + t^2 - t^3 \\
 &= -t + 3t^2 - t^3.
 \end{aligned}$$

Now this last one looks different from the first two, but it doesn't have a positive constant term (its constant term is zero); thus, we can normalize it by multiplying by $-t^{-1}$ to get

$$-t^{-1}(-t + 3t^2 - t^3) = 1 - 3t + t^2.$$

Colorings by Alexander Quandles. One advantage we have when coloring knots with finite Alexander quandles instead of quandles defined by operation tables is that we can use linear algebra to compute the set of quandle colorings of a knot or link. Specifically, the Alexander quandle presentation matrix A can be understood as the system of linear equations specifying colorings of K ; if X is a finite Alexander quandle, we can row-reduce the matrix A over X to find the solution space $\text{Hom}(\mathcal{Q}(K), X)$.

Example 76. Let X be the Alexander quandle $\Lambda_5/(t-3)$; let us find the set of X -colorings of the figure eight knot 4_1 by X . We could do this by making a table of possible colorings and checking which satisfy all of the crossing equations, but since X has an Alexander quandle structure we can instead use linear algebra. In $X = \Lambda_5/(t-3)$ we have $t = 3$, $1-t = 1-3 = -2 = 3$ and $-1 = 4$. We can then replace the t , $1-t$ and -1 values in the presentation matrix for the Alexander module of 4_1 :

$$\begin{bmatrix} -1 & t & 1-t & 0 \\ 1-t & 0 & -1 & t \\ -1 & 1-t & 0 & t \\ 0 & t & -1 & 1-t \end{bmatrix} \rightarrow \begin{bmatrix} 4 & 3 & 3 & 0 \\ 3 & 0 & 4 & 3 \\ 4 & 3 & 0 & 3 \\ 0 & 3 & 4 & 3 \end{bmatrix}.$$

We can then row-reduce this matrix over \mathbb{Z}_5 , i.e., using \mathbb{Z}_5 arithmetic rules, to find the space of solutions, i.e., the set of X -colorings of K .

$$\begin{aligned} \begin{bmatrix} 4 & 3 & 3 & 0 \\ 3 & 0 & 4 & 3 \\ 4 & 3 & 0 & 3 \\ 0 & 3 & 4 & 3 \end{bmatrix} &\sim \begin{bmatrix} 1 & 3 & 4 & 2 \\ 3 & 0 & 4 & 3 \\ 4 & 3 & 0 & 3 \\ 0 & 3 & 4 & 3 \end{bmatrix} \sim \begin{bmatrix} 1 & 3 & 4 & 2 \\ 0 & 1 & 2 & 2 \\ 4 & 3 & 0 & 3 \\ 0 & 3 & 4 & 3 \end{bmatrix} \\ &\sim \begin{bmatrix} 1 & 3 & 4 & 2 \\ 0 & 1 & 2 & 2 \\ 0 & 1 & 4 & 0 \\ 0 & 3 & 4 & 3 \end{bmatrix} \sim \begin{bmatrix} 1 & 3 & 4 & 2 \\ 0 & 1 & 2 & 2 \\ 0 & 0 & 2 & 3 \\ 0 & 3 & 4 & 3 \end{bmatrix} \sim \begin{bmatrix} 1 & 3 & 4 & 2 \\ 0 & 1 & 2 & 2 \\ 0 & 0 & 2 & 3 \\ 0 & 0 & 3 & 2 \end{bmatrix} \end{aligned}$$

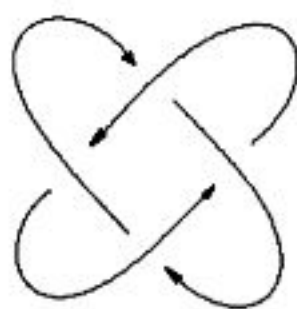
$$\begin{aligned}
& \sim \begin{bmatrix} 1 & 3 & 4 & 2 \\ 0 & 1 & 2 & 2 \\ 0 & 0 & 2 & 3 \\ 0 & 0 & 0 & 0 \end{bmatrix} \sim \begin{bmatrix} 1 & 3 & 4 & 2 \\ 0 & 1 & 2 & 2 \\ 0 & 0 & 1 & 4 \\ 0 & 0 & 0 & 0 \end{bmatrix} \sim \begin{bmatrix} 1 & 3 & 4 & 2 \\ 0 & 1 & 0 & 4 \\ 0 & 0 & 1 & 4 \\ 0 & 0 & 0 & 0 \end{bmatrix} \\
& \sim \begin{bmatrix} 1 & 3 & 0 & 1 \\ 0 & 1 & 0 & 4 \\ 0 & 0 & 1 & 4 \\ 0 & 0 & 0 & 0 \end{bmatrix} \sim \begin{bmatrix} 1 & 0 & 0 & 4 \\ 0 & 1 & 0 & 4 \\ 0 & 0 & 1 & 4 \\ 0 & 0 & 0 & 0 \end{bmatrix}.
\end{aligned}$$

Thus, the kernel has dimension 1 and thus is isomorphic to \mathbb{Z}_5 ; hence there are 5 colorings of 4_1 by X . Since we already know there are five constant colorings (where every arc gets the same color), these are all of the quandle colorings of 4_1 by X .

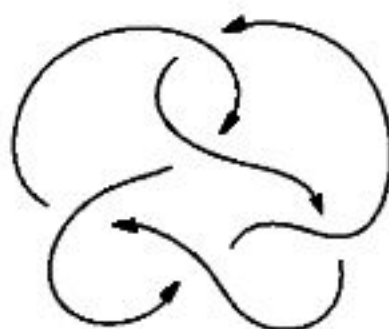
Exercises. 1. Complete the operation table for the Alexander quandle $A = \Lambda_3/(2 + t + t^2)$ from Example 72.

2. Find the operation table for the Alexander quandle $A = \Lambda_2/(1 + t + t^3)$.

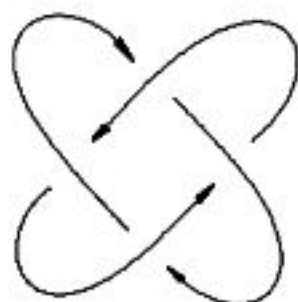
3. Compute the Alexander polynomial of the $(4, 2)$ torus link below.



4. Compute the Alexander polynomial of the knot 5_2 below, then do it again with a different choice of row and column eliminated.

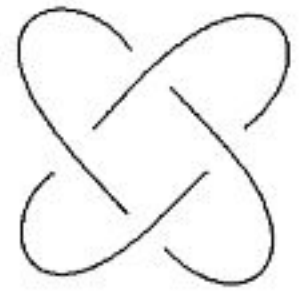


5. Find the set of quandle colorings of 6_1 by the Alexander quandle $\Lambda_3/(t-2)$ using row-reduction in \mathbb{Z}_3 .
6. Find the set of quandle colorings of the $(4, 2)$ -torus link



by the Alexander quandle $\Lambda_4/(t-3)$ using row-reduction in \mathbb{Z}_4 . Keep in mind that since 2 is not invertible in \mathbb{Z}_4 , you can only multiply rows by 1 and 3, not 2.

Chapter 4



Quandles and Groups

1. Fundamental Group

In this section we will introduce the notion of the fundamental group of a subset of \mathbb{R}^n . But first let's back up a little bit to the history of this mathematical notion. In 1895, the famous French mathematician Jules Henri Poincaré (29 April 1854–17 July 1912) associated to each topological space a group called the fundamental group of the space. More generally, Poincaré's research in geometry led to the abstract topological definition of homotopy and homology. This was the beginning of a new field of mathematics called algebraic topology. Poincaré was responsible for formulating the Poincaré conjecture, which was one of the most famous unsolved problems in mathematics until it was solved in 2002–2003 by Grigori Perelman.

The concept of *homotopy* allows us to define a useful equivalence of functions in general and paths in particular. It corresponds to a continuous deformation of one function to another. The precise statement is given by the following:

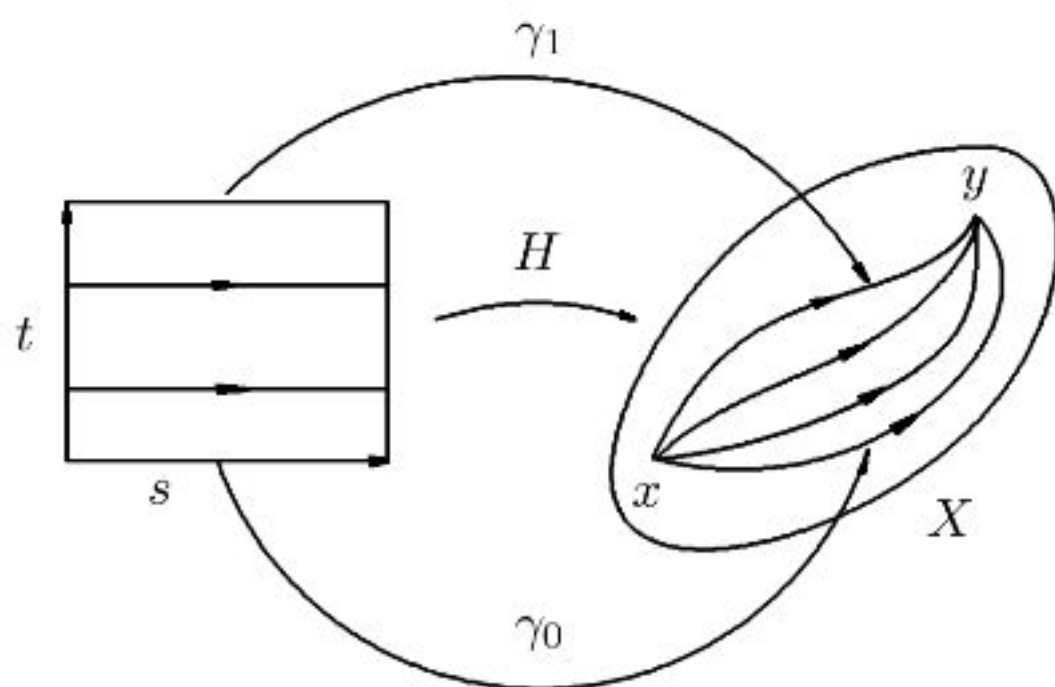
Definition 16. Let X be a subset of \mathbb{R}^n and let $x, y \in X$. A path in X from x to y is a continuous map γ from the interval $[0, 1]$ to X such that $\gamma(0) = x$ and $\gamma(1) = y$. Two paths γ_0 and γ_1 with endpoints fixed ($\gamma_0(0) = \gamma_1(0) = x$ and $\gamma_0(1) = \gamma_1(1) = y$), are said to be *path homotopic* if there is a continuous map $H : [0, 1] \times [0, 1] \rightarrow X$

satisfying

$$\begin{aligned} H(s, 0) &= \gamma_0(s), & 0 \leq s \leq 1, \\ H(s, 1) &= \gamma_1(s), & 0 \leq s \leq 1, \\ H(0, t) &= x, & 0 \leq t \leq 1, \\ H(1, t) &= y, & 0 \leq t \leq 1. \end{aligned}$$

We usually write $\gamma_0 \simeq \gamma_1$ to denote that γ_0 and γ_1 are path homotopic.

The map H can be thought of as a map from the interval $[0, 1]$ to the space of paths on X via the map $t \mapsto \gamma_t$ where $\gamma_t : [0, 1] \rightarrow X$ is given by $\gamma_t(s) = H(s, t)$. One can think of the variable t as a time parameter and the set $\{\gamma_t\}_{0 \leq t \leq 1}$ as a family of paths that moves continuously with t . To better understand the intuitive pictorial idea, imagine that the two curves γ_0 and γ_1 are made of rubber bands. Intuitively γ_0 is said to be *deformed* into γ_1 if by stretching and pulling the rubber band γ_0 can be continuously moved in the space X till it coincides with the rubber band γ_1 . The rule is that during the movement the rubber band must never be broken. The picture



gives the illustration of all this showing some intermediate paths between the initial path γ_0 and the terminal path γ_1 .

Example 77. For $n \geq 2$, and for all $x, y \in \mathbb{R}^n$, any two paths γ_0 and γ_1 from x to y are homotopic via the *linear homotopy*, $H(s, t) = (1 - t)\gamma_0(s) + t\gamma_1(s)$, where $0 \leq s, t \leq 1$.

The notion of path homotopy gives an equivalence relation on the set of paths in X from x to y (see exercise 1 below). The equivalence classes modulo this equivalence relation are called the *homotopy classes* of paths from x to y . The *homotopy class* of a path γ is denoted $[\gamma]$. Then by definition $[\gamma_0] = [\gamma_1]$ is equivalent to $\gamma_0 \simeq \gamma_1$.

Composition or product of paths. Let $x, y, z \in X$ and let α be a path from x to y and β be a path from y to z . Since $\alpha(1) = \beta(0)$, we can define the *composite* path $\alpha\beta$ by

$$(\alpha\beta)(s) = \begin{cases} \alpha(2s), & 0 \leq s \leq \frac{1}{2}, \\ \beta(2s - 1), & \frac{1}{2} \leq s \leq 1. \end{cases}$$

From this product, we see that if we have four paths $\gamma_1, \gamma_2, \gamma_3$ and γ_4 with $\gamma_1(1) = \gamma_3(0)$, $\gamma_2(1) = \gamma_4(0)$ and such that $\gamma_1 \simeq \gamma_2$ and $\gamma_3 \simeq \gamma_4$ then $\gamma_1\gamma_3 \simeq \gamma_2\gamma_4$. This allows us to define a product on the set of equivalence classes of paths. Precisely, we define $[\alpha][\beta] = [\alpha\beta]$. One can then easily check the following property of *associativity*, that is,

$$([\alpha][\beta])[\gamma] = [\alpha]([\beta][\gamma]).$$

For any $x \in X$, let c_x be the constant path at x , that is, the path given by $c_x(t) = x$ for $t \in [0, 1]$. One can easily see that if γ is a path from x to y then $[\gamma][c_y] = [\gamma]$ and $[c_x][\gamma] = [\gamma]$. Given a path γ , we can consider the path $\tilde{\gamma}$ given by $\tilde{\gamma}(t) = \gamma(1 - t)$ (the path going in the opposite direction). It is easy to see that $[\gamma][\tilde{\gamma}] = [c_x]$ and $[\tilde{\gamma}][\gamma] = [c_y]$.

Now if we specialize a bit into the notion of paths we obtain the notion of loops:

Definition 17. Let X be a subset of \mathbb{R}^n and let x be a fixed element of X . A loop based at x in X is a path with initial point and endpoint x . The set of loops based at x is denoted $\pi_1(X, x)$.

It is then clear that $\pi_1(X, x)$ with the operation of multiplication of homotopy classes is a group since multiplication is associative, the identity element is the constant map x and the inverse of the homotopy class of a loop γ is the homotopy class of the inverse loop γ^{-1} , i.e. γ with the opposite direction.

Definition 18. The group $\pi_1(X, x)$ of loops based at x , with the operation of multiplication of homotopy classes, is called the *fundamental group* of the space X based at x .

Remark 1. If in a subset $X \subset \mathbb{R}^n$, every two points x_0 and x_1 can be connected by a path, then the space X is called *path connected*. In this case, the two groups $\pi_1(X, x_0)$ and $\pi_1(X, x_1)$ are isomorphic.

It turns out that an ambient isotopy taking a knot K to another knot K' induces an isomorphism of the fundamental group of the knot complement $\mathbb{R}^3 \setminus K$, called the *knot group* of K , onto the fundamental group of $\mathbb{R}^3 \setminus K'$. In particular, two knots with different knot groups cannot be equivalent.

Exercises. 1. Prove that path homotopy is an equivalence relation on the set of paths of a subset X of \mathbb{R}^n .

2. Let γ a path in \mathbb{R} with $\gamma(0) = 0$ and $\gamma(1) = x \neq 0$. Let $\alpha : [0, 1] \rightarrow \mathbb{R}$ be the direct path sending t to tx . Prove that γ and α are homotopic.

3. Compute the fundamental group of the plane \mathbb{R}^2 .

4. Compute the fundamental group of the *unknot*, that is the circle

$$S^1 = \{(x, y) \in \mathbb{R}^2; \mid x^2 + y^2 = 1\}.$$

5. Given three paths γ_1 , γ_2 and γ_3 with $\gamma_1(1) = \gamma_2(0)$ and $\gamma_2(1) = \gamma_3(0)$. Prove that $(\gamma_1 \gamma_2) \gamma_3 \simeq \gamma_1 (\gamma_2 \gamma_3)$.

6. Prove that the product of homotopy classes of paths is associative, i.e. that

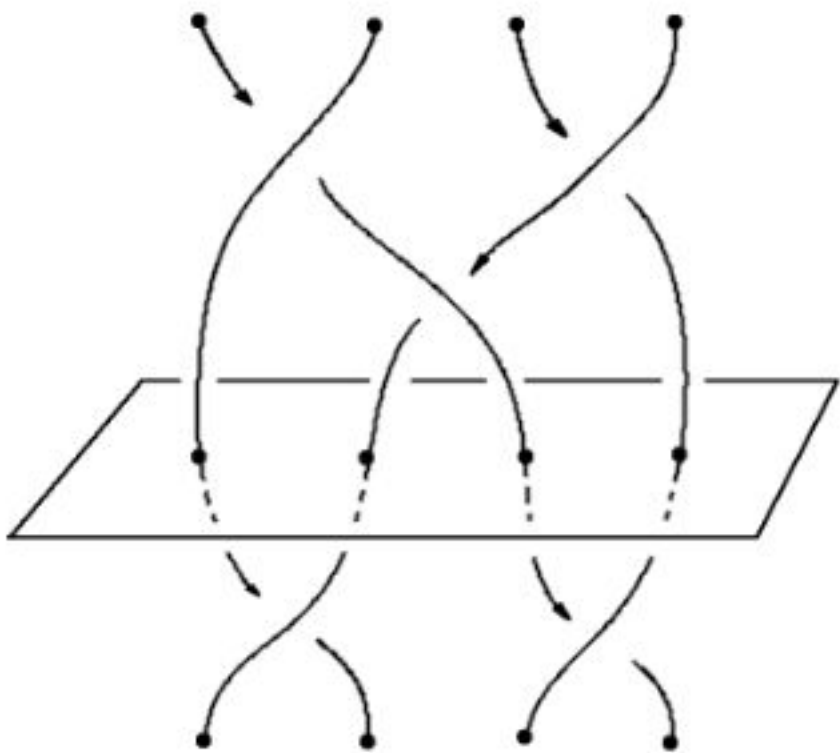
$$([\alpha][\beta])[\gamma] = [\alpha]([\beta][\gamma])$$

for paths α, β, γ .

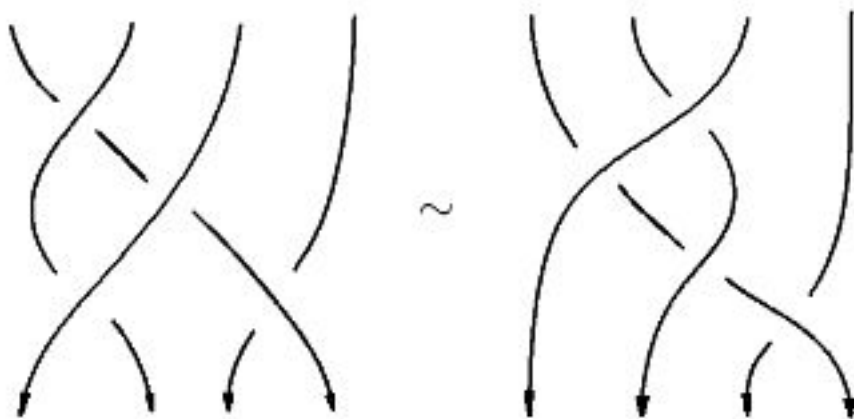
2. Braid Groups

In this section, we look at important relations between braids and links. It is the study of knots which motivated the study of braids. As we mentioned before in Chapter 1, an n -braid is a tangle with n inputs and n outputs which has no maxima and no minima in the

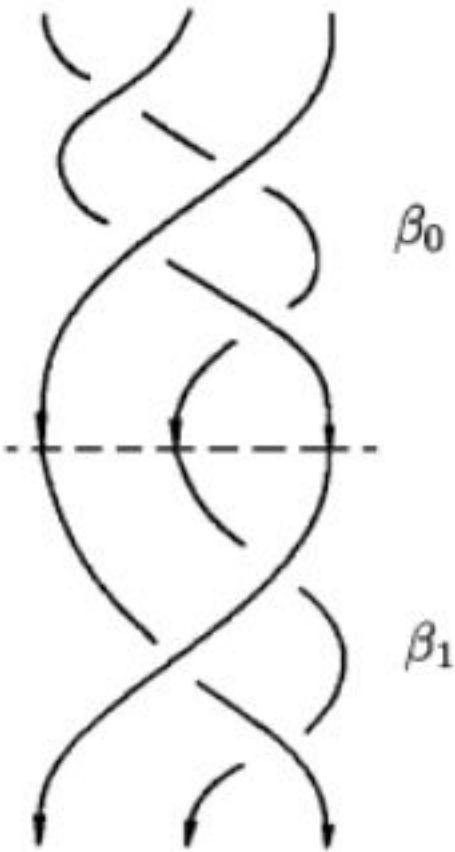
vertical direction. More precisely, if we think of the strings as coming down from a top horizontal plane to a bottom horizontal plane then each horizontal plane in between intersects the braid at exactly n points as shown in the picture:



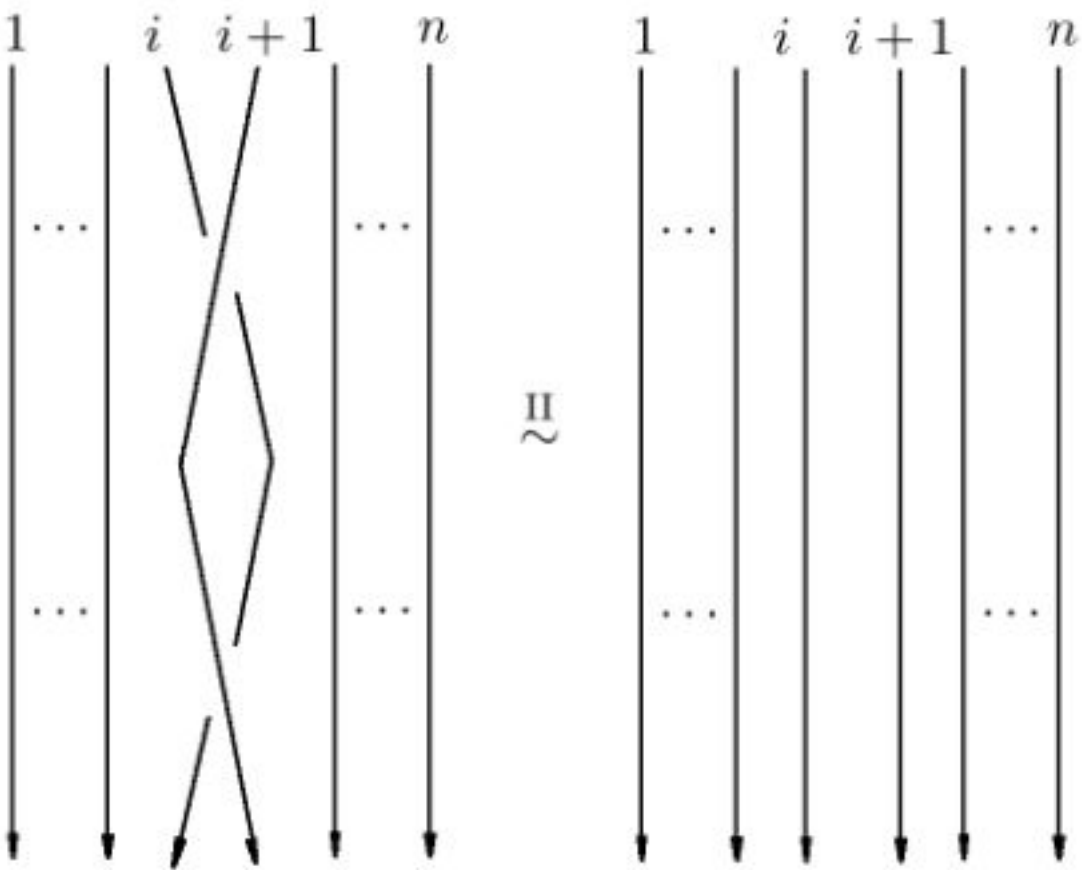
We consider braids with downward orientation. We can always arrange the strands in such a way that no two crossings of the braid occur at the same height. Two braids are called *isotopic* if one can be obtained from the other by a continuous deformation with the requirement that the top and the bottom endpoints of the braid are kept fixed and such that at any time an intersection with a horizontal plane is made exactly of n points. More formally, two braids β_0 and β_1 are isotopic if there exists a continuous map $H : \beta_0 \times [0, 1] \rightarrow \mathbb{R}^2 \times [0, 1]$ such that for all $t \in [0, 1]$, the continuous map $H_t : \beta_0 \rightarrow \mathbb{R}^2 \times [0, 1]$ sending $x \in \beta_0$ to $H(x, t)$ is an embedding whose image is a braid on n strings, where H_0 is the identity map from β_0 to itself and $H_1(\beta_0) = \beta_1$. The map H and the family of maps $\{H_t(\beta_0)\}_{0 \leq t \leq 1}$ are called an *isotopy* of β_0 into β_1 . The following are two isotopic 4-braids.



Product of Braids. Let β_0 and β_1 be two braids on n strings. The product $\beta_0 \beta_1$ is the braid obtained by stacking β_0 over β_1 :

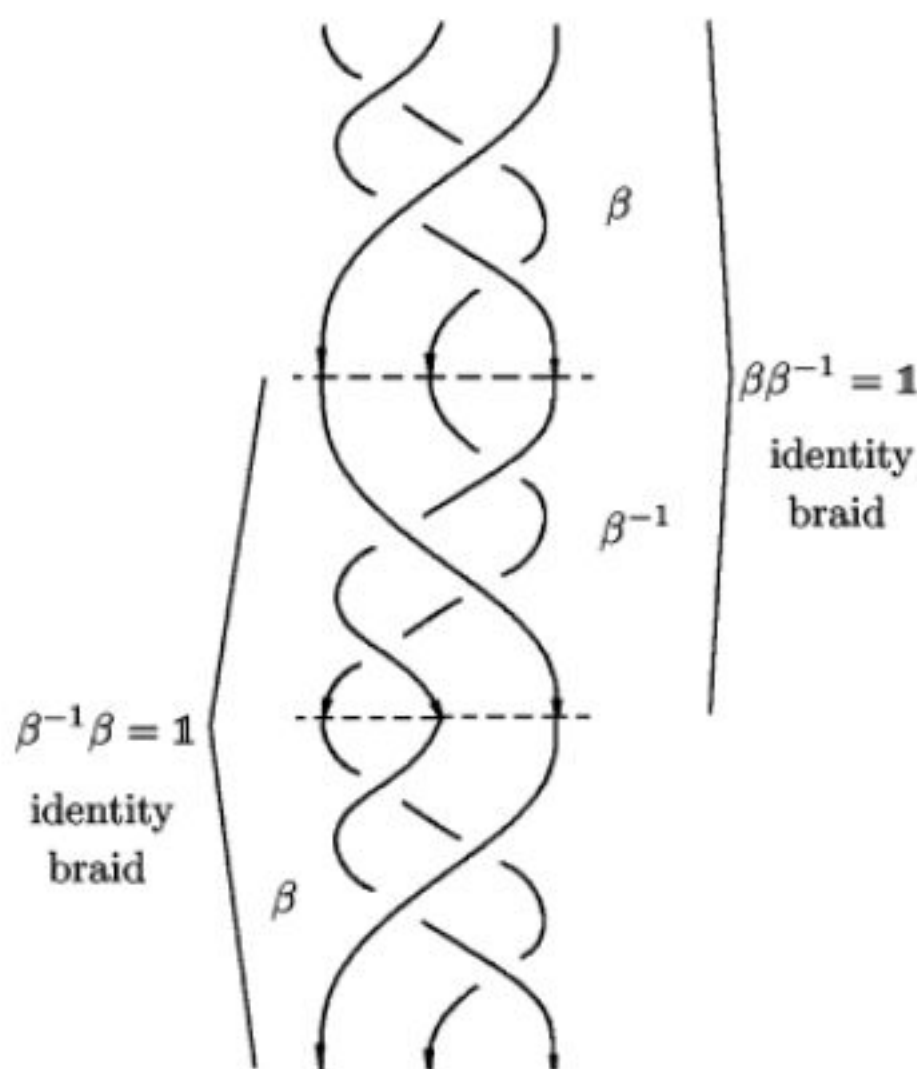


If β_0 and β'_0 are isotopic n -braids and β_1 and β'_1 are also isotopic n -braids, then the product $\beta_0 \beta_1$ is isotopic to $\beta'_0 \beta'_1$ as n -braids. This makes the notion of the product of braids a well-defined operation.



It is clear from the definition that the identity braid (the braid with vertical strings and no crossings) is the neutral element of this product. We know from Reidemeister move II that composing a positive crossing and a negative crossing gives the identity braid on two strings. This is the basic principle of constructing the inverse of a

given braid. Given a braid β , we slice it by finitely many planes in such a way that between every two consecutive planes we have exactly one crossing. We then construct the inverse of the braid by putting crossings with opposite signs starting from the bottom crossing and going up, as can be seen from the following picture.

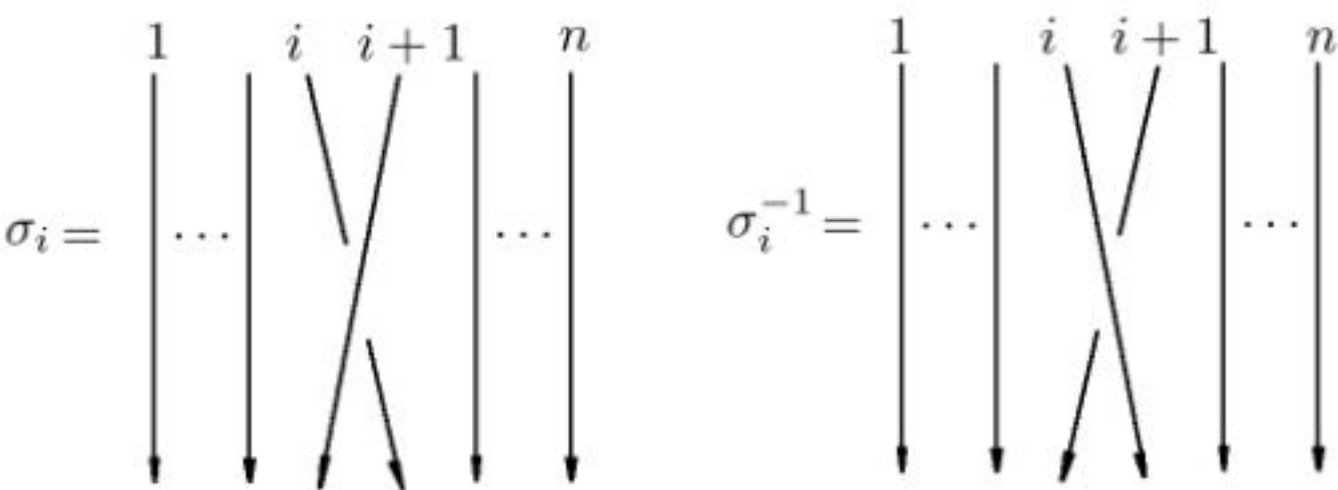


The braid group we just described using geometric braids can be characterized algebraically in the following sense. The *braid group* on n strands, denoted B_n , is the group generated by $(n - 1)$ generators, $\sigma_1, \sigma_2, \dots, \sigma_{n-1}$ subject to the *braid relations*:

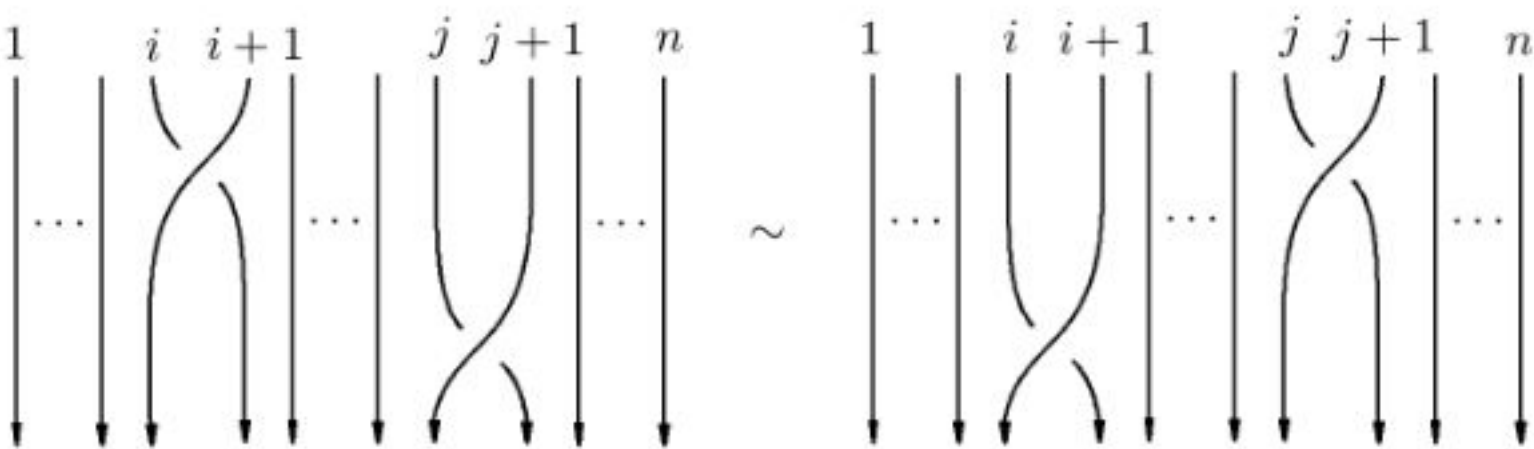
- (i) $\forall i, j$ where $1 \leq i, j \leq n - 1$, $\sigma_i \sigma_j = \sigma_j \sigma_i$ if $|i - j| \geq 2$, and
- (ii) $\forall i$ where $1 \leq i \leq n - 1$, $\sigma_i \sigma_{i+1} \sigma_i = \sigma_{i+1} \sigma_i \sigma_{i+1}$.

With n strings, the braid σ_i represents the braid with only one positive crossing between the i th and $(i + 1)$ th string (all the other strings go vertically down). Its inverse is the braid σ_i^{-1} with only a negative crossing between the i th and $(i + 1)$ th string as can be seen

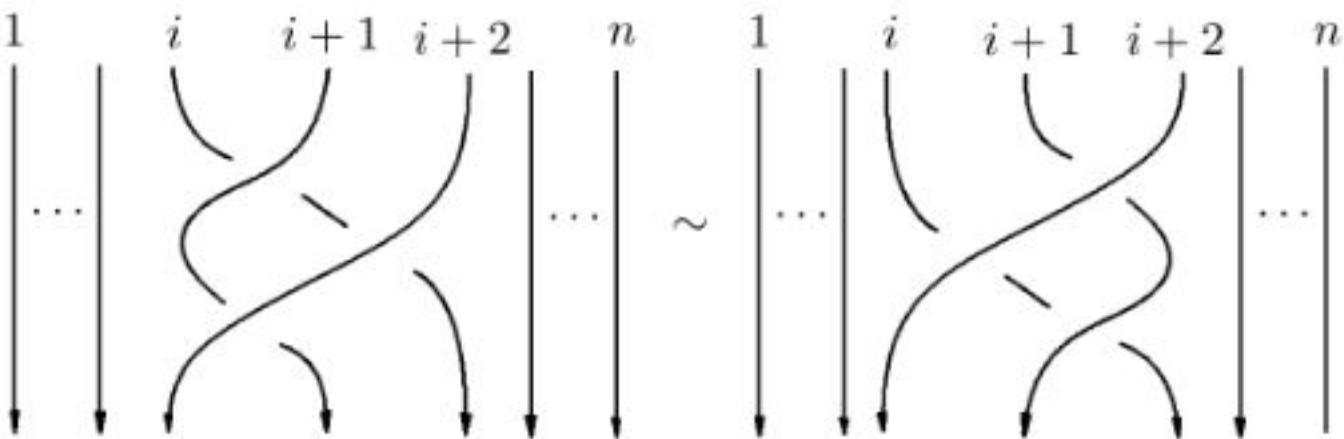
from the following picture:



The braid relations (i) and (ii) can be seen respectively from the following diagrammatic pictures:



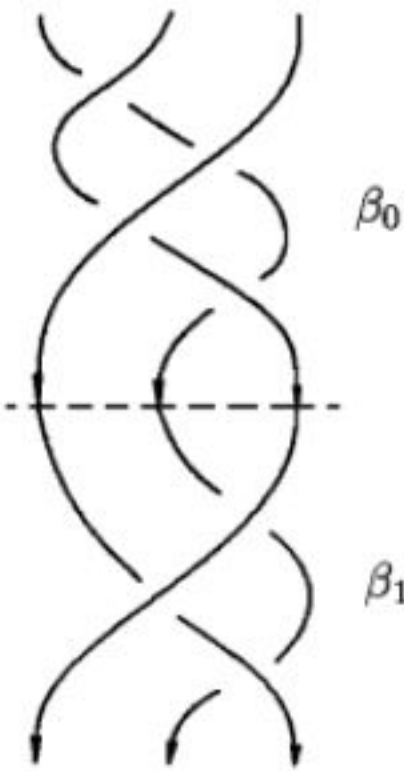
and



In particular, in terms of universal algebra we can define the n -string braid group with the group presentation

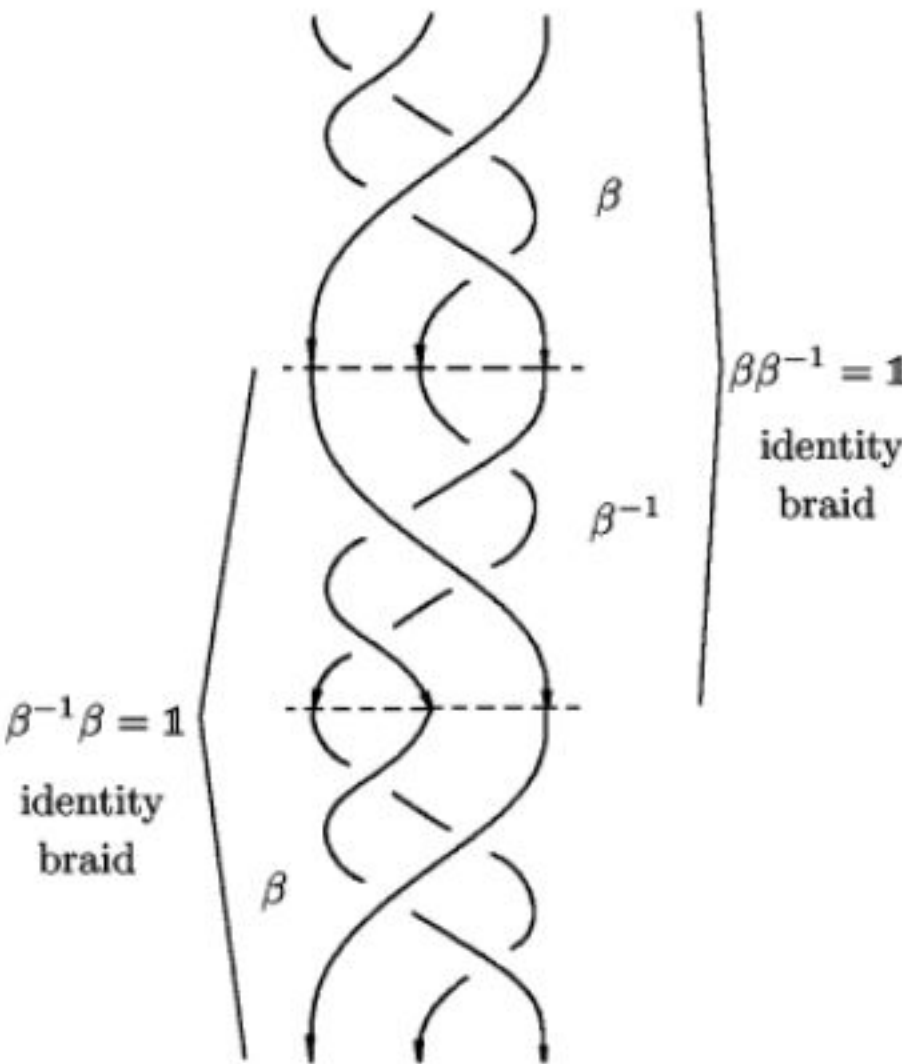
$$B_n = \left\langle \sigma_1, \dots, \sigma_{n-1} \mid \begin{array}{l} \sigma_i \sigma_{i+1} \sigma_i = \sigma_{i+1} \sigma_i \sigma_{i+1}, \\ \sigma_i \sigma_j = \sigma_j \sigma_i \text{ for } |i - j| \geq 2 \end{array} \right\rangle.$$

This description of braids by generators and relations makes the study of braids very convenient. For example, the braids β_0 and β_1



can be written as $\beta_0 = \sigma_1\sigma_2\sigma_1\sigma_2^{-1}$ and $\beta_1 = \sigma_2\sigma_1\sigma_2^{-1}$. Thus their product is the braid $\beta_0 \beta_1 = \sigma_1\sigma_2\sigma_1\sigma_2^{-1}\sigma_2\sigma_1\sigma_2^{-1}$ which is equivalent to the braid $\sigma_1\sigma_2\sigma_1^2\sigma_2^{-1}$.

In the figure



the braid β is given by

$$\beta = \sigma_1\sigma_2\sigma_1\sigma_2^{-1}$$

and its inverse is

$$\beta^{-1} = \sigma_2 \sigma_1^{-1} \sigma_2^{-1} \sigma_1^{-1}$$

as can be seen from the simple verification

$$\begin{aligned} \beta \beta^{-1} &= \sigma_1 \sigma_2 \sigma_1 \sigma_2^{-1} \sigma_2 \sigma_1^{-1} \sigma_2^{-1} \sigma_1^{-1} \\ &= \mathbb{1} \end{aligned}$$

(where $\mathbb{1}$ is the identity braid) and

$$\begin{aligned} \beta^{-1} \beta &= \sigma_2 \sigma_1^{-1} \sigma_2^{-1} \sigma_1^{-1} \sigma_1 \sigma_2 \sigma_1 \sigma_2^{-1} \\ &= 1. \end{aligned}$$

This description of braids by generators and relations is very practical in computing the set of colorings of a given braid or knot. It is easy to see from the definition that the braid group B_1 is the trivial group. Any braid on two strands is isotopic to a braid with m crossings. This integer m characterizes the given braid. Thus the group B_2 is isomorphic to the cyclic group \mathbb{Z} . Algebraically said, the group B_2 is generated by a single element σ_1 and no relations, thus it is isomorphic to \mathbb{Z} . The group B_3 is the group generated by two elements σ_1 and σ_2 with only the relation $\sigma_1 \sigma_2 \sigma_1 = \sigma_2 \sigma_1 \sigma_2$. We will see that this is the fundamental group of the trefoil knot complement.

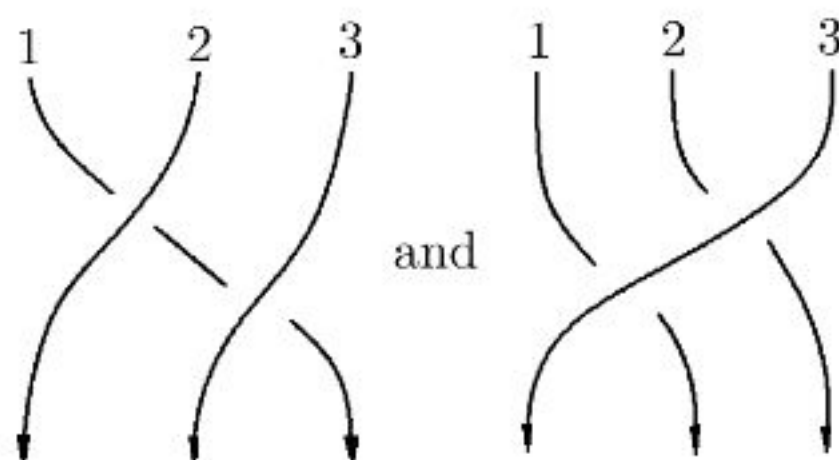
Mapping each σ_i of the braid group B_n to the transposition $\tau_i = [1, 2, \dots, i-1, i+1, i, i+2, \dots, n]$ switching i and $i+1$ in the symmetric group S_n on n letters gives a natural surjective group homomorphism from the braid group B_n to the symmetric group S_n . This is because the analogous braid relations (i) and (ii) are satisfied by the transpositions:

$$\tau_i \tau_{i+1} \tau_i = \tau_{i+1} \tau_i \tau_{i+1}.$$

Since every permutation is a product of transpositions, we have an onto group homomorphism $B_n \rightarrow S_n$ sending σ_i to τ_i .

We want to turn a link into a braid, and vice-versa. One direction is easier than the other. We obtain a link from a braid b by connecting the lower ends of the braid with the upper ends, denoted \hat{b} . The closure of a braid is usually taken to be oriented. Recall that all the strands of the braids are oriented from top to bottom.

Note that isotopic braids generate isotopic links, and that non-isotopic braids may generate an isotopic link; see for example the following figure,



in which the braids $\beta_0 = \sigma_1\sigma_2$ and $\beta_1 = \sigma_2\sigma_1$ have the same closure that is the unknot. We leave it as an exercise to check that β_0 and β_1 cannot be isotopic.

Because the closures of two braids with a different number of strands can give the same knot, we need the following definition. The *braid index* of a knot K , denoted $\text{braidind}(K)$, is the minimum number of strings needed to express K as a closed braid.

Closing a braid b with n strings gives a collection of closed, simple curves in \mathbb{R}^3 (i.e. that do not intersect each other), so the closure is a link. Obviously, the closure \hat{b} can have no more than n components, because each point A_i on the top plane connects with a unique point B_i on the bottom plane (at most n disconnected curves in \hat{b}). It follows that the closure \hat{b} of any braid $b \in B_n$ is a link with at most n components.

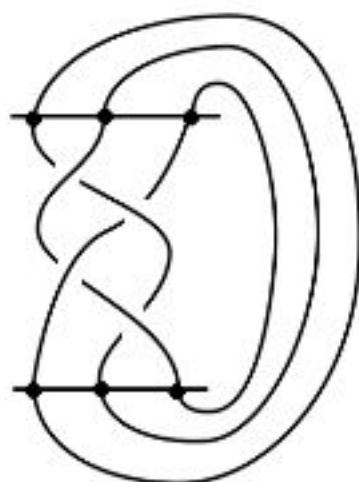
Given a link L , does there exist a braid b such that the closure of b is equivalent to L ?

In 1923 J. W. Alexander answered this question in the positive, as can be seen in the theorem below. There is an algorithm to construct the braid from the link, but it is *quite long and difficult*, even for links with few crossings.

Theorem 6 (Alexander's Theorem). *For any link, L , there exists an integer $n > 0$ and a braid $b \in B_n$ such that L is equivalent to \hat{b} .*

Example 78. The braid index of the right-handed trefoil knot is 2. The closure of the braid $\beta = \sigma_1^3$ is the right-handed trefoil. The left-handed trefoil is the closure of σ_1^{-3} .

Example 79. The braid index of the figure eight knot is 3. The closure of the braid $\beta = \sigma_1 \sigma_2^{-1} \sigma_1 \sigma_2^{-1}$ is the figure eight knot.



Example 80. The braid index of the knot 6_1 (see knot table in Chapter 1) is 4. The closure of the braid $\beta = \sigma_1^2 \sigma_2 \sigma_1^{-1} \sigma_3^{-1} \sigma_2 \sigma_3^{-1}$ is the knot 6_1 .

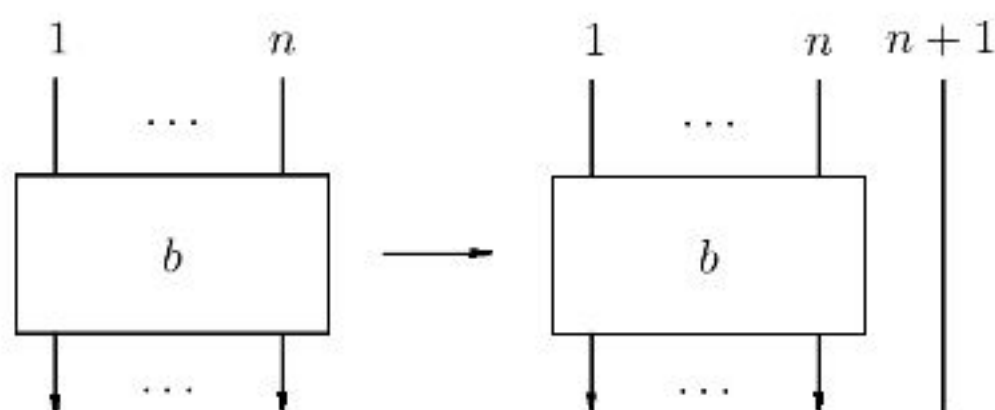
In a 1935 paper, A.A. Markov gave a proof of Markov's Theorem below.

Definition 19. Two braids are *Markov equivalent* if their closure gives the same oriented knot.

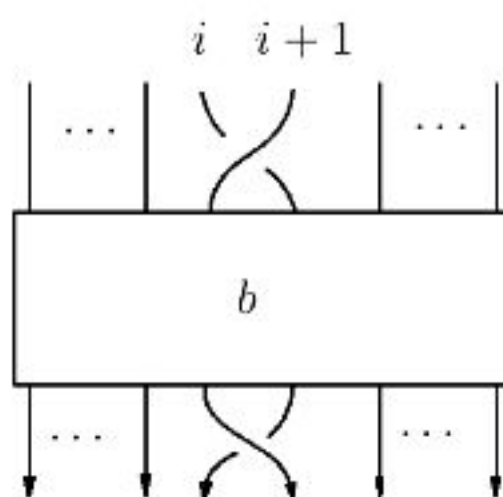
This requires two moves called, respectively, conjugation and stabilization:

- (1) $b \sim_M \sigma_i b \sigma_i^{-1}$,
- (2) $b \sim_M b \sigma_n$ or $b \sim_M b \sigma_n^{-1}$.

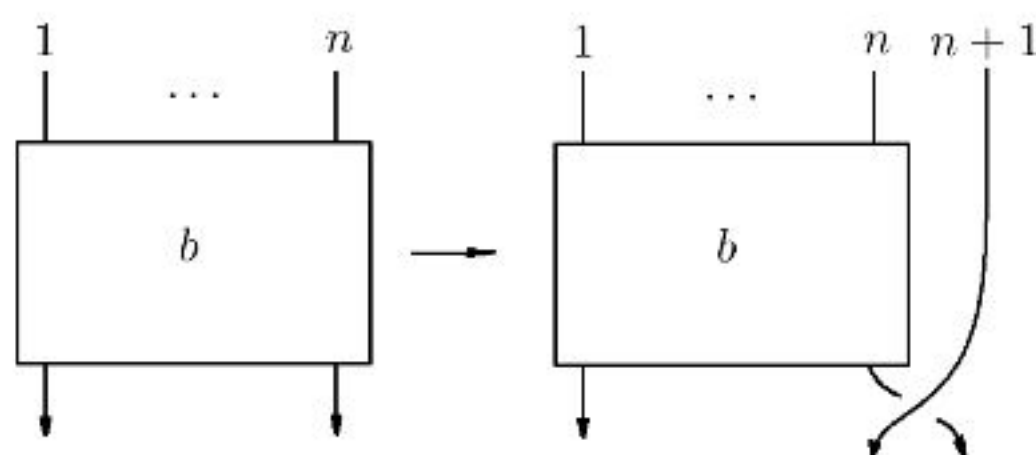
In this last relation we use the natural inclusion of the braid group B_n into the braid group B_{n+1} by adding a vertical string to the right of a braid b as can be seen from the following figure.



The conjugation move is given by the figure



The stabilization move is seen in the following figure:



Thus b is considered as an element of B_n while $b\sigma_n$ and $b\sigma_n^{-1}$ are considered as elements of B_{n+1} .

Theorem 7 (Markov's Theorem). *The closures of braids b and b' are isotopic links if and only if b' can be obtained from b by a sequence of Markov moves (conjugation and stabilization).*

Exercises. 1. Find another explicit example (than the one given above) of nonisotopic braids that generate isotopic links.

2. Check that the element $(\sigma_1\sigma_2\sigma_1)^2$ of the braid group B_3 commutes with σ_1 and with σ_2 and thus lies in the center $Z(B_3)$ of B_3 .

3. Prove that the relation between braids, $\beta_0 \sim \beta_1$ if and only if β_0 and β_1 are isotopic, is an equivalence relation.

4. Prove that the two braids $\beta_0 = \sigma_1\sigma_2$ and $\beta_1 = \sigma_2\sigma_1$ cannot be isotopic (Hint: use the natural mapping of the braid group B_3 to the symmetric group S_3).

5. A braid is called *pure* if for every k the k th strand on top ends at the k th position on the bottom; that is, a braid is pure if it induces the identity permutation in S_n . Prove that pure braids form a subgroup of B_n .

