STUDENT ARTICLE

4

The Knot Quandle

Eleanor Birrell†
Harvard University '09
Cambridge, MA 02138
ebirrell@fas.harvard.edu

Abstract

Mathematicians have been interested in knot theory, or the study of knots, since the early nineteenth century. However, despite this interest, some basic questions remain unanswered; for example there is no effective way to definitively determine whether or not two knots are the same. In this paper, we will look at a powerful but frequently overlooked knot invariant: the knot quandle. We will show that the knot quandle is a generalization of several more familiar invariants and that it is a complete invariant up to orientations. However, as we will see, determining whether two quandles are isomorphic is computationally intractable, which limits the utility of this otherwise powerful invariant.

4.1 Introduction

Knot theory is a subfield of mathematics that can be described simply as the study of knots, or embeddings of S^1 into $\mathbb{R}^3$. Since its beginning in the nineteenth century, knot theory has appealed to mathematicians for a variety of reasons. It contains