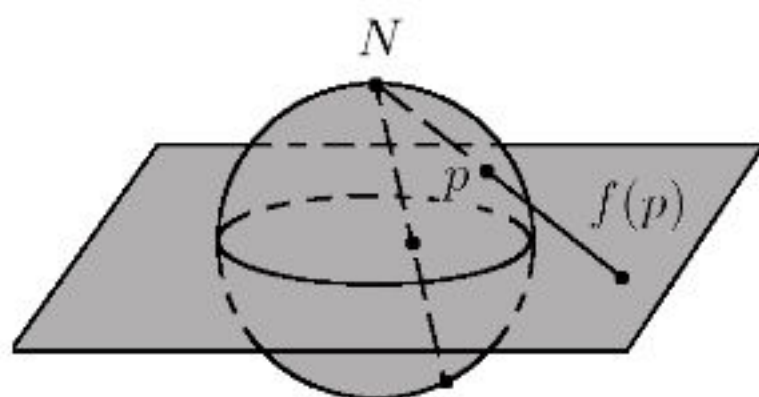
3. Knot Groups

Let us think about a knot in \mathbb{R}^3 . The space \mathbb{R}^3 is infinitely large, extending forever in three mutually perpendicular directions. In topology, however, size is a relative thing; for instance, an open interval like $(0, 1)$ is topologically the same as every other open interval, including infinite intervals like $(-\infty, \infty)$. For simplicity, we prefer to stick to sets which are finite in size, specifically those which have a property known as *compactness*¹.

It turns out that by adding a single point to \mathbb{R}^n , usually called “the point at infinity”, we get a compact space which is topologically the same as the set of all unit vectors in \mathbb{R}^{n+1} , known as the n -sphere:

$$S^n = \{ \vec{x} \in \mathbb{R}^{n+1} \mid \|\vec{x}\| = 1 \}.$$

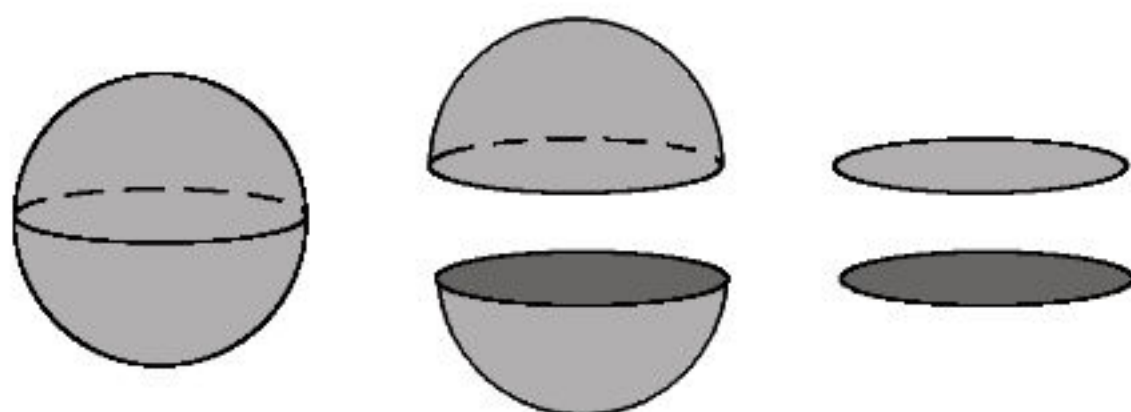
It is easier to see the correspondence in the $n = 2$ case since we can visualize both \mathbb{R}^2 and S^2 . The correspondence is called *stereographic projection*, and here’s how it works: think of \mathbb{R}^2 as the x - y plane in \mathbb{R}^3 and think of S^2 as the unit sphere in \mathbb{R}^3 , i.e., centered at the the origin and with radius 1.



Let us call the north pole of the sphere, i.e., the point $(0, 0, 1)$, N . Then given a point p on the sphere other than the north pole N , draw the line between N and p ; this line intersects the plane in a unique point $f(p)$. Conversely, given any point on the plane, the line joining

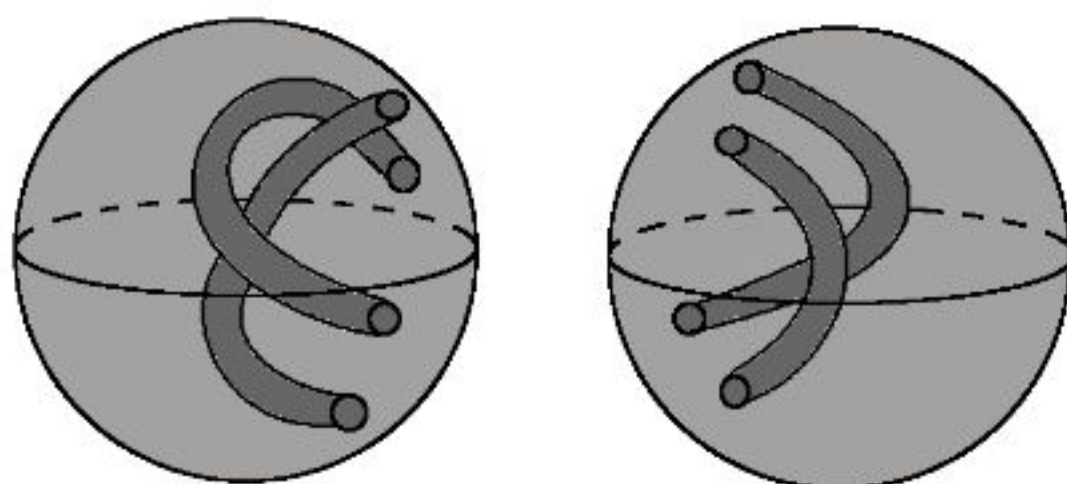
¹Formally, a set is *compact* if every time we can cover the set with open sets, it turns out that we only need a finite number of these open sets; for sets in \mathbb{R}^n , compactness is equivalent to being closed and bounded.

it and N intersects S^2 in a unique point. This correspondence maps the unit disk to the southern hemisphere, the origin to the south pole, and maps the northern hemisphere to the outside of the unit disk. In fact, we can think of the outside of the unit disc as another disc whose center is the point at infinity, corresponding to our north pole N . Thus, we can think of the sphere S^2 as the result of gluing two discs together along their boundary circles.



In a similar way, we can think of S^3 as a finite size version of \mathbb{R}^3 obtained by gluing together two solid balls along their boundary spheres, one inside and the other outside; the center of the outside ball is the north pole of S^3 , the “point at infinity” we add to \mathbb{R}^3 . Locally, S^3 looks exactly like \mathbb{R}^3 , but if you go far enough in the same direction you eventually come back to where you started.

Now, suppose we have a knot K inside of S^3 . The *knot complement* of K is the result of removing K from S^3 ; we might picture this as the result of drilling a K -shaped open tunnel out of S^3 . We can even draw a knot complement using the trick of dividing S^3 into two balls to be glued together:

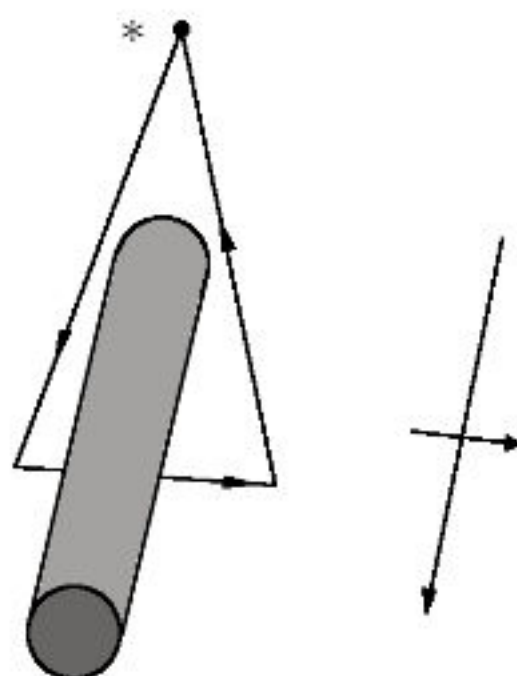


Of course, different choices of position for the dividing sphere will give us different pictures of the same knot complement. The dividing

sphere is known as a *Heegaard splitting*, an essential tool in the study of *3-manifolds*, sets of points which locally look like \mathbb{R}^3 . Note that the boundary of the knot complement is the knotted torus in the shape of K .

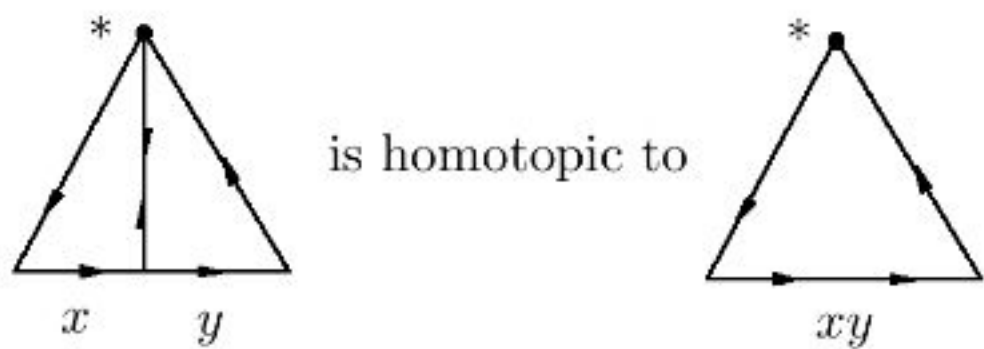
The *knot group* of a knot $K \subset S^3$ is the fundamental group of the knot complement $S^3 \setminus K$. Let us start with the case of the unknot. Choose a basepoint $*$ not on the knot. Then any loop based at $*$ either links the knot or does not. A loop which does not link the knot can be shrunk down to the basepoint in $S^3 \setminus K$ and thus is trivial. A loop which links K wraps around K an integer number of times; in fact, the number of times the loop wraps around K is exactly the linking number of the loop with K . (Negative linking number means wrapping around K backwards). It turns out that any two loops linking the unknot with the same linking number are homotopic, so the fundamental group of $S^3 \setminus K$ can be identified with the integers \mathbb{Z} .

More generally, a loop in $S^3 \setminus K$ is nontrivial in $\pi_1(S^3 \setminus K)$ if it links K . In particular, $\pi_1(S^3 \setminus K, *)$ is generated by loops which link each arc in a diagram of K exactly once. We usually indicate these on a diagram with a little arrow passing under the arc, which we can think of as the base of a triangular loop from a base point above the knot diagram.

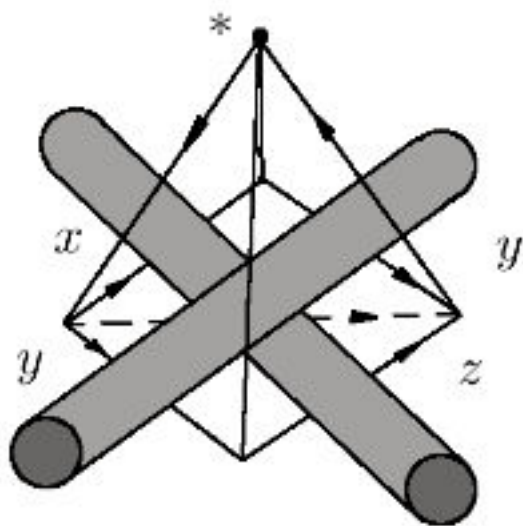


What is the relationship of these generators at a crossing? First, notice that if we have two triangular loops x and y which share a side forming the end of the first loop and the start of the second loop, then the product xy is homotopic to the loop given by the outside

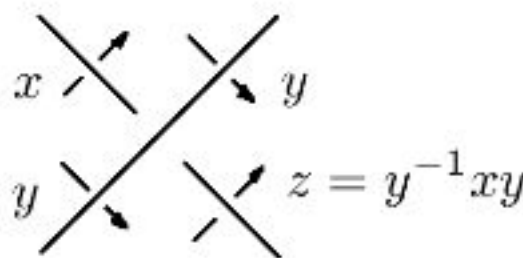
triangle, with homotopy given by following the first two-thirds of x , then just waiting at the midpoint for the last third of x and first third of y , then finishing y .



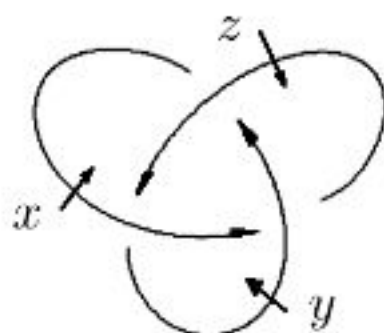
The four triangular loops around the crossing form the sides of a pyramid. In particular, the product of the loops yz is homotopic to the loop which goes diagonally across the pyramid (via the dashed line in the figure), which is homotopic to the product xy as shown.



Thus, given a knot diagram, we can get a presentation for the fundamental group of the knot complement by drawing a little arrow for each arc and getting a relation of the form $xy = yz$ or equivalently $z = y^{-1}xy$ at each crossing. This is called the *Wirtinger presentation* for the knot group.



Example 81. The trefoil knot K below has knot group with presentation $\pi_1(S^3 \setminus K) = \langle x, y, z \mid y = z^{-1}xz, x = y^{-1}zy, z = x^{-1}yx \rangle$.



Knot groups are always infinite; given any generator x of a knot group, the powers x^n for $n \in \mathbb{Z}$ are all distinct elements of the group. The only knot with an abelian knot group is the unknot; all the others have noncommuting generators.

Exercises. 1. Show that the following are two presentations of the same group (can you recognize which knot relates to this group):

$$\langle a, b \mid aba = bab \rangle$$

and

$$\langle x, y \mid x^2 = y^3 \rangle.$$

2. Recall that the braid group B_3 on three strands has the presentation $B_3 = \langle \sigma_1, \sigma_2 \mid \sigma_1\sigma_2\sigma_1 = \sigma_2\sigma_1\sigma_2 \rangle$. Use the natural mapping from the braid group B_3 to the symmetric group S_3 to deduce that the fundamental group of the trefoil is nonabelian. Furthermore deduce that the trefoil is not equivalent to the unknot.

3. Compute a presentation of the fundamental group of the figure eight knot. Map it by a group homomorphism to a symmetric group to deduce that the figure eight knot not equivalent to the unknot.

4. Draw a diagram of the connected sum of two left-handed trefoils (called *Granny* knot) and then find a presentation of its fundamental group.

5. Draw a diagram of the connected sum of a trefoil and its mirror image (called *square* knot) and then find a presentation of its fundamental group.

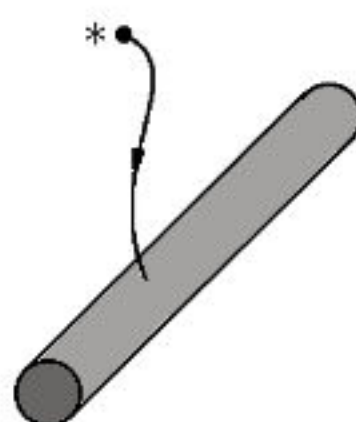
6. Show that there is no knot with knot group isomorphic to \mathbb{Z}_2 .

4. Knot Quandles

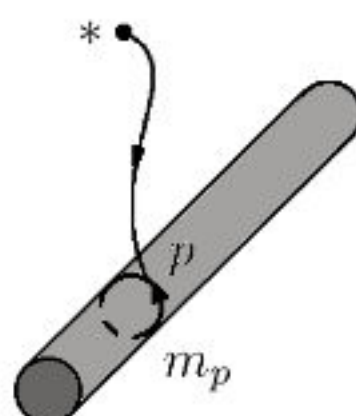
Let K be a knot in S^3 and choose a basepoint $* \in S^3 \setminus K$. The knot complement $S^3 \setminus K$ has boundary consisting of a knotted torus in the shape of K . We have already seen a combinatorial definition of the fundamental quandle $\mathcal{Q}(K)$ of K in terms of generators corresponding to arcs in the knot diagram with relations of the form $x \triangleright y = z$ at crossings. The fundamental quandle has a geometric interpretation as well, quite similar to the knot group but subtly different.

Elements of the fundamental quandle of a knot are homotopy classes not of loops but of *paths* from the basepoint to the boundary of the knot complement left by drilling out the K -shaped tunnel, with some restrictions:

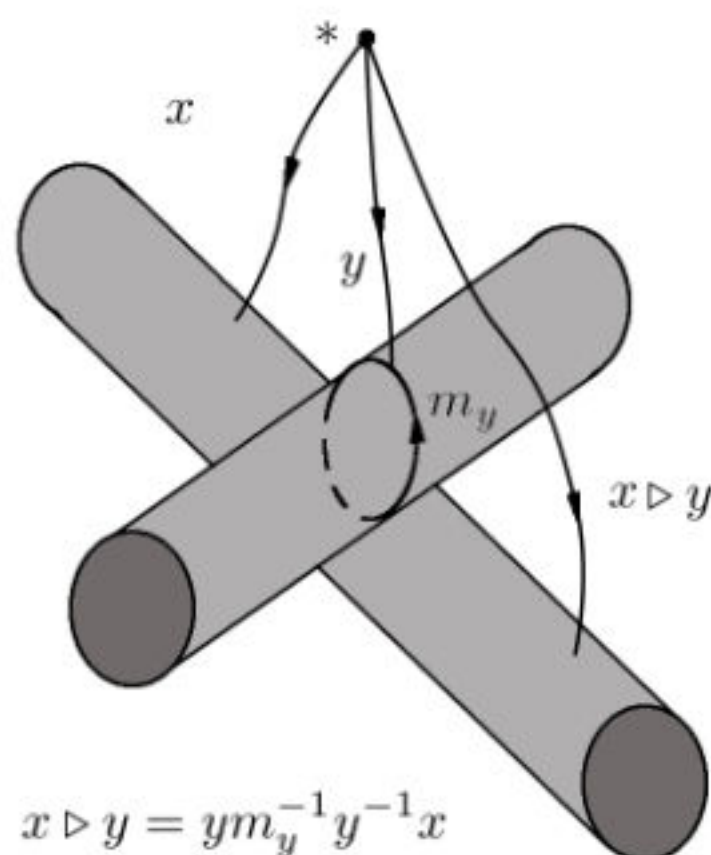
- The initial point of the path must stay fixed at the basepoint during the homotopy.
- The terminal point must stay on the boundary torus but can wander during the homotopy.



For every point p on the boundary of $S^3 \setminus K$ there is a circle on the boundary, unique up to isotopy, which links the original knot with linking number 1, which we call the *meridian* at p , denoted m_p . If p is the terminal point of a representative of the class of y , we will write m_y for m_p .



The quandle operation $x \triangleright y$ in the fundamental quandle of a knot is then given by first going along y , then going backward along the meridian at y and continuing backwards along y to the basepoint, then following x . We can see the homotopy by simply visualizing dragging the terminal point of x along the the boundary to the terminal point of $x \triangleright y$.

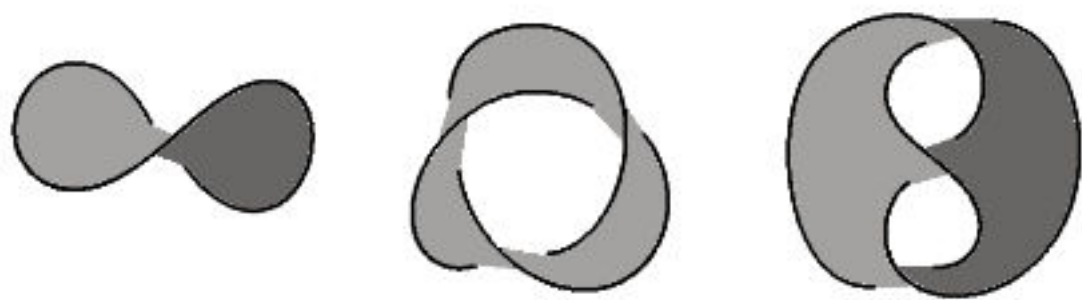


In his PhD dissertation [Joy82] in 1980, David Joyce proved that the fundamental quandle of a knot is a complete invariant up to reflection; that is, if K and K' have isomorphic fundamental quandles, then K is ambient isotopic to either K' or the reverse of the mirror image of K' . In a sense, this means that quandles really are knots translated into algebra; all other knot invariants should in principle be derivable from the fundamental quandle.

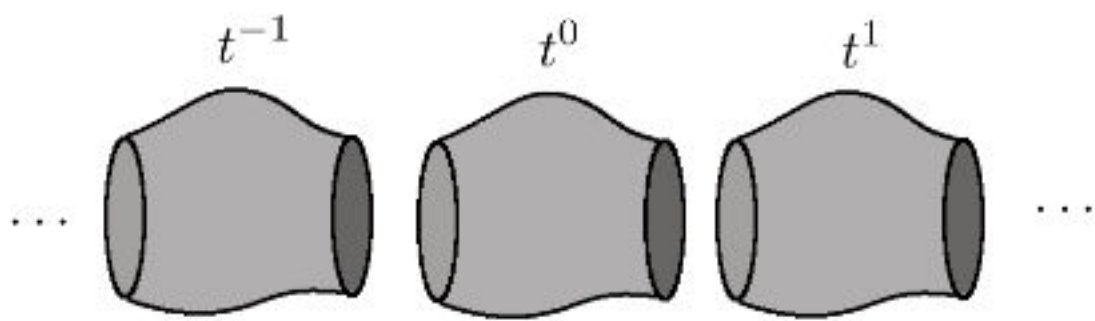
Alexander Quandles of Knots. Recall that the *Alexander quandle* $\mathcal{A}(K)$ of a knot is the fundamental quandle interpreted as an Alexander quandle. There is a geometric way to understand this as well, involving the *infinite cyclic cover* of the knot complement.

A *Seifert surface* for a knot K is an orientable surface (i.e., a surface with a well-defined top side and bottom side) with the knot K as its boundary. Note that not every surface with K as boundary is a Seifert surface; for example, the middle picture below shows a Möbius band with trefoil boundary, which is not a Seifert surface as

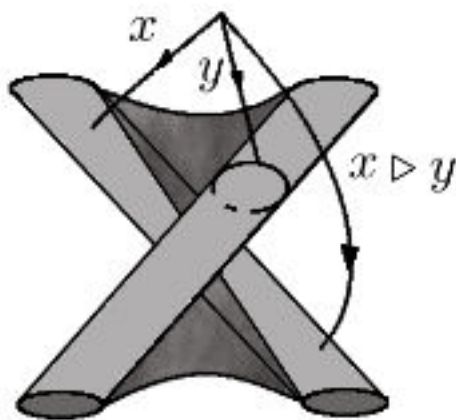
it has only one side.



Every knot has a Seifert surface – in fact, many; Seifert gave an algorithm showing how to construct a Seifert surface from a knot diagram. The fact that a Seifert surface has two sides means that if we cut the knot complement along the Seifert surface, the result will have two copies of the surface as its boundary, one from the top and one from the bottom. We can then place two copies of the cut-open knot complement with the top surface from one copy matching the bottom surface from the next copy and glue the surfaces together. If we repeat this with an infinite chain of copies of the knot complement indexed by positive and negative powers of t , the result is the infinite cyclic cover of $S^3 \setminus K$.



Locally, it looks just like a knot complement, but the Seifert surface acts like a kind of portal into the next copy; it's a bit like the “looking into infinity” resulting from two mirrors facing each other, except that each “mirror” leads to the next copy of the knot complement.



Elements of the Alexander quandle of a knot can be understood as \mathbb{Z} -linear combinations of homotopy classes of paths in the infinite cyclic cover from copies of the basepoint to the boundary torus with multiplication by t shifting a path into the next copy. In particular, we can visualize the Alexander quandle with a Seifert surface spanning a crossing acting like a portal into the next copy of the knot complement. Then the homotopy taking $x \triangleright y$ to $ym_y^{-1}y^{-1}x$ has $x \triangleright y$ and y on one side of the portal, m_y^{-1} takes us through the portal into the next copy, and we have t times $-y$ and x ; thus, we obtain

$$x \triangleright y = y - ty + tx = tx + (1 - t)y.$$

Exercises. 1. Compute a presentation of the fundamental quandle of the figure eight knot.

2. Draw a diagram to show that a path x from the basepoint to the boundary torus is homotopic to the path $x \triangleright x$ at a type I move.

3. Draw a picture to show the homotopy between x and $(x \triangleright y) \triangleright^{-1} y$ at a type II move.

4. Find an example of a Seifert surface for the figure eight knot.

5. Prove that the knot quandle of the Hopf link is the trivial quandle of two elements.

5. Augmented Quandles

We have seen before that quandles and groups are closely related. In Chapter 3 we considered two interesting groups coming from a given quandle: the group $\text{Aut}(X)$ of all automorphisms (self-homomorphisms that are bijective) of a quandle X and the group $\text{Inn}(X)$ generated by right multiplications β_x where $x \in X$. We have the map $\beta : X \rightarrow \text{Inn}(X)$ that sends x to β_x , where $\beta_x(y) = y \triangleright x$; this map is a quandle homomorphism from X to the conjugation quandle of the inner automorphism group. For any $x \in X$ the map β_x satisfies the equation

$$\beta_x(y \triangleright z) = \beta_x(y) \triangleright \beta_x(z)$$

which is another way of stating the third quandle axiom. In fact, we can think of the structure of a quandle as a kind of module with the elements of the automorphism group playing the role of scalars, an idea known as *augmented quandles*. To see how this works, we will need the concept of a *group action*.

Actions of Groups. Recall that if X is a set, then S_X denotes the set of permutations (bijections) of the set X . This is a group with composition of permutations. If $X = \{1, 2, \dots, n\}$, then we denote S_X by S_n .

Let G be a group and X be a set. A (left) action of G on X is a group homomorphism ϕ from G to S_X . The mapping $G \times X \rightarrow X$ sending (g, x) to $\phi(g)(x)$ (denoted gx) must then satisfy two axioms:

- (i) $ex = x$, for all x in X (where e is the identity element in the group G), and
- (ii) $g(hx) = (gh)x$ for all x in X and all elements g and h in G .

Note that $\phi(g^{-1}) = (\phi(g))^{-1}$, $\forall g \in G$, that is, the inverse of the map $x \mapsto gx$ is the map $x \mapsto g^{-1}x$. We also say that X is a G -set. For example, the group S_n acts on the set $\{1, 2, \dots, n\}$ with the group homomorphism ϕ being the identity. Let $GL_n(\mathbb{F})$ be the group of invertible n by n matrices with coefficients in \mathbb{F} , then $GL_n(\mathbb{F})$ acts on the space \mathbb{F}^n by matrix multiplication. Notice that any group G acts on itself by the map $G \times G \rightarrow G$ sending (g, h) to gh . For each $a \in G$, the map $\pi_a : G \rightarrow G$ sending x to axa^{-1} is called an inner automorphism of G . The set of all inner automorphisms of G is denoted by $\text{Inn}(G)$. It is a normal subgroup of $\text{Aut}(G)$, and more precisely, for every $a \in G, \rho \in \text{Aut}(G)$ we have the suspiciously familiar equation

$$\rho\pi_a\rho^{-1} = \pi_{\rho(a)}.$$

Augmented Quandles. Now let X be a quandle. The action of the group $\text{Inn}(X)$ on X can be generalized to define an action of a group G on a quandle X as a map $G \times X \rightarrow X$ sending (g, x) to gx such that

- (i) for all $g, h \in G$ and for all $x \in X$, $g(hx) = (gh)x$ and
- (ii) for all $g \in G$ and for all $x, y \in X$, $g(x \triangleright y) = (gx) \triangleright (gy)$.

For any quandle homomorphism f from X to itself, we have

$$(f\beta_x)(y) = f(y \triangleright x) = f(y) \triangleright f(x) = \beta_{f(x)}(f(y)).$$

This means that $\beta_{f(x)} = f\beta_x f^{-1}$. This identity and the relation $\beta_x(x) = x$ give rise to the notion of an *augmented quandle* by changing the group $\text{Inn}(X)$ to a general group G , which we call the *augmentation group*.

Definition 20. An *augmented quandle* consists of a pair (X, G) where X is a quandle, G is a group acting on X and an *augmentation* map $\epsilon : X \rightarrow G$ such that

- (i) for all x in X , $\epsilon(x)x = x$ and
- (ii) for all g in G and x in X , $\epsilon(gx) = g\epsilon(x)g^{-1}$.

Example 82. For a quandle X , there are two typical examples of augmented quandles, one with $G = \text{Aut}(X)$ and the other with $G = \text{Inn}(X)$, where the augmentation map ϵ in both cases is the map β mentioned above that sends x to β_x and the action is the natural one.

Example 83. For a knot K in S^3 , let $X = Q(K)$ be the fundamental quandle of the knot and $G = \pi_1(S^3 \setminus K)$ be the fundamental group of the knot complement $S^3 \setminus K$. The action of $\pi_1(S^3 \setminus K)$ on $X = Q(K)$ is by first going around a loop representing the element of G , then going down the path represented by the element of $Q(K)$. For $x \in X = Q(K)$, $\epsilon(x)$ is the homotopy class of the loop at the base point $*$ which traverses the arc, then the boundary of the disc counterclockwise, then the arc again back to the base point $*$. This gives an example of an augmented quandle: the knot quandle is an augmented quandle with augmentation group given by the knot group.

Before we give another example let us introduce another group associated to a quandle, the *enveloping* group G_x of a quandle X , also called the *associated group*. This is the group obtained by interpreting the quandle operation as a conjugation. It is given by the group presentation

$$G_x = \langle x \in X \mid (x \triangleright y)y^{-1}x^{-1}y \rangle$$

with all elements of X as generators. For a quandle (X, \triangleright) let $F(X)$ be the free group on the set X and consider the quotient

group $F(X)/N$ of $F(X)$ by the normal subgroup N generated by $(x \triangleright y)y^{-1}x^{-1}y$ where x and y belong to X . This group is denoted by G_X . We have an onto map $\epsilon : X \rightarrow G_X$ sending x to $[x]$. The enveloping group has a nice property (usually called the *universal property*) that for any group G and any given quandle homomorphism $\phi : (X, \triangleright) \rightarrow (G, *)$, where $g * h = hgh^{-1}$ for all $g, h \in G$, there exists a *unique* group homomorphism $\psi : G_X \rightarrow G$ such that $\psi \epsilon = \phi$.

Example 84. Any quandle can be thought of as an augmented quandle where the augmentation is the map $\epsilon : X \rightarrow G_X$.

Orbits, Stabilizers and Invariant Subspaces. The action of a group G on a set X gives an equivalence relation on X : $x \sim y$ if and only if there exists an element g in G such that $gx = y$. The equivalence class $[x]$ of an element x in X is called the *orbit* of x under the action of G (denoted by Gx), and the quotient set X/\sim is called the *space of orbits*. For example, the orbits of the action of G on itself by inner automorphisms are by definition the *conjugacy classes* of the group G (two elements x and y are in the same conjugacy class if there exists an element g in G such that $g^{-1}xg = y$). Another example is, if H is a subgroup (not necessarily normal) of a group G , then the orbits of the action of H on G ($(h, g) \mapsto hg, \forall h \in H$ and $g \in G$) are the classes of G modulo H . For $x \in X$, the set $G_x = \{g \in G, gx = x\}$ is a subgroup of G , called the *stabilizer* of x under the action of G . It is the pre-image of x under the surjective map $G \rightarrow Gx$ sending g to gx . Since $gx = hx$ is equivalent to $g^{-1}h \in G_x$ which is also equivalent to $[g] = [h]$ in the quotient space G/G_x , we then have a natural bijection $G/G_x \rightarrow Gx$ sending $[g]$ to gx . First one sees that $y \in G_{gx}$ is equivalent to $g^{-1}yg \in G_x$. This will make the mapping $G/G_x \rightarrow Gx$ sending $[g]$ to gx well defined since $[g] = [g']$ means $g' = gh$ for some $h \in H$ and then $g'x = (gh)x = g(hx) = gx$ because $h \in G_x$. This map is surjective by construction and $gx = g'x$ is equivalent to $g^{-1}g' \in G_x$, making it an injective mapping. We say that x is *invariant* under the action if for all g in G , we have $gx = x$, i.e. if the orbit of x is the singleton $\{x\}$ or $G_x = G$. If G is finite, then the cardinality of the orbit Gx is the quotient of the cardinality of G by the cardinality of the stabilizer G_x . If in addition, X is a

finite set, then the cardinality of X is given by the formula (called *class formula*)

$$|X| = \sum_{x \in A} \frac{|G|}{|G_x|}$$

where A is a subset of X containing one representative from each class.

Example 85. As an application, we prove that the cardinality of S_n is $n! = 2 \times 3 \times 4 \times \cdots \times n$ (n factorial). The group S_n acts on the set $\{1, \dots, n\}$; let us denote $|S_n| = t_n$. There is only one orbit for this action and the stabilizer of an element can be identified with S_{n-1} . Then $t_n = n t_{n-1}$ and thus $t_n = n!$.

Recall that if G is a group and H is a subgroup of G (not necessarily a normal subgroup) and we define a relation on G by $x \sim y$ if and only if $y^{-1}x \in H$, (meaning $x = yz$ for some $z \in H$), then it is straightforward to see that this relation is an equivalence relation and the equivalence class of g is $[g] = gH = \{gh, h \in H\}$. The classes form a partition of G and we denote by G/H the set of equivalence classes. Now if G is finite, then $|G/H| = \frac{|G|}{|H|}$ because the number of elements of each class equals $|H|$. As a consequence, we have the *Lagrange* theorem stating:

Theorem 8. *If G is a finite group and H is a subgroup of G , then $|H|$ divides $|G|$.*

Universal Quandles. Let X be a quandle and $G = \text{Aut}(X)$ its automorphism group. For any $\phi \in \text{Aut}(G)$, there is a quandle structure on the group G given by

$$g \triangleright h = h\phi(h^{-1}g)$$

which we might call a *universal quandle*. Note that if G is abelian, then this quandle is an Alexander quandle with $t = \phi$. For this

reason, we might think of this quandle structure on G as a “non-abelian Alexander quandle”. We verify the validity of two of the three quandle axioms below; the last one is left as an exercise.

(i) Consider $g \in G$. We have

$$\begin{aligned} g \triangleright g &= g\phi(g^{-1}g) \\ &= g\phi(1) \\ &= g(1) \\ &= g, \end{aligned}$$

(ii) $g \triangleright^{-1} h$ is given by $h\phi^{-1}(h^{-1}g)$, then we have

$$\begin{aligned} (g \triangleright h) \triangleright^{-1} h &= h\phi(h^{-1}g) \triangleright^{-1} h \\ &= h\phi^{-1}(h^{-1}h\phi(h^{-1}g)) \\ &= h\phi^{-1}(\phi(h^{-1}g)) \\ &= hh^{-1}g \\ &= g, \end{aligned}$$

and $(g \triangleright^{-1} h) \triangleright h = g$ is similar.

In his original 1982 paper introducing quandles [Joy82], David Joyce gave us the following construction which shows that *every* quandle is isomorphic to this type of quandle (or a slight generalization), hence the adjective “universal”.

Let X be a quandle with automorphism group $G = \text{Aut}(X)$ and let $p \in X$. As we have seen, the right multiplication map $\beta_p : X \rightarrow X$ with $\beta_p(q) = q \triangleright p$ is an automorphism of X , i.e., $\beta_p \in G$. Moreover, conjugation in G by β_p is an automorphism of G . Then G has quandle operation

$$x \triangleright y = y\beta_p y^{-1}x\beta_p^{-1}.$$

Define a map $e : G \rightarrow X$ (e for “evaluation”) by $e(g) = g(p)$. Then e is a homomorphism of quandles:

$$\begin{aligned}
 e(x \triangleright y) &= (x \triangleright y)(p) \\
 &= y(\beta_p(y^{-1}(x(\beta_p^{-1}(p)))))) \\
 &= y(\beta_p(y^{-1}(x(p \triangleright^{-1} p)))) \\
 &= y(\beta_p(y^{-1}(x(p)))) \\
 &= y(y^{-1}(x(p)) \triangleright p) \\
 &= y(y^{-1}(x(p))) \triangleright y(p) \\
 &= x(p) \triangleright y(p) \\
 &= e(x) \triangleright e(y).
 \end{aligned}$$

Now if X is *homogeneous*, i.e., if for every $p, p' \in X$ there is an automorphism $\phi : X \rightarrow X$ such that $\phi(p) = p'$, then e is surjective since every $p' \in X$ is $\phi(p)$ for some $\phi \in G$.

Next, let H be the stabilizer of p , i.e., the subgroup H of G such that the action of elements of H fix p :

$$H = \{\phi \in G \mid \phi(p) = p\}.$$

Then G is a union of cosets of H , $\phi H = \{\phi h : h \in H\}$. Moreover, this set of cosets G/H has quandle structure defined by

$$xH \triangleright yH = y\beta_p y^{-1}x\beta_p^{-1}H$$

and since $h(p) = p$ for all $h \in H$, e induces a surjective quandle homomorphism $\bar{e} : G/H \rightarrow X$ by

$$\bar{e}(\phi H) = \phi(H(p)) = \phi(p).$$

Finally, \bar{e} is injective since if $\bar{e}(xH) = \bar{e}(yH)$ then we have

$$x(H(p)) = y(H(p)) \iff y^{-1}(x(H(p))) = H(p) \iff y^{-1}(x(p)) = p,$$

so $y^{-1}x \in H$, and we have $yH = y(y^{-1}xH) = xH$. Hence, we have the following theorem:

Theorem 9 (Joyce, 1982). *If X is a homogeneous quandle with automorphism group G and H is the stabilizer of an element $p \in X$, then $X \cong \text{Aut}(G)/H$.*

If X is not homogeneous, then a similar construction yields a similar result, with the difference that we need to choose multiple p s. See [Joy82] for more.

Exercises. 1. Compute the inner group $\text{Inn}(R_3)$ of the dihedral quandle R_3 and then write all the details for the augmented quandle $(R_3, \text{Inn}(R_3))$.

2. Write all the details to show that

$$(X = \mathcal{Q}(3_1), G = \pi_1(S^3 \setminus 3_1))$$

is an augmented quandle, where \mathcal{Q} stands for the fundamental quandle and π_1 for the fundamental group.

3. Given a quandle operation table, explain how to identify the inner automorphism group.

4. Let $G = S_3$ be the group of permutations of 3 letters. Make the operation table for the quandle structure on G with $x \triangleright y = y\phi(y^{-1}x)$ with $\phi(x)$ given by conjugation by the transposition $[2, 1, 3]$.

5. Recall that $R_3 = \{1, 2, 3\}$ with quandle operation matrix

$$\begin{bmatrix} 1 & 3 & 2 \\ 3 & 2 & 1 \\ 2 & 1 & 3 \end{bmatrix}$$

is the dihedral quandle of 3 elements. Show that R_3 is homogeneous with automorphism group $G = S_3$. Then letting $\phi = [2, 1, 3]$ as in problem 4, find the stabilizer H of the element $p = 1$ and the operation table of G/H .

6. Let G be a group acting on a set X . Prove that the following relation on X is an equivalence relation: $x \sim y$ if and only if there exists an element g in G such that $gx = y$.

6. Quandles and Quasigroups

In this section, we will discuss the relation between quandles and some algebraic structures called quasigroups. More precisely, we will explain the relation between left and right distributive quasigroups

and the following types of quandles: Alexander, Latin and medial quandles. Two connections between quasigroups and quandles were established in [Smi92].

Self-distributivity appeared in 1929 in the work of Burstin and Mayer [BM29], where they studied quasigroups which are left and right distributive, i.e., satisfying

$$x \diamond (y \diamond z) = (x \diamond y) \diamond (x \diamond z)$$

and

$$(x \diamond y) \diamond z = (x \diamond z) \diamond (y \diamond z)$$

respectively. They proved that there are no distributive quasigroups of orders 2 or 6, observed that the group of automorphisms is transitive, and showed that such a quasigroup is idempotent.

Definition 21 ([Bru58]).

- (1) A *quasigroup* is a set Q , with a binary operation \diamond , such that for all $u \in Q$, the right “multiplication” β_u and left “multiplication” λ_u , by u , are both permutations.
- (2) If the operation \diamond has an identity element e in Q then the quasigroup is called a *loop* and is denoted (Q, \diamond, e) .

What does the definition of a quasigroup mean? The requirement that right and left multiplications are permutations means that the operation table for a quasigroup has no repeated elements in any row or column. Such a table is known as a *Latin square*. Moreover, in a quasigroup, we can “divide” from the right and from the left. In other words, the equation $x \diamond y = z$ has a solution in x , that is $x = \beta_y^{-1}(z)$ and the same equation $x \diamond y = z$ has a solution in y , that is $y = \lambda_x^{-1}(z)$. Sometimes we refer to this respectively as “division from the right” and “division from the left.”

Before giving examples, we mention the following charming story about the word “loop”. Here we quote from the article [Pf00] of Hala Orlik Pflugfelder,

It was at this point that the terminology of quasigroup theory underwent a historic change. It became apparent that it was necessary to distinguish

between two classes of quasigroups: those with and those without an identity element. A new name was needed to designate the system with identity. This occurred around 1942, among people of Albert's circle in Chicago, who coined the word "loop" after the Chicago Loop. For Chicago locals, the term "Loop" designated the main business area and the elevated train that literally made a loop around this part of the city. It was a brilliant choice in several senses. First, the word "loop" rhymes with "group". Second, it expresses a sense of closure. And third, it is short and simple, so that it could be easily adopted in other languages. Today, it is used in many languages, with slight variations: for example, DIE LOOP in German (first used by Pickert) and LUPA in Russian. The French are, of course, an original and nonconforming people, so in French it is LA BOUCLE.

We also note that "loops" in this sense should not be confused with the loops which form the elements of the fundamental group!

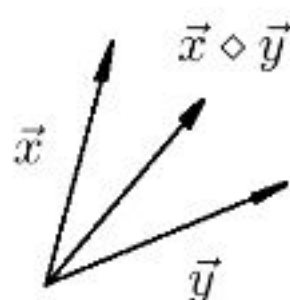
Example 86. Every group G is a loop since the equation $xy = z$ always has a solution for x , that is $x = zy^{-1}$, and also has a solution for y , that is $y = x^{-1}z$.

Example 87. The set \mathbb{Z} of integers with operation $x \diamond y = x - y$ is a quasigroup. Notice that this operation doesn't give a group structure on \mathbb{Z} since subtraction is nonassociative.

Example 88. The vector space \mathbb{R}^n with operation

$$\vec{x} \diamond \vec{y} = \frac{1}{2}(\vec{x} + \vec{y})$$

is a quasigroup.

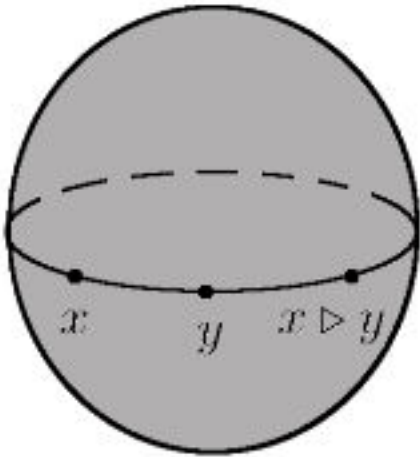


Example 89. The following operation table is a Latin square and thus makes $(\mathbb{Z}_6, *)$ a quasigroup.

$*$	0	1	2	3	4	5
0	3	1	4	5	2	0
1	0	4	3	2	5	1
2	1	5	0	4	3	2
3	4	3	2	1	0	5
4	5	2	1	0	4	3
5	2	0	5	3	1	4

One can see easily that this quasigroup is neither commutative nor associative and has no identity, thus it is not a group.

Example 90.



The kei in Example 56 can be generalized to any sphere

$$S^n = \{\vec{x} \in \mathbb{R}^n \mid ||\vec{x}|| = 1\};$$

with the operation

$$\vec{x} \diamond \vec{y} = 2(\vec{x} \cdot \vec{y})\vec{y} - \vec{x},$$

the sphere S^n is a quasigroup.

Definition 22. A subset S of a quasigroup (Q, \diamond) is called a *sub-quasigroup* of Q if S itself is a quasigroup with respect to the operation \diamond .

In other words, a sub-quasigroup S is a subset of Q which is closed under multiplication and division.

Example 91. The vector space \mathbb{Q}^n is sub-quasigroup of the quasigroup given in Example 88.

Definition 23. Let $(Q, *)$ and (Q', \diamond) be two quasigroups. The set $Q \times Q'$ with the operation $(x, y) \cdot (x', y') := (x * x', y \diamond y')$ is called the *direct product* of the quasigroups Q and Q' , analogous to the direct sum.

We leave it as an exercise to check that this operation gives a quasigroup structure on the Cartesian product of two quasigroups.

Definition 24. A map $f : Q \rightarrow Q'$ from a quasigroup $(Q, *)$ to a quasigroup (Q', \diamond) is called a *homomorphism* if for all $x, y \in Q$, we have

$$f(x * y) = f(x) \diamond f(y).$$

In quasigroup theory, the notion of homomorphism is often too strong, so it can be replaced by a notion of “homotopy” (do not confuse this notion of homotopy with the notion of homotopy of paths).

Definition 25. Let $f, g, h : Q \rightarrow Q'$ be three maps from a quasigroup $(Q, *)$ to a quasigroup (Q', \diamond) . The triple (f, g, h) is called a *homotopy* if for all $x, y \in Q$ we have

$$f(x) \diamond g(y) = h(x * y).$$

When f, g and h are all bijections then the triple (f, g, h) is called an *isotopy*.

This definition tells us that if we start with a quasigroup (Q', \diamond) and three bijections f, g, h from Q to Q' , then we can define a quasigroup structure on Q by

$$x * y = h^{-1}(f(x) \diamond g(y)), \forall x, y \in Q.$$

To see this, we check that left multiplication in Q is a bijection. The equation $x * y = x * z$ implies that $f(x) \diamond g(y) = f(x) \diamond g(z)$ since h^{-1} is injective. Now because (Q', \diamond) is a quasigroup and g is an injection we have $y = z$. This shows that left multiplication by any x is an injective mapping. Again since (Q', \diamond) is a quasigroup, then for a fixed $x \in Q$ and for any $y \in Q$, there exists $u \in Q'$ such that $f(x) \diamond u = h(y)$. Since g is a bijection there exists $v \in Q$ such that $u = g(v)$ thus $x * v = y$ and this shows the surjectivity of the left multiplication by x . It is similar to prove that right multiplication by any element is a bijection. This shows that isotopies are commonly

used to create quasigroups. We recover the definition of quasigroup homomorphism by setting $f = g = h$ in the previous definition.

Example 92. Consider (Q', \diamond) to be $(\mathbb{Z}_4, +)$. Let f be the bijection that permutes 0 with 1 and 2 with 3. Let g be the cycle permutation (0213) and h the cycle permutation (032) that fixes 1. We then obtain, using the previous definition, the quasigroup $(\mathbb{Z}_4, *)$ given by the following operation table:

$*$	0	1	2	3
0	0	2	3	1
1	3	0	1	2
2	1	3	2	0
3	2	1	0	3

Notice that the quasigroup $(\mathbb{Z}_4, *)$ is neither commutative nor associative and has no identity, even though $(\mathbb{Z}_4, +)$ is a group.

Example 93. A two element set $\{0, 1\}$ has exactly two quasigroup structures,

\diamond	0	1
0	0	1
1	1	0

and

\diamond	0	1
0	1	0
1	0	1

which are isotopic. Thus, up to isotopy, there is only one quasigroup of order 2 which is the group \mathbb{Z}_2 . Similarly, the only quasigroup of order 3 up to isotopy is the group \mathbb{Z}_3 , and the only two quasigroups of order 4 are the Klein group $\mathbb{Z}_2 \times \mathbb{Z}_2$ and the cyclic group \mathbb{Z}_4 .

Quasigroups differ from groups in the sense that they satisfy identities which usually conflict with associativity as can be seen from the previous examples. Distributive quasigroups have transitive groups of automorphisms (that is, for any two nonidentity elements x and y , there is an automorphism σ such that $\sigma(x) = y$) but the only group with this property is the trivial group. In [Ste57] it is shown that there are no right-distributive quasigroups whose order is twice an odd number, since in a right-distributive quasigroup we have

$$\beta_{y \diamond z} = \beta_z \beta_y \beta_z^{-1}$$

and the mapping $x \mapsto \beta_x$ is injective. In particular, a right-distributive quasigroup is a Latin quandle, so it follows that no quandle with $2(2k+1)$ elements can be Latin.

Definition 26 ([Bru58]). Let (M, \diamond) be a set with a binary operation. M is called a *Moufang loop* if it is a loop such that the binary operation satisfies the identity

$$(2) \quad (x \diamond y) \diamond (z \diamond x) = x \diamond ((y \diamond z) \diamond x).$$

The following are some other equivalent forms of the Moufang identity (2):

$$(3) \quad x \diamond (y \diamond (x \diamond z)) = ((x \diamond y) \diamond x) \diamond z,$$

$$(4) \quad z \diamond (x \diamond (y \diamond x)) = ((z \diamond x) \diamond y) \diamond x,$$

$$(5) \quad (x \diamond y) \diamond (z \diamond x) = (x \diamond (y \diamond z)) \diamond x,$$

As the name suggests, the Moufang identity is named for Ruth Moufang who discovered it in some geometrical investigations in 1935 (see for example [Mou33]). Like quandles, Moufang loops differ from groups in that they are generally not associative. The Moufang identities may be viewed as relaxed forms of associativity. We will see that the *smallest* nonassociative Moufang loop has order 12 and comes from the smallest nonabelian group.

Example 94. Any group is an associative loop and thus a Moufang loop. This follows directly from the definition.

Our next example uses the algebra of *quaternions*, invented by the Irish mathematician William Rowan Hamilton in 1843. It is similar to the algebra of complex numbers except that it is associative but not commutative. Precisely, any quaternion can be written uniquely in the form $q = a + bi + cj + dk$, where $a, b, c, d \in \mathbb{R}$. Thus the quaternions form a 4-dimensional vector space over the reals. The basis elements i, j and k satisfy the following multiplication properties: $i^2 = j^2 = k^2 = ijk = -1$. Thus multiplying any two elements $q = a + bi + cj + dk$ and $q' = a' + b'i + c'j + d'k$ gives

$$\begin{aligned} qq' &= (aa' - bb' - cc' - dd') + (ab' + ba' + cd' - dc')i \\ &\quad + (ac' + ca' + db' - bd')j + (ad' + da' + bc' - cb')k. \end{aligned}$$

Now to construct one of the simplest nonassociative loops of order 8, we need to modify slightly the multiplication of each basis element i, j, k with itself. We will have the following.

Example 95. The set $\{1, -1, i, -i, j, -j, k, -k\}$ with the the modified multiplication $i^2 = j^2 = k^2 = 1$ and all other products unchanged gives an example of a quasigroup. More precisely, it is a nonassociative loop of order 8 as can be seen from its operation table

\diamond	1	-1	i	$-i$	j	$-j$	k	$-k$
1	1	-1	i	$-i$	j	$-j$	k	$-k$
-1	-1	1	$-i$	i	$-j$	j	$-k$	k
i	i	$-i$	1	-1	k	$-k$	$-j$	j
$-i$	$-i$	i	-1	1	$-k$	k	j	$-j$
j	j	$-j$	$-k$	k	1	-1	i	$-i$
$-j$	$-j$	j	k	$-k$	-1	1	$-i$	i
k	k	$-k$	$-j$	j	$-i$	i	1	-1
$-k$	$-k$	k	j	$-j$	i	$-i$	-1	1

This example can be generalized to give examples of nonassociative Moufang loops.

Example 96. Let G be a group denoted multiplicatively and consider the set $G \times \{0, 1\}$ with multiplication given by $(g, 0)(h, 0) = (gh, 0)$, $(g, 0)(h, 1) = (hg, 1)$, $(g, 1)(h, 0) = (gh^{-1}, 0)$, and $(g, 1)(h, 1) = (h^{-1}g, 0)$. This set is denoted $M(G, 2)$. In 1974 Orin Chein [Che74] proved that $M(G, 2)$ is a nonassociative Moufang loop if and only if G is nonabelian group. Thus the smallest nonassociative Moufang loop is $M(D_3, 2)$ with order 12, where D_3 is the symmetry group of an equilateral triangle.

Theorem 10 (Moufang’s Theorem). *Let a, b, c be three elements in a commutative Moufang loop M for which the following relation holds:*

$$(a \diamond b) \diamond c = a \diamond (b \diamond c).$$

Then the subloop generated by $\{a, b, c\}$ is associative and hence is an abelian group.

Let (X, \diamond) be a right-distributive quasigroup. Then

$$(x \diamond x) \diamond x = (x \diamond x) \diamond (x \diamond x).$$

This implies that each element is idempotent and (X, \diamond) is then a Latin quandle. Fix $a \in X$ and define an operation, denoted $+$, on X by

$$x + y = \beta_a^{-1}(x) \diamond \lambda_a^{-1}(y).$$

Then $a + y = y$ and $y + a = y$, and $(X, +, a)$ is a loop. Therefore any right-distributive quasigroup satisfying one of the Moufang identities (3), (4), (5) and (2) is a Moufang loop. Notice that

$$\beta_a(x) + \lambda_a(y) = x \diamond y.$$

The Moufang loop is commutative if and only if

$$(6) \quad (u \diamond v) \diamond (w \diamond z) = (u \diamond w) \diamond (v \diamond z).$$

A set (X, \diamond) with a binary operation that satisfies equation (6) is said to be *medial* (Belousov [Bel60]) or *abelian* (Joyce [Joy82]). The Bruck-Toyoda theorem gives the following characterization of medial quasigroups: Given an Abelian group M , two commuting automorphisms f and g of M and a fixed element a of M , define an operation \diamond on M by

$$x \diamond y = f(x) + g(y) + a.$$

This quasigroup is called an *affine* quasigroup. It is easy to check that (M, \diamond) is a medial quasigroup; the Bruck-Toyoda theorem states that *every* medial quasigroup is of this form. That is, every medial quasigroup is isomorphic to a quasigroup defined from an abelian group in this way. Belousov gave the connection between distributive quasigroups and Moufang loops in the following way:

Theorem 11 ([Bel60]). *If (X, \diamond) is a distributive quasigroup, then for all $a \in X$, $(X, +, a)$ is a commutative Moufang loop.*

Latin quandles are right distributive quasigroups and left distributive Latin quandles are distributive quasigroups. Belousov's theorem tells us that if (X, \diamond) is a left-distributive Latin quandle then $(X, +)$ is a commutative Moufang loop. The Bruck-Slaby theorem tells us that (X, \diamond) is affine over a commutative Moufang loop, and thus medial. The smallest Latin quandle that is not left distributive is of order 15. It was found by David Stanovsky (see [Sta04], p. 29) using an automatic model builder SEM for all quasigroups satisfying left distributivity, but not mediality.

Example 97. In the survey paper [Gal88, p. 950], Galkin defines a type of quandle, later generalized in [CEH⁺13]. Let A be an abelian group, also regarded naturally as a \mathbb{Z} -module. Let $\mu : \mathbb{Z}_3 \rightarrow \mathbb{Z}$, $\tau : \mathbb{Z}_3 \rightarrow A$ be functions (not necessarily homomorphisms) satisfying $\mu(0) = 2$, $\mu(1) = \mu(2) = -1$, and $\tau(0) = 0$. Define a binary operation \triangleright on $\mathbb{Z}_3 \times A$ by

$$(x, a) \triangleright (y, b) = (2y - x, -a + \mu(x - y)b + \tau(x - y))$$

for $x, y \in \mathbb{Z}_3$ and $a, b \in A$. Then this operation \triangleright defines a quandle structure on $\mathbb{Z}_3 \times A$ called a *Galkin quandle*. One can see that for any abelian group A and $c_1, c_2 \in A$, $G(A, c_1, c_2)$ and $G(A, 0, c_2 - c_1)$ are isomorphic.

Each Galkin quandle $G(A, c)$ is *connected*, i.e. every element can be expressed as a quandle word starting with any other element, but not Latin unless A has odd order, and $G(A, c)$ is nonmedial unless $3A = 0$.

We conclude with the following properties relating distributivity and medality to quandles. Alexander quandles are left-distributive and medial. It is easy to check that for a finite Alexander quandle (M, t) with $t \in \text{Aut}(M)$, the following are equivalent:

- (1) (M, t) is connected,
- (2) $(1 - t)$ is an automorphism of M , and
- (3) (M, t) is Latin.

It was also proved by Toyoda [Toy41] that a Latin quandle is Alexander if and only if it is medial. As noted by Galkin, $G(\mathbb{Z}_5, 0)$ and $G(\mathbb{Z}_5, 1)$ are the smallest nonmedial Latin quandles and hence the smallest non-Alexander Latin quandles.

We note that medial quandles are left-distributive (by idempotency). It is proved in [CEH⁺13] that any left-distributive connected quandle is Latin. This implies, by Toyoda's theorem, that every medial connected quandle is Alexander and Latin. The smallest Latin quandles that are not left-distributive are the Galkin quandles of order 15. It is known that the smallest left-distributive Latin quandle that is not Alexander is of order 81, as proven by V. D. Belousov [Bel60].

Exercises. 1. Prove that all Alexander quandles are medial.

2. Let (X, \triangleright) be a set with binary operation that is medial. Let $x, y \in X$ and m and n be positive integers. Define x^n inductively by $x^{n+1} := x^n \triangleright x$. Prove that $(x \triangleright y)^n = x^n \triangleright y^n$, $(x^n)^m = (x^m)^n$ and thus $x^{nm} := (x^n)^m$ is well defined.

3. Let (X, \triangleright) be a set with binary operation satisfying $(x \triangleright y) \triangleright y = x$ and $x \triangleright (x \triangleright y) = y$ then prove that the operation \triangleright is commutative ($x \triangleright y = y \triangleright x$).

4. Let $f : Q \rightarrow Q'$ be a quasigroup homomorphism. Prove that for all $x, y \in Q$ we have the following two identities:

$$\beta_{f(y)}^{-1}(f(x)) = f(\beta_y^{-1}(x))$$

and

$$\lambda_{f(y)}^{-1}(f(x)) = f(\lambda_y^{-1}(x)).$$

5. Let $(Q, *, e)$ and (Q', \diamond, e') be two loops and $f : Q \rightarrow Q'$ be a quasigroup homomorphism. Prove that $f(e) = e'$.

6. Let $(Q, *)$ and (Q', \diamond) be two quasigroups. Prove that componentwise multiplication and division give a quasigroup structure on the Cartesian product $Q \times Q'$ of two quasigroups.

7. Let $f : Q \rightarrow Q'$ be a function between two quasigroups Q and Q' . Prove that f is a quasigroup homomorphism *if and only if* its graph $\Gamma = \{(x, y) \in Q \times Q'; f(x) = y\}$ is a sub-quasigroup of the quasigroup product $Q \times Q'$.

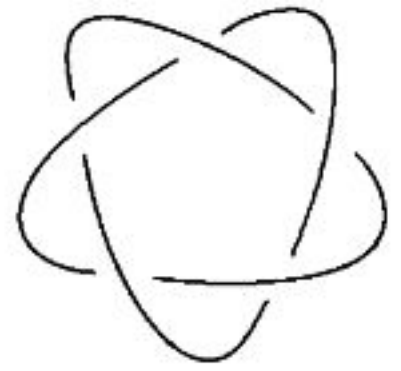
8. Prove that the isotopy relation is an equivalence relation on the set of quasigroups.

9. Consider the set \mathbb{Z}_3 of integers mod 3 with the binary operation $x \diamond y = 2x + 2y$. Prove that $(\mathbb{Z}_3, +)$, $(\mathbb{Z}_3, -)$ and (\mathbb{Z}_3, \diamond) are isotopic quasigroups.

10. Consider the set \mathbb{R} of real numbers with the binary operation $x \diamond y = \frac{1}{2}(x+y)$. Prove that $(\mathbb{R}, +)$ and (\mathbb{R}, \diamond) are isotopic quasigroups.

Chapter 5

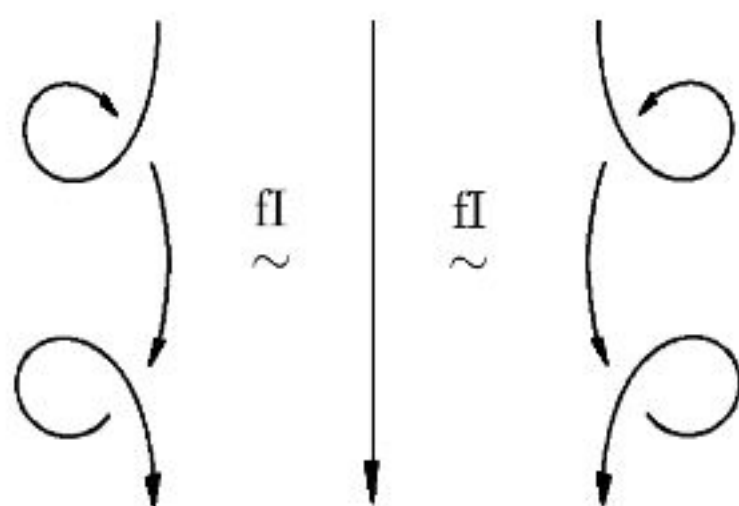
Generalizations of Quandles



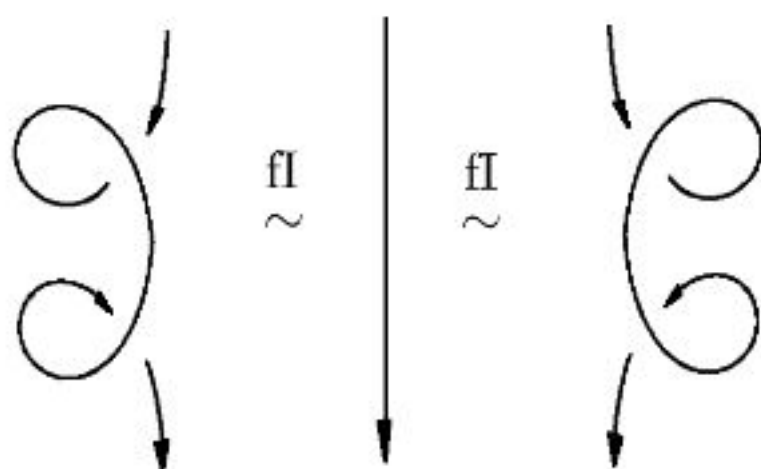
1. Racks

Recall that a *framed knot* is a knot with a framing curve on the the torus neighborhood of the knot which maps one-to-one onto the knot if we contract the torus down to its core. Alternatively, if we think of the torus neighborhood of the knot as a stack of discs with the points of the knot as centers of the discs, then the framing curve runs along the boundary of the torus and intersects each disc exactly once. Then two knots are *framed isotopic* if there is an ambient isotopy of one knot onto the other which takes the framing curve of the first knot to the framing curve of the second knot.

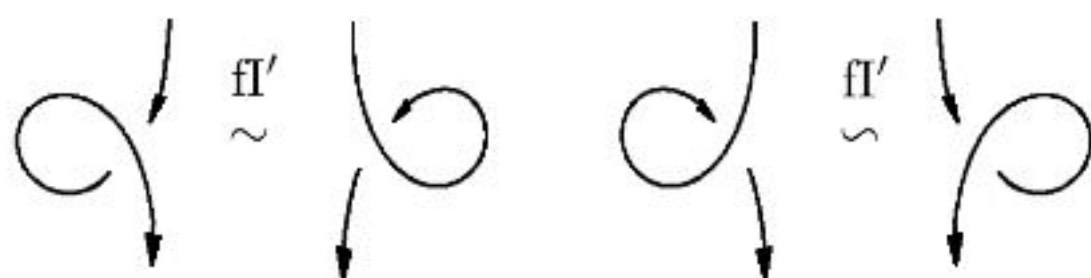
We can think of framed knots combinatorially as the result of changing the Reidemeister moves to replace the usual type I move with the *framed type I move*:



The framed Reidemeister move as we have drawn it above comes in the two pictured forms; in fact, there are two more equivalent versions which we can obtain from the pictured moves by pushing a kink all the way around the knot to the other side, a procedure which only requires type II and III moves:

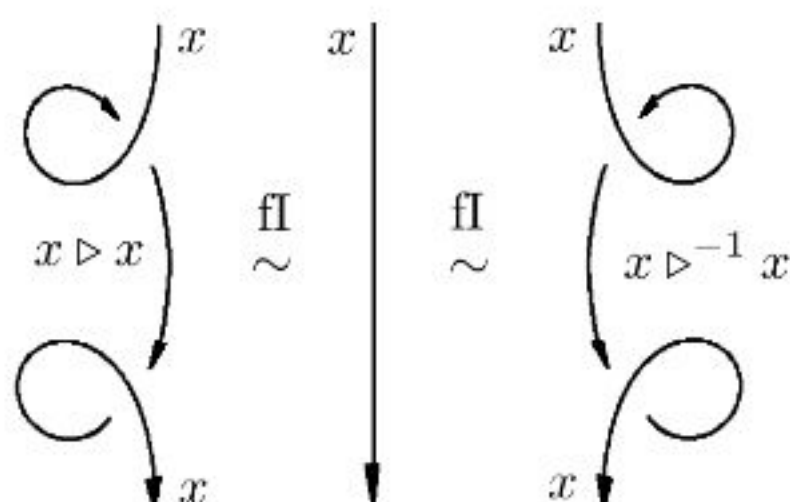


As we saw back in Chapter 1, in the presence of type II and III moves, these framed type I moves are equivalent to the alternative framed type I moves below:



What is the result of “quandlizing” the framed Reidemeister moves, i.e., replacing the quandle axioms coming from the usual oriented Reidemeister moves with the framed ones? Well, the second and third moves are the same, so those axioms are also the same.

The first quandle axiom, that $x \triangleright x = x$ for all x , is no longer required once we switch to framed isotopy. In fact, it turns out that the framed type I move imposes no conditions on the algebraic structure at all.



Thus, we have a new definition:

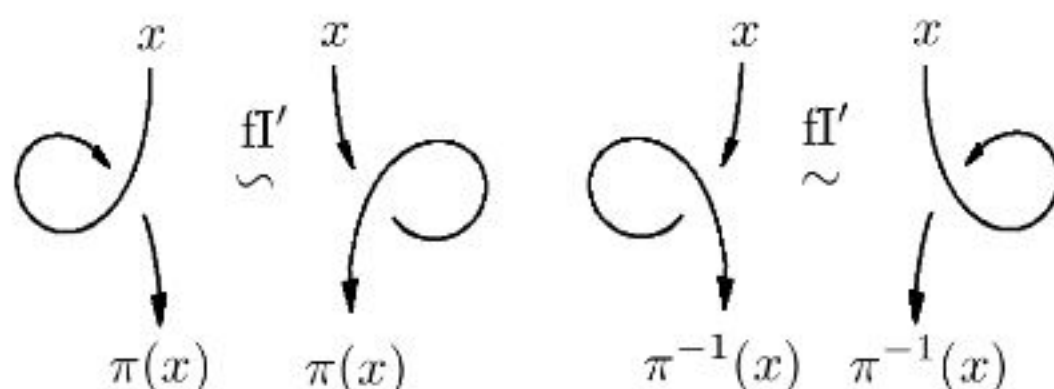
Definition 27. A *rack* is a set X with two binary operations $\triangleright, \triangleright^{-1} : X \times X \rightarrow X$ satisfying

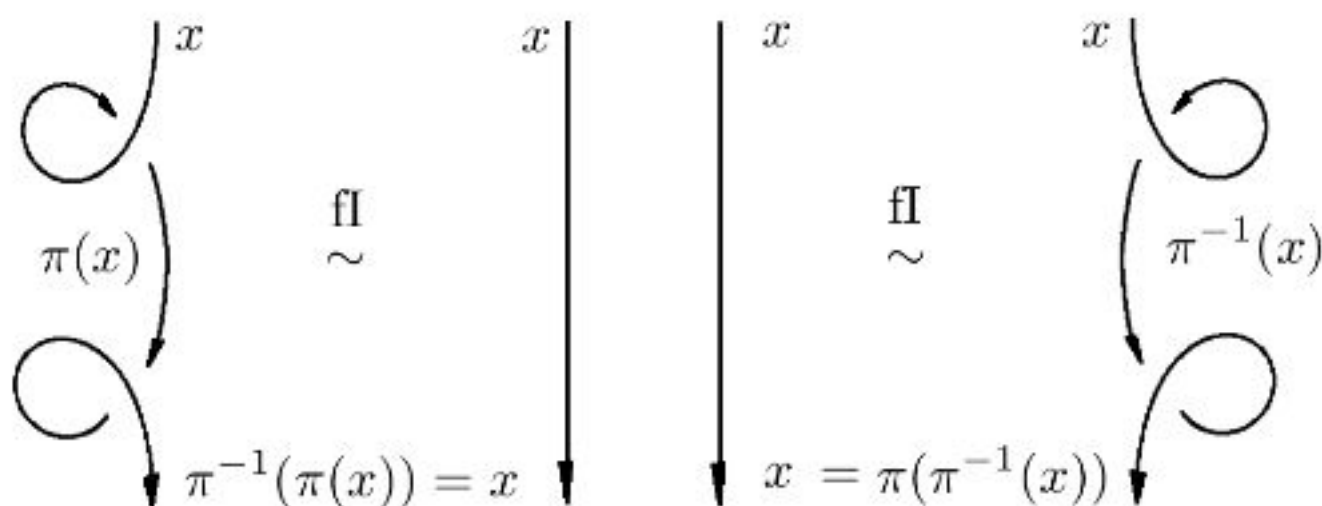
- $(x \triangleright y) \triangleright^{-1} y = x = (x \triangleright^{-1} y) \triangleright y$ and
- $(x \triangleright y) \triangleright z = (x \triangleright z) \triangleright (y \triangleright z)$.

Note that in some places in the literature, e.g. [FR92], $x \triangleright y$ is written as x^y .

We can think of racks as “almost-quandles” where some elements are not idempotent. The idempotency in a quandle comes from the Reidemeister type I move; specifically, we can think of the quandle axiom $x \triangleright x = x$ as the requirement that passing through a kink does not change the label.

In framed isotopy on the other hand, going through a kink is a bijective map $\pi : X \rightarrow X$ defined by $\pi(x) = x \triangleright x$ with inverse $\pi^{-1}(x) = x \triangleright^{-1} x$ known as the *kink map*. The alternate form of the framed isotopy moves show that the crossing sign at a kink determines whether the map is π or π^{-1} regardless of winding number.





Consider a finite rack X . Even if $\pi(x) \neq x$, if we go through enough kinks we eventually run out of new labels. For any $x \in X$, the *rank* of x is the smallest positive integer n such that $\pi^n(x) = x$. The least common multiple of all such n for all elements of X is known as the *rack characteristic* or *rack rank* of X . Equivalently, we can define N as the minimal integer $n \geq 1$ such that $\pi^n : X \rightarrow X$ is the identity map.

Example 98. Every quandle is a rack. In fact, we could define a quandle as a rack in which every element is idempotent, or as a rack in which the kink map is the identity, or as a rack of characteristic $N = 1$.

Example 99. Let X be a set and $\sigma : X \rightarrow X$ a bijection. Then X is a rack with operations

$$x \triangleright y = \sigma(x) \quad \text{and} \quad x \triangleright^{-1} y = \sigma^{-1}(x)$$

since we have

$$(x \triangleright y) \triangleright z = \sigma^2(x) = (x \triangleright z) \triangleright (y \triangleright z).$$

We call this a *constant action rack* since the action of y on x in $x \triangleright y$ does not vary with y but is constant as a function of y . In particular, the kink map π is just σ .

Example 100. Let $\ddot{\Lambda} = \mathbb{Z}[t^{\pm 1}, s]/(s^2 - s(1 - t))$ be the quotient of the set of polynomials with invertible variable t and noninvertible variable s where we set $s^2 = s(1 - t)$. Then any $\ddot{\Lambda}$ -module X is a rack (known as a (t, s) -rack) with rack operation

$$\vec{x} \triangleright \vec{y} = t\vec{x} + s\vec{y}.$$

We can easily verify that the rack axioms are satisfied: first, we have

$$\begin{aligned}\vec{x} \triangleright \vec{y} &= t\vec{x} + s\vec{y}, \\ \vec{x} \triangleright \vec{y} - s\vec{y} &= t\vec{x}, \\ t^{-1}(\vec{x} \triangleright \vec{y} - s\vec{y}) &= \vec{x},\end{aligned}$$

so we have $x \triangleright^{-1} y = t^{-1}(x - sy)$. Then for self-distributivity, we have

$$(\vec{x} \triangleright \vec{y}) \triangleright \vec{z} = t(t\vec{x} + s\vec{y}) + s\vec{z} = t^2\vec{x} + ts\vec{y} + s\vec{z}$$

while

$$\begin{aligned}(\vec{x} \triangleright \vec{z}) \triangleright (\vec{y} \triangleright \vec{z}) &= t(t\vec{x} + s\vec{z}) + s(t\vec{y} + s\vec{z}) \\ &= t^2\vec{x} + ts\vec{y} + (ts + s^2)\vec{z} \\ &= t^2\vec{x} + ts\vec{y} + (ts + s(1 - t))\vec{z} \\ &= t^2\vec{x} + ts\vec{y} + s\vec{z}\end{aligned}$$

as required. Since $\pi(\vec{x}) = \vec{x} \triangleright \vec{x} = t\vec{x} + s\vec{x} = (t + s)\vec{x}$, X has finite characteristic N if and only if $(t + s)^N = 1$ for some $N \geq 1$. For example, the (t, s) -rack $\ddot{\Lambda}^n$ of ordered n -tuples of $\ddot{\Lambda}$ has infinite characteristic.

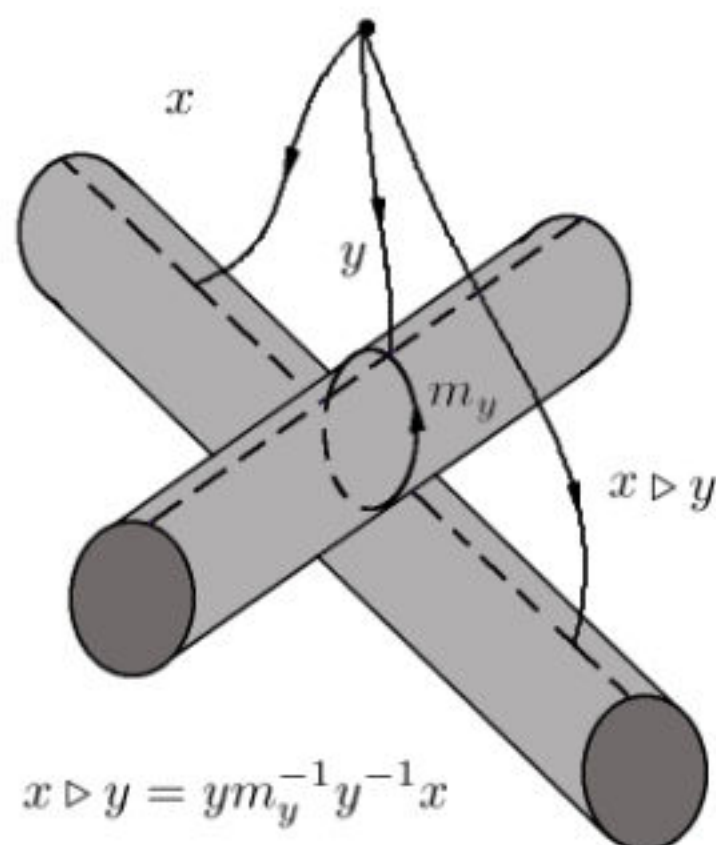
Example 101. Consider $X = \mathbb{Z}_n$. We can make X a (t, s) -rack by choosing values of t and s in \mathbb{Z}_n such that t and n have greatest common divisor 1 and such that $s^2 = s(1 - t)$. For example, in \mathbb{Z}_4 , we could set $t = 3$ and $s = 2$; then 3 is invertible in \mathbb{Z}_4 with $3^{-1} = 3$ (since $3(3) = 9 = 1$), and we have $s^2 = 2^2 = 4 = 0$ and $s(1 - t) = 2(1 - 3) = 2(-2) = -4 = 0$. This (t, s) -rack has operation table

\triangleright	0	1	2	3
0	0	2	0	2
1	3	1	3	1
2	2	0	2	0
3	1	3	1	3

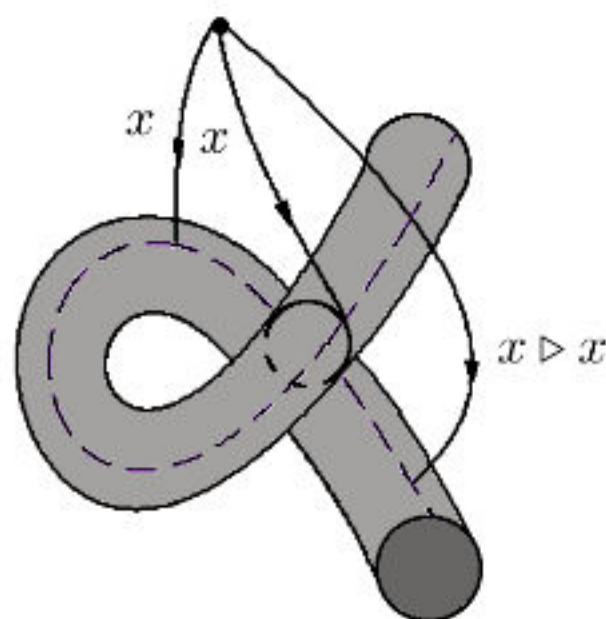
Rack Counting Invariant. As with quandles, each framed oriented knot K has a *fundamental rack* $\mathcal{R}(K)$ which we can define topologically or combinatorially. Let's consider the topological version first.

Recall that a framed knot can be understood as a pair consisting of a knot K and a framing curve F on the solid torus with K as the

core. Then as with the fundamental quandle, elements of the fundamental rack are homotopy classes of paths starting at a basepoint but this time ending at the framing curve, with the restriction that during the homotopy the starting point has to stay fixed at the basepoint and the endpoint has to stay on the framing curve. The rack operation is the same as in the quandle case:

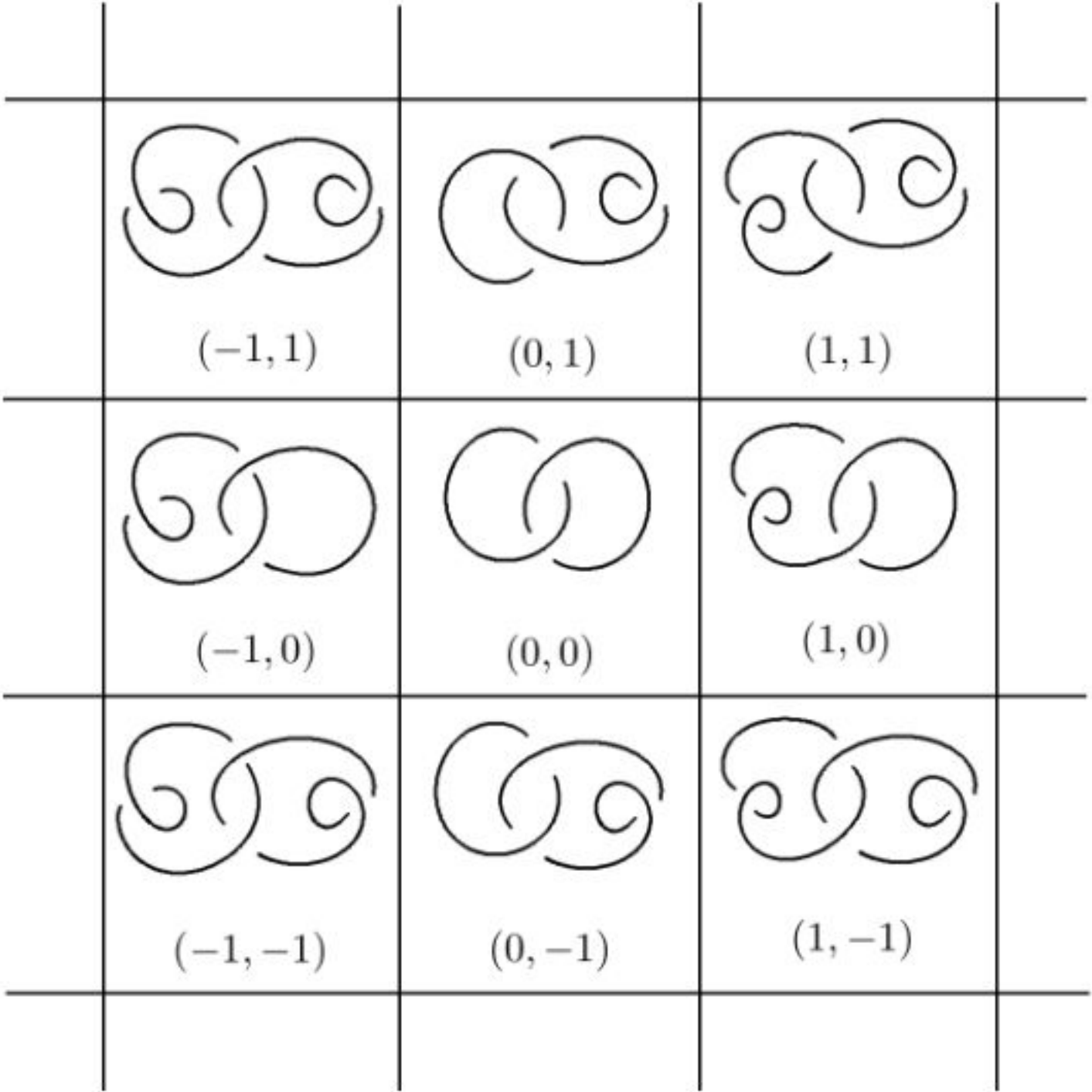


Using the blackboard framing, we can see why the quandle axiom $x \triangleright x = x$ works in the quandle case but fails in the rack case: the homotopy taking $x \triangleright x$ to x requires the terminal point to go around a meridian of the boundary torus of the knot complement, which takes it off the blackboard framing curve.



Let L be a link with c components. Each component has its own writhe or framing independent of the writhes of the other components.


Thus, a link with c components K_1, \dots, K_c has a *framing vector* $\vec{w} = (w_1, \dots, w_c)$ specifying the framing of each component. The set of all such framing vectors forms an *integral lattice*, that is, a copy of the set of all points in \mathbb{R}^n with integer coordinates, and for each writhe vector \vec{w} there is distinct framed version of L which we denote by $L_{\vec{w}}$.



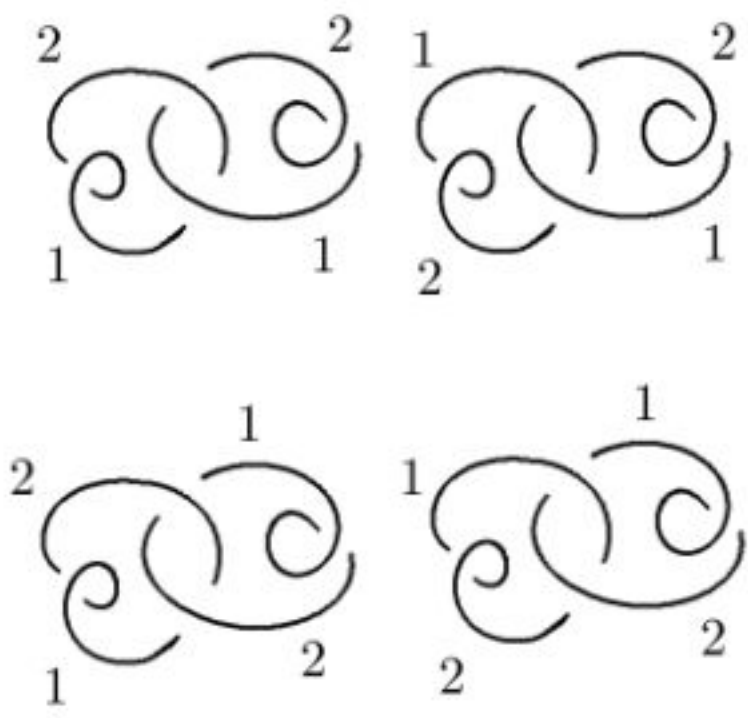
These framed links are all different in general, with different fundamental racks. Indeed, we can often distinguish links $L_{\vec{w}}$ and $L_{\vec{w}'}$ with different framing vectors $\vec{w} \neq \vec{w}'$ with rack counting invariants.

Example 102. The Hopf link with framing $\vec{w} = (0, 0)$ has no valid rack labelings by the rack $X = \{1, 2\}$ with operation table

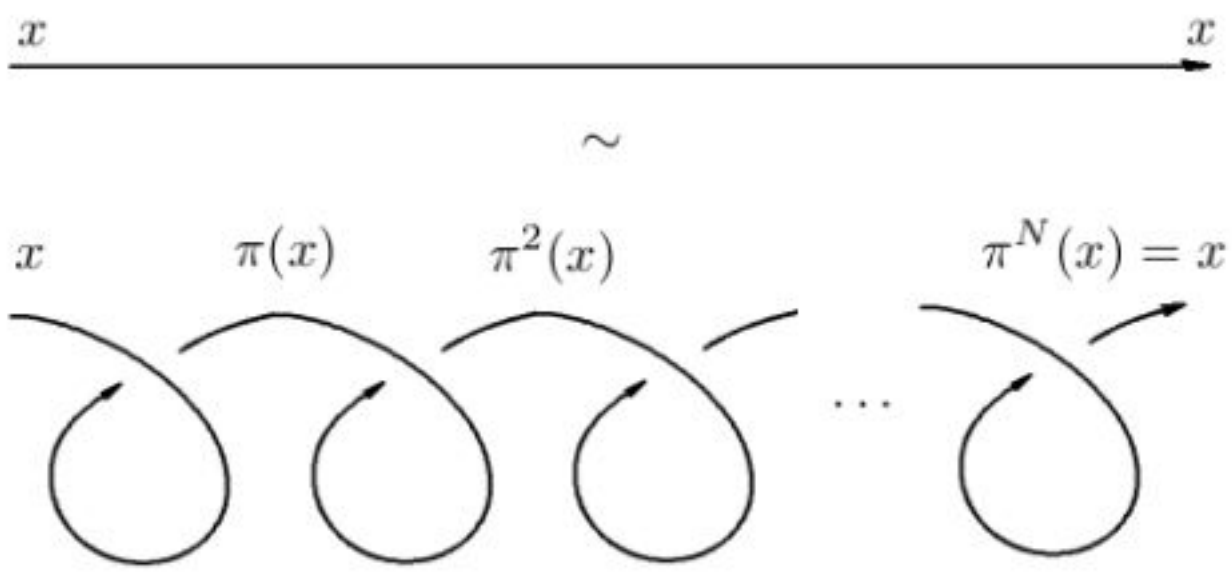
\triangleright	1	2
1	2	2
2	1	1

x  y

since for $x = 1$ we need $y = y \triangleright 1$ but $1 \triangleright 1 = 2$ and $2 \triangleright 1 = 1$, and the situation is the same for $x = 2$. However, if we change the framing vector to $\vec{w}' = (1, 1)$, then there are four X -labelings of L .



Now suppose our labeling rack X has finite characteristic N . Then if we have a rack labeling of a link L and we change the framing by doing an N -phone cord move



we get a unique valid X -labeling of the new framed link. In particular, if X is a rack of characteristic N and L and L' are framed links which are related by framed Reidemeister and N -phone cord moves, then $|\text{Hom}(\mathcal{R}(L), X)| = |\text{Hom}(\mathcal{R}(L'), X)|$.



Framed links with different framings which are equivalent by N -phone cord moves are still different framed links, even though they have the same number of labelings with respect to racks whose characteristic is a multiple of N .

Consider the integral lattice of framings of a link L and the numbers of X -colorings by a rack X with characteristic N . Since any framings of L differing by N -phone cord moves have the same number of labelings by X , the infinite lattice is tiled by copies of a tile with N labelings on a side. Further, the infinite integral lattice of framings of L is an invariant of the unframed link L . In fact, if we reduce the framing vectors mod N , we get a canonical tile of framing vectors \vec{w} corresponding to elements of $(\mathbb{Z}_N)^c$. Thus, if we add up the numbers of colorings of framings $L_{\vec{w}}$ of L over one complete tile, we get an invariant of L which we call the *integral rack counting invariant*

$$\Phi_X^{\mathbb{Z}}(L) = \sum_{\vec{w} \in (\mathbb{Z}_N)^c} |\text{Hom}(\mathcal{R}(L_{\vec{w}}), X)|.$$

Example 103. The rack in Example 102 has characteristic $N = 2$, so a tile of framing vectors is $\{(0, 0), (1, 0), (0, 1), (1, 1)\}$. As we have seen, the Hopf link has 4 X -labelings in framing $(1, 1)$ and no X -labelings in framing $(0, 0)$; the reader can verify that there are no X -labelings in framings $(1, 0)$ and $(0, 1)$. Then the integral rack counting invariant for the Hopf link is

$$\Phi_X^{\mathbb{Z}}(L) = \sum_{\vec{w} \in (\mathbb{Z}_2)^c} |\text{Hom}(\mathcal{R}(L_{\vec{w}}), X)| = 0 + 0 + 0 + 4 = 4.$$

Finally, we note that since a quandle is a rack of rank $N = 1$, if X is a quandle then the integral lattice of framings is tiled with a tile 1 element on a side, i.e., every framing has the same number of labelings. In particular, the new definition of $\Phi_X^{\mathbb{Z}}(L)$ when we think of a quandle X as rack coincides with our previous definition.

- Exercises.** 1. Identify all rack structures on a set with two elements.
2. Find two nonquandle racks with three elements and prove that your answers are not isomorphic.
3. Using the rack axioms, prove that $x \triangleright (x \triangleright^{-1} x) = x \triangleright^{-1} x$.
4. Let X be an *involutory rack*, i.e. a rack in which $\triangleright^{-1} = \triangleright$. Prove that X has rack characteristic $N = 1$ or $N = 2$.
5. Find the characteristic of the rack with operation matrix

$$\begin{bmatrix} 1 & 1 & 1 \\ 2 & 3 & 3 \\ 3 & 2 & 2 \end{bmatrix}.$$

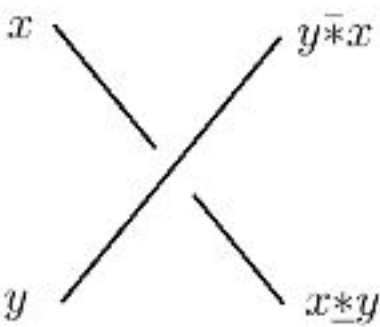
6. Compute the rack counting invariant for the figure eight knot with respect to the rack in problem 5.
7. Compute the rack labelings for the blackboard framing of the knot 5_1 as depicted in the knot table in Chapter 1 by the (t, s) -rack \mathbb{Z}_4 with $t = 3$ and $s = 3$ using row-reduction.

2. Bikei

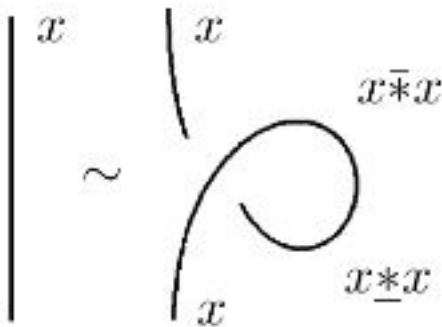
So far, all of the algebraic structures we have defined using the Reidemeister moves have used the assumption that each label corresponds to an arc in a knot diagram, i.e. a portion of the diagram running from one undercrossing point to the next. This is partly motivated by topology; after all, the fundamental quandle and group generators correspond to arcs in diagrams.

However, there's no reason we have to limit ourselves this way; in mathematics we can feel free to consider any ideas we can dream up and pursue their logical consequences. For example, instead of dividing a knot diagram only at undercrossings, we can divide it at both under- and over-crossings. Specifically, a *semiarc* is a portion of a knot or link diagram between two crossing points; if we imagine flattening the knot onto paper, semiarcs are the portions of the knot between the points where the flattened knot crosses itself.

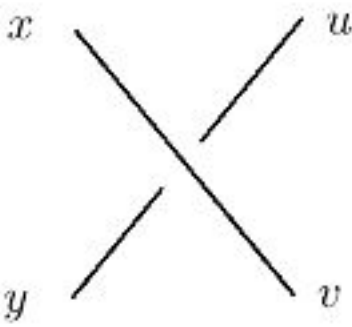
Let us start with the case of unoriented knots and links, like we did before. Now instead of one operation at a crossing, we have two:



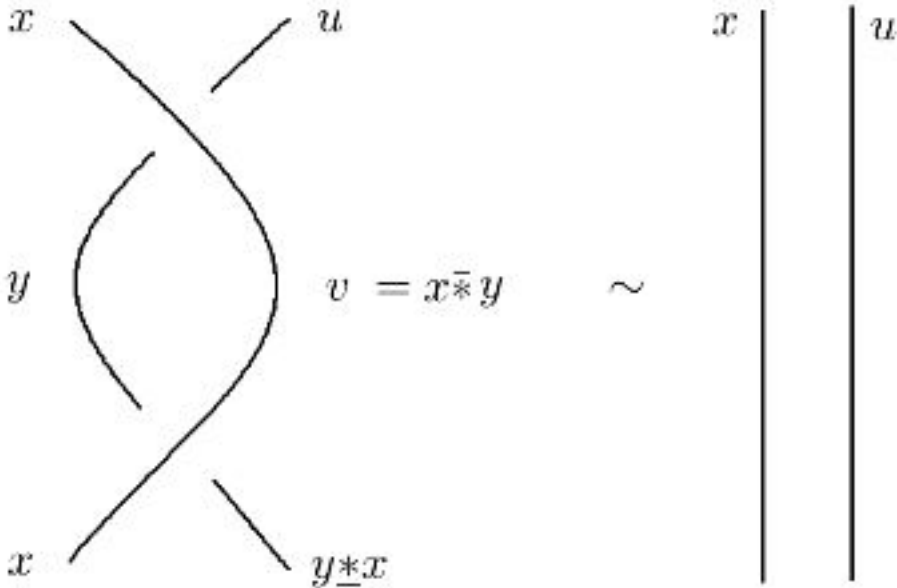
The Reidemeister I move says we need *equal self-actions*, $x*\bar{x} = x*\underline{x}$:



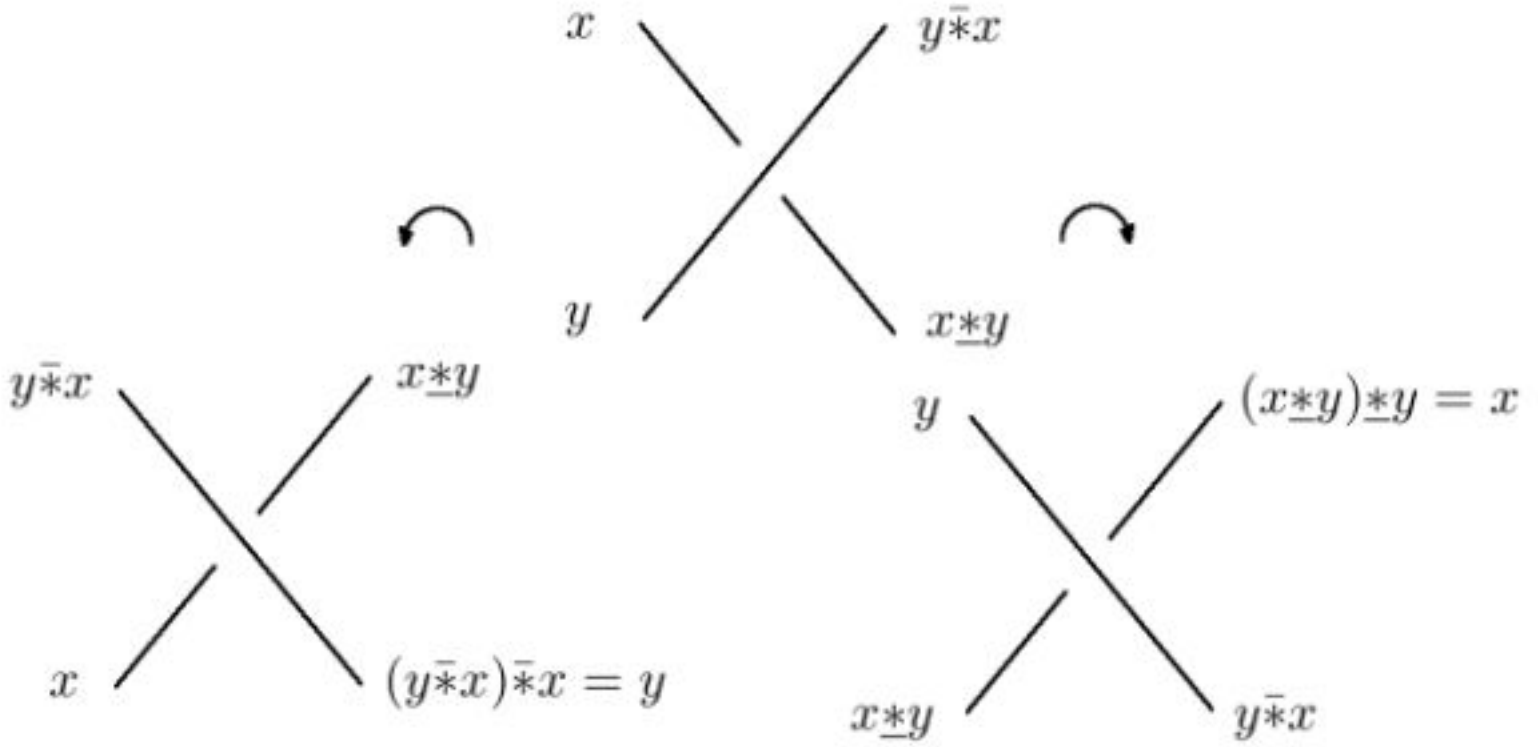
The Reidemeister II move says that the over- and under-crossing operations do not depend on which way the crossing is rotated: suppose we label the semiarcs on a rotated crossing as pictured.



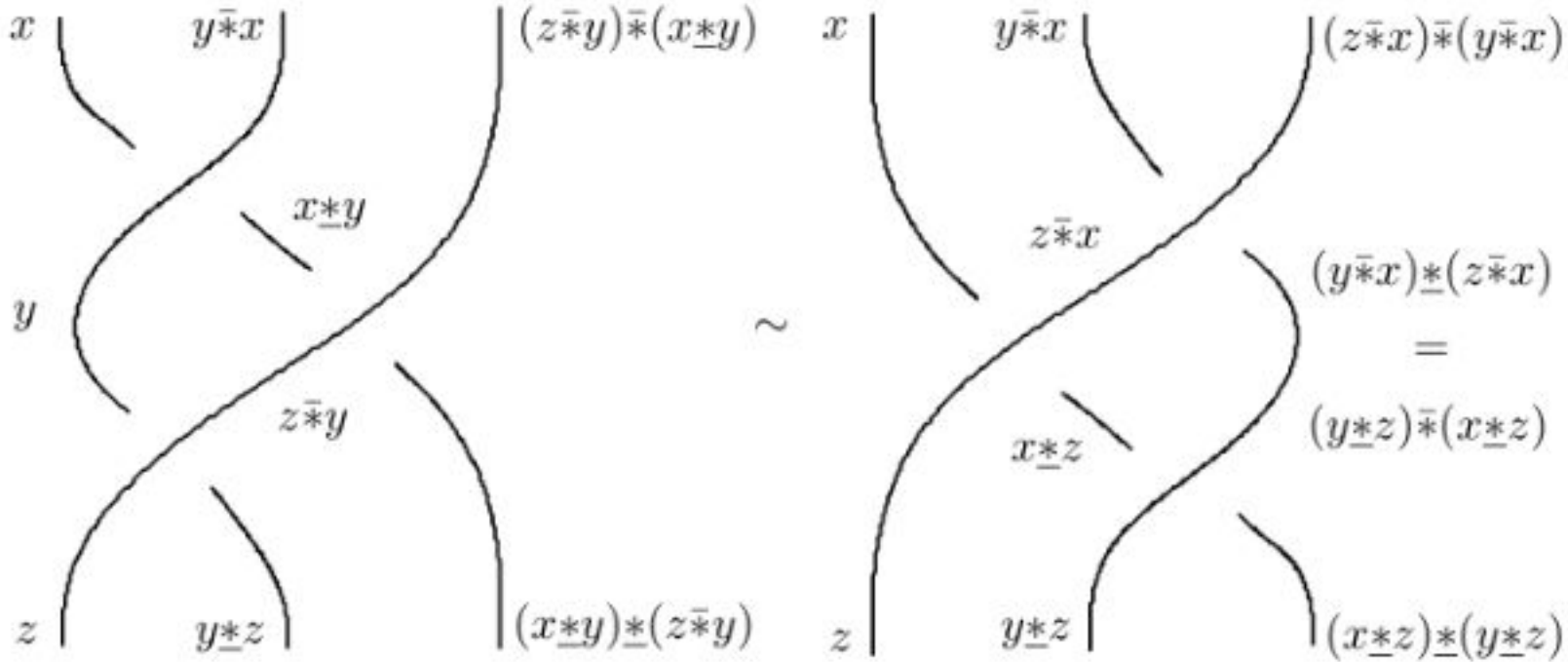
Then in move II we have $u = y*\underline{x}$ and $v = x*\bar{y}$.



In particular, rotating the crossing reveals that, much like the case of *kei*, the *bikei* operations are involutions:



Moreover, we must have $x*(y*x) = x*y$, $y*x = y*(x*y)$. $(x*y)*y = x$, and $(y*x)*x = y$. Finally, the Reidemeister III move gives us three conditions, each of which looks a little like the Moufang loop identity:



Thus, we need the *exchange laws*

$$\begin{aligned} (z*y)*x &= (z*x)*y, \\ (y*x)*z &= (y*z)*x, \\ (x*y)*z &= (x*z)*y. \end{aligned}$$

Definition 28. A *bikei* is a set X with two binary operations $\bar{*}, * : X \times X \rightarrow X$ such that for all $x, y, z \in X$, we have

(i)

$$x*x = x\bar{*}x,$$

(ii)

$$\begin{aligned}
x \underline{*} (y \bar{*} x) &= x \underline{*} y & \text{(ii.i)}, \\
x \bar{*} (y \underline{*} x) &= x \bar{*} y & \text{(ii.ii)}, \\
(x \underline{*} y) \underline{*} y &= x & \text{(ii.iii)}, \\
(x \bar{*} y) \bar{*} y &= x & \text{(ii.iv)}, \quad \text{and}
\end{aligned}$$

(iii)

$$\begin{aligned}
(x \bar{*} y) \bar{*} (z \underline{*} y) &= (x \bar{*} z) \bar{*} (y \bar{*} z) & \text{(iii.i)}, \\
(x \bar{*} y) \underline{*} (z \bar{*} y) &= (x \underline{*} z) \bar{*} (y \underline{*} z) & \text{(iii.ii)}, \\
(x \underline{*} y) \underline{*} (z \bar{*} y) &= (x \underline{*} z) \underline{*} (y \underline{*} z) & \text{(iii.iii)}.
\end{aligned}$$

Example 104. Every kei is a bikei with operations $x \underline{*} y = x \triangleright y$ and $x \bar{*} y = x$.

Example 105. The bikei axioms are symmetric with respect to interchanging the operations $\underline{*}$ and $\bar{*}$. Then if X is a bikei, there is a *dual* bikei X' with operations $\underline{*}', \bar{*}'$ defined by

$$x \underline{*}' y = x \bar{*} y \quad \text{and} \quad x \bar{*}' y = x \underline{*} y.$$

This duality can be visualized geometrically as the result of looking at the knot diagram from the other side of the paper.

Example 106. Let X be a set and $\sigma : X \rightarrow X$ any involution. Then X is a bikei with operations $x \underline{*} y = \sigma(x) = x \bar{*} y$. To see this, we simply verify that the definition satisfies the axioms:

(i)

$$x \underline{*} x = \sigma(x) = x \bar{*} x,$$

(ii)

$$\begin{aligned}
x \underline{*} (y \bar{*} x) &= \sigma(x) = x \underline{*} y, \\
x \bar{*} (y \underline{*} x) &= \sigma(x) = x \bar{*} y, \\
(x \underline{*} y) \underline{*} y &= \sigma^2(x) = x, \\
(x \bar{*} y) \bar{*} y &= \sigma^2(x) = x, \quad \text{and}
\end{aligned}$$

(iii)

$$\begin{aligned}
(x \bar{*} y) \bar{*} (z \underline{*} y) &= \sigma^2(x) = (x \bar{*} z) \bar{*} (y \bar{*} z), \\
(x \bar{*} y) \underline{*} (z \bar{*} y) &= \sigma^2(x) = (x \underline{*} z) \bar{*} (y \underline{*} z), \\
(x \underline{*} y) \underline{*} (z \bar{*} y) &= \sigma^2(x) = (x \underline{*} z) \underline{*} (y \underline{*} z),
\end{aligned}$$

as required. We call this a *constant action bikei*.

As with kei and quandles, there are finite and infinite bikei. For a finite bikei X , we can specify the bikei operations with operation tables. Since now we have two operations, it is often convenient to combine the two operation tables into a single block matrix. Thus, if we have a set $X = \{x_1, x_2, \dots, x_n\}$ with bikei operations $x_i \underline{*} x_j = x_k$ and $x_i \bar{*} x_j = x_h$, we can specify the operation tables with a block matrix $\left[\begin{array}{c|c} U & L \end{array} \right]$ where $U_{ij} = k$ and $L_{ij} = h$.

Example 107. The block matrix

$$\left[\begin{array}{cc|cc} 2 & 2 & 2 & 2 \\ 1 & 1 & 1 & 1 \end{array} \right]$$

defines operations $\underline{*}, \bar{*}$ on the set $X = \{x_1, x_2\}$ where we have $x_1 \underline{*} x_1 = x_2$, $x_2 \bar{*} x_1 = 1$, etc. To verify that these operations satisfy the bikei axioms, we must check all possible substitutions of x_1 and x_2 for x, y, z in the axioms. Let us content ourselves for the moment with verifying the exchange laws for $x = x_1, y = x_2, z = x_1$:

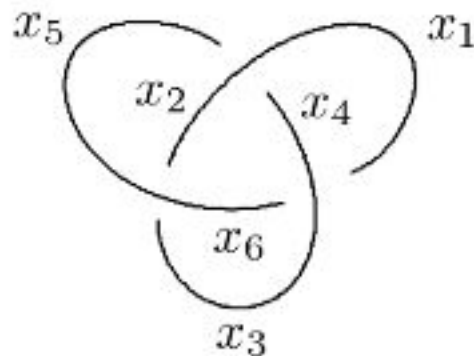
$$\begin{aligned} (x_1 \bar{*} x_2) \bar{*} (x_1 \underline{*} x_2) &= x_2 \bar{*} x_2 = x_1 = x_2 \bar{*} x_1 = (x_1 \bar{*} x_1) \bar{*} (x_2 \bar{*} x_1), \\ (x_1 \bar{*} x_2) \underline{*} (x_1 \bar{*} x_2) &= x_2 \underline{*} x_2 = x_1 = x_2 \bar{*} x_1 = (x_1 \underline{*} x_1) \bar{*} (x_2 \underline{*} x_1), \\ (x_1 \underline{*} x_2) \underline{*} (x_1 \bar{*} x_2) &= x_2 \underline{*} x_2 = x_1 = x_2 \underline{*} x_1 = (x_1 \underline{*} x_1) \underline{*} (x_2 \underline{*} x_1). \end{aligned}$$

Let L be an unoriented link with diagram D . We can define the *fundamental bikei* of L , $\mathcal{BK}(L)$, in a combinatorial way using universal algebra. First, choose a set of unique labels, say $S = \{x_1, \dots, x_n\}$, for the semiarcs of L . Next, the set of *bikei words* in S is the set of finite strings of the symbols $(,), \bar{*}, \underline{*}$ and the labels in S which make sense as operator expressions – $x_3 \underline{*} ((x_2 \bar{*} x_1) \underline{*} x_2)$ is a bikei word, while $((x_2 x_2) \underline{*})$ is not. More precisely, we can define the set $W(S)$ of bikei words in S recursively by the rules that

- $x \in S \Rightarrow x \in W$ and
- $x, y \in W \Rightarrow x \underline{*} y, x \bar{*} y \in W$.

Then the *fundamental bikei* of L , $\mathcal{BK}(L)$, is the set of equivalence classes of $W(S)$ under the equivalence relation determined by the crossing relations and the bikei axiom relations; for example, we have $x \underline{*} x \sim x \bar{*} x$ for any $x \in W(S)$, etc. As with quandles, we can specify the fundamental bikei of a knot or link with a presentation by

generators and relations; for example, the trefoil knot below has fundamental bikei presentation



$$\mathcal{B}(K) = \langle x_1, x_2, x_3, x_4, x_5, x_6 \mid \begin{array}{l} x_5 \underline{*} x_2 = x_4, \quad x_2 \bar{*} x_5 = x_1, \\ x_3 \underline{*} x_5 = x_2, \quad x_5 \bar{*} x_3 = x_6, \\ x_6 \underline{*} x_3 = x_1, \quad x_3 \bar{*} x_6 = x_4 \end{array} \rangle.$$

Then for example the bikei word $(x_6 \bar{*} x_2) \bar{*} (x_3 \underline{*} x_5)$ is equivalent to x_6 since we have

$$(x_6 \bar{*} x_2) \bar{*} (x_3 \underline{*} x_5) \sim (x_6 \bar{*} x_2) \bar{*} x_2 \sim x_6.$$

A subset S of a bikei X is a *subbikei* if whenever $x, y \in S$ we have $x \underline{*} y \in S$ and $x \bar{*} y \in S$. We can think of a subbikei as a smaller bikei embedded inside of X .

Let X and Y be bikei with operations $\underline{*}_X, \bar{*}_X$ and $\underline{*}_Y, \bar{*}_Y$ respectively. Then a map $f : X \rightarrow Y$ is a *bikei homomorphism* if for all $x, x' \in X$ we have

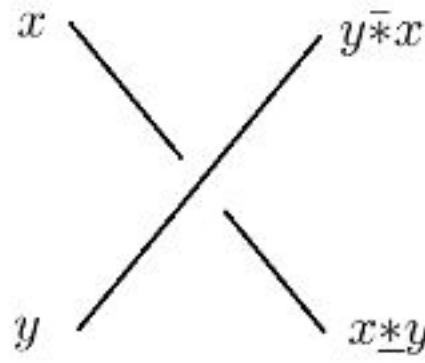
$$f(x \underline{*}_X x') = f(x) \underline{*}_Y f(x') \quad \text{and} \quad f(x \bar{*}_X x') = f(x) \bar{*}_Y f(x').$$

The set of all bikei homomorphisms $f : X \rightarrow Y$ is written $\text{Hom}(X, Y)$.

The Counting Invariant. Just like with kei and quandles, given a finite bikei we can define an invariant of unoriented knots by counting valid bikei colorings of knot or link diagrams. The number of such colorings is a link invariant by construction, since we set up the bikei axioms so that for each valid bikei coloring of a diagram before a Reidemeister move there is exactly one valid coloring of the diagram after the move. That is, we have

Theorem 12. *Let L be an unoriented link diagram, X be a finite bikei, and $\Phi_X^{\mathbb{Z}}(L)$ be the number of assignments of elements of X to*

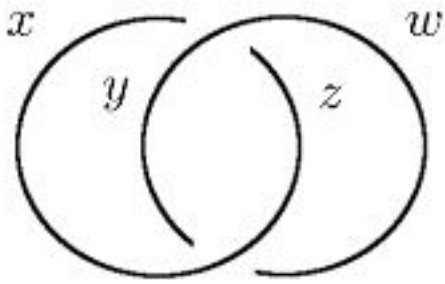
semiarcs in L such that at each crossing the labels satisfy the condition



Then $\Phi_X^{\mathbb{Z}}(L)$ is a link invariant.

As with quandles, we can think of a coloring of L by X as a homomorphism $f : \mathcal{BK}(L) \rightarrow X$ from the fundamental bikei of L to X . Thus, we might write $\Phi_X^{\mathbb{Z}}(L)$ as $|\text{Hom}(\mathcal{BK}(L), X)|$.

Example 108. Let $X = \mathbb{Z}_4$ and define $x\bar{*}y = 3x$ and $x*\underline{y} = x + 2y$. One can show (see exercise 6) that X is a bikei under these operations. Then the Hopf link L has bikei presentation



$$\langle x, y, z, w \mid \begin{array}{l} x*\underline{y} = z, \quad y\bar{*}x = w, \\ y*\underline{x} = w, \quad x\bar{*}y = z \end{array} \rangle$$

Then to compute the set of bikei labelings, we have a homogeneous system of linear equations over \mathbb{Z}_4 with coefficient matrix

$$\begin{bmatrix} 1 & 2 & 3 & 0 \\ 0 & 3 & 0 & 3 \\ 2 & 1 & 0 & 3 \\ 3 & 0 & 3 & 0 \end{bmatrix}.$$

Then after row-reduction over \mathbb{Z}_4 we have

$$\begin{bmatrix} 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 2 & 2 \\ 0 & 0 & 0 & 0 \end{bmatrix}.$$

Setting $x_3 = \alpha \in \mathbb{Z}_4$ and $x_4 = \beta$, we have $2\alpha + 2\beta = 2(\alpha + \beta) = 0$ so $\alpha + \beta \in \{0, 2\}$; if we let $\gamma \in \mathbb{Z}_2$, then $\alpha + \beta = 2\gamma$ and $\beta = 3\alpha + 2\gamma$. Thus, we have solution set $x_1 = 3\alpha$, $x_2 = \alpha + 2\gamma$ for $\alpha \in \mathbb{Z}_4$ and

$\beta \in \mathbb{Z}_2$ and the set of colorings is isomorphic to $\mathbb{Z}_4 \oplus \mathbb{Z}_2$. Hence we have counting invariant value $\Phi_x^{\mathbb{Z}}(L) = 8$.

- Exercises.** 1. Find all bikei structures on the set $X = \{x_1, x_2\}$.
2. Write a computer program in your favorite programming language to test whether a pair of square matrices define bikei operations.
3. Suppose V is an \mathbb{F} -vector space. Set

$$\begin{aligned}\vec{x} \bar{*} \vec{y} &= \alpha \vec{x} + \beta \vec{y} \quad \text{and} \\ \vec{x} \underline{*} \vec{y} &= \gamma \vec{x} + \delta \vec{y}.\end{aligned}$$

What conditions on $\alpha, \beta, \gamma, \delta \in \mathbb{F}$ are necessary and sufficient for $\underline{*}, \bar{*}$ to be bikei operations on V ?

4. Let $X = \mathbb{Z}_3$. What values $a, b, c, d \in \mathbb{Z}_3$ make X a bikei with operations $x \underline{*} y = ax + by$ and $x \bar{*} y = cx + dy$?
5. Prove that every link L of c components has bikei counting invariant $\Phi_X^{\mathbb{Z}}(L) = 2^c$ with respect to the bikei with operation matrix

$$\left[\begin{array}{cc|cc} 2 & 2 & 2 & 2 \\ 1 & 1 & 1 & 1 \end{array} \right]$$

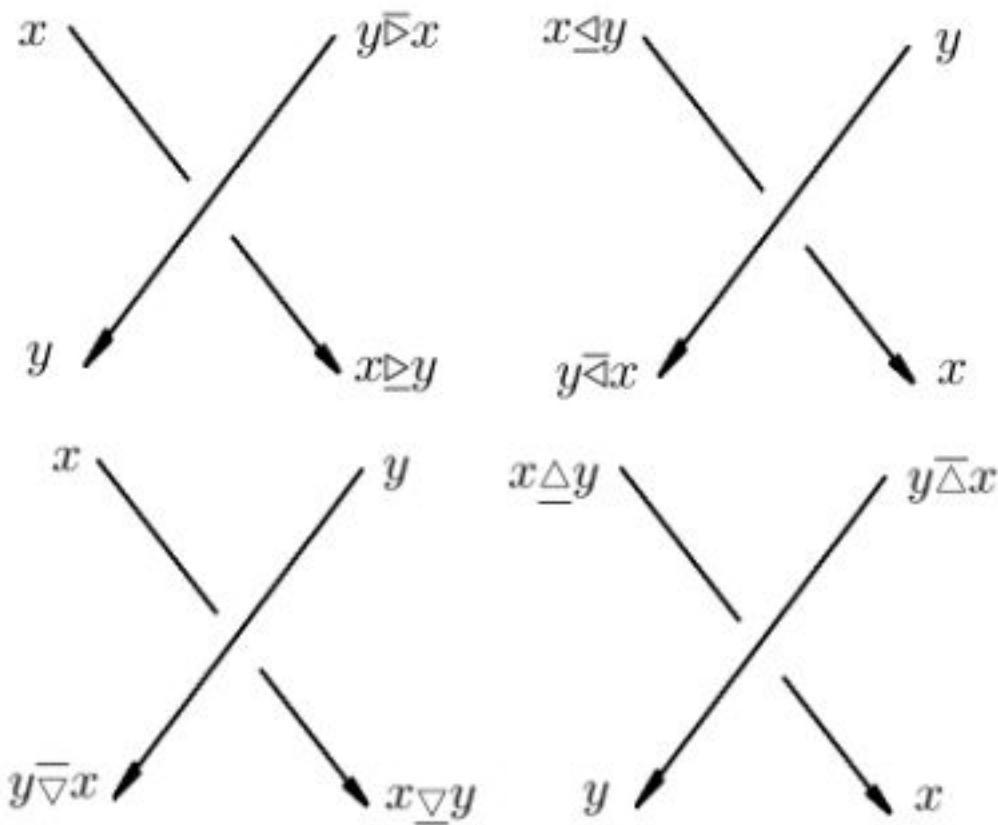
6. Verify that the operations in example 108 satisfy the bikei axioms.

3. Biracks and Biquandles

What happens when we take the bikei style of coloring semiarcs rather than arcs and apply it in the oriented and framed oriented cases? Historically, the framed oriented case of semiarc-coloring algebraic structures was considered in 1994 [**FRS95**] before bikeis. Let's see how it works.

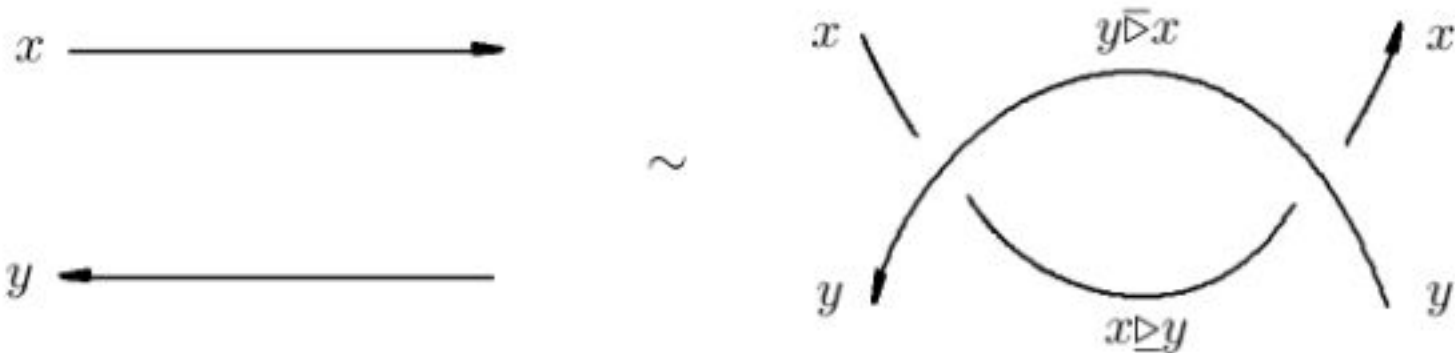
Let us start by naming the various operations. At a positively oriented crossing there are eight possible operations on neighboring

semiarcs on each other as pictured



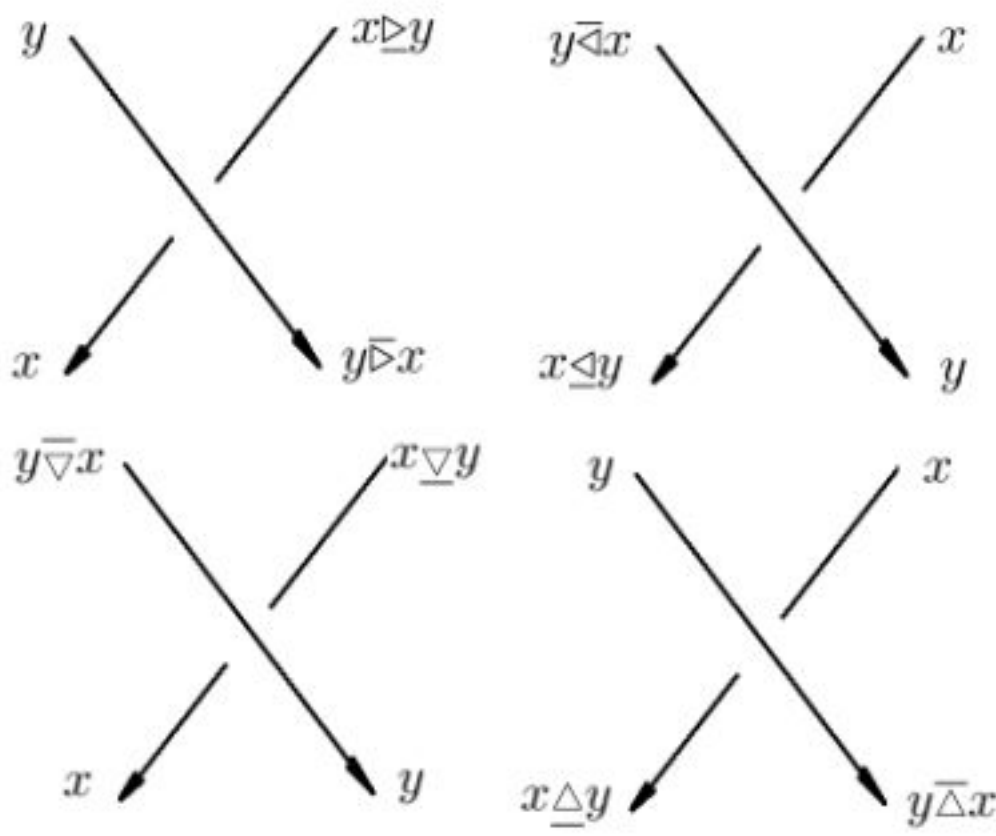
together with eight more operations at negatively oriented crossings. By considering the requirements of Reidemeister equivalence, we can reduce this list of sixteen operations substantially.

First, we want to make sure our new algebraic structure can be used for counting invariants, which means we must ensure that for every coloring of a diagram before a move, there must be a *unique* corresponding coloring after the move. In particular, in both the oriented and framed cases, the Reidemeister I, II and framed I moves require the coloring operations to satisfy the *adjacent pairs* rule: the colors on any two sides of a crossing must determine the colors on the other two sides. For example, the move



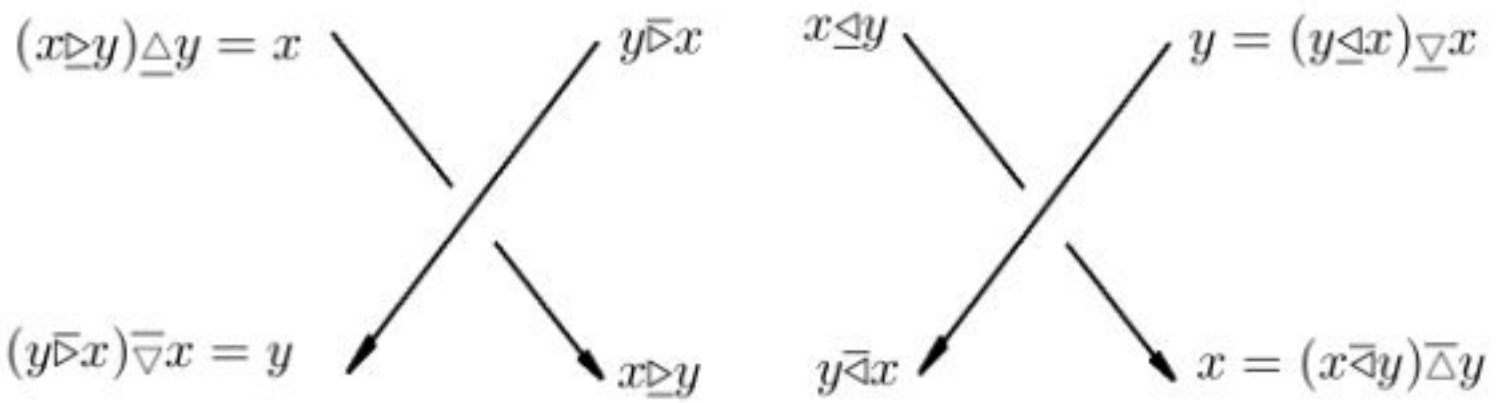
requires that for each pair $(x, y) \in X \times X$, there must be a unique pair $(y\overline{\triangleright}x, x\triangleright y)$ if we are to have a unique coloring after the move, and there is a move for each pair of operations. In particular, we can think in terms of maps of pairs: the map $H : X \times X \rightarrow X \times X$ defined

by $H(x, y) = (y \rhd x, x \rhd y)$ must be injective. Indeed, the type II move requires that H is bijective, with inverse given by the corresponding map of pairs at the negative crossing. This implies that we don't have independent operations at positive and negative crossings; rather, the operations at negative crossings are the pairwise inverses of the operations at the positive crossings. Thus, we really have only eight operations:



and switching the sign of the crossing just switches the position of the colors.

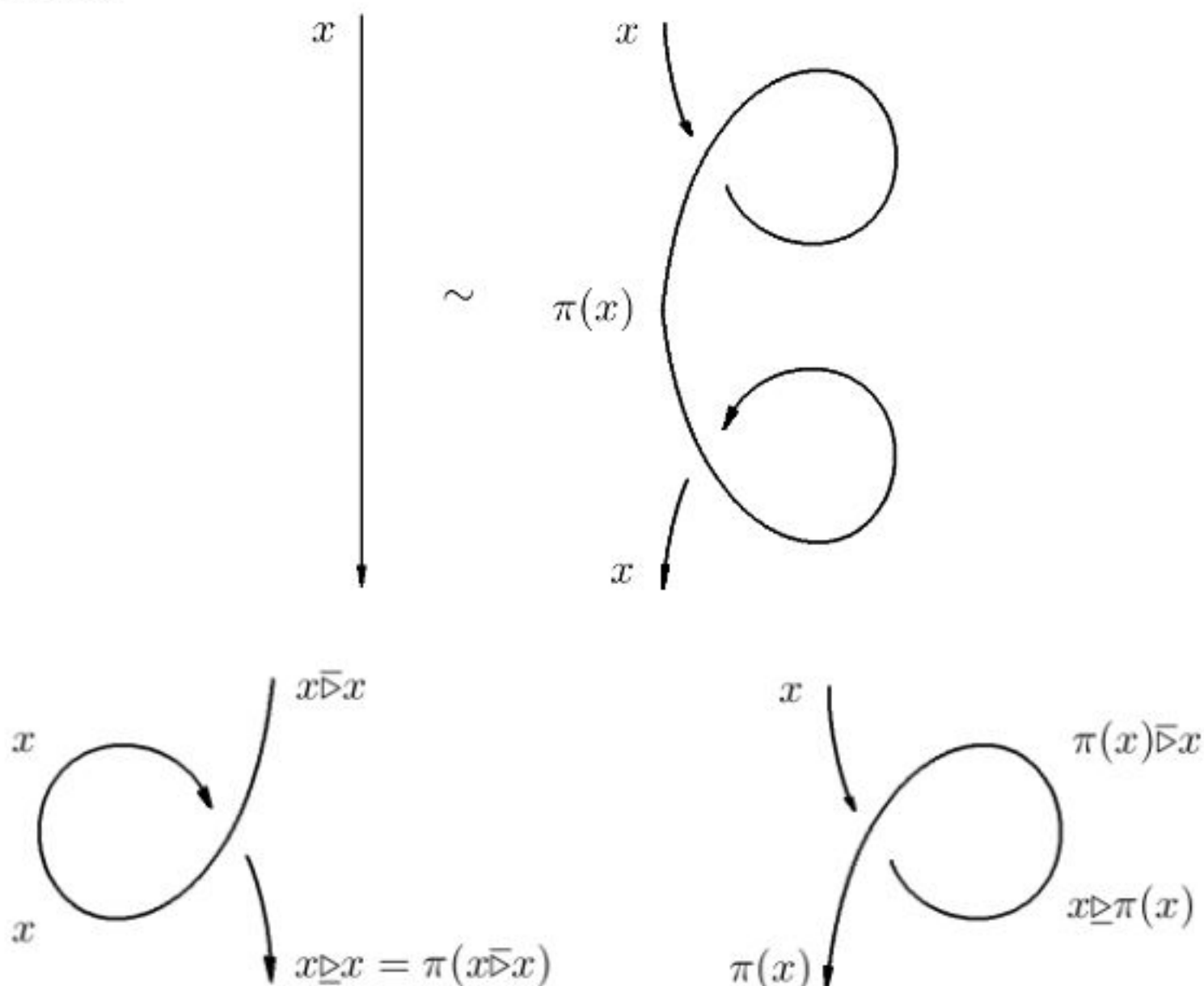
Moreover, these eight operations are not all independent; indeed, the adjacent pairs rule implies that the operations are right-invertible with the vertical operations expressible as right inverses of the horizontal operations.



In particular, we have $\triangle = \rhd^{-1}$, $\bar{\nabla} = \bar{\rhd}^{-1}$, $\bar{\triangle} = \bar{\triangle}^{-1}$ and $\nabla = \lhd^{-1}$. Finally, the operations $\bar{\triangle}$ and \lhd can be interpreted as the components of the inverse of the of pairs $H(x, y) = (y \rhd x, x \rhd y)$; thus, we have

two independent operations $\rhd, \bar{\rhd} : X \times X \rightarrow X$ which are right-invertible and satisfy the additional requirement that the map of pairs $H(x, y) = (y \bar{\rhd} x, x \rhd y)$ is invertible.

The framed Reidemeister I moves require an invertible kink map $\pi : X \rightarrow X$ with the conditions that $\pi(x \bar{\rhd} x) = x \rhd x$ and $\pi(x) \bar{\rhd} x = x \rhd \pi(x)$.



Finally, as in the bikei case, we have exchange laws between the operations \rhd and $\bar{\rhd}$:

$$\begin{aligned} (x \rhd y) \rhd (z \rhd y) &= (x \rhd z) \rhd (y \bar{\rhd} z), \\ (x \rhd y) \bar{\rhd} (z \rhd y) &= (x \bar{\rhd} z) \rhd (y \bar{\rhd} z), \\ (x \bar{\rhd} y) \bar{\rhd} (z \bar{\rhd} y) &= (x \bar{\rhd} z) \bar{\rhd} (y \rhd z). \end{aligned}$$

Thus, we can formally state our definition:

Definition 29. A *birack* is a set X with right-invertible operations $\rhd, \bar{\rhd} : X \times X \rightarrow X$ and a bijection $\pi : X \rightarrow X$ satisfying for all $x, y, z \in X$,

$$(i) \quad \pi(x \bar{\rhd} x) = x \rhd x \text{ and } \pi(x) \bar{\rhd} x = x \rhd \pi(x).$$

- (ii) The map of pairs $H(x, y) = (y \bar{\triangleright} x, x \triangleright y)$ is invertible.
- (iii) The exchange laws:

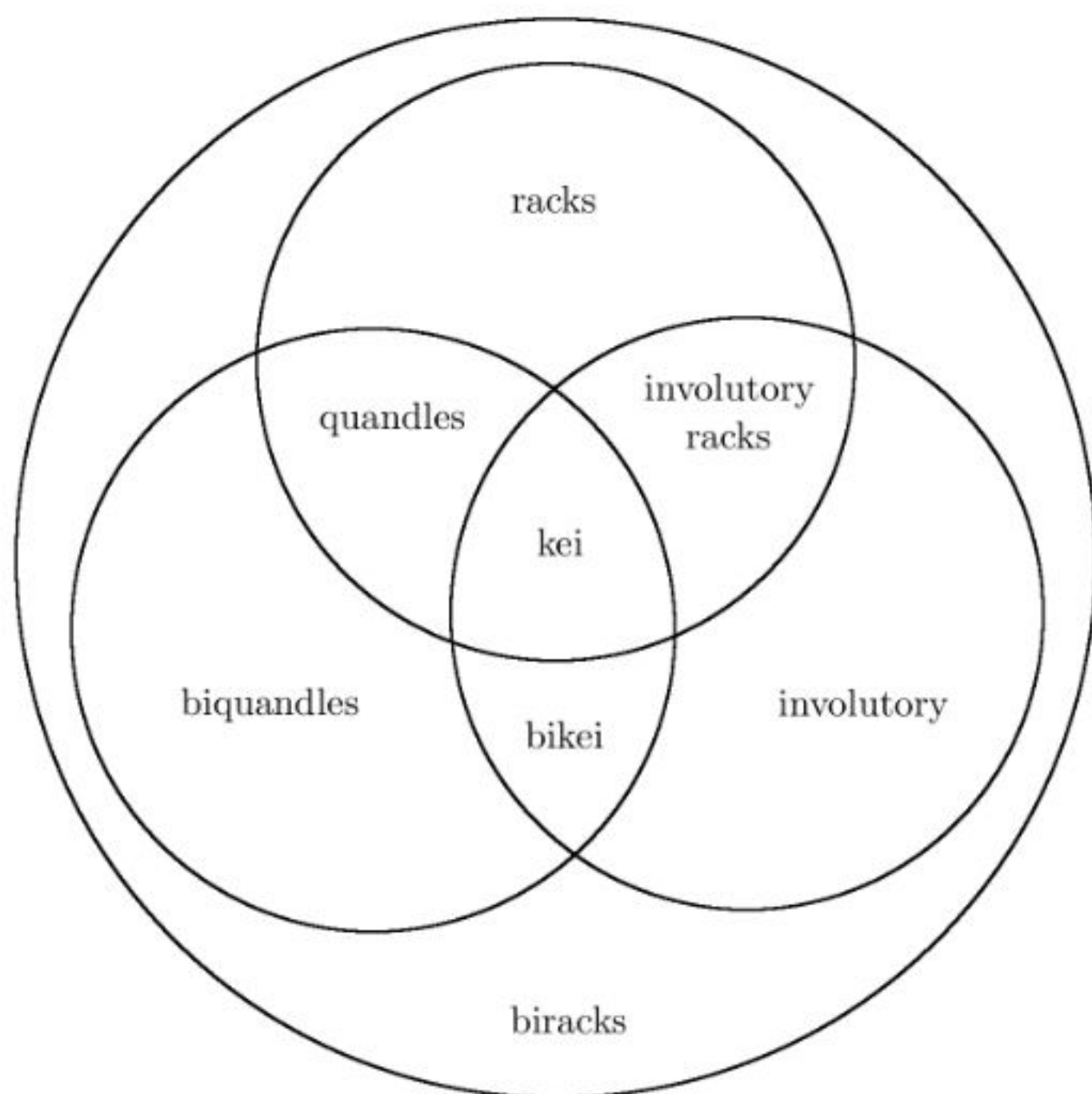
$$\begin{aligned} (x \triangleright y) \triangleright (z \triangleright y) &= (x \triangleright z) \triangleright (y \bar{\triangleright} z), \\ (x \triangleright y) \bar{\triangleright} (z \triangleright y) &= (x \bar{\triangleright} z) \triangleright (y \bar{\triangleright} z), \\ (x \bar{\triangleright} y) \bar{\triangleright} (z \bar{\triangleright} y) &= (x \bar{\triangleright} z) \bar{\triangleright} (y \triangleright z). \end{aligned}$$

As in the rack case, if X is a finite set, there is an integer N such that $\pi^N : X \rightarrow X$ is the identity map; the smallest such positive N is the *characteristic* of X . A birack of characteristic $N = 1$ is called a *biquandle*.

Example 109. Every quandle, kei, or rack is a birack with operations $x \bar{\triangleright} y = x$ and $x \triangleright y = x \triangleright y$.

Example 110. A bikei is a birack with $x \bar{\triangleright} y = x \bar{*} y$ and $x \triangleright y = x \underline{*} y$.

We illustrate the relationship between these algebraic structures with the following Venn diagram from [AN12]:



Example 111. Let X be a set and $\sigma, \tau : X \rightarrow X$ bijections which commute, i.e., such that $\sigma(\tau(x)) = \tau(\sigma(x))$. Then X is a birack with operations $x \rhd y = \sigma(x)$ and $x \bar{\rhd} y = \tau(x)$ and $\pi(x) = \tau^{-1}(\sigma(x))$. A birack of this type is called a *constant action birack*. Let us verify the axioms:

(i) We have

$$\pi(x \bar{\rhd} x) = \tau^{-1}(\sigma(\tau(x))) = \tau^{-1}\tau\sigma(x) = \sigma(x) = x \rhd x$$

and

$$\pi(x) \bar{\rhd} x = \tau(\tau^{-1}(\sigma(x))) = \sigma(x) = x \rhd \pi(x).$$

(ii) The map $H(x, y) = (\tau(y), \sigma(x))$ has inverse map

$$H^{-1}(x, y) = (\sigma^{-1}(y), \tau^{-1}(x)), \quad \text{and}$$

(iii)

$$\begin{aligned} (x \rhd y) \rhd (z \rhd y) &= \sigma^2(x) &= (x \rhd z) \rhd (y \bar{\rhd} z), \\ (x \rhd y) \bar{\rhd} (z \rhd y) &= \sigma(\tau(x)) = \tau(\sigma(x)) &= (x \bar{\rhd} z) \rhd (y \bar{\rhd} z), \\ (x \bar{\rhd} y) \bar{\rhd} (z \bar{\rhd} y) &= \tau^2(x) &= (x \bar{\rhd} z) \bar{\rhd} (y \rhd z). \end{aligned}$$

As with bikei, we can represent a birack structure on a finite set $X = \{x_1, x_2, \dots, x_n\}$ with a pair of operation tables for \rhd and $\bar{\rhd}$ encoded as matrices U and L such that $U_{ij} = k$ where $x_k = x_i \rhd x_j$ and $L_{ij} = h$ where $x_h = x_i \bar{\rhd} x_j$. Then for example the constant action birack structure on $X = \{x_1, x_2, x_3, x_4\}$ with operations given by the permutations $\sigma = [2, 3, 4, 1]$ and $\tau = [4, 1, 2, 3]$ has operation matrix

$$\left[\begin{array}{cccc|cccc} 2 & 2 & 2 & 2 & 4 & 4 & 4 & 4 \\ 3 & 3 & 3 & 3 & 1 & 1 & 1 & 1 \\ 4 & 4 & 4 & 4 & 2 & 2 & 2 & 2 \\ 1 & 1 & 1 & 1 & 3 & 3 & 3 & 3 \end{array} \right].$$

This birack has kink map $\pi = [3, 1, 4, 2]$ and characteristic $N = 2$ since $\pi^2 = \text{Id}$.

Maps of Pairs. An alternative way to state the definition of a birack is in terms of maps of pairs. In much of the literature in which the properties of biracks were worked out, this was the notation used. Precisely, given a set X , let $\Delta(x) = (x, x)$. Then we say that an

invertible map of pairs $B : X \times X \rightarrow X \times X$ given by $B(x, y) = (B_1(x, y), B_2(x, y))$ is a *birack map* if it satisfies the conditions:

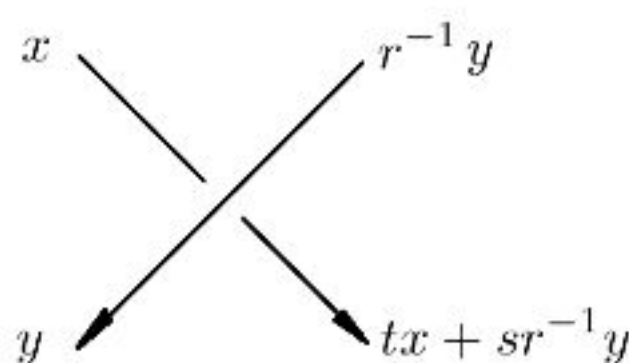
- (i) There is a unique invertible map $H : X \times X \rightarrow X \times X$ satisfying $H(x, B_1(x, y)) = (y, B_2(x, y))$ for all $x, y \in X$.
- (ii) The map $\pi(x) = (H\Delta)_2(H\Delta)_1^{-1}$ is a bijection where $(H\Delta)_j$ is the j th component of $(H\Delta)$.
- (iii) B satisfies the *set-theoretic Yang-Baxter equation*

$$(B \times I)(I \times B)(B \times I) = (I \times B)(B \times I)(I \times B)$$

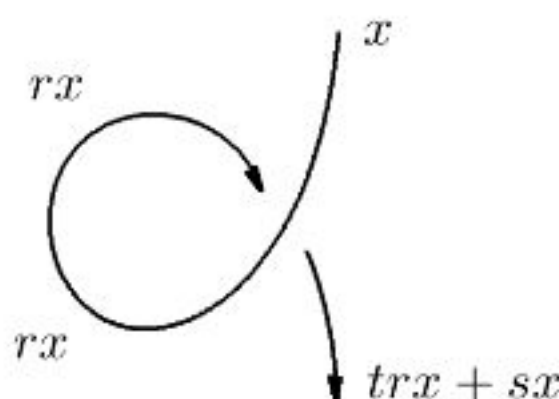
where $I : X \rightarrow X$ is the identity map $I(x) = x$.

The components of $B(x, y)$ are often written (y_x, x^y) . In our notation, the map $B(x, y)$ is given by $(y \bar{\triangleright} x, x \underline{\triangleright} y)$.

(t, s, r) -Biracks. An interesting example of a birack or biquandle structure is the (t, s, r) -birack. Let $\bar{\Lambda} = \mathbb{Z}[t^{\pm 1}, s, r^{\pm 1}]/(s^2 - s(1 - tr))$ be the quotient of the set of polynomials with integer coefficients in a variable s and two invertible variables t and r modulo the condition that s^2 equals $(1 - tr)s$. Then the operations $x \underline{\triangleright} y = tx + sr^{-1}y$ and $y \bar{\triangleright} x = r^{-1}y$ define a birack structure on any $\bar{\Lambda}$ -module A .



The kink map of a (t, s, r) -birack is given by $\pi(x) = (tr + s)x$.



If s is invertible, then $s^2 = (1 - tr)s$ says $s = 1 - tr$ so we have $\pi(x) = (tr + s)x = (tr + 1 - tr)x = x$ and our birack is a biquandle, called an *Alexander biquandle*.

We have seen special cases of (t, s, r) -biracks already:

- A (t, s, r) -birack with $r = 1$ is a (t, s) -rack;
- A (t, s, r) -birack with $r = 1$ and $s = 1 - t$ is an Alexander quandle

Let $X = \mathbb{Z}_n$. We can give X the structure of a (t, s, r) -birack by choosing elements t, s, r in \mathbb{Z}_n such that t and r are invertible in \mathbb{Z}_n (this happens if t and r are *coprime* to n , that is, if t and r have greatest common divisor 1 with n) and such that $s^2 = s(1 - tr)$. If s is also invertible, then we have an Alexander biquandle.

Example 112. For instance, take $X = \mathbb{Z}_3$ with $t = 1$, $s = 2$ and $r = 2$. Then $\gcd(1, 3) = 1$, $\gcd(2, 3) = 1$, and $s^2 = 2^2 = 4 = 1$ while $s(1 - tr) = 2(1 - 1(2)) = 2(-1) = -2 = 1$ so the operations $x \underline{\triangleright} y = x + 2y$ and $x \overline{\triangleright} y = 2x$ define a (t, s, r) -birack structure on \mathbb{Z}_3 . The operation tables are given by

$\underline{\triangleright}$	0	1	2		$\overline{\triangleright}$	0	1	2
0	0	2	1	and	0	0	0	0
1	1	0	2		1	2	2	2
2	2	1	0		2	1	1	1

This birack has kink map $\pi(x) = (tr + s)x = (1(2) + 2)x = 4x = x$, so $\pi(x) = x$ and X is an Alexander biquandle.

Example 113. Now consider $X = \mathbb{Z}_4$ and set $t = 3$, $s = 3$ and $r = 3$. Then $\gcd(3, 4) = 1$ and $s^2 = 3^2 = 9 = 1$ while $s(1 - tr) = 3(1 - 3(3)) = 3(-8) = 0$ so the operations $x \underline{\triangleright} y = 3x + 2y$ and $x \overline{\triangleright} y = 3x$ define a (t, s, r) -birack structure on \mathbb{Z}_4 . This birack has kink map $\pi(x) = (tr + s)x = (3(3) + 2)x = 17x = x$; then $\pi^2(x) = 3(3x) = 9x = x$, so X is a birack of characteristic $N = 2$.

Counting Invariants. As with racks, each framing of a knot or link can have potentially different numbers of colorings by a birack X , but if X has characteristic N then N -phone cord equivalent framings of the same knot or link will have the same number of X -colorings, which can be interpreted as birack homomorphisms from the *fundamental*

birack of L , $\mathcal{BR}(L)$, to X . Thus, we can define the *basic birack counting invariant* $\Phi_X^B(L_{\vec{w}})$ of a link L of c components with framing vector \vec{w} to be the number of birack homomorphisms $|\text{Hom}(\mathcal{BR}(L_{\vec{w}}), X)|$ from the fundamental birack of $L_{\vec{w}}$ to X , and then the *integral birack counting invariant* of the unframed link L is the sum of these basic counting invariants over a complete tile of framings mod N :

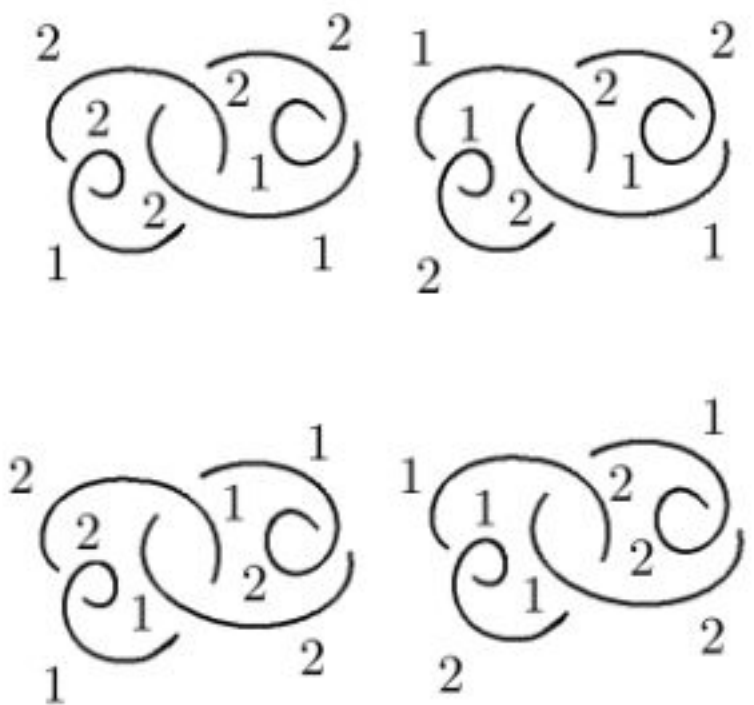
$$\Phi_X^{\mathbb{Z}}(L) = \sum_{\vec{w} \in (\mathbb{Z}_N)^c} |\text{Hom}(\mathcal{BR}(L_{\vec{w}}), X)|.$$

If $N = 1$ and X is a biquandle, then a tile of framings is a single framing and the integral counting invariant is the same as the basic counting invariant.

Example 114. Consider the Hopf link L and the constant action birack defined by the operation matrix

$$\left[\begin{array}{cc|cc} 1 & 1 & 2 & 2 \\ 2 & 2 & 1 & 1 \end{array} \right].$$

Noticing that $x \rhd y = x$ and $x \rhd y = \sigma(x)$ where $\sigma = [2, 1]$, we can think of colorings by X as colorings in which the colors stay the same when going under a crossing and switch from 1 to 2 or 2 to 1 when going over a crossing. Then $\pi(1) = 2$ and $\pi(2) = 1$, so $\pi^2 = \text{Id}$ and X has characteristic $N = 2$. Then to compute $\Phi_X^{\mathbb{Z}}(L)$, we need to look at colorings of framings of L in $(\mathbb{Z}_2)^2 = \{(0, 0), (0, 1), (1, 0), (1, 1)\}$. There are no valid colorings of L with writhe vectors $(0, 0)$, $(0, 1)$ or $(1, 0)$, but there are four with writhe vector $(1, 1)$:



Thus we have $\Phi_X^{\mathbb{Z}}(L) = 0 + 0 + 0 + 4 = 4$.

- Exercises.**
1. Compute the integral birack counting invariant for the knot 6_1 with blackboard framing as shown in the knot table in Chapter 1 with respect to the (t, s, r) -birack \mathbb{Z}_3 with $t = 1, s = 2, r = 2$.
 2. Show that the (t, s, r) -birack operations satisfy the exchange laws.
 3. Compute the counting invariant of the trefoil knot with respect to the Alexander biquandle \mathbb{Z}_3 with $t = 2$ and $r = 1$ using row-reduction over \mathbb{Z}_3 .
 4. Find the characteristic of the (t, s, r) -birack $X = \mathbb{Z}_8$ with $t = 1, s = 4$ and $r = 5$.
 5. Compute the counting invariant of the figure eight knot with respect to the (t, s, r) -birack $X = \mathbb{Z}_4$ with $t = 1, s = 2$ and $r = 3$.

Chapter 6



Enhancements

1. Basic Enhancements

An *enhancement* of a counting invariant is another stronger invariant from which we can recover the counting invariant. Many of the knot invariants defined and studied using quandles and related algebraic structures can be understood as enhancements of a counting invariant.

Let X be a finite quandle. Recall that the quandle counting invariant $\Phi_X^{\mathbb{Z}}(L)$ of a link L counts colorings of the arcs in L with elements of X , which we can also understand as quandle homomorphisms $f : \mathcal{Q}(L) \rightarrow X$ from the fundamental quandle of L to X . In fact, it's not just the *number* of quandle colorings of a diagram of L which is invariant under Reidemeister moves; it's the *set* of homomorphisms $\text{Hom}(\mathcal{Q}(L), X)$.



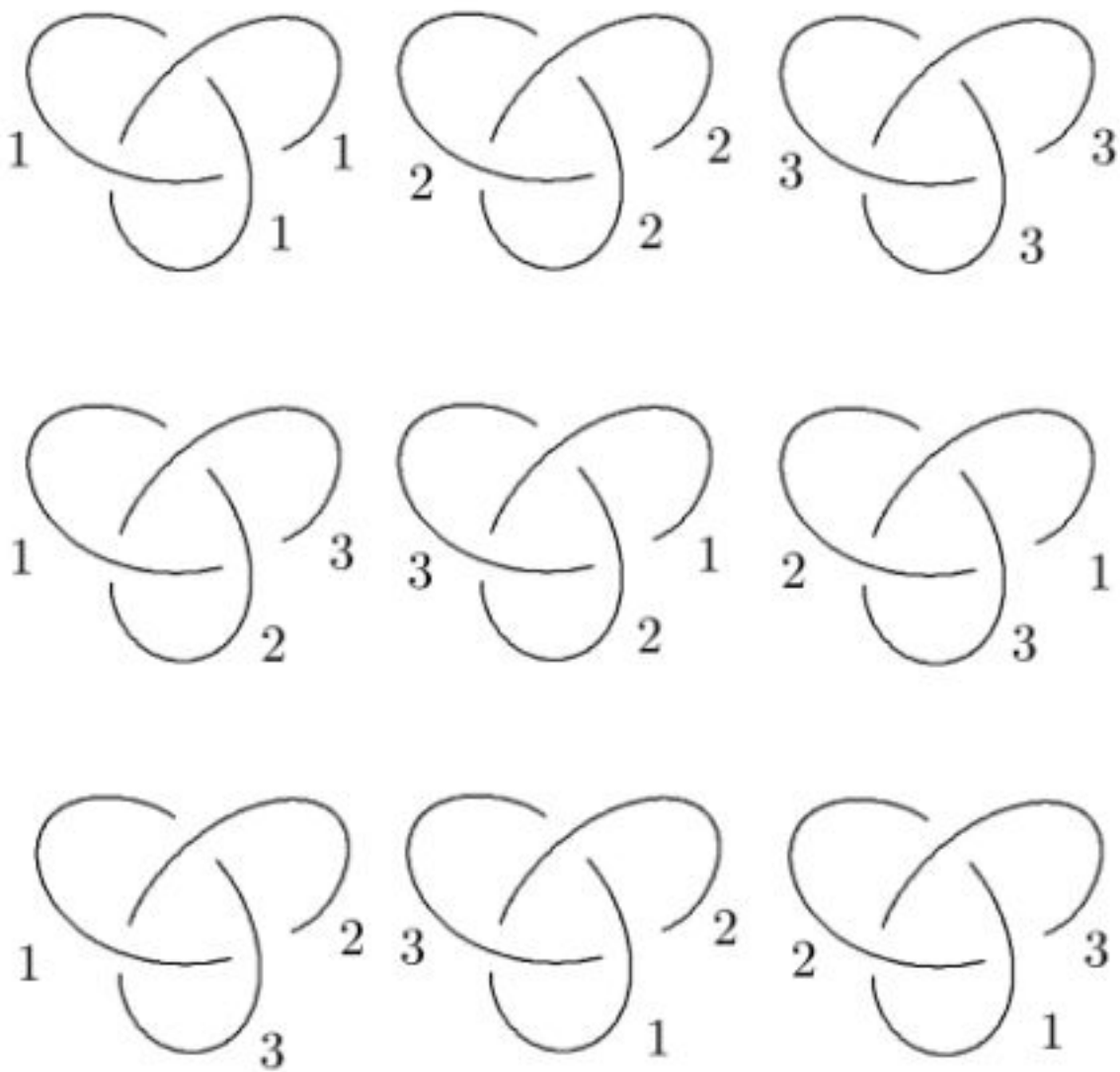
While a homomorphism $f : \mathcal{Q}(L) \rightarrow X$ can be represented as, for instance, a vector specifying an element of X for each arc in a diagram of L , such a representation depends on the diagram we have chosen for L . Thus, we should more properly think of a homomorphism $f : \mathcal{Q}(L) \rightarrow X$ as an equivalence class of colorings of diagrams of L .

Now, suppose we have an invariant ϕ not just of links L but of X -colored links; that is, something we can compute from an X -colored

link L such that ϕ is the same before and after doing Reidemeister moves to L with the corresponding X -colorings. Then instead of just adding up the number of colorings, we can collect the ϕ values for each coloring. Since we might have repeated ϕ values, we need to make a *multiset* or *set with multiplicities*, where each element has an associated multiplicity (or equivalently, we just allow repeated elements in our set, like $\{1, 1, 1, 2, 3, 3, \}$). The total number of ϕ values then tells us the number of colorings, so we can recover the counting invariant; however, different links may have different combinations of ϕ values which we can use to tell links apart even if they have the same counting invariant value.

Image Enhancement. Consider the set of kei colorings of the trefoil knot 3_1 by the three-element kei with the operation table below.

\triangleright	1	2	3
1	1	3	2
2	3	2	1
3	2	1	3



Looking at the colorings, we notice that three of the colorings are different from the other six. In terms of tricolorings, three of these are *trivial* tricolorings and six are nontrivial. As we have seen, a nontrivial tricoloring is really a *surjective* kei homomorphism. Thus what is distinguishing these kei colorings of the trefoil is the cardinality of the image subkei: three have one-element images while six have three-element images. In particular, the image subkei

$$\text{Im}(f) = \{x \in X \mid x = f(a) \text{ for some } a \in \mathcal{K}(L)\}$$

of a kei homomorphism $f : \mathcal{K}(L) \rightarrow X$ is an invariant of X -colorings of L , and hence so is its cardinality $|\text{Im}(f)|$. Then we have a new invariant enhancing the counting invariant consisting of the multiset of cardinalities of image subkeis over the set of kei colorings of L , which we call the *image enhancement multiset* of L with respect to X :

$$\Phi_X^{\text{Im}, M}(L) = \{|\text{Im}(f)| \mid f \in \text{Hom}(\mathcal{K}(L), X)\}.$$

If the elements of our multiset M are numbers, we can encode the multiset conveniently as a polynomial (or as an infinite series if M is countably infinite) by making the multiplicities into coefficients and the elements into exponents of a dummy variable u , resulting in a function of u known as a *generating function*. For example, the multiset

$$M = \{0, 0, 1, 1, 1, 2, 3, 3, 4\}$$

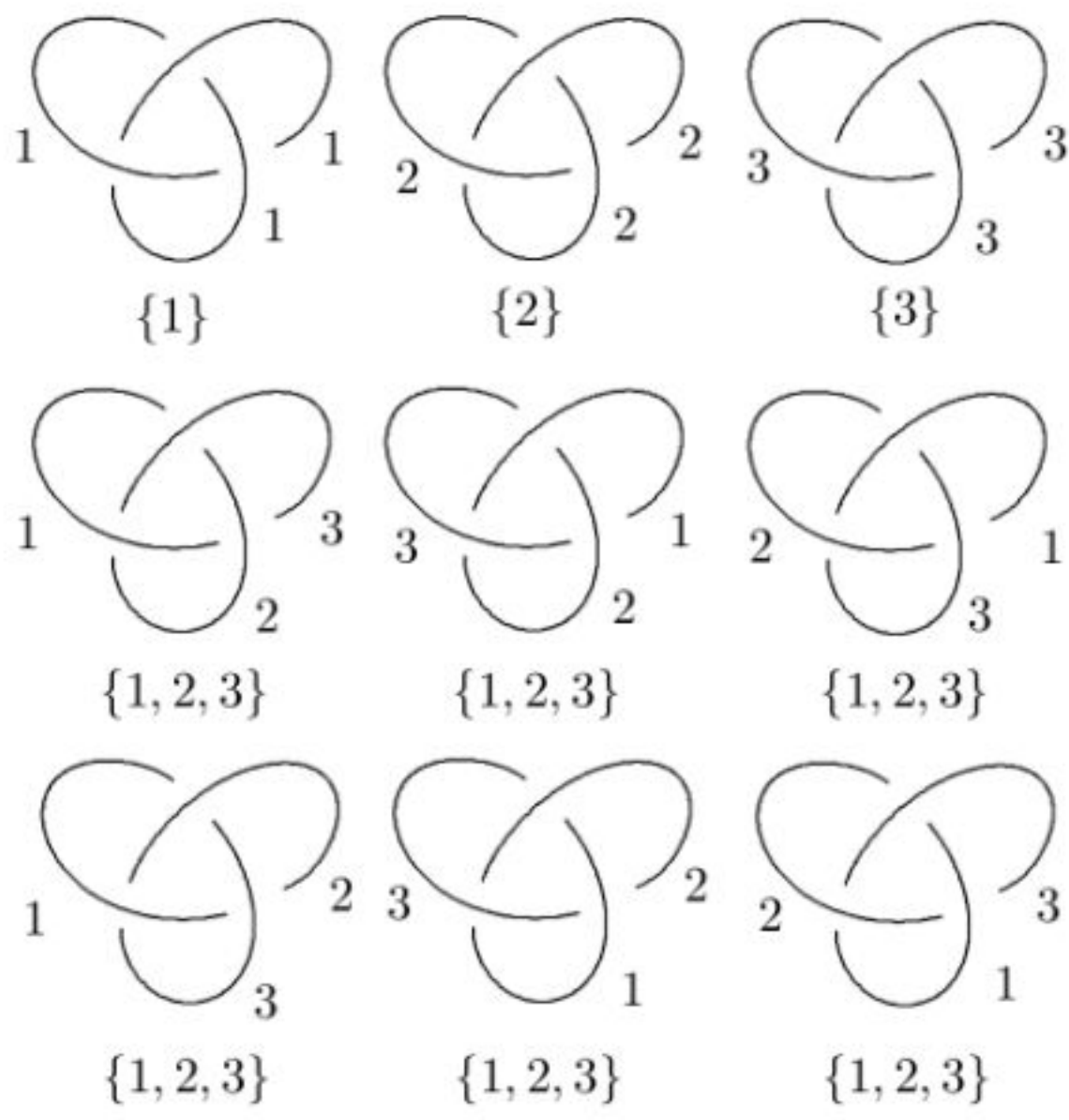
has generating function

$$2 + 3u + u^2 + 2u^3 + u^4.$$

The image enhancement idea works for all of the types of knots and links (unoriented, oriented, framed) and all of the types of appropriate coloring objects (kei, quandle, group, rack, bikei, biquandle or birack). Thus, given a kei X we can define the *image enhancement polynomial*

$$\Phi_X^{\text{Im}}(L) = \sum_{f \in \text{Hom}(\mathcal{K}(L), X)} u^{|\text{Im}(f)|}.$$

Example 115. The trefoil knot 3_1 has nine colorings by the 3-element Takasaki kei $X = \mathbb{Z}_3$ with $x \triangleright y = 2x + 2y$ with image subkeis as depicted.



Then the image enhancement multiset invariant is

$$\Phi_X^{\text{Im}, M}(3_1) = \{1, 1, 1, 3, 3, 3, 3, 3, 3\}$$

and the image enhancement polynomial is $\Phi_X^{\text{Im}}(3_1) = 3u + 6u^3$.

Writhe Enhancement. In the definition of the rack and birack counting invariants, we sum the numbers of colorings of a diagram of L over a tile of framing vectors whose side length is the rack or birack characteristic N . One easy way to enhance this invariant is to keep track of which colorings come from which framing.

Suppose we have a rack or birack of characteristic N . We can form a multiset enhancement of the rack or birack counting invariant by making a multiset of ordered pairs of the number of colorings in each framing together with the framing vector, i.e.,

$$\Phi_X^{W, M} = \{(|\text{Hom}(\mathcal{BR}(L_{\vec{w}}), X)|, \vec{w}) \mid \vec{w} \in (\mathbb{Z}_N)^c\}.$$

Alternatively, we can get a multivariable polynomial by defining dummy variables q_1, \dots, q_c and converting framing vectors $\vec{w} = (w_1, \dots, w_c)$ into monomials $q_1^{w_1} \dots q_c^{w_c}$, which we denote by $q^{\vec{w}}$. Then the *writhe enhancement polynomial* is

$$\Phi_X^W(L) = \sum_{\vec{w} \in (\mathbb{Z}_N)^c} |\text{Hom}(\mathcal{BR}(L_{\vec{w}}), X)| q^{\vec{w}}.$$

Example 116. Let X be the nontrivial two-element rack, i.e., the constant action rack $X = \{1, 2\}$ defined by the bijection $\sigma(1) = 2$ and $\sigma(2) = 1$. Then as we have seen, the Hopf link L has 4 X -colorings with writhe vector $\vec{w} = (1, 1)$ and no colorings with writhe vectors $\vec{w} \in \{(0, 0), (0, 1), (1, 0)\}$. Thus, we have

$$\Phi_X^W(L) = 0q_1^0q_2^0 + 0q_1^0q_2^1 + 0q_1^1q_2^0 + 4q_1^1q_2^1 = 4q_1q_2.$$

On the other hand, the unlink of two components L' has 4 colorings with writhe vector $\vec{w} = (0, 0)$ and no colorings with writhe vector $\vec{w} \in \{(0, 1), (1, 0), (1, 1)\}$, so the unlink has invariant

$$\Phi_X^W(L') = 4q_1^0q_2^0 + 0q_1^0q_2^1 + 0q_1^1q_2^0 + 0q_1^1q_2^1 = 4.$$

In particular, the Hopf link and the two-component unlink have the same rack counting invariant value $\Phi_x^{\mathbb{Z}}(L) = \Phi_x^{\mathbb{Z}}(L') = 4$ but are distinguished by their writhe enhancements.

Homomorphism Enhancements. Another basic way to enhance a counting invariant is to select a surjective homomorphism $g : X \rightarrow Y$ and use it to divide the set of X colorings of L into disjoint subsets. As with the image enhancement, we will use the case of kei for simplicity, but the same idea works with quandles, racks, bikei, biquandles and biracks as well.

Let X and Y be finite kei and suppose that $g : X \rightarrow Y$ is a surjective kei homomorphism. Then if L is a link and $f \in \text{Hom}(\mathcal{K}(L), X)$ is an X -coloring of L , then $gf : \mathcal{K}(L) \rightarrow Y$ defined by $gf(x) = g(f(x))$ is a Y -coloring of L .

We can define an equivalence relation \sim on $\text{Hom}(\mathcal{K}(L), X)$ by setting $f \sim f'$ if $gf = gf'$:

- $gf = gf$ so \sim is reflexive,
- $gf = gf'$ implies $gf' = gf$ so \sim is symmetric, and

- $gf = gf'$ and $gf' = gf''$ implies $gf = gf''$, so \sim is transitive.

Then the *homomorphism enhancement* of the kei counting invariant with respect to $g : X \rightarrow Y$ is the partition of the set of X -colorings by the equivalence relation \sim into the set of equivalence classes

$$\Phi_g^M(L) = \{[f] \mid f \in \text{Hom}(\mathcal{K}(L), X), gf \sim gf'\}.$$

We can also formulate a polynomial version of the homomorphism enhancement:

$$\Phi_g(L) = \sum_{h \in \text{Hom}(\mathcal{K}(L), Y)} u^{|\{f \in \text{Hom}(\mathcal{K}(L), X) \mid fg=h\}|}.$$

We can think of this enhancement as an enhancement of the counting invariant with respect to X (grouping together X -colorings which project to the same Y -coloring) or with respect to Y (for each Y -coloring, the number of X -colorings it lifts to is an invariant of Y -colored isotopy) since we can recover both $\Phi_Y^{\mathbb{Z}}(L)$ as the sum of coefficients and $\Phi_X^{\mathbb{Z}}(L)$ as the sum of the products of each term's exponent times its coefficient.

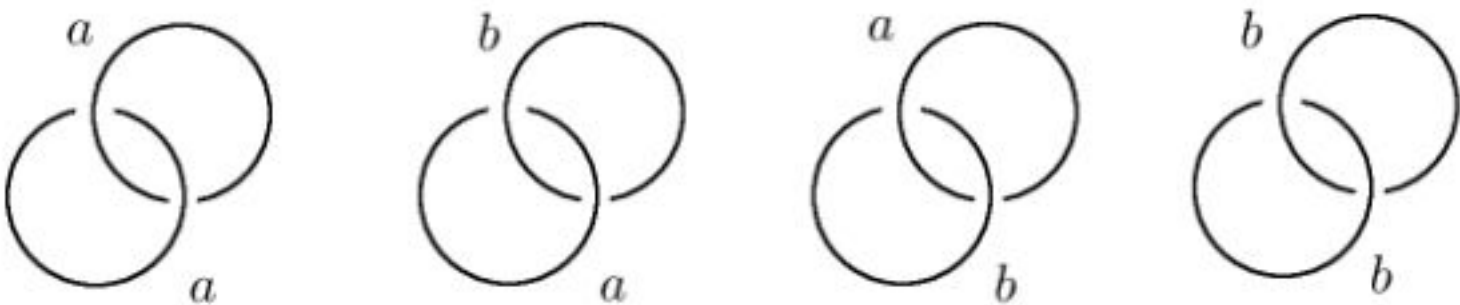
Example 117. Consider the kei X and Y with operation tables

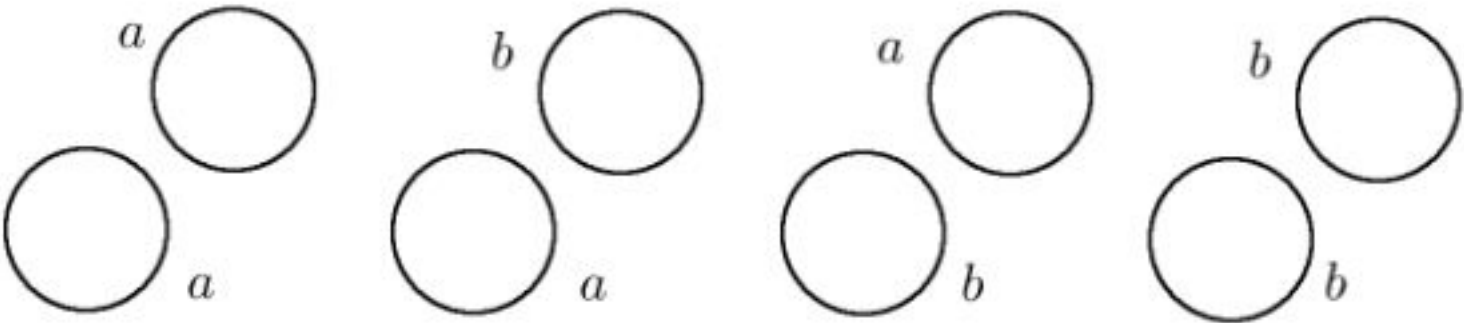
\triangleright_X	1	2	3
1	1	1	2
2	2	2	1
3	3	3	3

and

\triangleright_Y	a	b
a	a	a
b	b	b

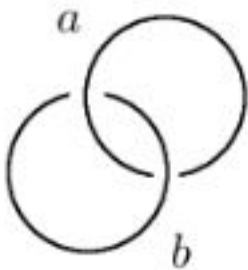
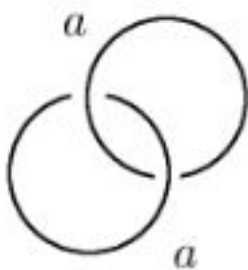
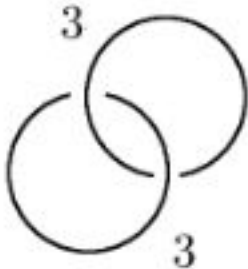
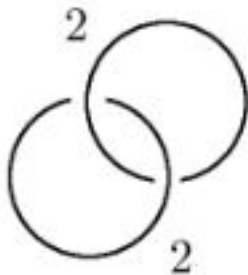
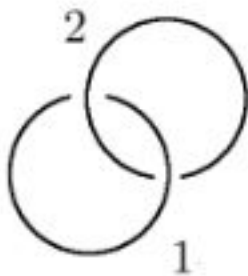
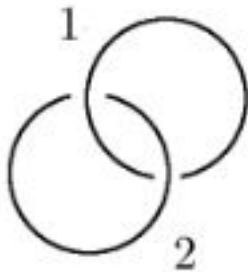
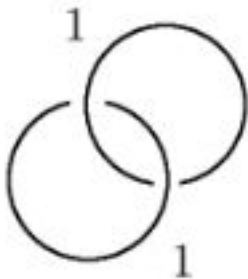
Then the map $f : X \rightarrow Y$ defined by $f(1) = f(2) = a$ and $f(3) = b$ is a kei homomorphism. The Hopf link L and unlink of two components U_2 both have $\Phi_Y^{\mathbb{Z}}(L) = \Phi_Y^{\mathbb{Z}}(U_2) = 4$ colorings by Y as depicted.

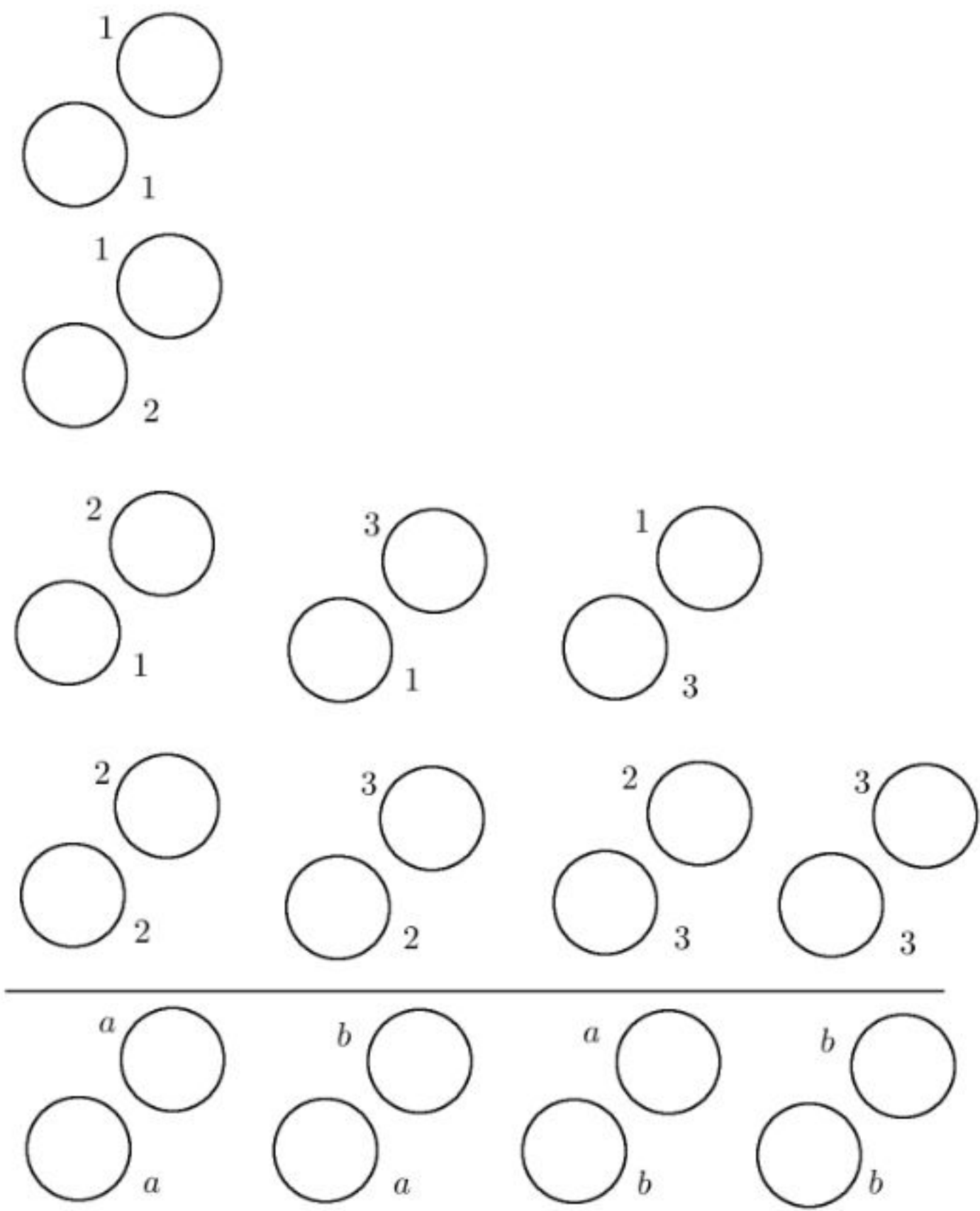




However, these colorings lift to different numbers of colorings by X , and we have

$$\Phi_g(L) = 2u^0 + u^1 + u^4 \neq \Phi_g(U_2) = u^1 + 2u^2 + u^4.$$





Exercises. 1. Compute the image enhancement polynomial for the figure eight knot with respect to the quandle with operation matrix

$$\begin{bmatrix} 1 & 3 & 4 & 2 \\ 4 & 2 & 1 & 3 \\ 2 & 4 & 3 & 1 \\ 3 & 1 & 2 & 4 \end{bmatrix}.$$

2. Compute the writhe enhancement polynomial for the Hopf link with respect to the (t,s) -rack \mathbb{Z}_4 with $t = 1$ and $s = 2$.

3. Compute the writhe enhancement polynomial for the Hopf link with respect to the (t, s, r) -birack \mathbb{Z}_4 with $t = 1$, $s = 2$ and $r = 3$.

4. Let X and Y be the quandles defined by the operation matrices

$$X = \begin{bmatrix} 1 & 4 & 4 & 1 \\ 2 & 2 & 2 & 2 \\ 3 & 3 & 3 & 3 \\ 4 & 1 & 1 & 4 \end{bmatrix} \quad Y = \begin{bmatrix} a & a \\ b & b \end{bmatrix}.$$

Show that the map $f : X \rightarrow Y$ defined by $f(1) = f(4) = a$ and $f(2) = f(3) = b$ is a quandle homomorphism and compute the homomorphism enhancement for the Hopf link.

5. Compute the homomorphism enhancement for the trefoil with respect to the homomorphism $f : \mathbb{Z}_3 \oplus \mathbb{Z}_3 \rightarrow \mathbb{Z}_3$ given by $f(x, y) = x$ where the quandle structures are given by

$$(x, y) \triangleright (u, v) = (2u - x, 2v - y) \quad \text{and} \quad x \triangleright u = 2u - x.$$

2. Structure Enhancements

Many of the examples we have seen of kei, quandles and their generalizations are not just kei or quandles but have additional algebraic structure. Alexander quandles are also Λ -modules; conjugation quandles are also groups; symplectic quandles are also vector spaces. In many cases we can exploit this extra structure to enhance the counting invariant.

Recall that to enhance a counting invariant, we can find an invariant ϕ of X -colored diagrams; then the multiset of ϕ -values gives us a potentially stronger invariant from which we can recover the counting invariant as the multiset's cardinality. As we have seen, we frequently find it useful to convert the multiset into a polynomial by taking a *generating function*, i.e., a polynomial in a variable u with elements as powers and multiplicities as coefficients.

Symplectic Quandle Enhancement. Let X be a *symplectic quandle*, i.e., a vector space V with an antisymmetric bilinear form $\langle, \rangle :$

$V \times V \rightarrow \mathbb{F}$, which is a quandle under the operations

$$\begin{aligned}\vec{x} \triangleright \vec{y} &= \vec{x} + \langle \vec{x}, \vec{y} \rangle \vec{y}, \\ \vec{x} \triangleright^{-1} \vec{y} &= \vec{x} - \langle \vec{x}, \vec{y} \rangle \vec{y}.\end{aligned}$$

Now suppose we have an X -coloring of a knot or link L . Each of the elements of X is a vector, so we can consider the subspace of X spanned by the elements coloring the arcs in X . We must be careful here: to get a subspace which is invariant under X -colored Reidemeister moves, we must take the subspace spanned by the full image subquandle of X generated by the arc colors, not just the space spanned by the arc colors themselves. Then we can use the dimension of the subspace of X spanned by the image of f for each quandle homomorphism $f : \mathcal{Q}(L) \rightarrow X$ as our multiset elements to obtain the *symplectic quandle enhancement multiset*

$$\Phi_X^{\text{Symp}, M}(L) = \{\dim(\text{Im}(f)) \mid f \in \text{Hom}(\mathcal{Q}(L), X)\}$$

and the *symplectic quandle enhancement polynomial*

$$\Phi_X^{\text{Symp}}(L) = \sum_{f \in \text{Hom}(\mathcal{Q}(L), X)} u^{\dim(\text{Im}(f))}.$$

If X is a finite vector space, we can alternatively use $\phi(f) = |\text{Im}(f)|$ in place of $\dim(\text{Im}(f))$.

Example 118. Let $X = (\mathbb{Z}_2)^2 = \{(0, 0), (1, 0), (0, 1), (1, 1)\}$, the four element vector space over \mathbb{Z}_2 , and let

$$\langle (x_1, x_2), (y_1, y_2) \rangle = \begin{bmatrix} x_1 & x_2 \end{bmatrix} \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = x_1 y_2 + x_2 y_1.$$

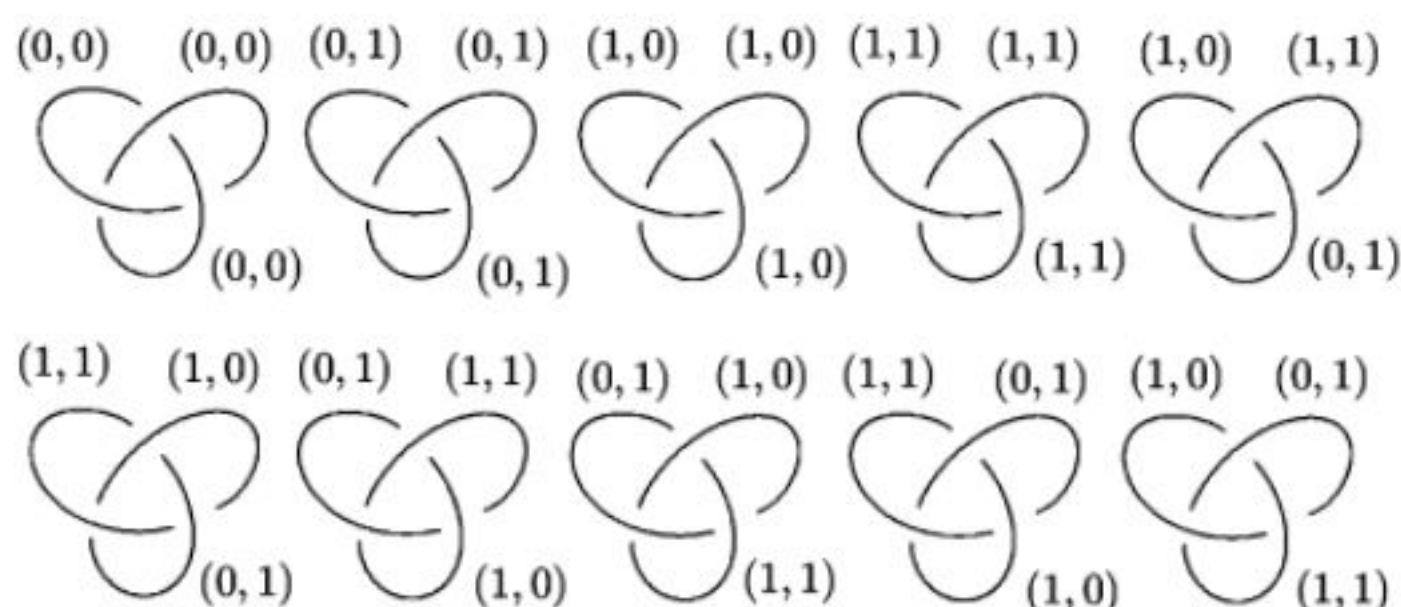
Then X is a symplectic quandle with operation

$$\begin{aligned}(x_1, x_2) \triangleright (y_1, y_2) &= (x_1, x_2) + (x_1 y_2 + x_2 y_1)(y_1, y_2) \\ &= (x_1 + x_1 y_1 y_2 + x_2 y_1^2, x_2 + x_1 y_2^2 + x_2 y_1 y_2).\end{aligned}$$

X has the operation table

\triangleright	(0, 0)	(0, 1)	(1, 0)	(1, 1)
(0, 0)	(0, 0)	(0, 0)	(0, 0)	(0, 0)
(0, 1)	(0, 1)	(0, 1)	(1, 1)	(1, 0)
(1, 0)	(1, 0)	(1, 1)	(1, 0)	(0, 1)
(1, 1)	(1, 1)	(1, 0)	(0, 1)	(1, 1)

Then for instance the trefoil knot 3_1 has 10 X -colorings as depicted



The multiset version of the invariant is then

$$\begin{aligned}\Phi_X^{\text{Symp}, M}(3_1) &= \{\dim(\mathbb{Z}_2[(0,0)], \dim(\mathbb{Z}_2[(0,1)]), \dim(\mathbb{Z}_2[(1,0)]), \\ &\quad \dim(\mathbb{Z}_2[(1,1)]), 6 \times \dim(\mathbb{Z}_2[(0,1), (1,0), (1,1)])\} \\ &= \{0, 1, 1, 1, 2, 2, 2, 2, 2, 2\}\end{aligned}$$

or in polynomial form,

$$\Phi_X^{\text{Symp}}(3_1) = u^0 + 3u^1 + 6u^2 = 1 + 3u + 6u^2.$$

Module Enhancements. Several types of knot-coloring structures also have a module structure over a ring R , e.g. Alexander quandles and biquandles, (t, s) -racks and (t, s, r) -biracks, etc. For each of these, we can enhance the counting invariant by setting $\phi(f)$ equal to the cardinality of the submodule spanned by $\text{Im}(f)$ if X is finite or by setting $\phi(f)$ equal to the rank of the R -submodule spanned by $\text{Im}(f)$ if X is infinite, obtaining invariants

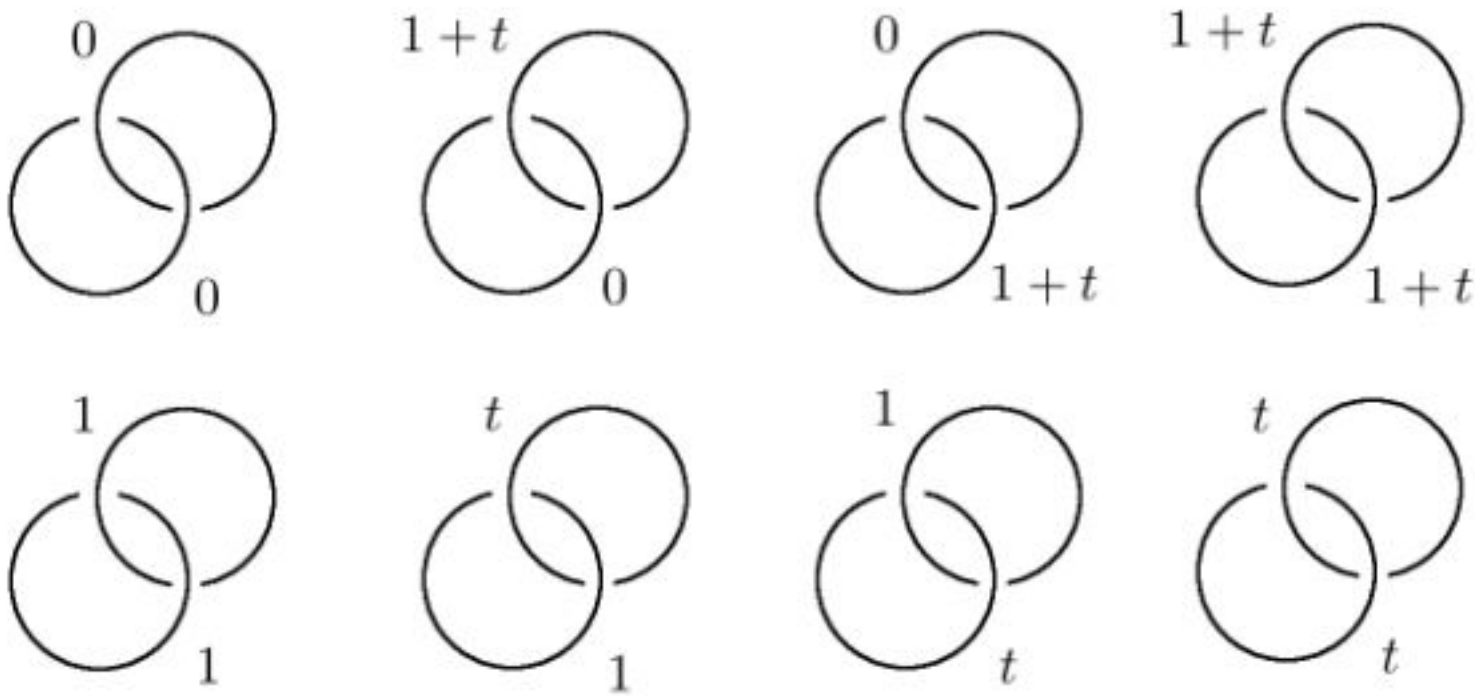
$$\begin{aligned}\Phi_X^{\text{Mod}, M}(L) &= \{ |R[\text{Im}(f)]| \mid f \in \text{Hom}(\mathcal{Q}(L), X) \} \\ \Phi_X^{\text{Mod}}(L) &= \sum_{f \in \text{Hom}(\mathcal{Q}(L), X)} u^{|R[\text{Im}(f)]|}.\end{aligned}$$

Example 119. Let X be the Alexander quandle $\Lambda_2/(1+t^2)$ from Example 73. Then X can be identified with the set $\{0, 1, t, 1+t\}$

where $t^2 = 1$, and X has the operation table

\triangleright	0	1	t	$1+t$
0	0	$1+t$	$1+t$	0
1	t	1	1	t
t	1	t	t	1
$1+t$	$1+t$	0	0	$1+t$

Then the Hopf link L has eight quandle colorings by X as depicted:



Closing the sets of arc colors under the quandle operation, we have

$$\{\{0\}, \{1\}, \{t\}, \{1+t\}, 2 \times \{0, 1+t\}, 2 \times \{1, t\}\}.$$

Then to get the submodules spanned by each of these sets, recall that a submodule is closed under addition and scalar multiplication. Then the Λ_2 -submodules spanned by the image subquandles are

$$\begin{aligned} &\{\{0\}, \{0, 1, t, 1+t\}, \{0, 1, t, 1+t\}, \{0, 1+t\}, \\ &\quad 2 \times \{0, 1+t\}, 2 \times \{0, 1, t, 1+t\}\} \end{aligned}$$

for submodule enhancement invariants

$$\begin{aligned} \Phi_X^{\text{Mod}, M}(L) &= \{1, 2, 2, 2, 4, 4, 4, 4\}, \\ \Phi_X^{\text{Mod}}(L) &= u + 3u^2 + 4u^4. \end{aligned}$$

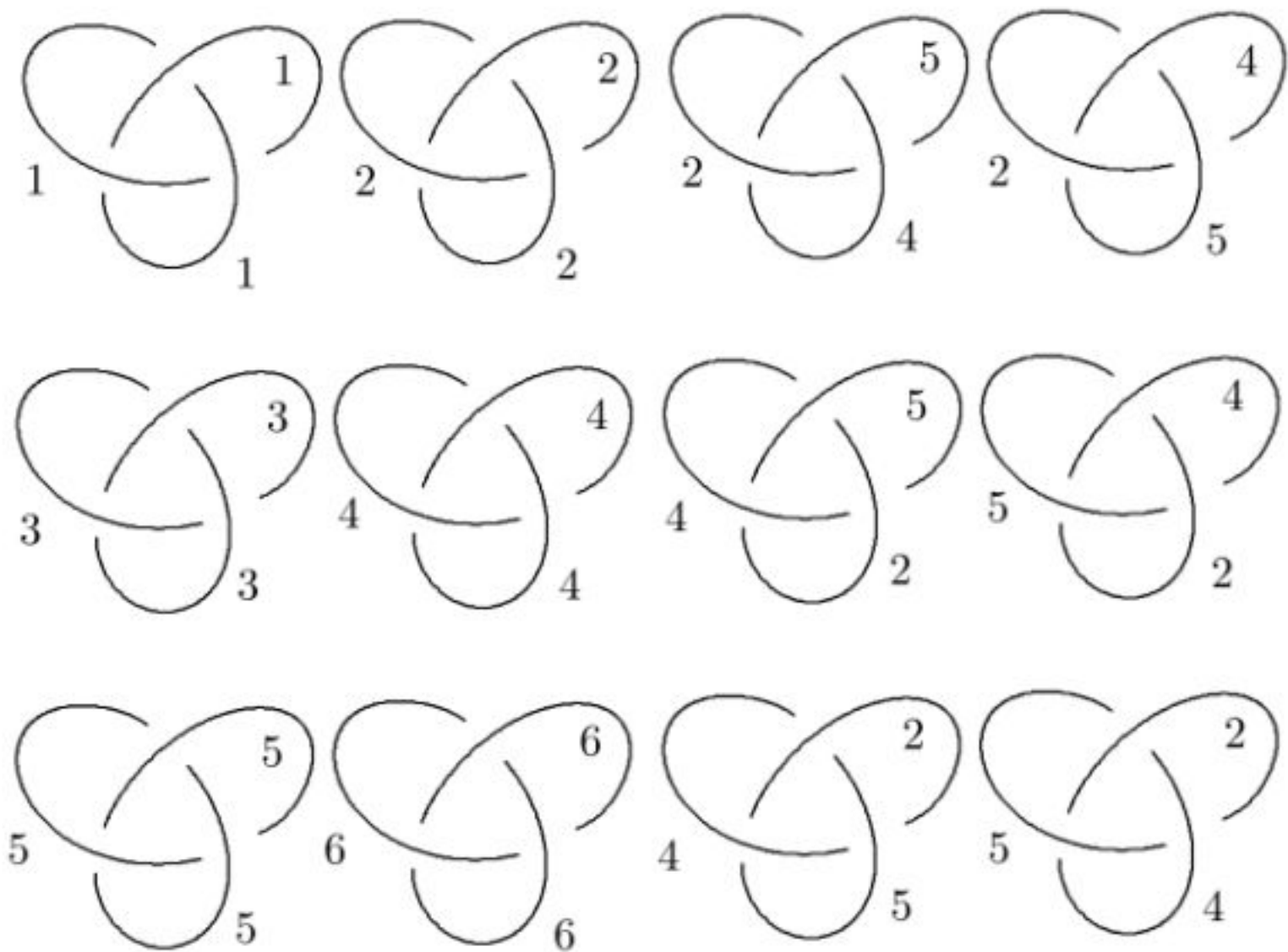
Group Enhancements. For our last structure enhancements, recall that a group G is a quandle under conjugation $x \triangleright y = y^{-1}xy$ and a kei under the core operation $x \triangleright y = yx^{-1}y$. We can use the group structure to enhance the counting invariant in several ways.

First, if X is a conjugation quandle or core quandle, for each X -coloring we can take ϕ to be the cardinality of the subgroup of X generated by the image subquandle. Let's see how it works with an example:

Example 120. Let $X = \{x_1, x_2, x_3, x_4, x_5, x_6\}$ be the conjugation quandle of the group D_3 of symmetries of an equilateral triangle. Then X has group and quandle operation tables

\cdot	1	2	3	4	5	6		\triangleright	1	2	3	4	5	6
1	1	2	3	4	5	6		1	1	1	1	1	1	1
2	2	1	4	3	6	5		2	2	2	5	5	4	4
3	3	5	6	2	4	1	and	3	3	6	3	6	6	3
4	4	6	5	1	3	2		4	4	5	2	4	2	5
5	5	3	2	6	1	4		5	5	4	4	2	5	2
6	6	4	1	5	2	3		6	6	3	6	3	3	6

The trefoil knot 3_1 has 12 colorings by X as depicted:



Then each of the monochromatic colorings has a singleton image subquandle, while the other six have image subquandle $\{2, 4, 5\}$ isomorphic to the three-element Takasaki kei. Recall that the subgroup

generated by a subset of a group is the closure of the set under multiplication and inverses; then we have the following subgroups:

$$\{\{1\}, \{1, 2\}, \{1, 4\}, \{1, 5\}, 2 \times \{1, 3, 6\}, 6 \times \{1, 2, 3, 4, 5, 6\}\}$$

so the multiset invariant is

$$\Phi_X^{\text{Subg}, M}(3_1) = \{1, 2, 2, 2, 3, 3, 6, 6, 6, 6, 6, 6\}$$

with polynomial version

$$\Phi_X^{\text{Subg}}(3_1) = u + 3u^2 + 2u^3 + 6u^6.$$

Another enhancement involving groups uses the observation that in the operation table of a quandle (or kei, rack, bikei, biquandle or birack), the columns are always permutations. Recall that the set of permutations of n things forms a group of $n!$ elements called the *symmetric group on n letters*, denoted S_n . Then instead of associating the subgroup of X generated by $\text{Im}(f)$, we can associate the subgroup of S_n generated by the columns of the operation table(s) of X corresponding to the elements of $\text{Im}(f)$. We call these groups the *column groups* of the subquandles $\text{Im}(f)$, denoted $\mathcal{CG}(\text{Im}(f))$.

Note that for kei, quandles, and racks the permutations represented by the columns in the operation tables are inner automorphisms, so for these structures the column group is a subgroup of the inner automorphism group; for bikei, biquandle and biracks, the column group elements are generally not automorphisms.

Example 121. Let us continue Example 120 and find the column group enhancement. Each of the singleton image subquandles determines a *cyclic subgroup*, i.e., a subgroup consisting of powers of a single permutation σ . Such a subgroup is isomorphic to \mathbb{Z}_n where n is the smallest integer greater than zero such that $\sigma^n = 1$. For example, x_2 has column permutation $[1, 2, 6, 5, 4, 3]$, and the composition of σ with itself is $\sigma^2 = [1, 2, 3, 4, 5, 6]$, the identity permutation. Thus, the column group of the image subquandle $\{x_2\}$ is a copy of \mathbb{Z}_2 . The image subquandle $\{2, 4, 5\}$ has column group generated by the permutations $[1, 2, 6, 5, 4, 3]$, $[1, 5, 6, 4, 2, 3]$, $[1, 4, 6, 2, 5, 3]$ which turns out to be isomorphic to S_3 itself. We end up with column groups as

listed:

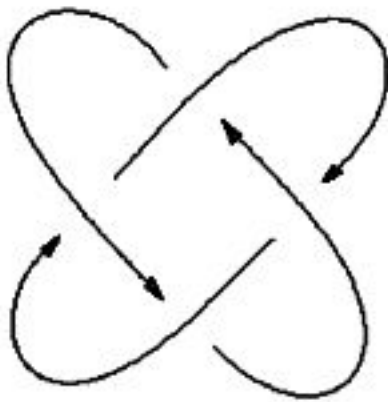
$\text{Im}(f)$	$\mathcal{CG}(\text{Im}(f))$
$\{1\}$	$\mathbb{Z}_1 = \{0\}$
$\{2\}$	\mathbb{Z}_2
$\{3\}$	\mathbb{Z}_3
$\{4\}$	\mathbb{Z}_2
$\{5\}$	\mathbb{Z}_2
$\{6\}$	\mathbb{Z}_3
$\{2, 4, 5\}$	S_3

Then we have column group enhancements

$$\begin{aligned}\Phi_X^{\mathcal{CG},M}(3_1) &= \{1, 2, 2, 2, 3, 3, 6, 6, 6, 6, 6, 6\} \quad \text{and} \\ \Phi_X^{\mathcal{CG}}(3_1) &= u + 3u^2 + 2u^3 + 6u^6.\end{aligned}$$

Exercises. 1. Find the symplectic quandle enhancement polynomial for the knot 6_1 with respect to the symplectic quandle in Example 118.

2. Find the module enhancement polynomial for the $(4, 2)$ -torus link



with respect to the Alexander quandle \mathbb{Z}_4 with $t = 3$.

3. Find the module enhancement polynomial for the Hopf link with respect to the (t, s, r) -birack \mathbb{Z}_4 with $t = 1$, $s = 2$ and $r = 3$.

4. Find the subgroup enhancement for the figure eight knot with respect to the conjugation quandle of the dihedral group D_4 (the symmetry group of a square).

5. Find the column group enhancement polynomial for the Hopf link with coloring kei with operation matrix

$$\begin{bmatrix} 1 & 1 & 1 & 2 \\ 2 & 2 & 2 & 1 \\ 3 & 3 & 3 & 3 \\ 4 & 4 & 4 & 4 \end{bmatrix}.$$

3. Quandle Polynomials

One of the differences between quandles and groups is that unlike groups, quandles have no identity element; instead, each element acts as its own identity element. Another way to say this is that in a group, there is a single element which acts trivially on everything else, while in a quandle the trivial action is distributed throughout the quandle. The *quandle polynomial* is a way of quantifying this distribution of trivial action as a two-variable polynomial.

Let X be a finite quandle. For each element $x \in X$, let $r(x)$ be the number of elements of X which act trivially on x , i.e. the set

$$r(x) = |\{y \in X \mid x \triangleright y = x\}|$$

and let $c(x)$ be the set of elements of X on which x acts trivially, i.e.

$$c(x) = |\{y \in X \mid y \triangleright x = y\}|.$$

In terms of the quandle's operation table, $r(x)$ counts the number of x s in row x and $c(x)$ counts how many entries in the column of x equal their row number.

Example 122. Consider the quandle X with operation table

\triangleright	1	2	3	4
1	1	1	1	1
2	3	2	2	3
3	2	3	3	2
4	4	4	4	4

Then $r(1) = 4$ and $r(2) = 2$ since row 1 has four 1s and row 2 has only two 2s. Similarly, $c(1) = 2$ and $c(2) = 4$ since column 1 has only two entries equal to their row numbers (namely, rows 1 and 4) but column 2 has all four entries equal to their row numbers.

For every element $x \in X$, we have a pair $(r(x), c(x))$ of integers. We can express this data conveniently as a polynomial in two variables which we call the *quandle polynomial* of X :

$$p(X) = \sum_{x \in X} t^{r(x)} s^{c(x)}.$$

Example 123. The quandle X with operation table

\triangleright	1	2	3	4
1	1	1	1	1
2	3	2	2	3
3	2	3	3	2
4	4	4	4	4

has the following $r(x)$ and $c(x)$ values:

x	$r(x)$	$c(x)$
1	4	2
2	2	4
3	2	4
4	4	2

and thus quandle polynomial

$$p(X) = t^4 s^2 + t^2 s^4 + t^2 s^4 + t^4 s^2 = 2t^4 s^2 + 2t^2 s^4.$$

Now, suppose $S \subset X$ is a subquandle of X . Then as a stand-alone quandle, S has its own quandle polynomial $p(S)$, but we can also form the *subquandle polynomial* of S as a subquandle of X by summing the contributions of the elements of S to $p(X)$:

$$p(S \subset X) = \sum_{x \in S} t^{r(x)} s^{c(x)}$$

where $c(x)$ and $r(x)$ are computed from the operation table of X . These subquandle polynomials carry information not just about the isomorphism type of S but also about how S is embedded in X , quite appropriate since knot theory is all about how certain objects are embedded inside other objects.

Example 124. The quandle X with operation table

\triangleright	1	2	3	4
1	1	1	2	2
2	2	2	1	1
3	3	3	3	3
4	4	4	4	4

has the following $r(x)$ and $c(x)$ values:

x	$r(x)$	$c(x)$
1	2	4
2	2	4
3	4	2
4	4	2

Then the subquandles $S_1 = \{1, 2\}$ and $S_2 = \{3, 4\}$ are both trivial quandles on two elements and thus both have quandle polynomial $p(S_1) = p(S_2) = 2t^2s^2$, but their subquandle polynomials are different, reflecting the fact that they are embedded in X in different ways:

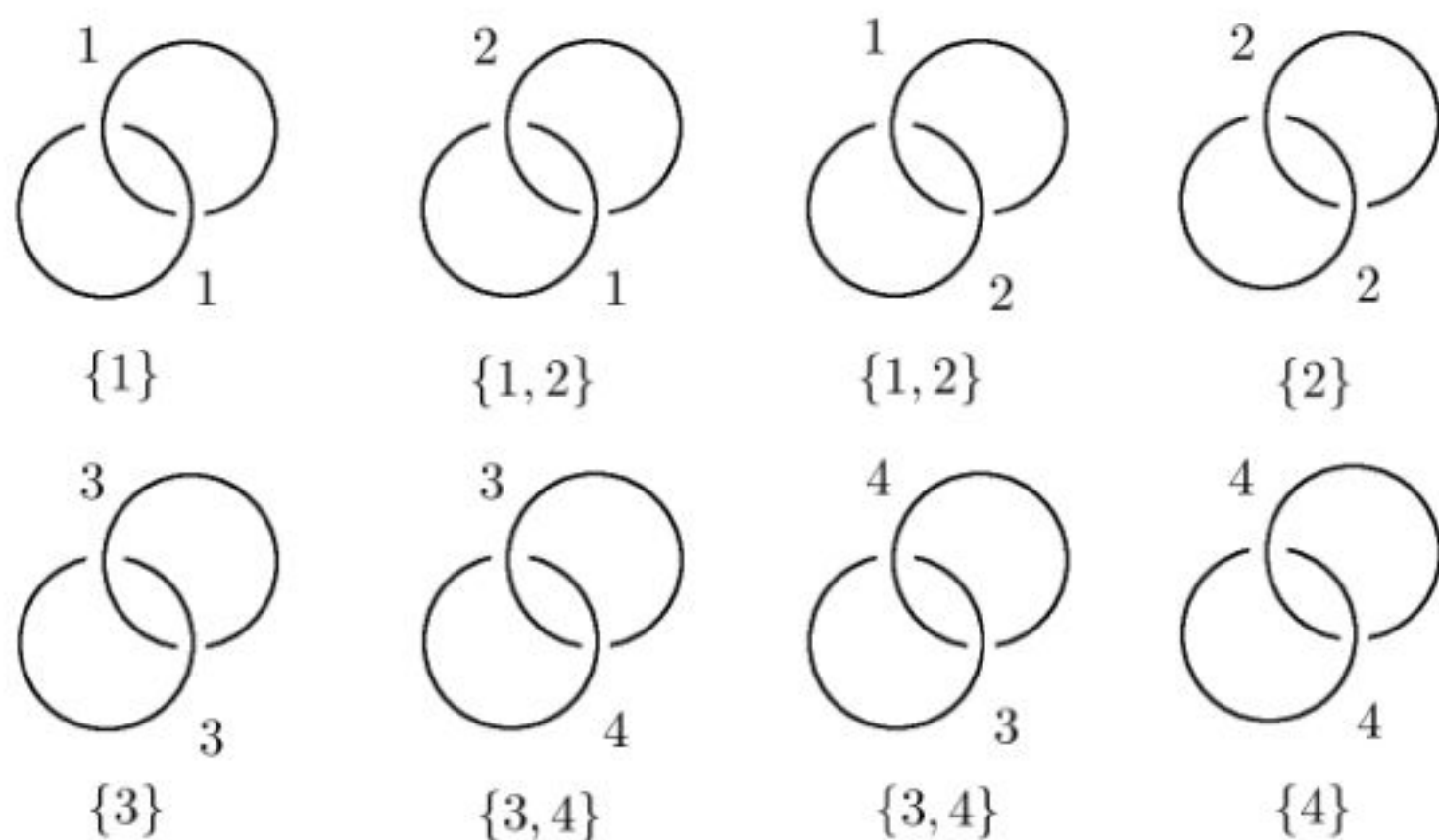
$$p(S_1 \subset X) = 2t^2s^4 \neq p(S_2 \subset X) = 2t^4s^2.$$

We can use subquandle polynomials to get a multiset-valued enhancement of the quandle counting invariant (or indeed, a further enhancement of the image enhancement invariant) by collecting for each $f \in \text{Hom}(\mathcal{Q}(L), X)$ the subquandle polynomial of the image subquandle. That is, the *subquandle polynomial enhancement* is the multiset

$$\Phi_X^p(L) = \{p(\text{Im}(f) \subset X) \mid f \in \text{Hom}(\mathcal{Q}(L), X)\}.$$

Example 125. Consider the Hopf link L and the kei $X = \{1, 2, 3, 4\}$ with the operation table in Example 124. There are eight X -colorings

of L as depicted.



Then we have subquandle polynomial enhancement

$$\begin{aligned}
 \Phi_X^p(L) &= \{p(\{1\} \subset X), p(\{2\} \subset X), p(\{3\} \subset X), p(\{4\} \subset X), \\
 &\quad 2 \times p(\{1, 2\} \subset X), 2 \times p(\{3, 4\} \subset X)\} \\
 &= \{t^2 s^4, t^2 s^4, t^4 s^2, t^4 s^2, 2 \times 2t^2 s^4, 2 \times 2t^4 s^2\} \\
 &= \{2 \times t^2 s^4, 2 \times t^4 s^2, 2 \times 2t^2 s^4, 2 \times 2t^4 s^2\}.
 \end{aligned}$$

Exercises. 1. Prove that if $\sigma : X \rightarrow Y$ is a quandle isomorphism, then $p(X) = p(Y)$.

2. Prove that a Latin quandle of n elements always has quandle polynomial $p(X) = nts$.

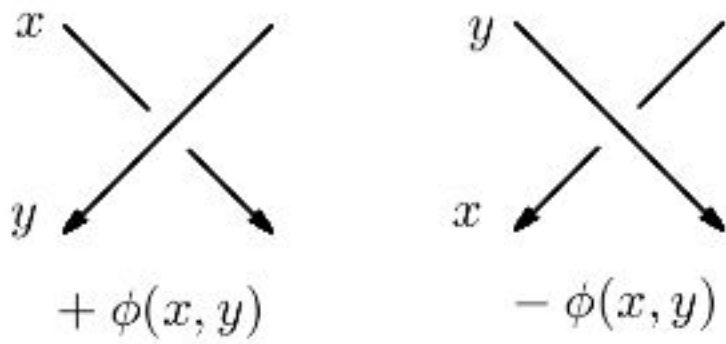
3. Compute the subquandle polynomial enhancement invariant for the figure eight knot with respect to the Alexander quandle $X = \Lambda_2/(1 + t + t^2)$.

4. Prove that there is no quandle with quandle polynomial $3t^2 s^2$.

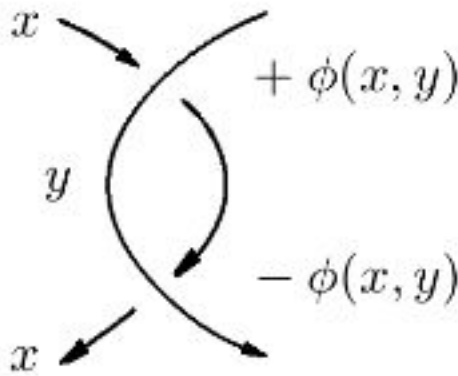
5. Define an enhancement of the rack counting invariant using subrack polynomials and compute an example.

4. Quandle Cocycle Enhancements

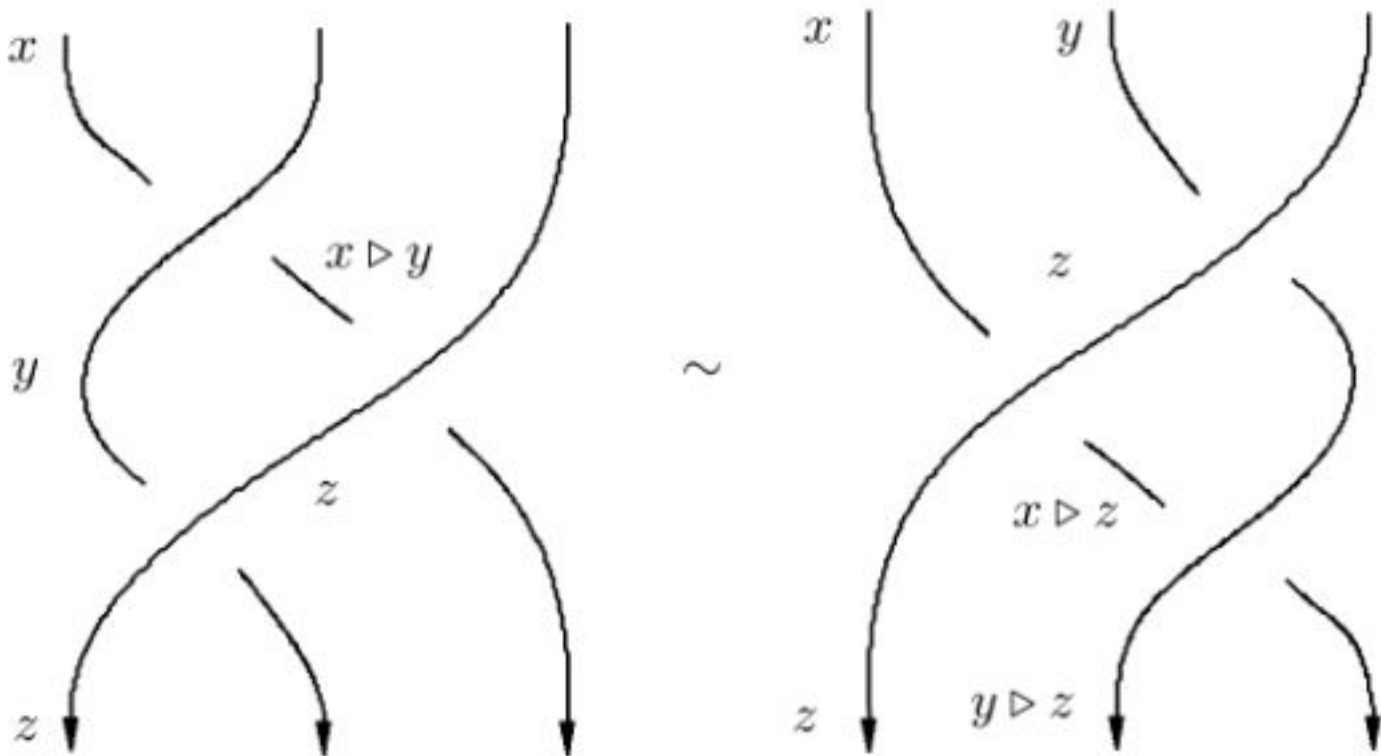
Let (X, \triangleright) be a finite quandle. We would like to define an enhancement of the quandle counting invariant $\Phi_X^{\mathbb{Z}}$ by defining a function $\phi : X \times X \rightarrow \mathbb{Z}$ where each crossing contributes an amount $\pm\phi(x, y)$ depending on its quandle coloring as shown:



This labeling rule has the advantage that the contributions from the two crossings at a type II move cancel out, like in the linking number case:



We can then ask what kind of function ϕ gives us an invariant total sum called a *Boltzmann weight* under the other Reidemeister moves. For the type III move, we have



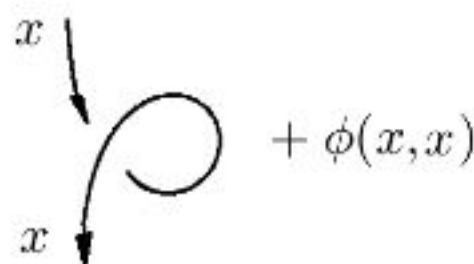
and thus we need

$$\phi(x, y) + \phi(y, z) + \phi(x \triangleright y, z) = \phi(x, z) + \phi(y, z) + \phi(x \triangleright z, y \triangleright z)$$

or, simplifying,

$$\phi(x, y) + \phi(x \triangleright y, z) = \phi(x, z) + \phi(x \triangleright z, y \triangleright z).$$

The type I move says we need $\phi(x, x) = 0$ for all x .



What kinds of functions ϕ , if any, satisfy these conditions and how can we find them? It turns out, such a function ϕ is precisely a cocycle in a cohomology space H^2 associated to the quandle X .

Quandle Cohomology. Let (X, \triangleright) be a finite quandle. For each integer $n \geq 1$, consider $\mathbb{Z}[X^n]$, the \mathbb{Z} -module with basis given by ordered n -tuples of elements of X . Then $\mathbb{Z}[X^n]$ has elements of the form $\sum \alpha(x_1, \dots, x_n)$.



Despite looking like familiar vectors, we cannot do the usual operations within components on these vectors – these are formal linear combinations, not vectors in \mathbb{Z}^n . For example, if $X = \{0, 1, 2\}$ is the dihedral quandle on three elements, then we can add $3(1, 2) + 2(1, 2)$ in $C_2(X)$ to get $5(1, 2)$, but this **not** equal to $(5, 10)$ since our quandle X does not have a 5 or 10. Similarly, we cannot add $(1, 1) + (1, 0)$ to get $(2, 1)$ since we are not working in \mathbb{Z}^2 or even $\mathbb{Z}_3 \oplus \mathbb{Z}_3$.

Now in order to ensure that the Reidemeister I condition is satisfied, we want $\phi(x, x) = 0$ for all $x \in X$. For each $n \geq 2$, the submodule of $\mathbb{Z}[X^n]$ generated by basis vectors of the form (x_1, x_2, \dots, x_n) where some $x_k = x_{k+1}$ is called the *degenerate* submodule, denoted $C_n^D(X)$. Then we define $C_n(X)$ to be the quotient module $C_n(X) =$

$\mathbb{Z}[X^n]/C_n^D(X)$. In practice, we can simply set vectors with repeated neighboring entries equal to the zero vector.

Next, let $C^n = \{\phi : C_n \rightarrow \mathbb{Z}\}$ be the set of linear transformations from C_n to \mathbb{Z} . For example, if $X = \{1, 2, 3\}$, then $C_2(X)$ has basis

$$\{\vec{b}_1 = (1, 2), \vec{b}_2 = (1, 3), \vec{b}_3 = (2, 1), \\ \vec{b}_4 = (2, 3), \vec{b}_5 = (3, 1), \vec{b}_6 = (3, 2)\}$$

so a typical element of $C_2(X)$ is a formal linear combination of these, e.g.,

$$\vec{v} = 3(1, 3) - 2(2, 1) + (2, 3),$$

which we can write as a column vector

$$\vec{v} = \begin{bmatrix} 0 \\ 3 \\ -2 \\ 1 \\ 0 \\ 0 \end{bmatrix}.$$

In particular, if we think of the elements of $C_2(X)$ as column vectors, then the elements of C^2 can be identified with row vectors of the same size, with function evaluation given by matrix multiplication. That is, we evaluate the function defined by a row vector at a column vector by taking the dot product. Row vectors considered as linear transformations of column vectors are sometimes called *dual vectors* or *covectors*; we can think of cohomology as homology of covectors.

Example 126. Consider the dihedral quandle on three elements X . Then $C_2(X)$ has basis $\{\vec{b}_1 = (1, 2), \dots, \vec{b}_6 = (3, 2)\}$ and is isomorphic to \mathbb{Z}^6 . We can identify $C_2(X)$ with the set of 6×1 column vectors with entries in \mathbb{Z} . Then $C^2(X)$ is the set of linear maps from $C_2(X)$ to \mathbb{Z} ; each such linear map can be expressed as the matrix product of a 1×6 row vector with our 6×1 input vector. For instance, the linear transformation

$$f(\alpha_1 \vec{b}_1 + \alpha_2 \vec{b}_2 + \alpha_3 \vec{b}_3 + \alpha_4 \vec{b}_4 + \alpha_5 \vec{b}_5 + \alpha_6 \vec{b}_6) = 3\alpha_1 - 2\alpha_3 + \alpha_6$$

can be identified with the row vector

$$f = [3 \quad 0 \quad -2 \quad 0 \quad 0 \quad 1] \in C^2(X).$$

Then evaluating f on

$$\vec{v} = 3(1, 2) + 2(2, 1) - (3, 1) + (3, 2) = \begin{bmatrix} 3 \\ 0 \\ 2 \\ 0 \\ -1 \\ 1 \end{bmatrix}$$

yields

$$\begin{aligned} f(\vec{v}) &= \begin{bmatrix} 3 & 0 & -2 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 3 \\ 0 \\ 2 \\ 0 \\ -1 \\ 1 \end{bmatrix} \\ &= 3(3) + 0(0) + 2(-2) + 0(0) + 0(-1) + 1(1) \\ &= 9 - 4 + 1 = 6. \end{aligned}$$

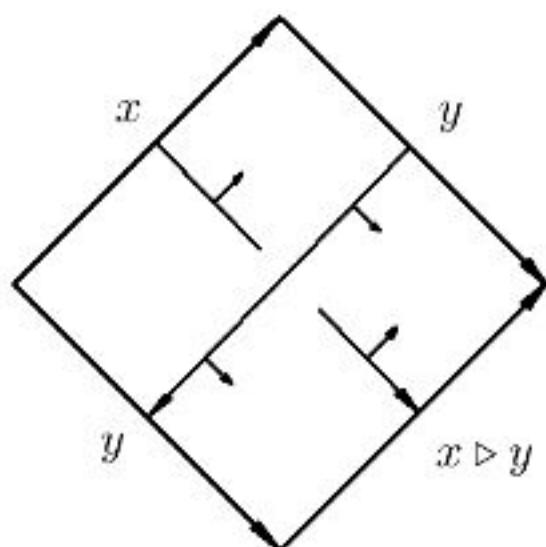
We can also think of $C^n(X)$ in terms of *characteristic functions*. Let X be a quandle and A be an abelian group (for simplicity assume that A is a finite cyclic group \mathbb{Z}_n or the infinite cyclic group \mathbb{Z}). The set of functions from $\mathbb{Z}[X^n]$ to A is generated by the characteristic functions denoted χ_x where $x \in X^n$. This function is defined by

$$\chi_x(y) = \begin{cases} 1 & x = y, \\ 0 & x \neq y, \end{cases}$$

on basis elements x, y ; that is, $\chi_x(y) = 1$ if $x = y$ and $\chi_x(y) = 0$ if $x \neq y$. The advantage of these functions is that we can write any function f from X^n to A uniquely as $f = \sum_{x \in X^n} \lambda_x \chi_x$.

To make this a cochain complex, we need differentials $d^n : C^{n-1} \rightarrow C^n$. Let's start by explaining how the differential maps are defined in low dimensions and then give the general formula. The second

differential $d^2 : C^1 \rightarrow C^2$ can be thought of as the result of pre-composing $f : C^1 \rightarrow \mathbb{Z}$ with the boundary of the 2-dimensional “preferred square” in the following figure



given by

$$\partial_2(x, y) = (x) - (x \triangleright y).$$

That is, for any $f : C_1(X) \rightarrow \mathbb{Z}$, we have $d^2 f = f \partial_2 : C_2 \rightarrow \mathbb{Z}$, so that

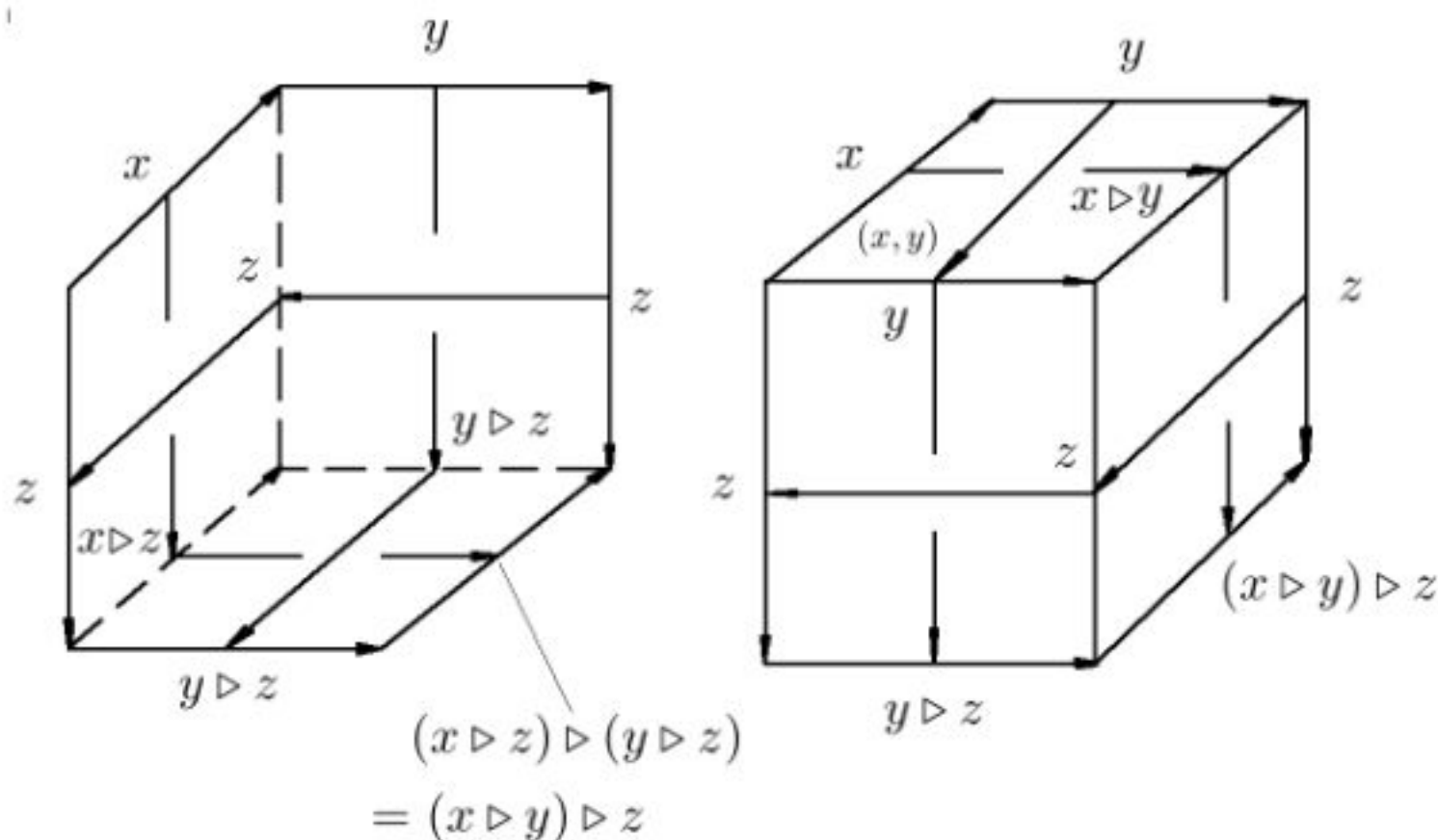
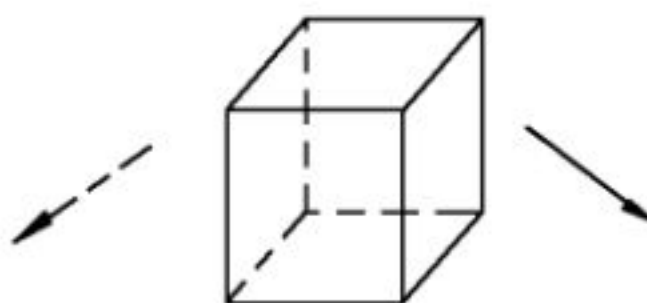
$$d^2 f \left(\sum \alpha(x, y) \right) = \sum \alpha(f(x) - f(x \triangleright y))$$

The third differential is given by precomposition with

$$\partial_3(x, y, z) = (x, z) - (x \triangleright y, z) - (x, y) + (x \triangleright z, y \triangleright z).$$

This formula comes from the boundary of a cube as can be seen from the following figure. The front faces of the cube on the left side of the figure give (x, y) , (y, z) and $(x \triangleright y, z)$ while the back faces on the right side of the figure give (x, z) , (y, z) and $(x \triangleright z, y \triangleright z)$. The faces of the cube are oriented in such a way that a face and its opposite have opposite orientations, giving us a consistent orientation of the cube. Then the face (x, y) is opposite to $(x \triangleright z, y \triangleright z)$, the face (y, z) is opposite to itself and the face $(x \triangleright y, z)$ is opposite to (x, z) . Since the pair (y, z) cancels as it appears once with positive sign and once

with negative sign we obtain the formula.



The fourth differential is given by precomposition with

$$\begin{aligned} \partial_4(x_1, x_2, x_3, x_4) &= (x_1, x_3, x_4) - (x_1 \triangleright x_2, x_3, x_4) - (x_1, x_2, x_4) \\ &\quad + (x_1 \triangleright x_3, x_2 \triangleright x_3, x_4) + (x_1, x_2, x_3) - (x_1 \triangleright x_4, x_2 \triangleright x_4, x_3 \triangleright x_4) \end{aligned}$$

The cochain complex for quandle cohomology is

$$\dots \xleftarrow{\partial_{n+1}} C^n \xleftarrow{d^n} \dots \xleftarrow{d^3} C^2 \xleftarrow{d^2} C^1 \xleftarrow{d^1=0} C^0,$$

where C_n can be thought of as the A -module with basis elements (x_1, x_2, \dots, x_n) , with $x_i \in X$, $x_i \neq x_{i+1}$, C^n is the space of linear maps from C_n to \mathbb{Z} , and the differential $d^n : C^{n-1} \rightarrow C^n$ is precomposition with the boundary map ∂_n given by

$$\begin{aligned} \partial_n(x_1, x_2, \dots, x_n) &= \sum_{i=2}^n (-1)^i [(x_1, \dots, x_{i-1}, x_{i+1}, \dots, x_n) \\ &\quad - (x_1 \triangleright x_i, \dots, x_{i-1} \triangleright x_i, x_{i+1}, \dots, x_n)] \end{aligned}$$

for $n \geq 2$ and $\partial_n = 0$ for $n \leq 1$. In particular, for an element of C_n thought of as a column vector, there is a corresponding function $f \in C^n$ obtained by transposing the column vector to get a row vector. The matrix of the differential d^{n-1} is then given by the transpose of the matrix of ∂_n .

More formally, we have

Definition 30. Let (X, \triangleright) be a quandle and A be an abelian group. A function $f : X \rightarrow A$ that satisfies the condition

$$f(x) - f(x \triangleright y) = 0$$

for all x, y in X is a *quandle 1-cocycle*.

Example 127. If the quandle X is trivial ($x \triangleright y = x, \forall x, y \in X$), then from definition 30 we see that any function f from X to A is a 1-cocycle.

Definition 31. A function $\phi : X \times X \rightarrow A$ such that for all x, y and z in X , the conditions

$$\begin{aligned} \phi(x, x) &= 0 \quad \text{and} \\ \phi(x, y) + \phi(x \triangleright y, z) &= \phi(x, z) + \phi(x \triangleright z, y \triangleright z) \end{aligned}$$

are satisfied is a *2-cocycle* of the quandle X with coefficients in A .

Example 128. It is straightforward to see that the function

$$\phi(x, y) = g(x) - g(x \triangleright y)$$

for any function g from X to A , satisfies the conditions of Definition 31. This function ϕ is called a *trivial 2-cocycle* (or *coboundary*).

Example 129. Consider the dihedral quandle R_3 . We will show in this example that every 2-cocycle $\Phi : X \times X \rightarrow \mathbb{Z}$ with coefficients in \mathbb{Z} is a coboundary. First we write

$$\Phi = \sum_{x, y \in R_3} \lambda_{(x, y)} \chi_{(x, y)}.$$

By substituting this expression of Φ in the equation (31), we obtain $\lambda_{(x, x)} = 0$ for all $x \in R_3$, and

$$\lambda_{(x, y)} + \lambda_{(x \triangleright y, z)} = \lambda_{(x, z)} + \lambda_{(x \triangleright z, y \triangleright z)}.$$

Now we write $R_3 = \{0, 1, 2\}$ with $x \triangleright y = 2y - x \pmod 3$, and substitute the values 0, 1, 2 for all possibilities of the variables x, y, z . We obtain after simplification the following equations

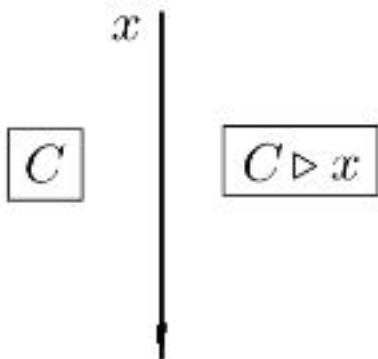
$$\begin{aligned} \lambda_{(0,0)} = \lambda_{(1,1)} = \lambda_{(2,2)} &= 0, \\ \lambda_{(0,1)} + \lambda_{(2,1)} &= 0, \\ \lambda_{(1,0)} + \lambda_{(2,0)} &= 0, \\ \lambda_{(0,2)} + \lambda_{(1,2)} &= 0, \\ \lambda_{(0,2)} + \lambda_{(2,1)} - \lambda_{(2,0)} &= 0. \end{aligned}$$

Again by substitution we can write the function f in the following form,

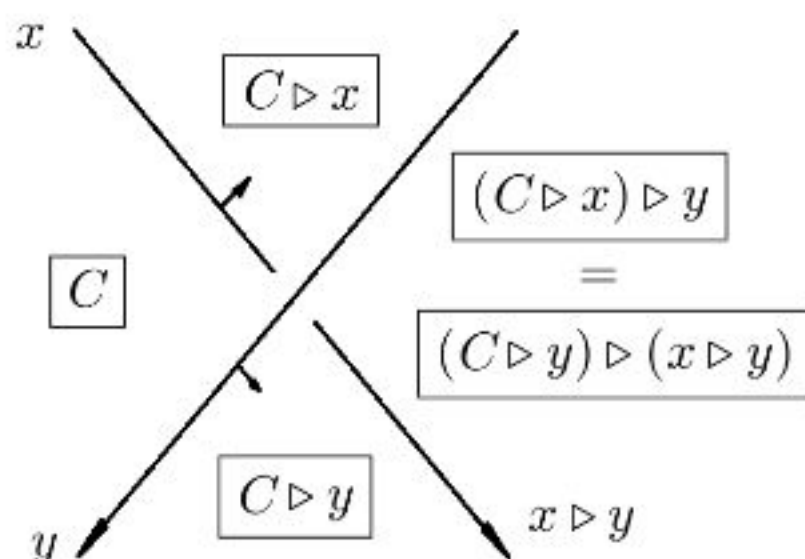
$$\begin{aligned} f &= \lambda_{(0,1)}[\chi_{(0,1)} - \chi_{(2,1)} + \chi_{(0,2)} - \chi_{(1,2)}] \\ &\quad + \lambda_{(1,0)}[\chi_{(1,0)} - \chi_{(2,0)} + \chi_{(0,2)} - \chi_{(1,2)}], \\ &= \lambda_{(0,1)}\delta(\chi_0) + \lambda_{(1,0)}\delta(\chi_1), \end{aligned}$$

making it a coboundary. This proves that every 2-cocycle is a coboundary in the cohomology of R_3 .

We can also get Boltzmann weights using 3-cocycles by considering *region colorings*, where in addition to each arc in the knot diagram getting a quandle element, each region between crossings also has an element of X assigned according to the condition



Then self-distributivity in X implies that the region coloring is well defined.

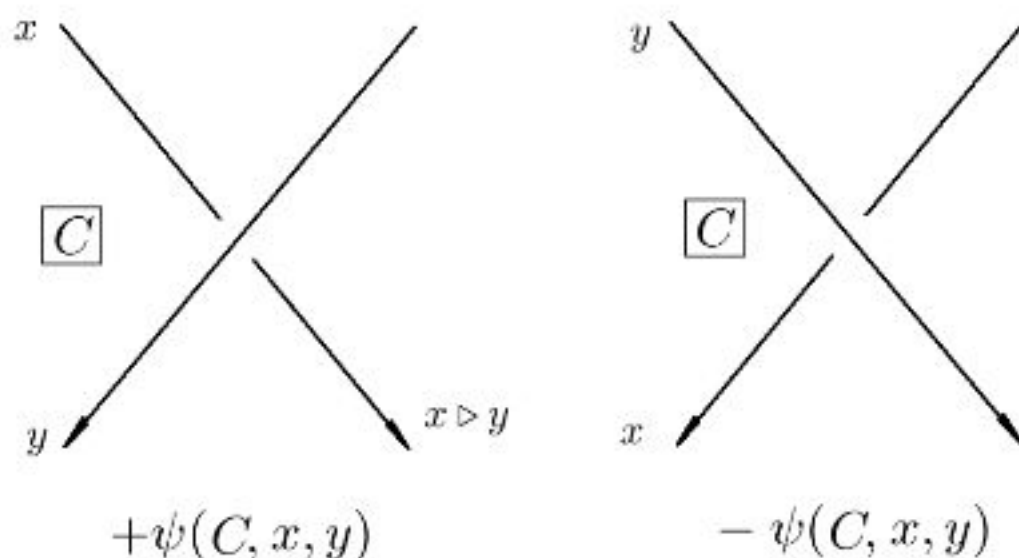


Definition 32. A function $\psi : X \times X \times X \rightarrow A$ such that for all $x, y, z, w \in X$ we have

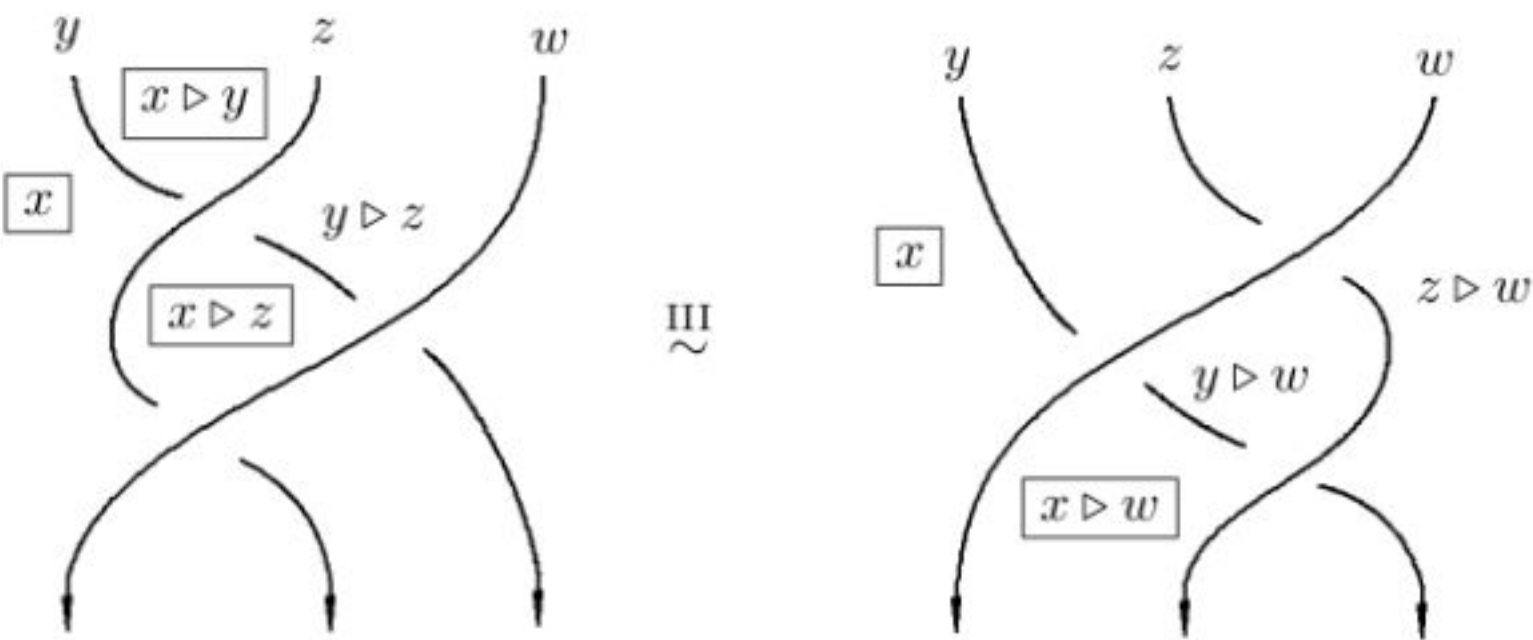
$$\begin{aligned} \psi(x, x, y) &= 0, \\ \psi(x, y, y) &= 0, \quad \text{and} \\ \psi(x, y, z) + \psi(x, z, w) &= \psi(x \triangleright y, z, w) + \psi(x, y, w) \\ &\quad + \psi(x \triangleright z, y \triangleright z, w) \quad + \psi(x \triangleright w, y \triangleright w, z \triangleright w) \end{aligned}$$

is a *quandle 3-cocycle*.

This definition can be understood in terms of the Reidemeister move III and region colorings. Given a region coloring of a knot, we can associate a row vector of three elements $\psi(C, x, y)$ to each crossing:



Then the three-cocycle condition comes from the third Reidemeister move:



Quandle Cocycle Enhancements. In [CJK⁺03], Scott Carter, Daniel Jelsovsky, Seiichi Kamada, Laurel Langford and Masahico Saito defined an enhancement of the quandle counting invariant using quandle cocycles which led to new results about knotted surfaces in 4-space; in particular, certain knotted spheres in 4-space are distinct from their reversed-orientation versions.

Let K be a knot, X a finite quandle, A an abelian group, and $\phi : X \times X \rightarrow A$ a 2-cocycle. Then the *quandle 2-cocycle enhancement*, also called the *State-Sum invariant*, of the knot K is the sum over all quandle colorings of K by X of expressions of the form $u^{\sum \pm \phi(x,y)}$ where the Boltzmann weight $\sum \pm \phi(x,y)$ is the sum of all the crossing weights,

$$\Phi_X^\phi(K) = \sum_{f \in \text{Hom}(\mathcal{Q}(K), X)} u^{\sum \pm \phi(x,y)}.$$

In the literature, it is common to write the abelian group A multiplicatively, e.g. writing u^n instead of nu for $n \in \mathbb{Z}$; this amounts to skipping the multiset step and going directly to the polynomial version of the invariant. With this style of notation, we have

$$\Phi_X^\phi(K) = \sum_{\mathcal{C}} \prod_{\tau} \phi(x,y)^{\epsilon(\tau)}$$

where the product is taken over all crossings of the given diagram, the sum is over all possible colorings and $\epsilon(\tau)$ is the sign of the crossing τ .

Similarly, if ψ is a 3-cocycle, then we have 3-cocycle enhancement

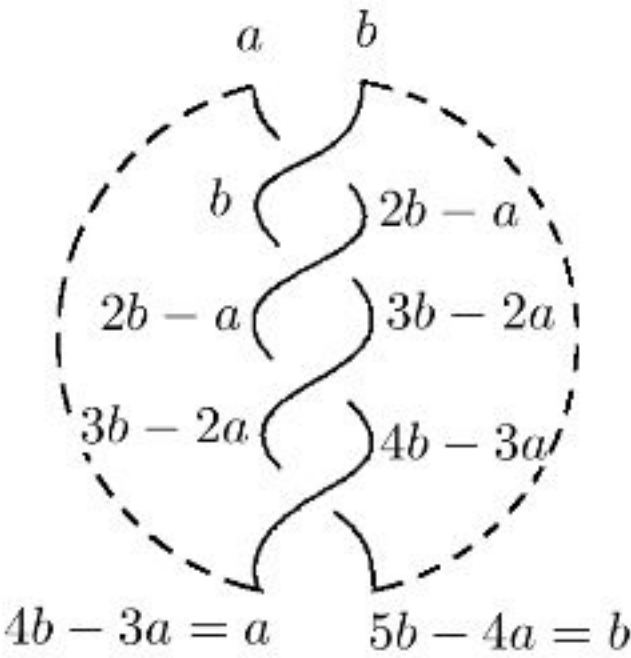
$$\Phi_X^{\psi,M}(K) = \sum_{f \in RC(K,X)} u^{\sum \pm \phi(x,y,z)}$$

or written multiplicatively,

$$\Phi_X^{\psi}(K) = \sum_{f \in RC(K,X)} \prod_{\tau} \phi(x,y,z)^{\epsilon(\tau)}$$

where $RC(L,X)$ is the set of region colorings of K by X . It was shown in [CJK⁺03] that these are knot invariants. We now consider some examples.

Example 130 ([CJK⁺03]). In this example we compute the quandle cocycle invariant of the torus link $T(4,2)$. As can be seen from the following figure the torus link $T(4,2)$ colors by the dihedral quandle $R_4 = \{0, 1, 2, 3\}$ where $i \triangleright j = 2j - i \pmod 4$. In fact, there are 16 possible colorings of $T(4,2)$ by R_4 since each pair (a,b) of elements of R_4 determines a coloring as can be seen below.



Let A be the group of integers denoted multiplicatively as

$$\mathbb{Z} = \{\dots, u^{-2}, u^{-1}, u^0, u, u^2, \dots\}.$$

Let ψ be the 2-cocycle of R_4 with coefficient in the integers \mathbb{Z} given by

$$\psi(a,b) = \begin{cases} u, & \text{if } (a,b) = (0,1) \text{ or } (a,b) = (0,3), \\ 1, & \text{otherwise.} \end{cases}$$

To compute the quandle cocycle invariant, we need to compute the contribution to it from all the 16 colorings, and then add all those

results. From the figure each pair (a, b) in $R_4 \times R_4$ contributes

$$\psi(a, b)\psi(b, 2b - a)\psi(2b - a, 3b - 2a)\psi(3b - 2a, a).$$

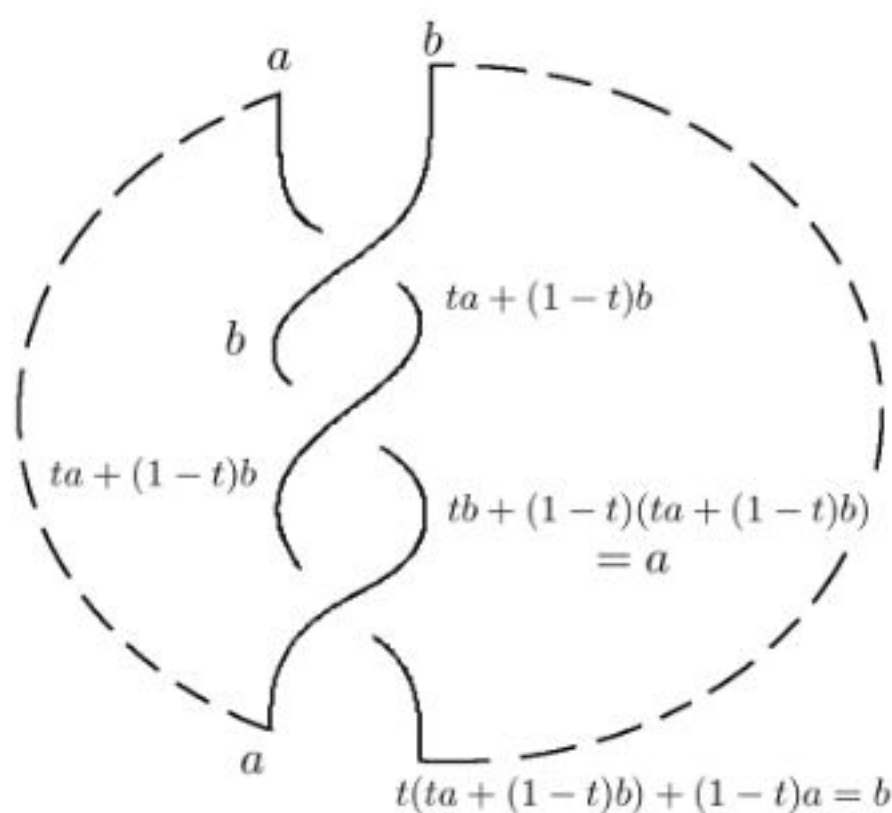
This implies that the eight pairs (a, b) with $a + b$ being *odd*, each contribute u . All other pairs each contribute 1 so that the cocycle invariant for this link is

$$\Phi_x^\phi(T(4, 2)) = 8 + 8u.$$

Example 131 ([CJK⁺03]). Let X be the Alexander quandle $X = \{0, 1, t, 1+t\}$ of polynomials in t with \mathbb{Z}_2 coefficients in which whenever we have t^2 we replace it by $t+1$, usually denoted by $\Lambda_2/(t^2+t+1)$. In this example we compute the colorings of the trefoil knot 3_1 , the knot 8_5 and the torus knot $T(5, 2)$ by this quandle X . We then compute the quandle cocycle invariant for each of them using the following two cocycle with coefficients in the two elements group $A = \mathbb{Z}_2 = \{1, u\}$ where $u^2 = 1$ (again we are using multiplicative notation for A). Let ψ be the 2-cocycle of X with coefficients in the integers $A = \mathbb{Z}_2$ given by

$$\psi(a, b) = \begin{cases} u, & \text{if } (a, b) \in \{(0, 1), (1, 0), (1+t, 0), (0, 1+t), \\ & (1, 1+t), (1+t, 1)\} \\ 1, & \text{otherwise.} \end{cases}$$

Consider the trefoil knot 3_1 . To compute the quandle cocycle enhancement, we need to compute the contribution to it from all the 16 colorings, and then add all those results. From the figure

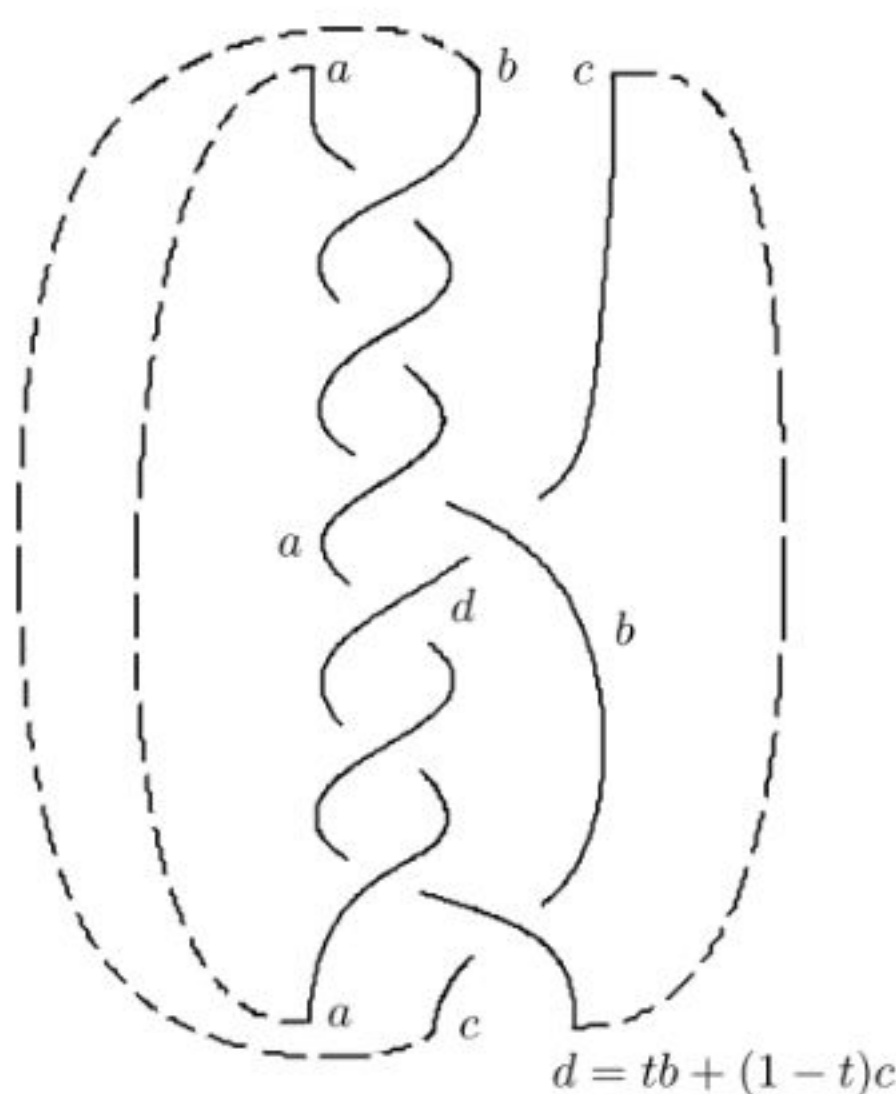


each pair (a, b) in X contributes

$$\psi(a, b)\psi(b, ta + (1 - t)b)\psi(ta + (1 - t)b, a).$$

Each pair of the form (a, a) contributes 1 to the invariant while all other pairs each contribute u making the State-Sum equal $4 + 12u$. A similar computation gives the same result for all of the following knots: $4_1, 7_2, 7_3, 8_1, 8_4, 8_{11}$ and 8_{13} .

Now consider the knot 8_5 . We use the braid form of this knot, that is 8_5 is the closure of the three strand braid $\sigma_1^3\sigma_2^{-1}\sigma_1^3\sigma_2^{-1}$. From the following figure

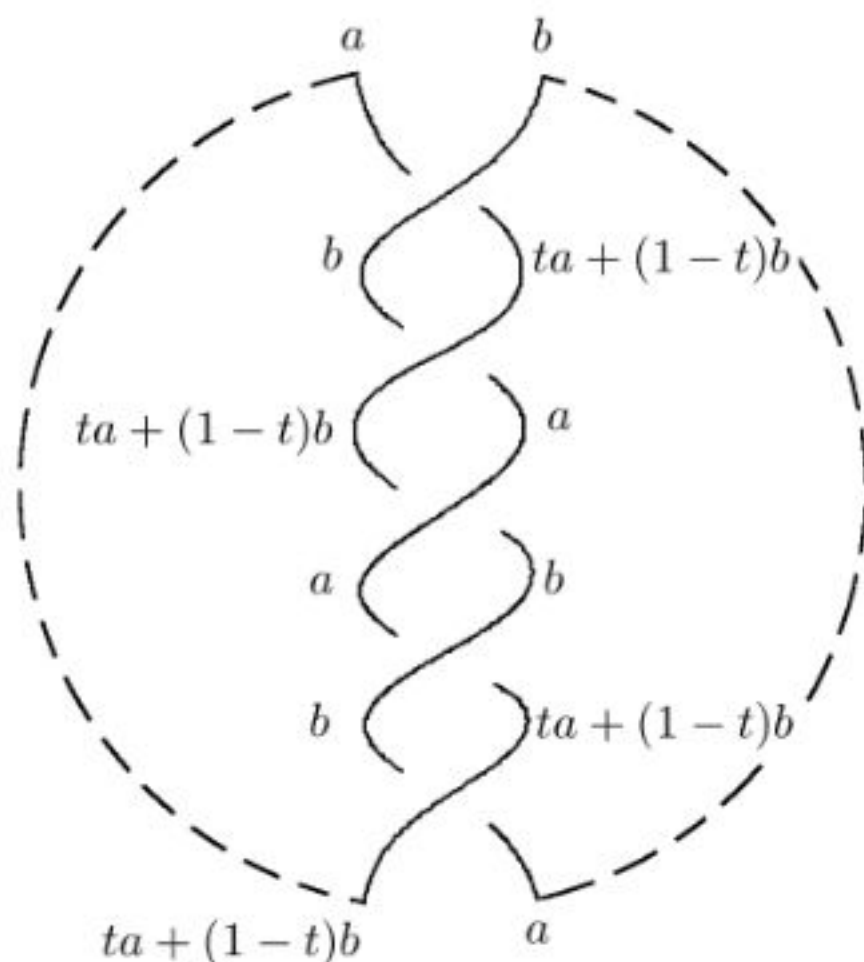


we see that a triple input (a, b, c) on the top colors the knot if and only if $b = c$. Thus the coloring triples are all of the form (a, b, b) . Those with $a = b$, each contribute 1 to the State-Sum while those with $a \neq b$ each contribute

$$[\psi(a, b)\psi(b, ta + (1 - t)b)\psi(ta + (1 - t)b, a)\psi(b, b)]^2 = u^2 = 1.$$

Thus the State-Sum invariant of the knot 8_5 is equal to 16. A similar computation gives the same result for all of the following knots: $8_{10}, 8_{15}, 8_{19}, 8_{20}$, and 8_{21} .

Finally, consider the torus knot $T(5, 2)$. From the following figure we see that the only colorings of the torus knot $T(5, 2)$ by the quandle X are the trivial colorings given by pairs of the form (a, a) .



Each coloring contributes 1 since $\psi(a, a) = 1$. Thus the State-Sum invariant of the torus knot $T(5, 2)$ is equal to 4.

Exercises. 1. Find all region colorings of the trefoil knot 3_1 by the dihedral quandle R_3 .

2. Let X be the Alexander quandle \mathbb{Z}_4 with $t = 3$. Compute $\partial_3(1, 2, 1)$.

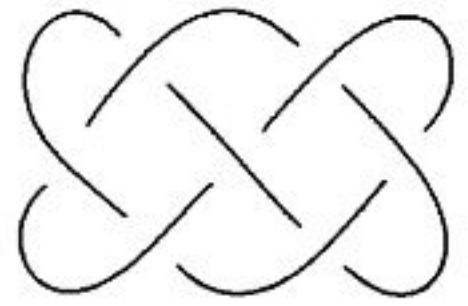
3. Find the matrix for the differential d^2 for the trivial quandle on three elements.

4. Prove that $d^3 d^2 = 0$ for the trivial quandle T_2 .

5. Let X be the Alexander quandle \mathbb{Z}_4 with $t = 3$. Show that the function $\phi(x, y) = (x - y)^4$ is a 2-cocycle and that $\psi(x, y, z) = (x - y)^4(y - z)^4$ is a 3-cocycle. These are examples of *Mochizuki cocycles* [Moc03, Moc05].

6. Compute the State-Sum invariants Φ_x^ϕ and Φ_x^ψ for the Hopf link with the respect to X, ϕ and ψ from problem 5.

Chapter 7

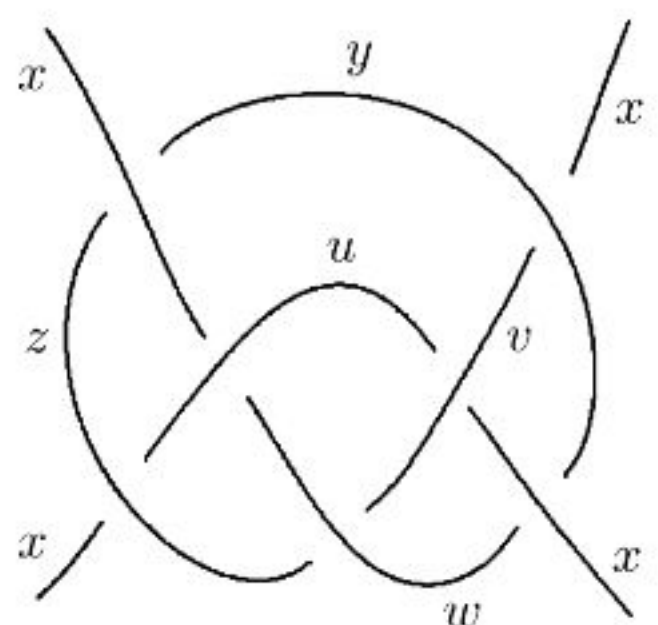
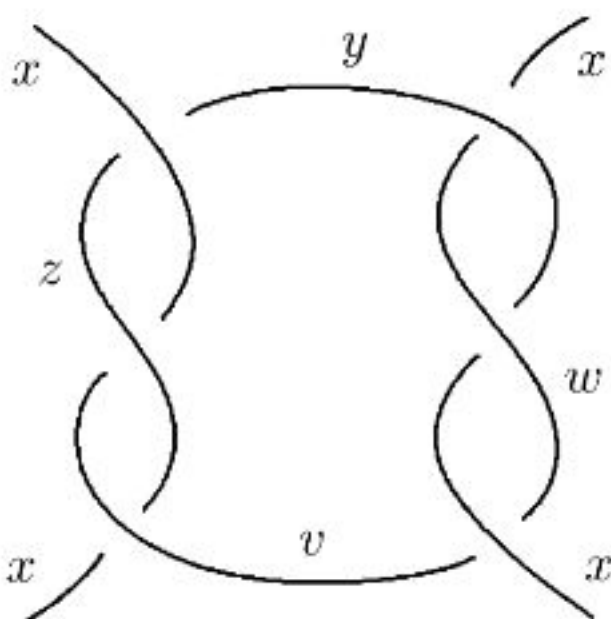


Generalized Knots and Links

1. Colorings of Tangles and Embeddings

In this section we will consider $(2, 2)$ -tangles, define their colorings and their quandle cocycle invariants. We then give a criterion for when a $(2, 2)$ -tangle embeds in a knot or link. As a consequence we use the cocycle invariant in terms of multisets to prove that some tangles cannot be embedded in certain knots in the table in Chapter 1.

A $(2, 2)$ -*tangle* is a portion of a knot or link with two fixed inputs and two fixed outputs. See the following figures called, respectively, tangle 6_2 and tangle 7_{17} . The list of all prime $(2, 2)$ -tangles with up to seven crossings can be found in [KSS03]



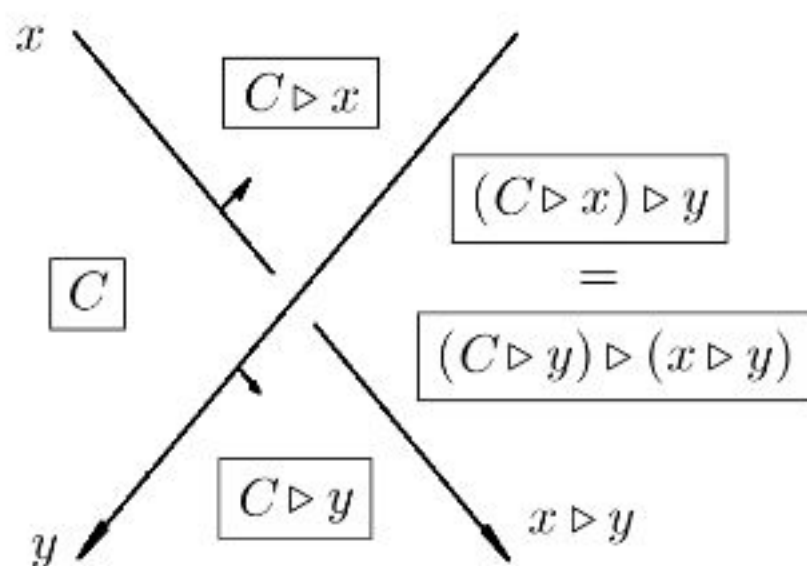
For this section's applications, we will need the multiset versions of the quandle 2-cocycle and 3-cocycle enhancements from the last chapter. Recall that a multiset is a set in which we allow repeated entries. For example, $\{a, a, a, b, b\}$ represents a multiset (S, m) where $S = \{a, b\}$, $m(a) = 3$ and $m(b) = 2$. This is also denoted by $\{3 \times a, 2 \times b\}$.

We briefly recall the definition of the cocycle invariant. Let f be a coloring of a knot diagram K by a finite quandle X . The *Boltzmann weight* $B(f, \tau) = B_\phi(f, \tau)$ at a crossing τ of K is then defined by $B(f, \tau) = \epsilon(\tau)\phi(x_\tau, y_\tau)$, where the pair (x_τ, y_τ) consists of the source colors at τ and $\epsilon(\tau)$ is the sign (± 1) of the crossing τ as before. Then the 2-cocycle invariant $\Phi_X^\phi(K)$ in multiset form can be expressed by

$$\Phi_X^{\phi, M}(K) = \left\{ \sum_{\tau} B(f, \tau) \mid f \in \text{Hom}(\mathcal{Q}(K), X) \right\}.$$

where $\text{Hom}(\mathcal{Q}(K), X)$ is the set of X -colorings of K .

Let f be a coloring of arcs and regions of a given diagram K . Specifically, for a coloring f , there is a coloring of regions that extends f as depicted.



Suppose that two regions R_1 and R_2 are separated by an arc colored by y , and the *normal vector* of the arc, obtained by rotating the direction vector counterclockwise 90 degrees, points from R_1 to R_2 . If R_1 is colored by C , then R_2 receives the color $C \triangleright y$. Let (C, x, y) (called the *ordered triple of colors* at a crossing τ) be the colors near a crossing τ such that C is the color of the region (called the source region) from which both normal vectors of the over- and under-arc point, x is the color of the under-arc (called the source under-arc)

from which the normal vector of the over-arc points, and y is the color of the over-arc as depicted above.

Let $\psi : X \times X \times X \rightarrow A$ be a quandle 3-cocycle, which we recall can be regarded as a function satisfying

$$\begin{aligned} &\psi(x, z, w) - \psi(x, y, w) + \psi(x, y, z) - \psi(x \triangleright y, z, w) \\ &+ \psi(x \triangleright z, y \triangleright z, w) - \psi(x \triangleright w, y \triangleright w, z \triangleright w) = 0, \quad \forall x, y, z, w \in X, \end{aligned}$$

and $\psi(x, x, y) = 0 = \psi(x, y, y), \forall x, y \in X$. We define a new Boltzmann weight at the crossings in a region colored diagram by

$$B(f, \tau) = \epsilon(\tau) \phi(C, x, y).$$

The 3-cocycle invariant is defined in a similar way to the 2-cocycle invariant as the multiset

$$\Phi_X^{\psi, M}(K) = \left\{ \sum_{\tau} B(f, \tau) \mid f \in RCT(T, X) \right\}$$

where $RC(T, X)$ denotes the set of colorings of the regions of K by X .

As we have seen, if the quandle X is finite, the invariant as a multiset can also be expressed as a polynomial: if a given multiset of group elements is $\{m_1 \times g_1, \dots, m_\ell \times g_\ell\}$, then we use the polynomial notation $m_1 u^{g_1} + \dots + m_\ell u^{g_\ell}$ where u is a formal symbol. For example, the multiset value of the invariant for a trefoil with the Alexander quandle $X = \Lambda_2/(t^2 + t + 1)$ with the same coefficient group $A = X$ and a certain 2-cocycle is $\{4 \times (0), 12 \times (t + 1)\}$, and is denoted by $4 + 12u^{(t+1)}$, where we use the convention $u^0 = 1$ and exponential rules apply.

For computing the invariants, one needs an explicit formula for cocycles. Polynomial cocycles were used first in [Moc03], and investigated closely including higher dimensional cocycles in [AS09].

We will use quandle cocycle invariants as obstructions to embedding tangles in knots. We must first define cocycle invariants for tangles.

Definition 33. Let T be a tangle and X be a quandle. A *boundary-monochromatic coloring* is a coloring of the arcs in a diagram of T to X satisfying the same quandle coloring condition as for knot diagrams

at each crossing, such that the (four) boundary points of the tangle diagram receive the same element of X . Region colorings of a tangle diagram are defined in a similar manner to the knot case.

Note that a tangle has a fundamental quandle analogous to that of a knot, with generators for each arc and quandle relations at the crossings. As with knots, the set of quandle colorings of T by a finite quandle X can be understood as the set of quandle homomorphisms $\text{Hom}(\mathcal{Q}(T), X)$ from the fundamental quandle of T to X , and the set of boundary-monochromatic colorings is a subset of this.

Denote by $\text{Col}_x(T)$ and $\text{Col}_X(T)$ the set of boundary-monochromatic colorings of T with the boundary color $x \in X$ and the set of all boundary-monochromatic colorings, respectively. Let

$$\Phi(T, x) = \left\{ \sum_{\tau} \epsilon(\tau) B(f, \tau) \mid f \in \text{Col}_x(T) \right\}.$$

Then the cocycle invariant for a tangle T is equal to

$$\Phi_X^\phi(T) = \bigcup_{x \in X} \Phi(T, x).$$

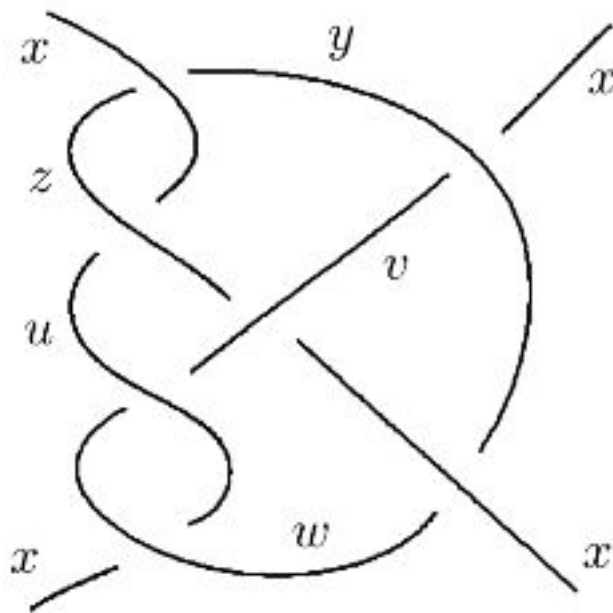
The invariants for region colorings are defined in a similar manner, by taking the sum over all colorings of regions as well as colorings of diagrams.

It can be proved in a way similar to the case of a knot that the number of colorings $|\text{Col}_X(T)|$ does not depend on the choice of a diagram of T . If a diagram D_1 of T has a coloring \mathcal{C}_1 , and a diagram D_2 is obtained from D_1 by a Reidemeister move, then there is a unique coloring \mathcal{C}_2 of D_2 induced from \mathcal{C}_1 , such that the colors stay the same except where the move is performed. Given two diagrams D_1 and D_2 of a tangle T , there is one-to-one correspondence between the set of colorings of D_1 and the set of colorings of D_2 and the cocycle invariant is well defined.

The following example collects tangles in the tangle table [**KSS03**] that have nontrivial boundary monochromatic colorings by some Alexander quandles. Specifically, variables x_i , $i = 1, 2, \dots$, are assigned to the arcs of tangle diagrams. Coloring conditions of the form $x_k = tx_i + (1 - t)x_j$ are imposed at crossings, giving rise to a

system of linear equations with coefficients in Λ that is solved to find which Alexander quandles give nontrivial colorings of the tangles.

Example 132. The following tangle 7_{13} is colored nontrivially by the dihedral quandle R_5 .

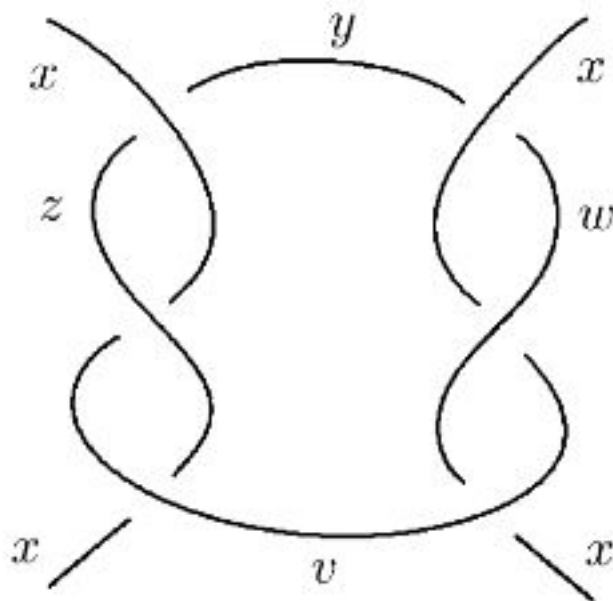


To see this, one writes the equations at all seven crossings,

$$\begin{aligned} 2x - y = z, \quad 2z - x = u, \quad 2u - v = w, \quad 2w - u = x, \\ 2v - z = x, \quad 2y - v = x, \quad 2x - y = w. \end{aligned}$$

A straightforward substitution gives, for example, the equation $5(u - y) = 0$ giving a nontrivial coloring of tangle 7_{13} by R_5 .

Example 133. The following tangle 6_3 with orientations NW in, SW out is colored nontrivially by the Alexander quandle $\Lambda_p/(t^2 - t + 1)$, where p is prime. Here NW and SW stand for northwest and southwest.



We explain briefly how the computation of colorings works for the tangle 6_3 with orientations NW in, SW in. We color it by an

Alexander quandle and let x be the color of the boundary arcs. Let y, z, v and w be the colors of the arcs as depicted in the figure above. From the crossings adjacent to the NW, SW, SE, NE endpoints, respectively, we obtain the relations:

$$\begin{cases} tz + (1 - t)x = y, \\ tx + (1 - t)v = z, \\ tx + (1 - t)v = w, \\ tw + (1 - t)x = y. \end{cases}$$

From the remaining two crossings, we have the following equations:

$$\begin{cases} tv + (1 - t)z = x, \\ tv + (1 - t)w = x. \end{cases}$$

These equations imply that $z = w$. A substitution gives the equation $(t^2 - t + 1)(v - z) = 0$, and it follows that the quandle $\Lambda_p/(t^2 - t + 1)$ colors the tangle $T(6_3)$ nontrivially.

Now we give the following theorem which gives us the conditions for a tangle to embed in a link.

Theorem 13. *Let T be a tangle and X a quandle. Suppose T embeds in a link L . Then we have the inclusion $\Phi_\phi(T) \subset_m \Phi_\phi(L)$.*

This theorem allows us to tell when a tangle does not embed in some knots and this is when the tangle cocycle invariant is not a sub-multiset of the cocycle invariant of the knot.

Example 134. The tangle $T(6_2)$ with the orientation of the NW arc inward and the SW arc outward does not embed in the knots in the table up to 8 crossings except, possibly, for 8_{18} .

Example 135. The knots in the table up to 8 crossings in which the tangle $T(6_3)$ embeds are exactly 8_{10} and 8_{20} . Here, the orientation of the tangle is such that the NW endpoint is oriented inward and the SW endpoint is oriented outward.

Exercises. 1. For the tangles $6_2, 6_3, 7_{13}, 7_{17}$ and 7_{18} (see [KSS03]), find the dihedral quandles which color them nontrivially.

2. Find the Alexander quandles which color nontrivially the following tangles $6_2, 6_3, 7_{13}, 7_{17}$ and 7_{18} with orientation NW inward and SW inward.

3. Redo the previous exercise by considering different orientations for NW and SW.

4. Given that

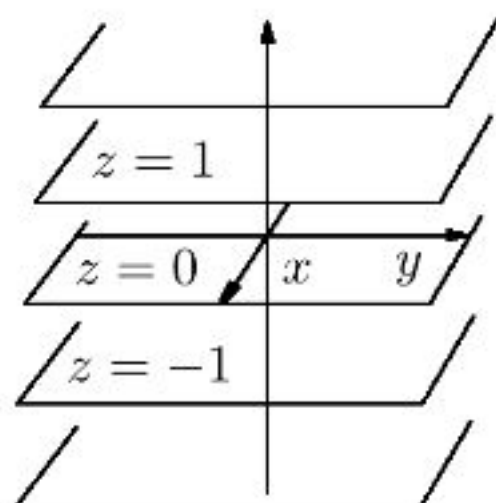
- the knot 7_4 and the tangle 6_2 are both colorable nontrivially by the Alexander quandle $X = \Lambda_3/(t^2 - t + 1)$,
- X has 3-cocycle $\psi(x, y, z) = (x - y)(y - z)^3$, and
- the quandle 3-cocycle invariant of the knot 7_4 with respect to ψ is $243 + 486u^{(t+1)}$.

Compute the quandle 3-cocycle invariant of the tangle 6_2 and deduce the fact that the tangle 6_3 with orientation NW inward and SW outward does not embed in the knot 7_4 .

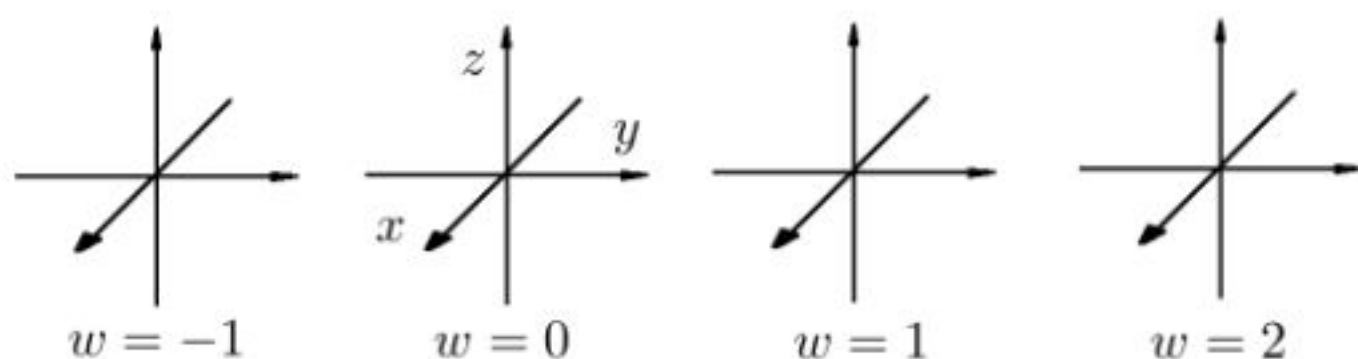
2. Surface Knots

We have seen how simple closed curves can be knotted in three-dimensional space \mathbb{R}^3 and its compact version S^3 . What about higher dimensions? It turns out that any two simple closed curves in \mathbb{R}^4 are ambient isotopic. We can see why this is so by thinking about knot diagrams.

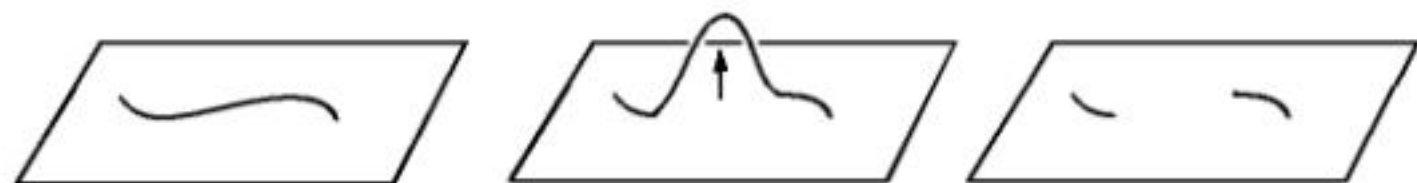
First, recall that we can conceptualize \mathbb{R}^3 as a stack of copies of \mathbb{R}^2 (called *planes*) indexed by a third variable:



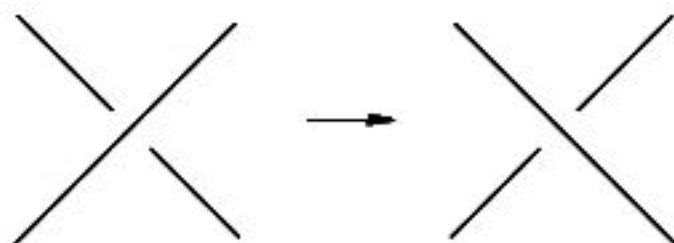
Similarly, we can think of \mathbb{R}^4 as a stack of copies of \mathbb{R}^3 (called *hyperplanes*) indexed by a fourth variable:



If we have a curve in a plane in \mathbb{R}^3 and we move part of it upward in the z direction, the part which moves up disappears from the original plane, even though it is still one connected continuous curve in \mathbb{R}^3 :



Now, if we have a knot in \mathbb{R}^4 with a crossing, we can move the overcrossing strand upward in the w direction, resulting the move portion disappearing from our original hyperplane. It hasn't vanished, though; the knot is still one connected continuous curve in \mathbb{R}^4 . We can then move the undercrossing strand to a higher z -position in the original hyperplane, and finally move the original overstrand back into its original position in the original hyperplane. The result is a *crossing change*.



It turns out that crossing changes are unknotting moves – if you are allowed to change crossings, you can unknot any knot. Thus, simple closed curves in \mathbb{R}^4 are all unknotted.

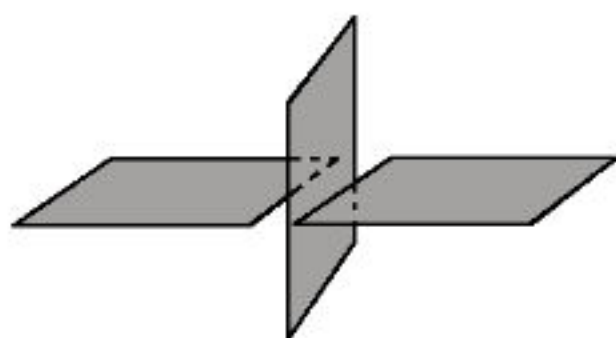
The problem is not that there are no knots in four dimensions; the problem is that for one-dimensional curves in four-dimensional space, the *codimension*, i.e., the difference between the dimension of

the ambient space in which the knot lives and the dimension of the knot itself, is too low. It turns out that for nontrivial knotting, we want the codimension to be 2. Thus, we have knotted 1-dimensional curves in a 3-dimensional ambient space, and we can have knotted 2-dimensional surfaces in a 4-dimensional ambient space.

Surface knot theory is a more complex topic than the theory of knotted curves for several reasons above and beyond our obvious inability to physically manipulate knotted surfaces. Unlike the \mathbb{R}^3 case in which there is only one kind of object to knot (namely, simple closed curves), there are infinitely many distinct types of surfaces – spheres, tori, tori with two holes, Klein bottles and more.

How can we possibly understand knotted surfaces in \mathbb{R}^4 ? Even though knotted curves need three dimensions of ambient space, they almost fit in two dimensions – in fact, we can put every knot in the plane with just a little bit of the z direction for crossings. Similarly, surfaces need four dimensions for (tame) knotting, but we can fit a knotted surface almost entirely in a single hyperplane with just a bit of thickness in the w direction for crossings.

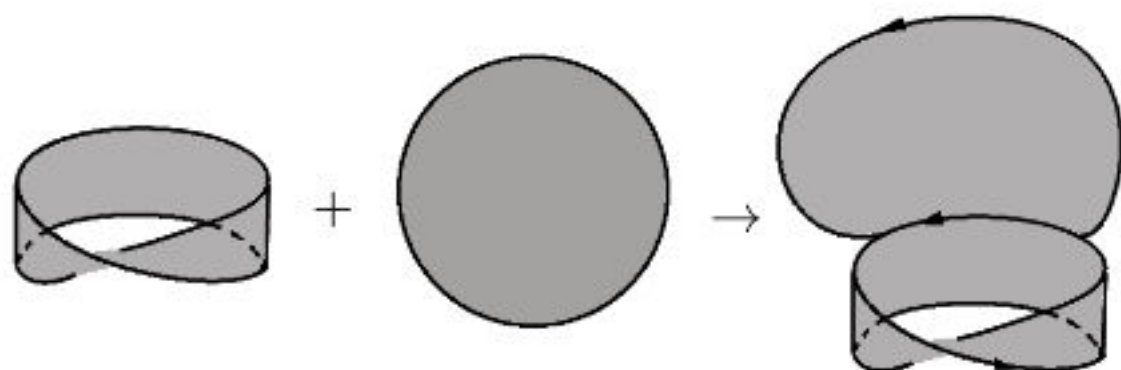
Where a knot diagram consists of arcs which meet at crossings with the undercrossing strand drawn broken to indicate crossing under in an invisible third dimension, a *knotted surface diagram* has sheets which meet along crossing curves, and we draw the undercrossing sheet broken to indicate crossing under in the invisible fourth dimension.



Indeed, there are surfaces which do not fit in three dimensions but require a fourth dimension, such as the *real projective plane* $\mathbb{R}P^2$ and the *Klein bottle*. The real projective plane can be defined formally as the quotient set of $\mathbb{R}^3 \setminus \vec{0}$ under the equivalence relation $\vec{x} \sim \alpha\vec{x}$ for $\alpha \neq 0$, that is, the set of lines through the origin in \mathbb{R}^3 . To identify the

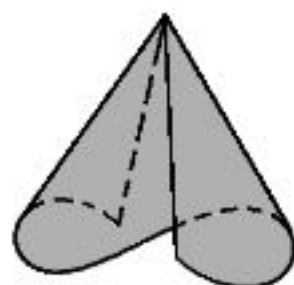
quotient space, we can restrict our attention to only the unit sphere, since every line through the origin goes through the sphere. In fact, each such line goes through the sphere in exactly two opposite points (called *antipodes*), so we can restrict to just the northern hemisphere – and we can now see why this quotient set is a surface.

However, along the equator we still have two points in each equivalence class. Thus, we can think of $\mathbb{R}P^2$ as the result of taking the northern hemisphere (which is topologically a disk) and gluing the points along the equator to their antipodes. This is hard to visualize, because at some point we experience a “brain breaking” sensation indicating that the thing we’re trying to visualize doesn’t fit in the kind of three-dimensional space we can visualize. We can see it perhaps a little better by imagining removing a disk from the interior of the large disk and just gluing along the boundary; then we get a *Möbius band*:

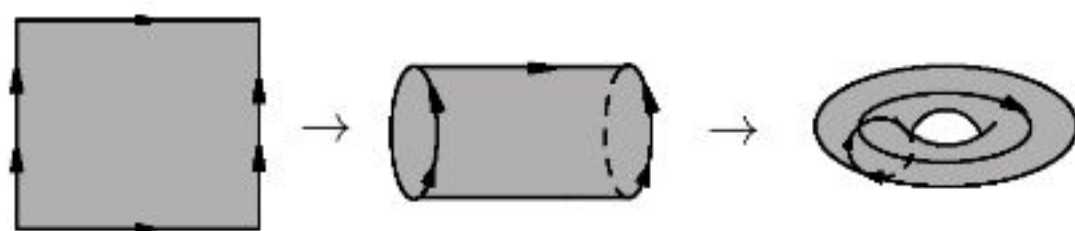


Thus, $\mathbb{R}P^2$ can be understood as the result of gluing a disk onto a Möbius band along their boundary circles, like two sides of a zipper, to form a seamless surface with no boundary.

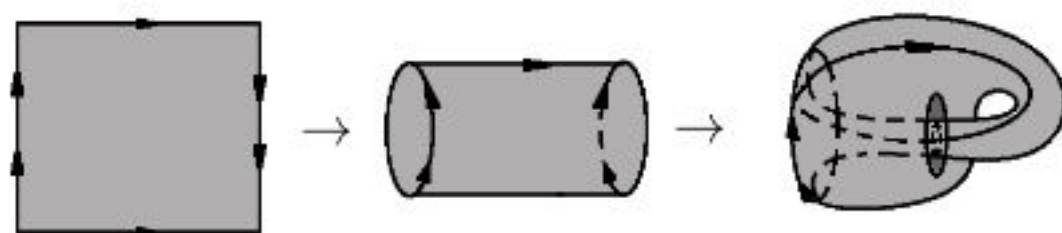
The real projective plane is sometimes called a *cross-cap*, because it can be understood as the result of “capping off” a knotted surface portion called a *cusp*:



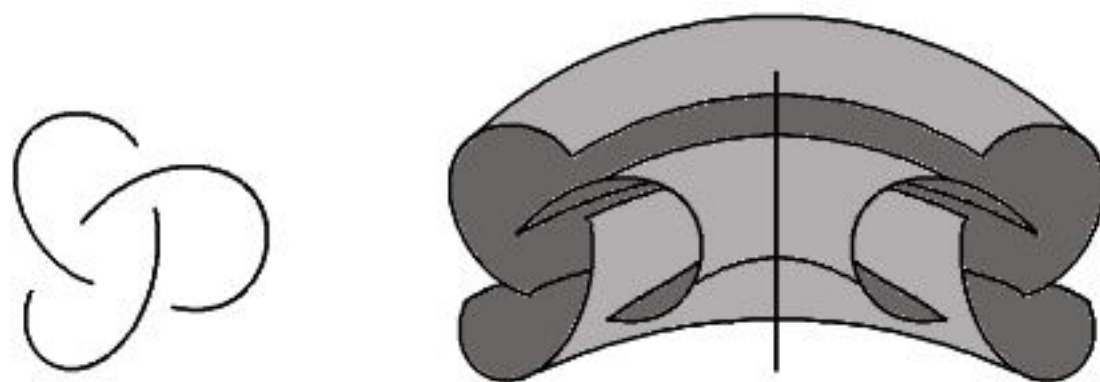
The Klein bottle can be understood as the result of removing discs from two projective planes and gluing the resulting Möbius bands together (or, just gluing two Möbius bands together along their boundaries). More often, we describe the Klein bottle as a quotient space of a rectangle: if we start with a rectangle and glue the top edge to the bottom edge and the left side to the right side like in the old video game *Asteroids*, the result is a torus:



On the other hand, if we repeat the procedure but reverse the direction of the circle at the last step, we get a Klein bottle.

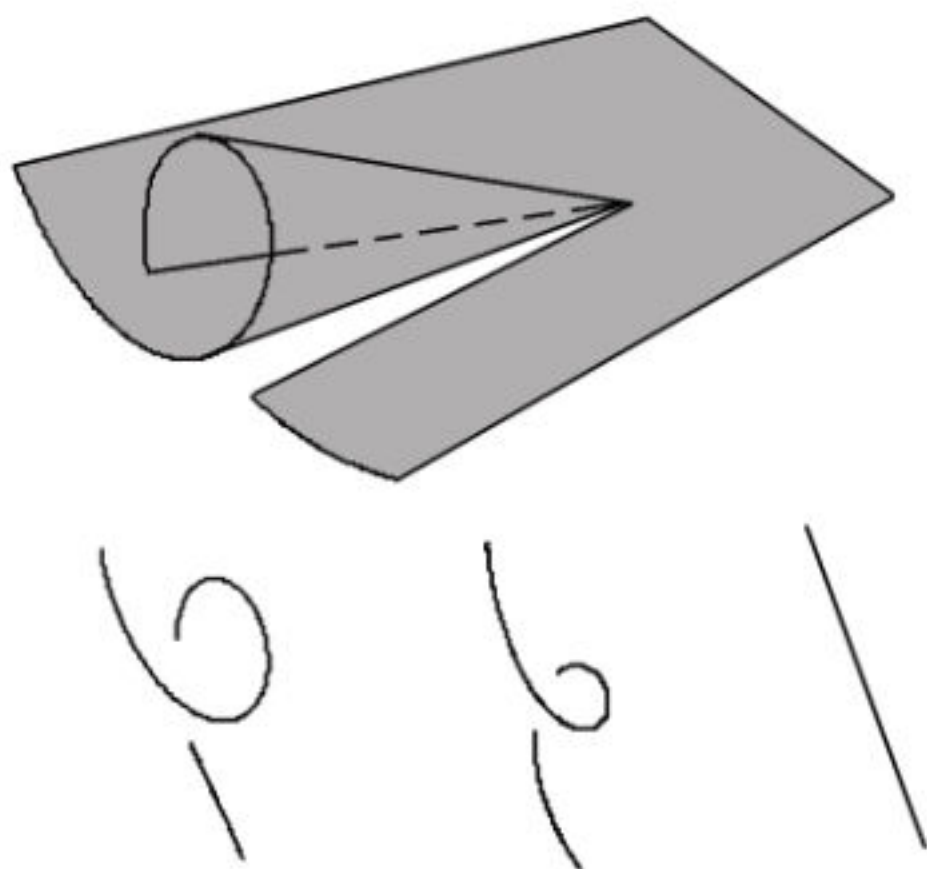


Knotted surfaces arise in many different ways. Given any knotted circle K in \mathbb{R}^3 , we can spin the knot K about an axis in \mathbb{R}^4 , letting K sweep out a knotted surface analogous to the surfaces of revolution we see in single-variable calculus:

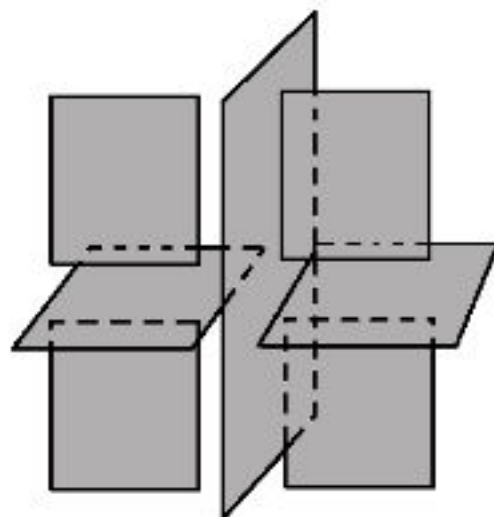


As a variation on spinning, we can *twist-spin* knots, where K rotates one or more times about another axis as it spins around the primary axis of revolution.

Every Reidemeister move sequence on knotted curves in \mathbb{R}^3 can be understood as a knotted surface, where we think of time as the w direction. In particular, slicing a knotted surface with a hyperplane yields a knotted curve; if we think of \mathbb{R}^4 as a stack of \mathbb{R}^3 s, then each slice is a frame of a movie. Indeed, we can represent knotted surfaces with *movie diagrams*, sequences of knot diagrams representing slices of a surface knot at different z values. For example, the Reidemeister I move corresponds to a cusp:



In addition to cusps and intersecting sheets, knotted surfaces can also contain *triple points*, points where three sheets all come together at a single point, with one sheet on top in the z direction, one in the middle and one on the bottom.

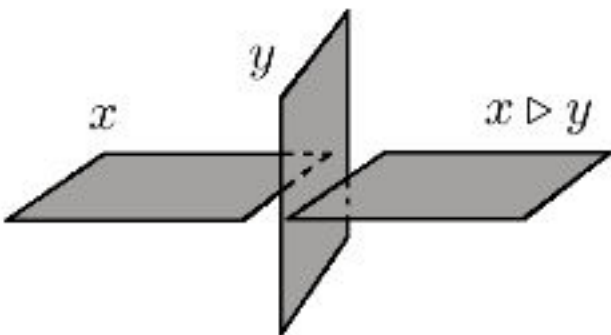


Just as there are Reidemeister moves for knotted curves in \mathbb{R}^3 , there are moves on knotted surface diagrams called *Roseman moves*

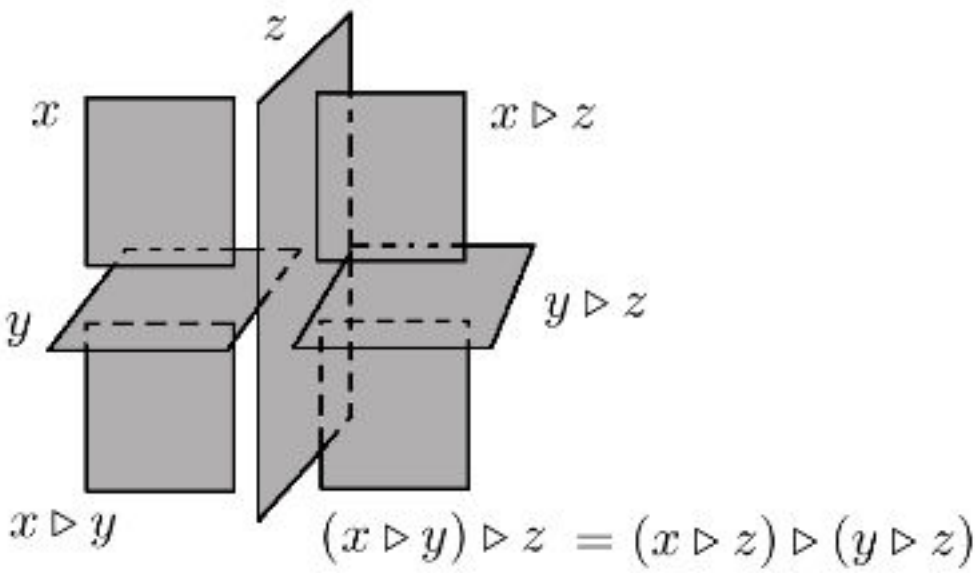
with the property that two knotted surface diagrams represent ambient isotopic surfaces in \mathbb{R}^4 if and only if the diagrams are related by Roseman moves.

Quandles and Surface Knots. As with knotted curves in \mathbb{R}^3 , we can use quandle-based invariants to distinguish knotted surfaces; since the Roseman moves can be expressed as Reidemeister moves on movie diagrams, quandle colorings are preserved by Roseman moves. We will look briefly at kei and quandle colorings; bikei, biquandle, rack and birack based invariants and their enhancements exist and are topics of ongoing research.

Suppose we have a finite kei X . We can color a knotted surface S by assigning an element of X to each sheet in a diagram of S . Then when a sheet labeled x crosses under a sheet labeled y , the result is a sheet labeled $x \triangleright y$.

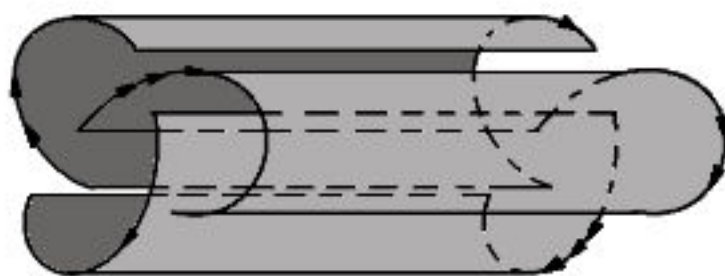


For the coloring to make sense at a triple point, we need the third kei axiom:



One can then verify that given a kei labeling of a knotted surface diagram before a move, there is a unique kei labeling of the diagram after the move, and hence counting invariants are defined for knotted surface diagrams.

Example 136. We can use the 3-element Takaski kei to distinguish the spun trefoil from the trefoil spun with a third of a twist:



This knotted surface diagram has one single sheet, so it can only have monochromatic colorings, where the original spun trefoil has the same set of nine colorings as the trefoil.

Exercises. 1. Draw a knotted surface diagram for the spun figure eight knot.

2. Let $S^3 = \{(x, y, z, t), x^2 + y^2 + z^2 + t^2 = 1\}$ be the 3-sphere in \mathbb{R}^4 . In order to get some understanding of this sphere, let's look at some cross-sections of it. Draw and describe series of the cross-sections for the following values of t : $t = -1, -\frac{3}{4}, -\frac{1}{2}, -\frac{1}{4}, 0, \frac{1}{4}, \frac{1}{2}, \frac{3}{4}$ and $t = 1$. How does the cross-section change as t goes from -1 to zero? How does the cross-section change as t goes from zero to 1 ?

3. Draw the knotted surface corresponding to the Reidemeister II move.

4. Draw the result of spinning the Hopf link with a $1/2$ twist.

5. Let K be the result of spinning a 5_1 knot with a $1/5$ twist. Use a quandle coloring argument to show that this knotted surface is distinct from the nontwisted spun 5_1 .

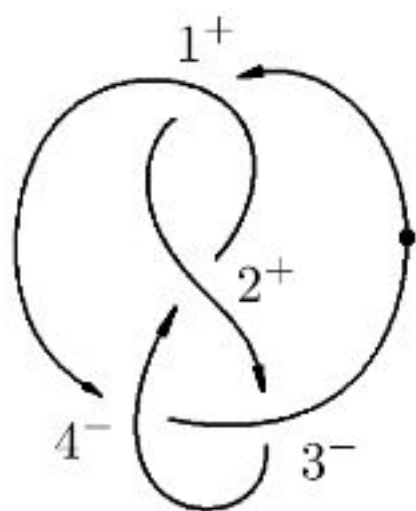
3. Virtual Knots

Knot theory is a very visual form of mathematics in that a lot of the information involved is in the form of pictures and visual diagrams. This is a great advantage for many of us since it makes knot theory easier to understand than some other less visualizable areas of mathematics, and indeed there are many computations we can do and theorems we can prove in knot theory entirely through pictures.

Suppose we want to write computer code to do computations of knot invariants so we don't have to check thousands of potential labelings for validity by hand. To do this, we need some more computer-friendly ways of representing knots in place of our usual pictures. One way to do this is with *Gauss codes*.

Suppose we have an oriented knot K . We start by selecting a basepoint on one of the semiarcs of K and numbering the crossings, noting for each crossing whether it is a positive or negative crossing.

To make the *signed Gauss code* for K , we start at the basepoint and travel around K following the given orientation; each time we go through a crossing, we want to write down the crossing number and sign together with whether we are passing over or under. For example, in the figure eight knot below, we have signed Gauss code $U1^+O2^+U3^-O4^-U2^+O1^+U4^-O3^-$.

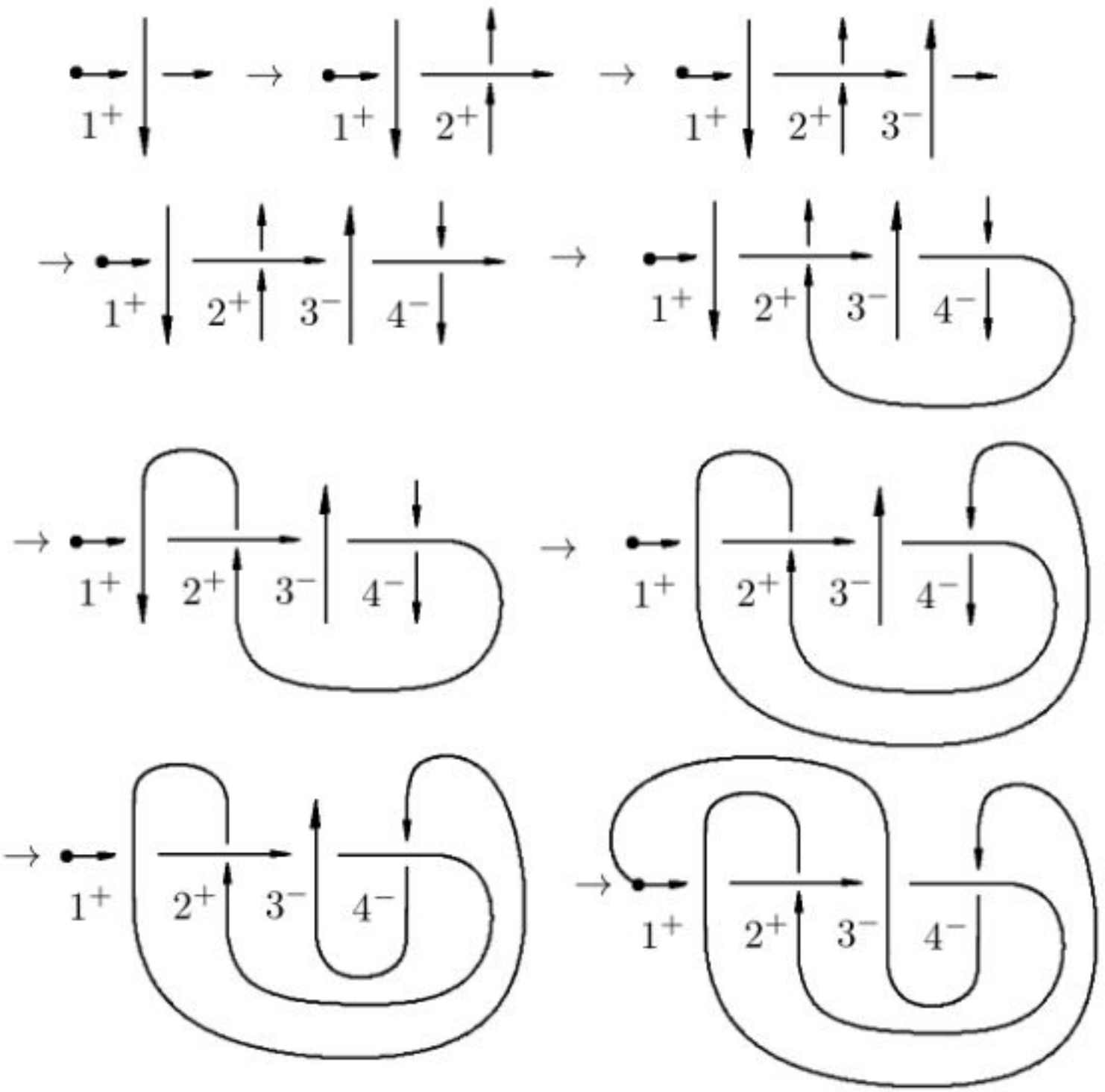


Choosing a different basepoint results in a cyclic permutation of the Gauss code, moving part of the code from the end to the beginning, e.g.,

$$\begin{array}{c}
 U1^+O2^+U3^-O4^-U2^+O1^+U4^-U3^- \\
 \updownarrow \\
 U3^-U1^+O2^+U3^-O4^-U2^+O1^+U4^- \\
 \updownarrow \\
 \vdots
 \end{array}$$

Gauss codes are useful since as strings of text rather than pictures, they are well-suited for use in computer code as well as in text-only forms of communication. Given a signed Gauss code, we can reconstruct the knot diagram up to isotopy on S^2 by drawing

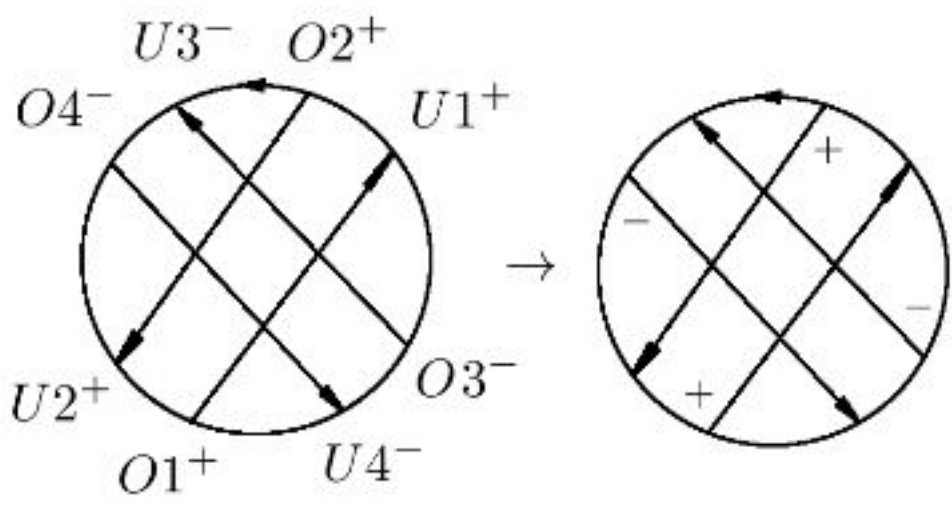
and connecting the crossings as instructed by the Gauss code:



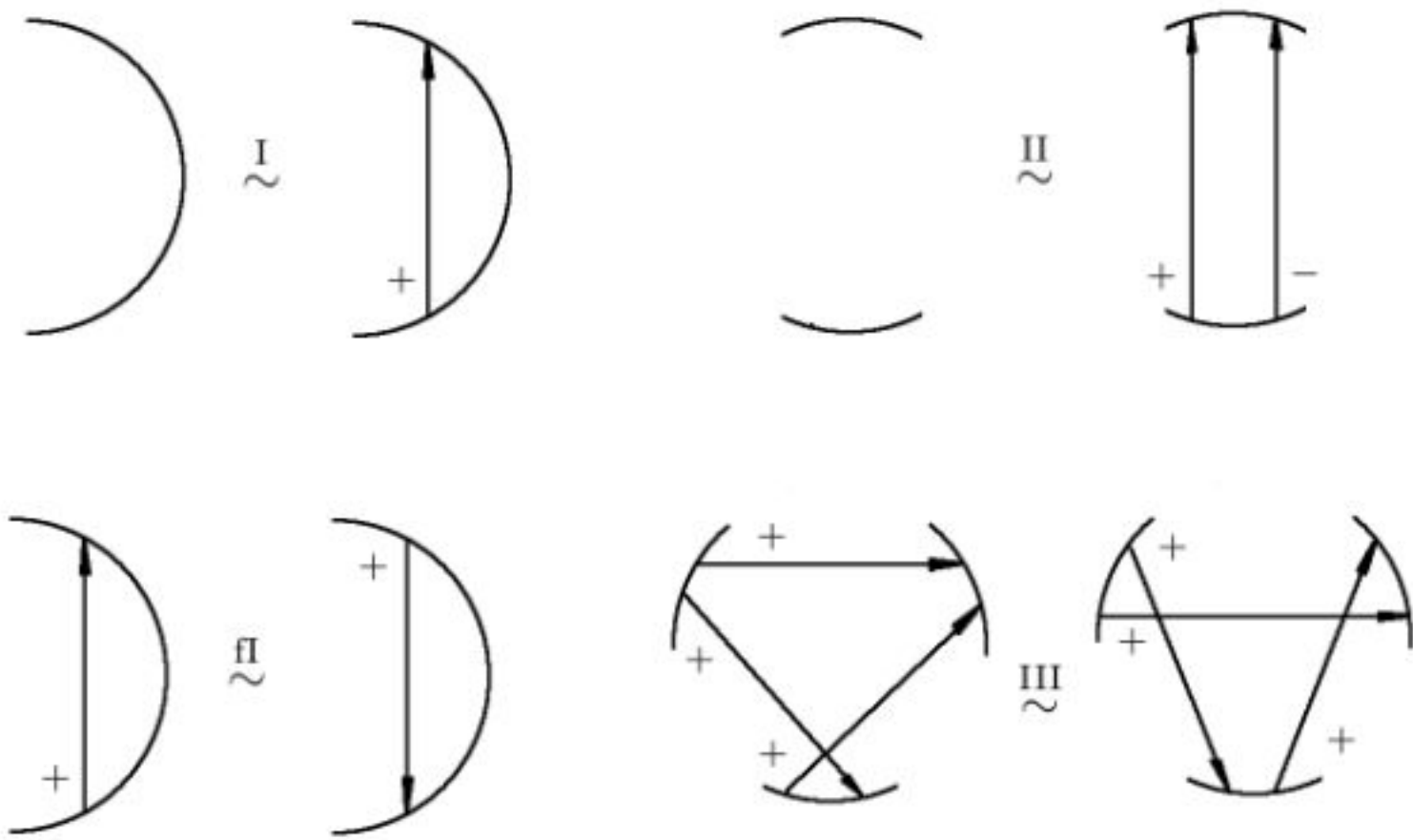
Notice that at the last step, we made a choice to put the last semiarc on the top of the diagram; we could instead have gone around the bottom of the diagram, which is equivalent to dragging the strand around the back of the sphere (if we think of our knot as drawn on the sphere).

Reidemeister moves change Gauss codes, but they do so in controlled ways. It is perhaps easiest to see how this works by looking at *Gauss diagrams*: in a Gauss diagram, we write a Gauss code counterclockwise around a circle. Each crossing appears twice along the circle; we connect the over and under label for each crossing with an

arrow pointing “in the direction of gravity”, i.e., toward the undercrossing. We then label each arrow with the crossing sign.



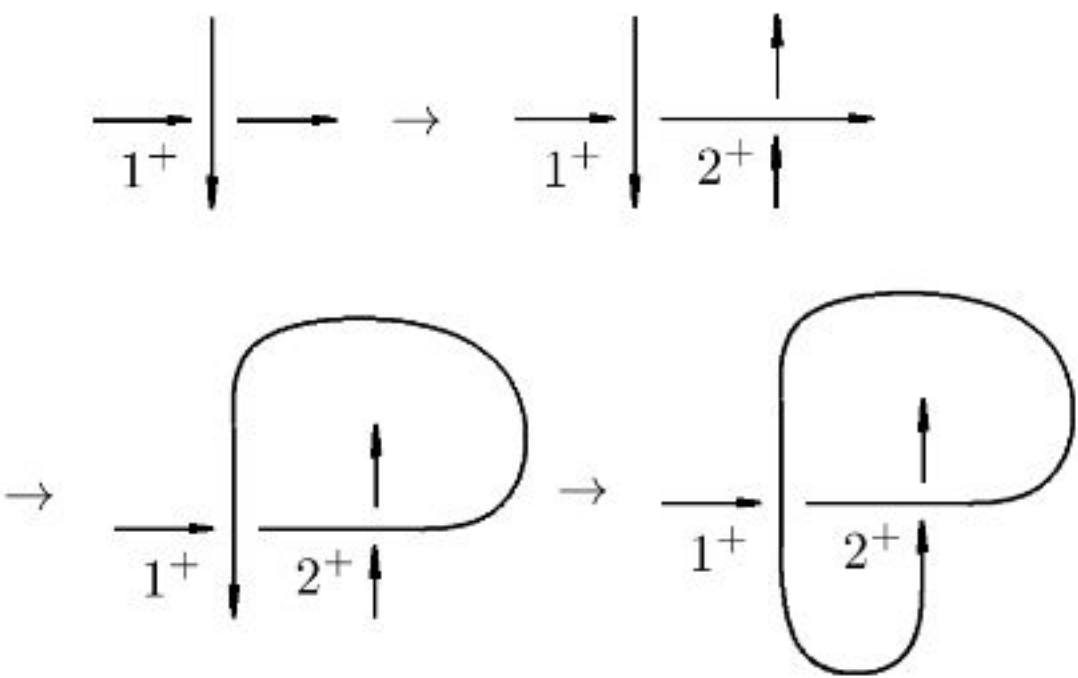
The Reidemeister moves change Gauss diagrams by inserting or deleting signed arrows in certain ways and by sliding signed arrow heads and tails past each other in certain ways:



Since we can think of a knot as an equivalence class of knot diagrams under the equivalence relation generated by the Reidemeister moves, we can now think of knots as equivalence classes of signed Gauss codes or Gauss diagrams under the equivalence relation generated by the Gauss code/diagram Reidemeister moves (note that we have only depicted a subset of the moves above; for example, there

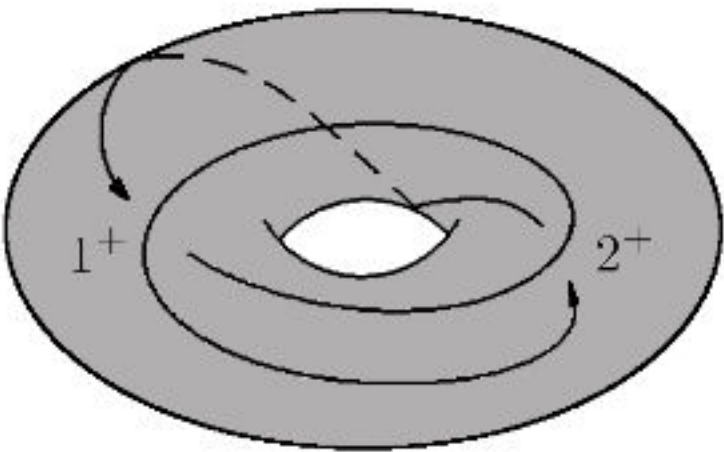
are type III moves with two positive and one negative crossing, etc.) together with the cyclic permutation equivalence mentioned above.

However, there is a slight problem with this idea: consider the Gauss code $U1^+O2^+O1^+U2^+$. As a Gauss code, it certainly has an equivalence class in the quotient set of Gauss codes modulo Reidemeister equivalence. What happens when we try to draw this knot?

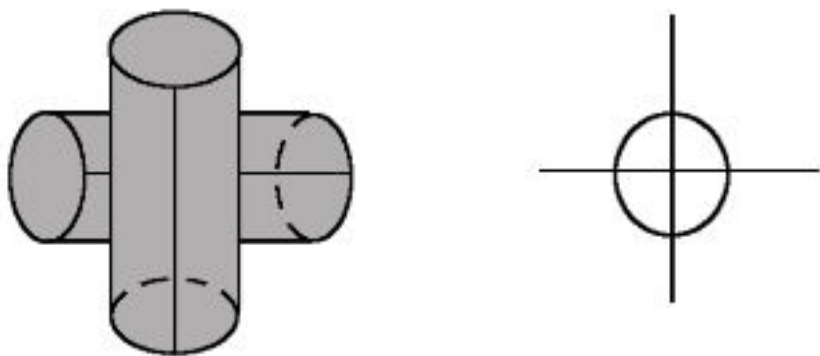


We need to join the ends but we can't because they are in different regions of the plane. It is tempting to say “let's just put in another crossing”, but notice that the Gauss code is supposed to already include all the crossings.

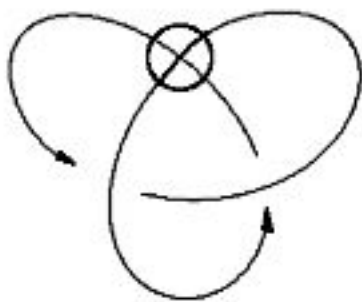
How can we connect the ends to make a knot when they are in separate regions of the plane? One idea will be familiar to fans of science fiction: the knot should go through a wormhole. That is, we can connect the regions of the plane with a bridge that avoids the rest of the knot. The net effect is that instead of drawing our knot diagram on the plane, we draw it on a torus. Then it turns out we can complete the diagram without needing extra crossings [KK00].



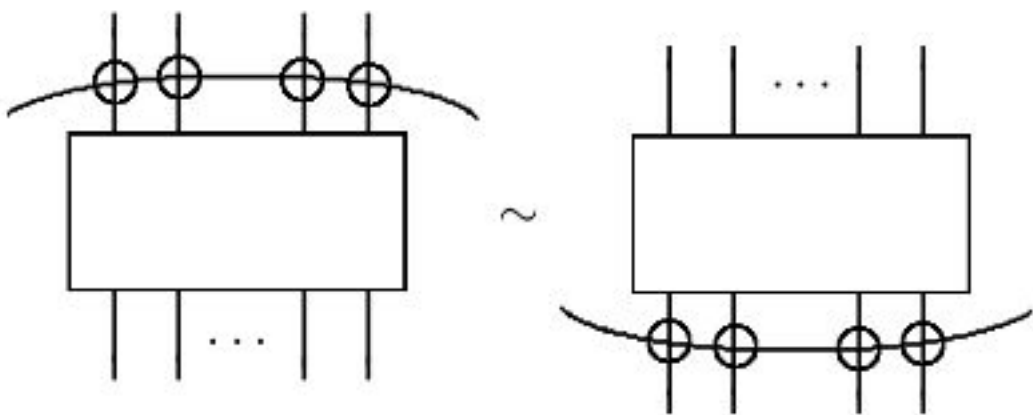
Another idea, introduced in [Kau99], is to go ahead and include a new crossing, but since the new crossing isn't in the Gauss code, make it a new type of crossing called a *virtual crossing*.



We can think of the virtual crossing as the result of squashing the torus on which the diagram really lives into the plane.

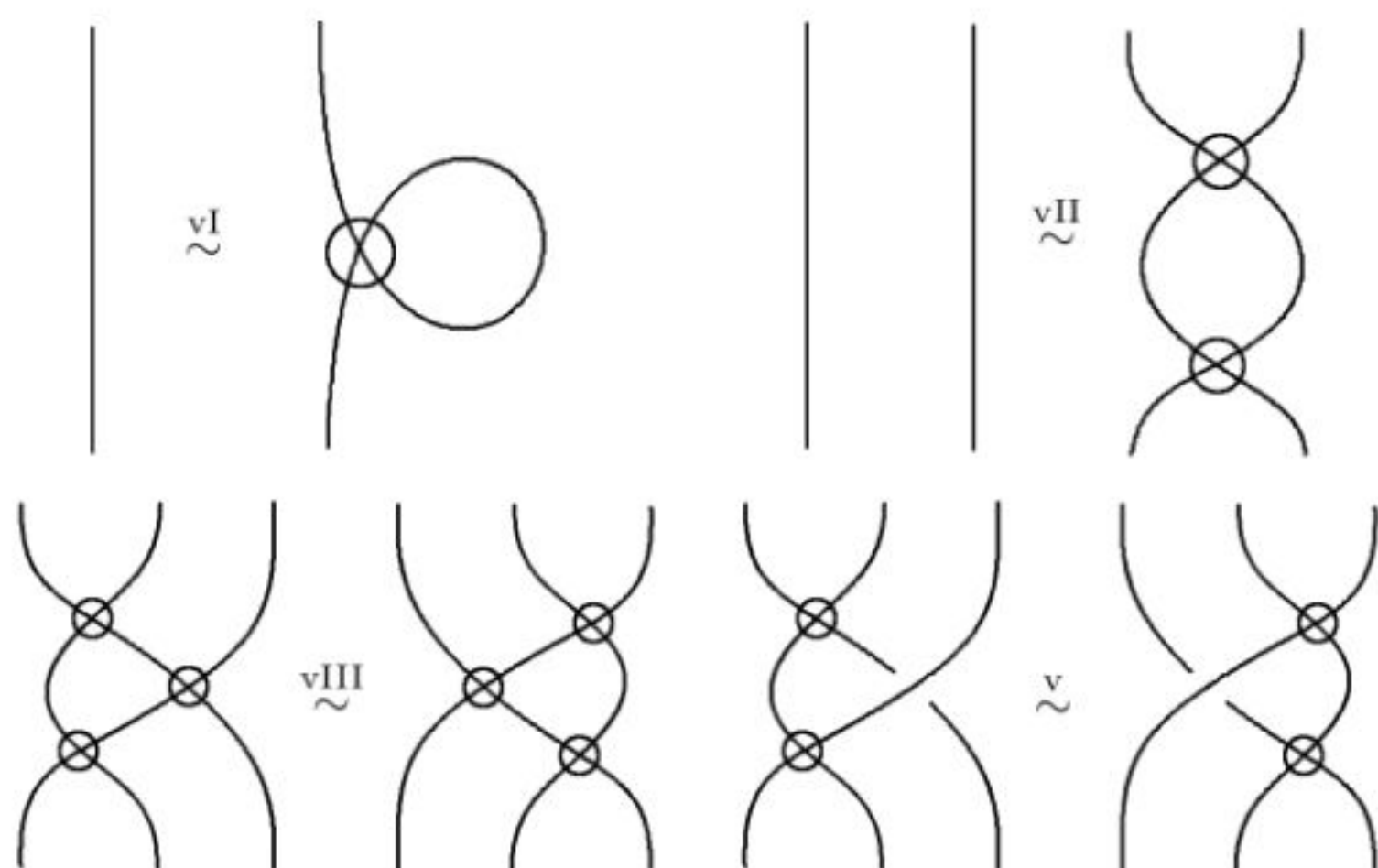


Since the virtual crossings do not appear in the Gauss codes, any arc which has only virtual crossings can be replaced with any other arc with same endpoints and only virtual crossings along its interior. This is known as the *detour move*:



Inside the box can be any tangle, including both classical and virtual crossings. The detour move breaks down into four new Reidemeister

moves depending on what we put in the box.

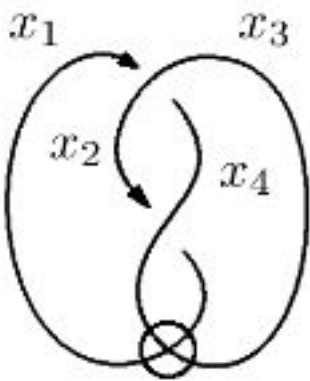


Thus, we can define a *virtual knot* as either:

- an equivalence class of Gauss codes/diagrams under the equivalence relation generated by the Gauss code Reidemeister moves and cyclic permutations, or
- an equivalence class of virtual knot diagrams under the equivalence relation generated by the classical and virtual Reidemeister moves I, II, III, vI, vII, vIII and v.

Invariants of Virtual Knots. Many invariants of classical knots extend to invariants of virtual knots by virtue of being defined locally from information at crossings rather than globally involving all of the ambient space. For example, quandle colorings work just as well for knots on surfaces as for knots on the plane; we simply ignore the virtual crossings when determining the colorings. The same is true for coloring by kei, racks, bikei, biquandles and biracks for the appropriate types of virtual knots. Note that for framed virtual knots we replace the classical type I move with the framed type I move, but still keep the usual virtual vI move.

Example 137.



The *virtual trefoil* knot is the only nontrivial virtual knot with only two classical crossings; it is the one we saw above drawn on the torus. Let us prove that it is nontrivial using biquandle labelings. Consider the biquandle X defined by the operation matrix

$$\left[\begin{array}{ccc|ccc} 2 & 1 & 3 & 2 & 2 & 2 \\ 1 & 3 & 2 & 3 & 3 & 3 \\ 3 & 2 & 1 & 1 & 1 & 1 \end{array} \right].$$

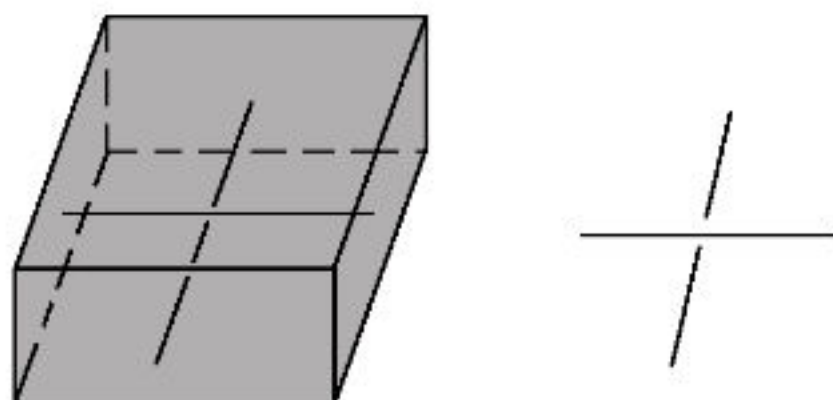
The unknot has three labelings by X (namely, the circle with labels “1”, “2” and “3”), but the virtual trefoil has no valid X -labeling: once we choose labels for x_1 and x_2 , these determine labels for $x_3 = x_2 \bar{\triangleright} x_1$ and $x_4 = x_1 \triangleright x_2$, and we will have a valid labeling only if $x_4 = x_3 \bar{\triangleright} x_2$ and $x_1 = x_2 \triangleright x_3$. We can then check all nine possibilities:

x_1	x_2	$x_3 = x_2 \bar{\triangleright} x_1$	$x_4 = x_1 \triangleright x_2$	$x_4 = x_3 \bar{\triangleright} x_2?$	$x_1 = x_2 \triangleright x_3?$
1	1	2	2	$2 \neq 3$	$1 = 1$
1	2	3	1	$1 = 1$	$1 \neq 2$
1	3	1	3	$3 \neq 2$	$1 \neq 3$
2	1	2	1	$1 \neq 3$	$2 \neq 1$
2	2	3	3	$3 \neq 1$	$2 = 2$
2	3	1	2	$2 = 2$	$2 \neq 3$
3	1	2	3	$3 = 3$	$3 \neq 1$
3	2	3	2	$2 \neq 1$	$3 \neq 2$
3	3	1	1	$1 \neq 2$	$3 = 3$

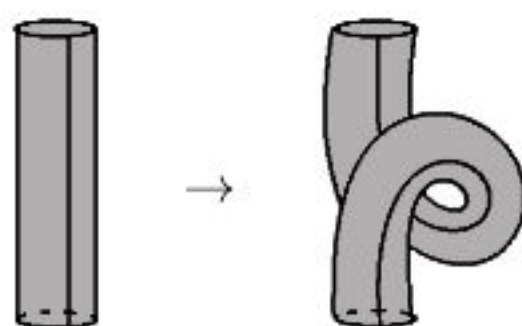
Since any sequence of virtual moves taking the unknot to the virtual trefoil would preserve the three biquandle colorings of the unknot, there cannot be any such sequence and the virtual trefoil is not equivalent to the unknot.

Virtual knots and knots on surfaces. What could be the meaning of a knot diagram on a surface? As we have previously observed,

even though knots require three dimensions, they're really almost two-dimensional, requiring only a little bit of thickness for crossings. Thus, instead of thinking of knots as living in \mathbb{R}^3 , we can think of knots as living in $\mathbb{R}^2 \times [-\epsilon, \epsilon]$, a *thickened plane*.

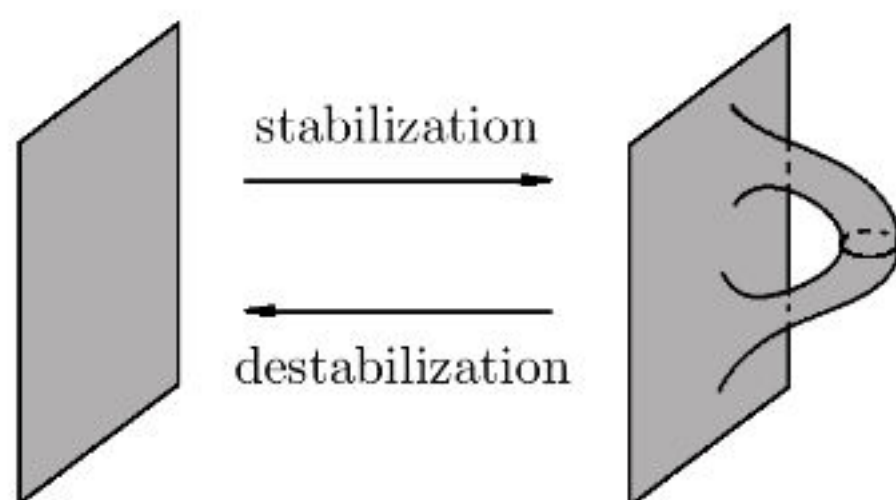


A knot drawn on a surface corresponds to a simple closed curve in a thickened version of the surface. For example, a thickened sphere is like the peel of an orange, while a thickened torus is like the frosting on a donut. Classical Reidemeister moves on a virtual knot correspond to isotopy of the knot within thickened surface. Virtual Reidemeister moves can correspond to movement of the thickened surface pre-squashing into the plane:

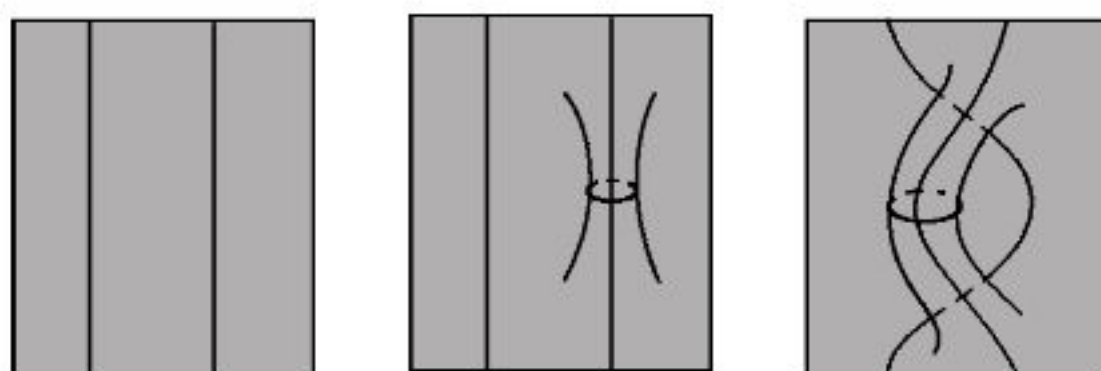


But the virtual moves can also involve changes to the supporting surface on which the knot diagram is drawn, namely adding or removing the “wormholes” (technically known as *handles*) that allow strands with virtual crossings to avoid other strands. Formally, the process is called *stabilization* if a handle is added and *destabilization* if a handle

is removed.



Then some virtual Reidemeister moves can involve adding or removing handles in addition to moving the handles around before squashing into the plane to make the virtual knot diagram.

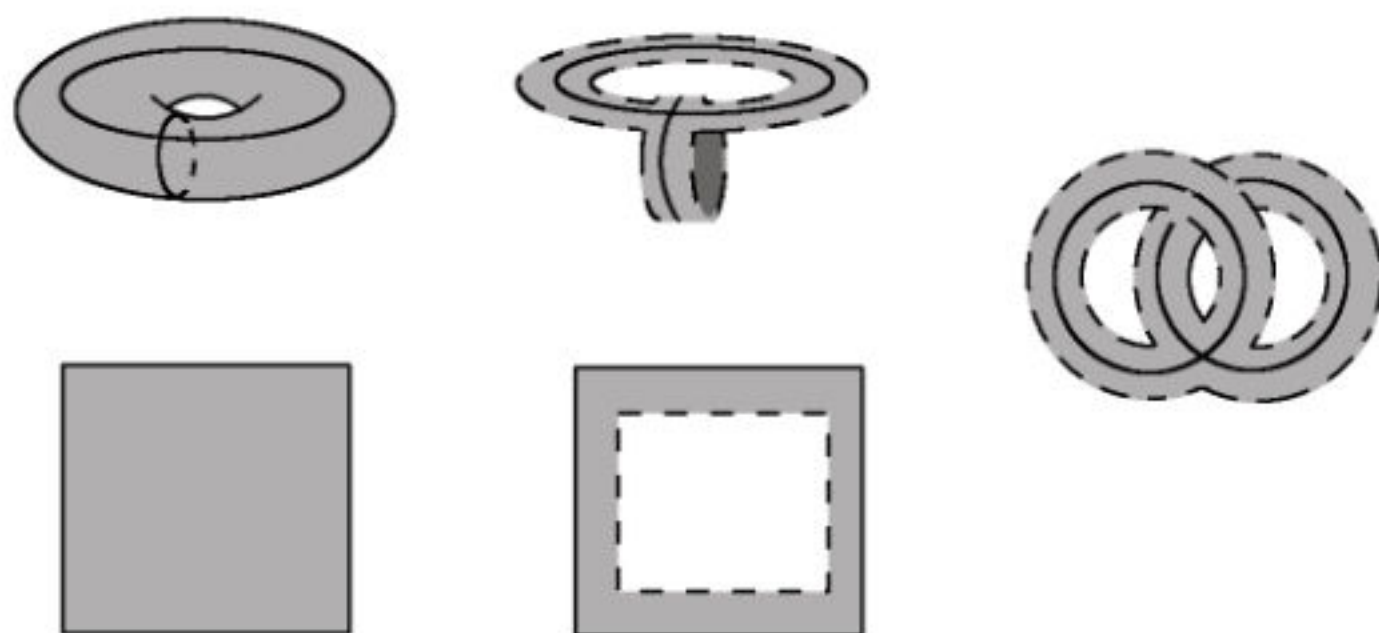


Thus, a virtual knot can be understood geometrically as a simple closed curve in a thickened surface up to ambient isotopy and stabilization of the surface.

Twisted Virtual Knots. A virtual knot can be understood as an equivalence class of knot diagrams drawn on surfaces which can be obtained from the sphere by stabilization moves (adding handles) with ordinary or *classical* knot theory being the special case of knots drawn on the sphere S^2 . What about knot diagrams drawn on other types of surfaces like projective planes or Klein bottles?

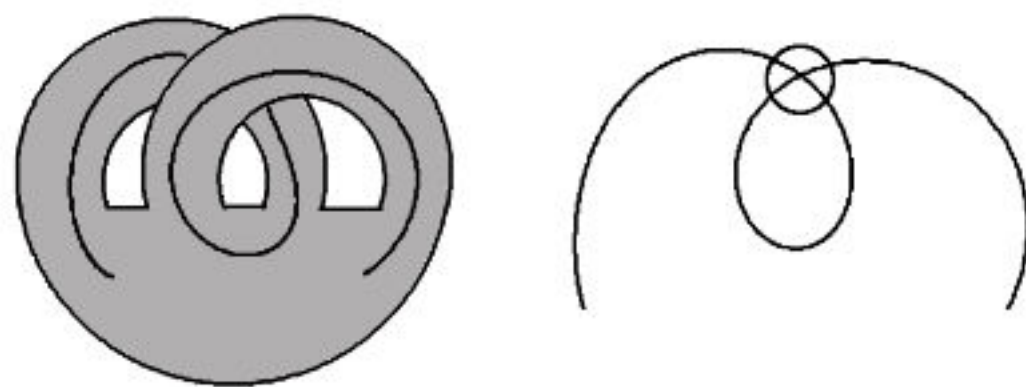
One of the great triumphs in the early history of topology was the classification of compact surfaces, which says that compact surfaces are classified by two numbers, the *genus* and the *cross-cap number*. The genus of a surface is the number of handles it has, while the cross-cap number is the number of projective planes or Möbius bands we can cut out from the surface.

We have already seen that removing a disc from a projective plane yields a Möbius band. What happens to a torus when we remove a disc? The result is two crossed bands:

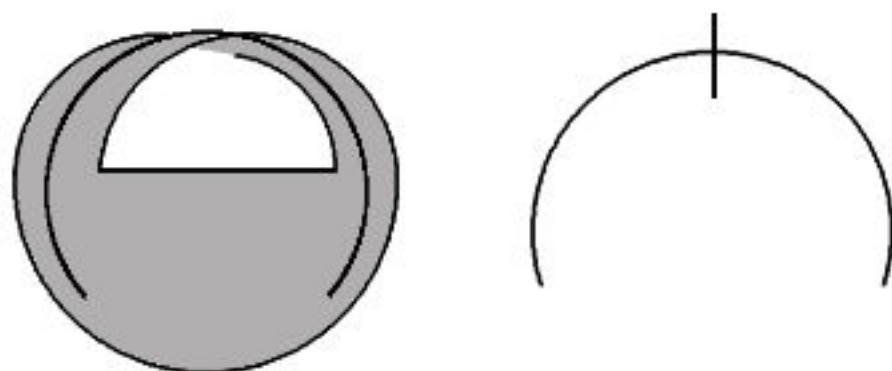


The classification theorem for compact surfaces says that every compact surface can be identified topologically with a surface obtained from a disk by attaching g pairs of crossed bands and c Möbius bands (where g is the genus and c is the cross-cap number), then gluing a disk along the outside boundary circle.

We have already seen that virtual crossings represent genus in the underlying surface:



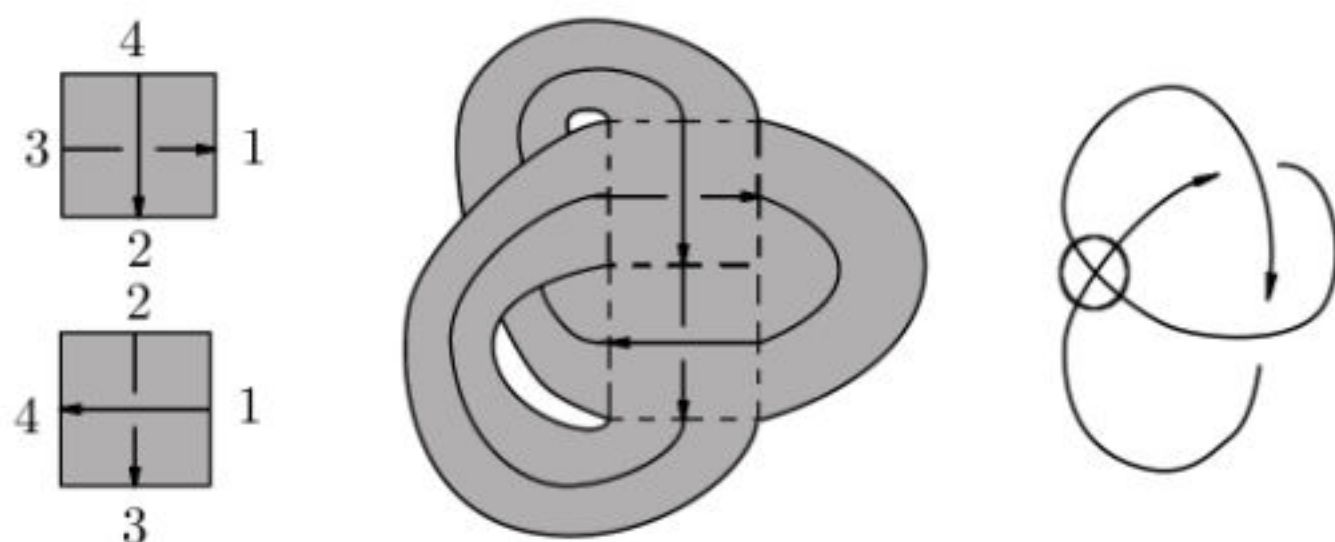
When our knot goes through a Möbius band, we mark this with a “twist bar”.



If a surface has cross-cap number $c > 0$, then the surface is not orientable. An *orientation* on a surface is a consistent choice of normal vector, or equivalently a consistent choice of top side vs. bottom side. A Möbius band is famously one-sided and thus has no top or bottom side; if we try to choose a consistent normal vector (more precisely, a nonvanishing continuous normal vector field), we find that sliding the normal vector along the center line of the Möbius band reverses the normal vector. Indeed, any closed path on a surface that reverses the normal vector is called an *orientation reversing path*, and a surface is orientable only if it has no orientation reversing paths.

Then a *twisted virtual knot* is simple closed curve in a thickened compact surface which may or may not be orientable; an arc with a twist bar is an orientation reversing path [Bou08].

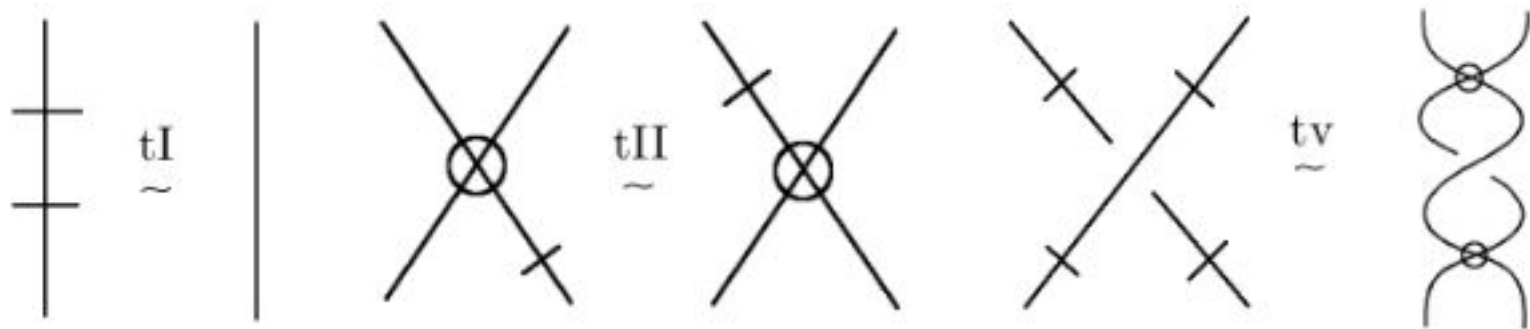
Twisted virtual knot theory is a natural extension of *abstract knot theory*, in which we think of a knot as a collection of crossings to be connected by bands. For example, the virtual trefoil can be built from two crossings with gluing information specified along the edges of the squares containing the crossings.



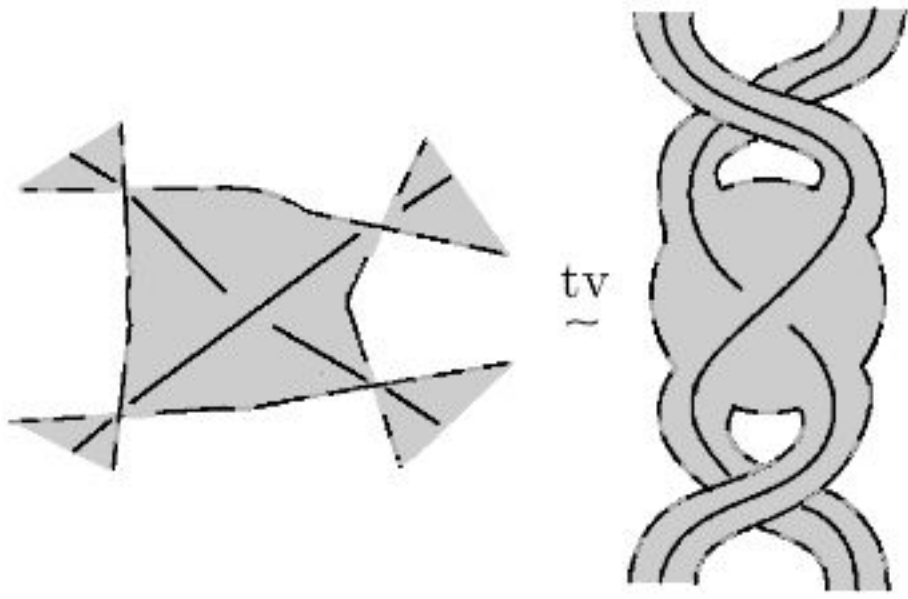
The result is a surface with boundary consisting of circles which we can then close by gluing on disks. Two abstract knots are equivalent if they are related by Reidemeister moves on the resulting surface.

A little thought then shows how the twist bars should interact with classical and virtual crossings. In addition to the classical and

virtual Reidemeister moves, we have the *twisted Reidemeister moves*.

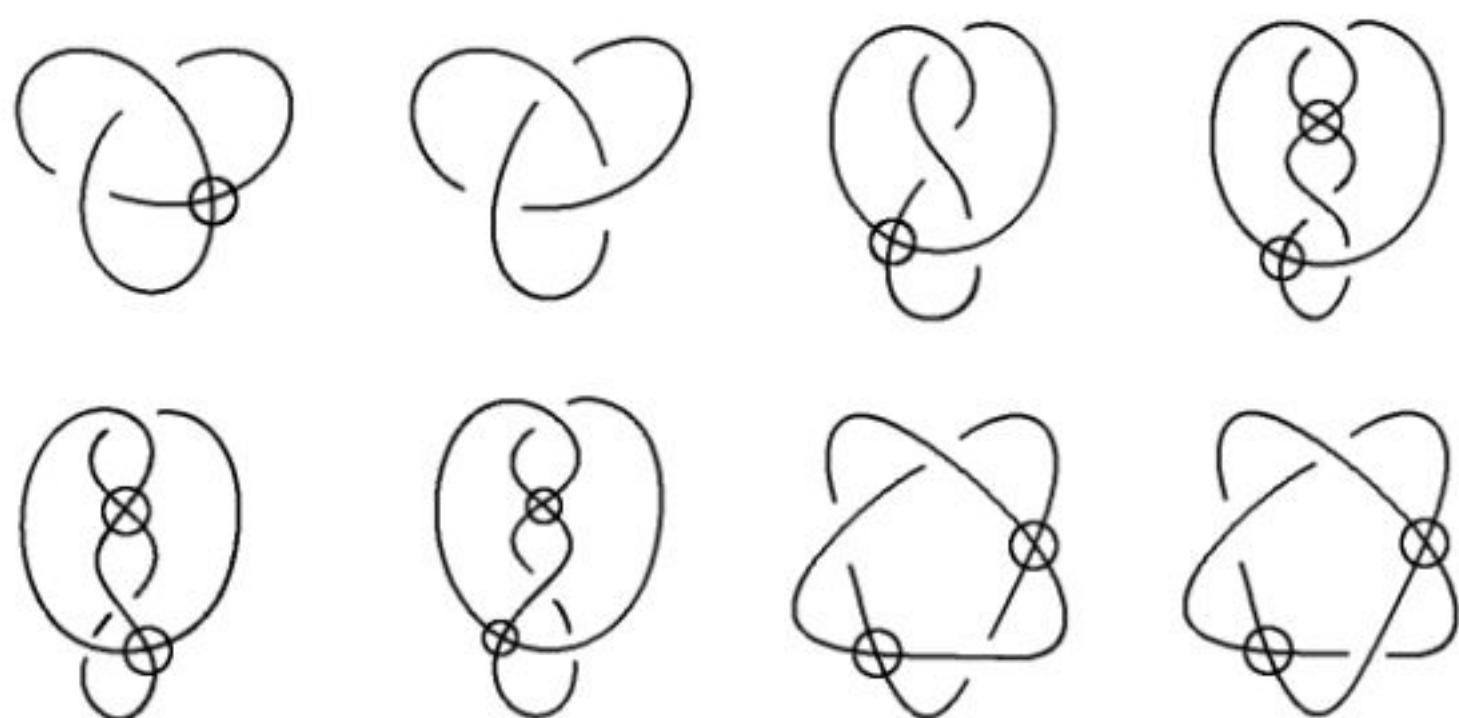


The first move says that two twists on an arc cancel (since a band with two half-twists is orientable), the second move says we can move a twist past a pair of crossed bands, and the last one says if we have a crossing with all four bands twisted, we can flip it over to replace the twists with two flat crossed bands:



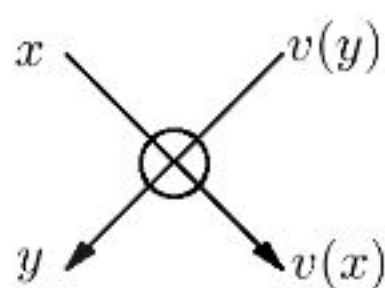
Virtual Knot Table. We conclude this section with a (small) table of virtual knots with two or three classical crossings. While there are only eight knots here, note that the corresponding table for classical knot includes only the trefoil. There are 108 distinct virtual knots

with 4 classical crossings; see [BN] for more.



Exercises. 1. Prove that if a virtual knot has only virtual crossings, it is unknotted.

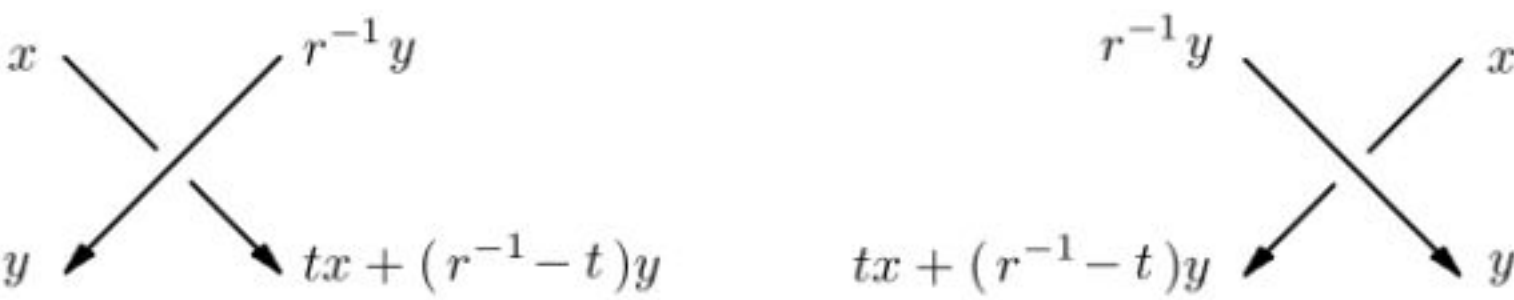
2. A *virtual quandle* is an algebraic structure with a quandle operation at classical crossings and new operations at virtual crossings. Suppose we divide our virtual knot at classical undercrossings and virtual crossings and define a map $v : X \rightarrow X$ as pictured:



Then prove that the new algebraic structure respects the virtual isotopy moves iff v is a quandle automorphism [KM05].

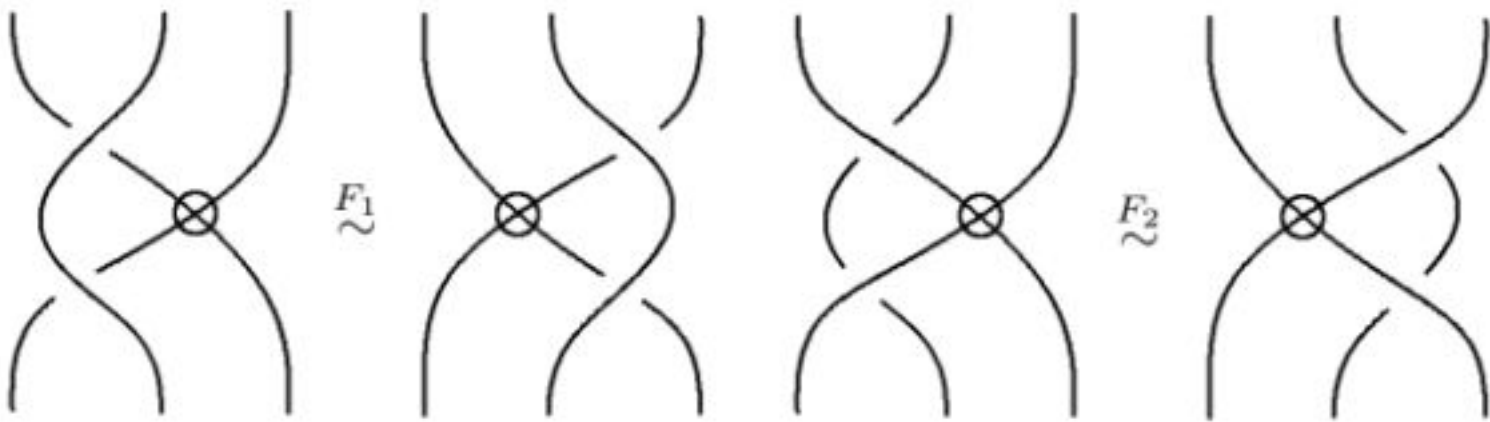
3. The *virtual Alexander polynomial* of a virtual knot is the determinant of the presentation matrix of the virtual knot's fundamental Alexander biquandle, obtained analogously to the presentation matrix for the Alexander module but with the following labelings at

crossings:



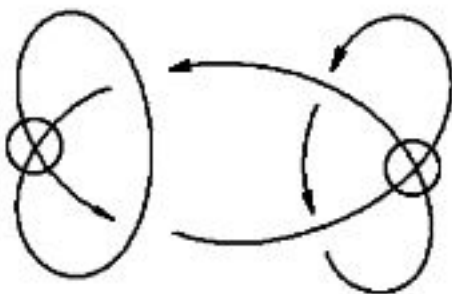
Compute the virtual Alexander polynomial for each of the virtual knots with 3 or fewer classical crossings [KR03].

4. There are some moves that seem like they could be legitimate moves, but in fact are not allowed because they change the Gauss code of the diagram. These are called *forbidden moves*:



Show that the virtual trefoil can be unknotted if you're allowed to use both forbidden moves in addition to the classical and virtual Reidemeister moves.

- 5. Show that the virtual trefoil has trivial fundamental quandle.
- 6. The knot



is called the *Kishino knot*. Show that it has trivial fundamental quandle, but is distinguished from the unknot by the counting invariant

with coloring biquandle

$$\left[\begin{array}{cccc|cccc} 3 & 2 & 1 & 4 & 3 & 1 & 4 & 2 \\ 1 & 4 & 3 & 2 & 2 & 4 & 1 & 3 \\ 4 & 1 & 2 & 3 & 1 & 3 & 2 & 4 \\ 2 & 3 & 4 & 1 & 4 & 2 & 3 & 1 \end{array} \right].$$

7. Define a notion of *twisted kei* by including an operation $x \rightarrow T(x)$ at twist bars. Show that there is only one twisted kei structure on the set $\{1, 2\}$ of two elements.

8. The knot



is called the *onefoil*. Use your answer from problem 7 to prove that the onefoil is not equivalent to the unknot.

Bibliography

- [AB26] J. W. Alexander and G. B. Briggs, *On types of knotted curves*, Ann. of Math. (2) **28** (1926/27), no. 1-4, 562–586. MR1502807
- [Ada04] C. C. Adams, *The knot book*, American Mathematical Society, Providence, RI, 2004. An elementary introduction to the mathematical theory of knots, Revised reprint of the 1994 original. MR2079925 (2005b:57009)
- [AER⁺08] K. Ameer, M. Elhamdadi, T. Rose, M. Saito, and C. Smudde, *Tangle embeddings and quandle cocycle invariants*, Experiment. Math. **17** (2008), no. 4, 487–497. MR2484432 (2009k:57003)
- [AG03] N. Andruskiewitsch and M. Graña, *From racks to pointed Hopf algebras*, Adv. Math. **178** (2003), no. 2, 177–243. MR1994219 (2004i:16046)
- [AN12] S. Aksoy and S. Nelson, *Bikei, involutory biracks and unoriented link invariants*, J. Knot Theory Ramifications **21** (2012), no. 6, 1250045, 13. MR2903175
- [AS09] K. Ameer and M. Saito, *Polynomial cocycles of Alexander quandles and applications*, J. Knot Theory Ramifications **18** (2009), no. 2, 151–165. MR2507921 (2010g:57009)
- [Bel58] V. D. Belousov, *Les quasi-groupes transitifs et distributifs*, Ukrain. Mat. Ž. **10** (1958), no. 1, 13–22. MR0095892 (20 #2390)
- [Bel60] V. D. Belousov, *The structure of distributive quasigroups*, Mat. Sb. (N.S.) **50** (92) (1960), 267–298. MR0120304 (22 #11059)
- [Bel67] V. D. Belousov, *Osnovy teorii kvazigrupp i lup*, Izdat. “Nauka”, Moscow, 1967. MR0218483 (36 #1569)
- [BF04] S. Budden and R. Fenn, *The equation $[B, (A - 1)(A, B)] = 0$ and virtual knots and links*, Fund. Math. **184** (2004), 19–29. MR2128040 (2005k:57007)
- [BF08] S. Budden and R. Fenn, *Quaternion algebras and invariants of virtual knots and links. II. The hyperbolic case*, J. Knot Theory Ramifications **17** (2008), no. 3, 305–314. MR2402508 (2009d:57006b)
- [BF11] A. Bartholomew and R. Fenn, *Biquandles of small size and some invariants of virtual and welded knots*, J. Knot Theory Ramifications **20** (2011), no. 7, 943–954. MR2819176 (2012e:57004)

- [BM29] C. Burstin and W. Mayer, *Distributive gruppen von endlicher ordnung.*, Journal für die reine und angewandte Mathematik **160** (1929), 111–130.
- [BN02] D. Bar-Natan, *On Khovanov's categorification of the Jones polynomial*, Algebr. Geom. Topol. **2** (2002), 337–370 (electronic). MR1917056 (2003h:57014)
- [BN] D. Bar-Natan, *The knot atlas* http://katlas.org/wiki/Main_Page.
- [Bol37] G. Bol, *Gewebe und gruppen*, Math. Ann. **114** (1937), no. 1, 414–431. MR1513147
- [Bou08] M. O. Bourgoïn, *Twisted link theory*, Algebr. Geom. Topol. **8** (2008), no. 3, 1249–1279. MR2443243 (2009g:57017)
- [Bri88] E. Brieskorn, *Automorphic sets and braids and singularities*, Braids (Santa Cruz, CA, 1986), 1988, pp. 45–115. MR975077 (90a:32024)
- [Bru58] R. H. Bruck, *A survey of binary systems*, Ergebnisse der Mathematik und ihrer Grenzgebiete. Neue Folge, Heft 20. Reihe: Gruppentheorie, Springer Verlag, Berlin-Göttingen-Heidelberg, 1958. MR0093552 (20 #76)
- [CCES08a] J. S. Carter, A. S. Crans, M. Elhamdadi, and M. Saito, *Cohomology of categorical self-distributivity*, J. Homotopy Relat. Struct. **3** (2008), no. 1, 13–63. MR2395367 (2010b:16069)
- [CCES08b] J. S. Carter, A. S. Crans, M. Elhamdadi, and M. Saito, *Cohomology of the adjoint of Hopf algebras*, J. Gen. Lie Theory Appl. **2** (2008), no. 1, 19–34. MR2368886 (2009a:16064)
- [CCE⁺08] J. S. Carter, A. S. Crans, M. Elhamdadi, E. Karadayi, and M. Saito, *Cohomology of Frobenius algebras and the Yang-Baxter equation*, Commun. Contemp. Math. **10** (2008), no. suppl. 1, 791–814. MR2468364 (2009i:16018)
- [CEGN14] J. Cenicerros, M. Elhamdadi, M. Green, and S. Nelson, *Augmented biracks and their homology*, Internat. J. Math. **25** (2014), no. 9, 1450087, 19. MR3266530
- [CEGS05] J. S. Carter, M. Elhamdadi, M. Graña, and M. Saito, *Cocycle knot invariants from quandle modules and generalized quandle homology*, Osaka J. Math. **42** (2005), no. 3, 499–541. MR2166720 (2006d:57017)
- [CEH⁺13] W. E. Clark, M. Elhamdadi, X.-d. Hou, M. Saito, and T. Yeatman, *Connected quandles associated with pointed abelian groups*, Pacific J. Math. **264** (2013), no. 1, 31–60. MR3079760
- [CENS03] J. S. Carter, M. Elhamdadi, M. A. Nikiforou, and M. Saito, *Extensions of quandles and cocycle knot invariants*, J. Knot Theory Ramifications **12** (2003), no. 6, 725–738. MR2008876 (2004g:57020)
- [CES02] J. S. Carter, M. Elhamdadi, and M. Saito, *Twisted quandle homology theory and cocycle knot invariants*, Algebr. Geom. Topol. **2** (2002), 95–135 (electronic). MR1885217 (2003a:57019)
- [CES04] J. S. Carter, M. Elhamdadi, and M. Saito, *Homology theory for the set-theoretic Yang-Baxter equation and knot invariants from generalizations of quandles*, Fund. Math. **184** (2004), 31–54. MR2128041 (2005k:57009)
- [CESS06] J. S. Carter, M. Elhamdadi, M. Saito, and S. Satoh, *A lower bound for the number of Reidemeister moves of type III*, Topology Appl. **153** (2006), no. 15, 2788–2794. MR2248382 (2007d:57009)
- [CESY14] W. E. Clark, M. Elhamdadi, M. Saito, and T. Yeatman, *Quandle colorings of knots and applications*, J. Knot Theory Ramifications **23** (2014), no. 6, 1450035, 29. MR3253967
- [Che74] O. Chein, *Moufang loops of small order. I*, Trans. Amer. Math. Soc. **188** (1974), 31–51. MR0330336 (48 #8673)

- [CHN13] A. S. Crans, A. Henrich, and S. Nelson, *Polynomial knot and link invariants from the virtual biquandle*, J. Knot Theory Ramifications **22** (2013), no. 4, 134004, 15. MR3055555
- [CJK⁺03] J. S. Carter, D. Jelsovsky, S. Kamada, L. Langford, and M. Saito, *Quandle cohomology and state-sum invariants of knotted curves and surfaces*, Trans. Amer. Math. Soc. **355** (2003), no. 10, 3947–3989. MR1990571 (2005b:57048)
- [CN09] J. Cenicerós and S. Nelson, *Virtual Yang-Baxter cocycle invariants*, Trans. Amer. Math. Soc. **361** (2009), no. 10, 5263–5283. MR2515811 (2010k:57021)
- [CN12] J. Cenicerós and S. Nelson, *(t, s) -racks and their link invariants*, Internat. J. Math. **23** (2012), no. 3, 1250001, 19. MR2902281
- [CN13] J. Cenicerós and S. Nelson, *Twisted virtual biracks and their twisted virtual link invariants*, Topology Appl. **160** (2013), no. 2, 421–429. MR3003341
- [CN14] E. Cody and S. Nelson, *Polynomial birack modules*, Topology Appl. **173** (2014), 285–293. MR3227223
- [CNS12] A. S. Crans, S. Nelson, and A. Sarkar, *Enhancements of rack counting invariants via dynamical cocycles*, New York J. Math. **18** (2012), 337–351. MR2928580
- [COP71] O. Chein and H. Orlik-Pflugfelder, *The smallest Moufang loop*, Arch. Math. (Basel) **22** (1971), 573–576. MR0297914 (45 #6966)
- [CS98] J. S. Carter and M. Saito, *Knotted surfaces and their diagrams*, Mathematical Surveys and Monographs, vol. 55, American Mathematical Society, Providence, RI, 1998. MR1487374 (98m:57027)
- [CSW⁺09] J. S. Carter, D. S. Silver, S. G. Williams, M. Elhamdadi, and M. Saito, *Virtual knot invariants from group biquandles and their cocycles*, J. Knot Theory Ramifications **18** (2009), no. 7, 957–972. MR2549477 (2010i:57026)
- [CW] J. C. Conway and G. C. Wraith, *correspondence*.
- [EMR12] M. Elhamdadi, J. Macquarrie, and R. Restrepo, *Automorphism groups of quandles*, J. Algebra Appl. **11** (2012), no. 1, 1250008, 9. MR2900878
- [EN12] M. Elhamdadi and S. Nelson, *N -degeneracy in rack homology and link invariants*, Hiroshima Math. J. **42** (2012), no. 1, 127–142. MR2952076
- [FJSK04] R. Fenn, M. Jordan-Santana, and L. Kauffman, *Biquandles and virtual links*, Topology Appl. **145** (2004), no. 1-3, 157–175. MR2100870 (2005h:57015)
- [FR92] R. Fenn and C. Rourke, *Racks and links in codimension two*, J. Knot Theory Ramifications **1** (1992), no. 4, 343–406. MR1194995 (94e:57006)
- [FRS95] R. Fenn, C. Rourke, and B. Sanderson, *Trunks and classifying spaces*, Appl. Categ. Structures **3** (1995), no. 4, 321–356. MR1364012 (96i:57023)
- [FT07] R. Fenn and V. Turaev, *Weyl algebras and knots*, J. Geom. Phys. **57** (2007), no. 5, 1313–1324. MR2289648 (2008k:57007)
- [Gal79] V. M. Galkin, *Left distributive finite order quasigroups*, Mat. Issled. **51** (1979), 43–54, 163. Quasigroups and loops. MR544332 (81d:20064)
- [Gal88] V. M. Galkin, *Quasigroups*, Algebra. Topology. Geometry, Vol. 26 (Russian), 1988, pp. 3–44, 162. Translated in J. Soviet Math. **49** (1990), no. 3, 941–967. MR978392 (89k:20103)
- [HN10] A. Henrich and S. Nelson, *Semiquandles and flat virtual knots*, Pacific J. Math. **248** (2010), no. 1, 155–170. MR2734169 (2011h:57015)
- [Jon85] V. F. R. Jones, *A polynomial invariant for knots via von Neumann algebras*, Bull. Amer. Math. Soc. (N.S.) **12** (1985), no. 1, 103–111. MR766964 (86e:57006)

- [Joy82] D. Joyce, *A classifying invariant of knots, the knot quandle*, J. Pure Appl. Algebra **23** (1982), no. 1, 37–65. MR638121 (83m:57007)
- [Kau91] L. H. Kauffman, *Knots and physics*, Series on Knots and Everything, vol. 1, World Scientific Publishing Co. Inc., River Edge, NJ, 1991. MR1141156 (93b:57010)
- [Kau99] L. H. Kauffman, *Virtual knot theory*, European J. Combin. **20** (1999), no. 7, 663–690. MR1721925 (2000i:57011)
- [KK00] N. Kamada and S. Kamada, *Abstract link diagrams and virtual knots*, J. Knot Theory Ramifications **9** (2000), no. 1, 93–106. MR1749502 (2001h:57007)
- [KM05] L. H. Kauffman and V. O. Manturov, *Virtual biquandles*, Fund. Math. **188** (2005), 103–146. MR2191942 (2006k:57015)
- [KR03] L. H. Kauffman and D. Radford, *Bi-oriented quantum algebras, and a generalized Alexander polynomial for virtual links*, Diagrammatic morphisms and applications (San Francisco, CA, 2000), 2003, pp. 113–140. MR1973514 (2004c:57013)
- [KSS03] T. Kanenobu, H. Saito, and S. Satoh, *Tangles with up to seven crossings*, Proceedings of the Winter Workshop of Topology/Workshop of Topology and Computer (Sendai, 2002/Nara, 2001), 2003, pp. 127–140. MR2023113 (2004j:57006)
- [LN09] D. Lam and S. Nelson, *An isomorphism theorem for Alexander biquandles*, Internat. J. Math. **20** (2009), no. 1, 97–107. MR2488716 (2009m:57023)
- [Mat82] S. V. Matveev, *Distributive groupoids in knot theory*, Mat. Sb. (N.S.) **119**(161) (1982), no. 1, 78–88, 160. MR672410 (84e:57008)
- [Moc03] T. Mochizuki, *Some calculations of cohomology groups of finite Alexander quandles*, J. Pure Appl. Algebra **179** (2003), no. 3, 287–330. MR1960136 (2004b:55013)
- [Moc05] T. Mochizuki, *The 3-cocycles of the Alexander quandles $\mathbb{F}_q[T]/(T - \omega)$* , Algebr. Geom. Topol. **5** (2005), 183–205 (electronic). MR2135551 (2006e:57019)
- [Mou33] R. Moufang, *Alternativkörper und der Satz vom vollständigen Vierseit (D_9)*, Abh. Math. Sem. Univ. Hamburg **9** (1933), no. 1, 207–222.
- [Nel01] S. Nelson, *Unknotting virtual knots with Gauss diagram forbidden moves*, J. Knot Theory Ramifications **10** (2001), no. 6, 931–935. MR1840276 (2002c:57009)
- [Nel03] S. Nelson, *Classification of finite Alexander quandles*, Proceedings of the Spring Topology and Dynamical Systems Conference, 2003, pp. 245–258. MR2048935 (2005b:57027)
- [Nel08] S. Nelson, *A polynomial invariant of finite quandles*, J. Algebra Appl. **7** (2008), no. 2, 263–273. MR2417045 (2009b:57029)
- [Nel11] S. Nelson, *Generalized quandle polynomials*, Canad. Math. Bull. **54** (2011), no. 1, 147–158. MR2797975
- [Nel14] S. Nelson, *Link invariants from finite racks*, Fund. Math. **225** (2014), 243–258. MR3205572
- [NN08] E. A. Navas and S. Nelson, *On symplectic quandles*, Osaka J. Math. **45** (2008), no. 4, 973–985. MR2493966 (2009m:20098)
- [NR14] S. Nelson and V. Rivera, *Quantum enhancements of involutory birack counting invariants*, J. Knot Theory Ramifications **23** (2014), no. 7, 1460006, 15. MR3265399
- [NV06] S. Nelson and J. Vo, *Matrices and finite biquandles*, Homology, Homotopy Appl. **8** (2006), no. 2, 51–73. MR2246021 (2007j:57009)

- [NW11] S. Nelson and R. Wieghard, *Link invariants from finite Coxeter racks*, J. Knot Theory Ramifications **20** (2011), no. 9, 1247–1257. MR2844806
- [NW13] S. Nelson and E. Watterberg, *Birack dynamical cocycles and homomorphism invariants*, J. Algebra Appl. **12** (2013), no. 8, 1350049, 14. MR3092526
- [Pei80] C. S. Peirce, *On the Algebra of Logic*, Amer. J. Math. **3** (1880), no. 1, 15–57. MR1505245
- [Pf00] H. O. Pflugfelder, *Historical notes on loop theory*, Comment. Math. Univ. Carolin. **41** (2000), no. 2, 359–370. Loops'99 (Prague). MR1780877 (2001h:01030)
- [Pf90] H. O. Pflugfelder, *Quasigroups and loops: introduction*, Sigma Series in Pure Mathematics, vol. 7, Heldermann Verlag, Berlin, 1990. MR1125767 (93g:20132)
- [Smi92] J. D. H. Smith, *Quasigroups and quandles*, Discrete Math. **109** (1992), no. 1-3, 277–282. Algebraic graph theory (Leibnitz, 1989). MR1192389 (93k:20104)
- [Sta04] D. Stanovský, *Left distributive left quasigroups*, Ph.D. Thesis, 2004.
- [Ste57] S. K. Stein, *On the foundations of quasigroups*, Trans. Amer. Math. Soc. **85** (1957), 228–256. MR0094404 (20 #922)
- [Tak42] M. Takasaki, *Abstraction of symmetric transformations*, Tohoku Math. J. **49** (1942), no. 145-207, 43.
- [Toy41] K. Toyoda, *On axioms of linear functions*, Proc. Imp. Acad. Tokyo **17** (1941), 221–227. MR0014105 (7,241g)
- [Vla10] J. Vlachy, *Small left distributive quasigroups*, Ph.D. Thesis, 2010.
- [Whi44] H. Whitney, *The self-intersections of a smooth n -manifold in $2n$ -space*, Annals of Mathematics (1944), 220–246.
- [Win84] S. K. Winker, *Quandles, knot invariants, and the n -fold branched cover*, Ph.D. Thesis, 1984.

Index

- \mathbb{Z}_n , 42
- abelian, 56
- Alexander module, 99
- Alexander polynomial, 96, 100, 101
- Alexander quandle, 91, 96
- ambient isotopy, 1
- amphichiral, 4
- augmented quandle, 129, 130
- automorphism, 29
- automorphism group, 93
- basis, 49, 61
- bikei, 159
- braid, 3
- braid group, 113
- braid index, 117
- braid relations, 113
- cardinality, 84
- chiral, 3
- closure of a braid, 117
- coboundary, 67
- cochain complex, 67
- cocycle, 67
- cohomology, 67
- coloring, 103
- column space, 51
- commutative ring, 48, 56
- congruence, 41
- conjugation quandle, 90
- connected quandle, 94
- connected sum, 7
- coordinates, 49
- Coxeter kei, 76
- crossing relation, 78
- crossing sign, 12
- cycle notation, 55
- cyclic kei, 75
- degenerate submodule, 193
- determinant, 100
- detour move, 225
- differential, 67
- dihedral group, 56, 64, 91
- dihedral quandle, 75, 91
- direct product, 59
- direct sum, 43
- dual quandle, 95
- enveloping group, 130
- equivalence class, 38
- equivalence classes, 78
- equivalence relation, 37
- faithful quandle, 95
- field, 47
- finitely generated, 60
- framed isotopy, 6, 14
- framing, 5, 13
- free, 44
- free abelian group, 61

- free group, 63
- free kei, 78
- fundamental group, 62, 110
- fundamental kei, 78
- fundamental quandle, 92, 126

- G. B. Briggs, 10
- Gauss code, 221
- generators, 60
- Grigori Perelman, 107
- group, 54
- group action, 129

- Henri Poincaré, 107
- homomorphism, 29, 57
- homomorphism enhancement, 178
- homotopy, 107
- homotopy class, 109

- ideal, 100
- image, 51, 57
- image enhancement, 175
- infinite cyclic cover, 126
- inner automorphism group, 93
- integers mod n , 42
- involutory quandle, 89
- isomorphism, 29, 57
- isotopic quasigroups, 139

- J. W. Alexander, 10, 101
- Jones polynomial, 19

- kei counting invariant, 84
- kei homomorphism, 81
- kei presentation, 79
- kernel, 51, 57
- kink map, 149
- Kishino knot, 234
- Klein bottle, 215, 217
- Klein group, 140
- knot, 1
- knot complement, 125
- knot diagram, 2, 9
- knot invariant, 17
- knot quandle, 92
- knot type, 1
- knotted surface diagram, 215
- Kurt Reidemeister, 10

- Latin quandle, 94
- Latin square, 138
- Laurent polynomial, 98
- linear transformation, 49
- link, 2
- linking number, 18
- local move, 10
- longitude, 5

- Möbius band, 231
- medial, 143
- medial quandle, 94
- meridian, 5, 125, 126, 152
- module, 49
- module enhancement, 183
- Moufang loop, 141
- Moufang's Theorem, 142

- normal subgroup, 58
- null space, 51

- obverse, 3
- operation, 27
- operation table, 28
- orbit, 94
- oriented knot, 4

- partition, 38
- path connected, 110
- path homotopy, 107
- permutation, 30
- planar isotopy, 10
- point at infinity, 121
- presentation, 46, 62
- pretzel knot, 15
- prime knot, 7
- principal ideal, 100
- projective plane, 215

- quandle, 89
- quandle homomorphism, 92
- quandle polynomial, 188
- quasigroup, 136
- quaternion, 141
- quotient group, 58
- quotient set, 39
- quotient vector space, 44

- rack characteristic, 150
- rack rank, 150
- Ralph Fox, 22

rank, 61
reduced echelon form, 50
Reidemeister moves, 10
relation, 37
reverse, 4
ring, 48

Seifert surface, 126
simple quandle, 95
Smith normal form, 51
solution space, 51
stereographic projection, 120
subgroup, 57
symmetric group, 55
symplectic quandle, 181

Takasaki kei, 75
tangle, 2
tangle coloring, 207
Tietze moves, 79
torsion, 44
transitive action, 140
transvection group, 94
ts-rack, 150

universal algebra, 45

vector space, 47
virtual crossing, 225
virtual knot, 226
virtual trefoil, 227

writhe, 12

From prehistory to the present, knots have been used for purposes both artistic and practical. The modern science of Knot Theory has ramifications for biochemistry and mathematical physics and is a rich source of research projects for undergraduate and graduate students and professionals alike. Quandles are essentially knots translated into algebra.



Courtesy of Jim Hoste

This book provides an accessible introduction to quandle theory for readers with a background in linear algebra. Important concepts from topology and abstract algebra motivated by quandle theory are introduced along the way. With elementary self-contained treatments of topics such as group theory, cohomology, knotted surfaces and more, this book is perfect for a transition course, an upper-division mathematics elective, preparation for research in knot theory, and any reader interested in knots.

ISBN 978-1-4704-2213-4



9 781470 422134

STML/74



For additional information
and updates on this book, visit
www.ams.org/bookpages/stml-74

AMS on the Web
www.ams.org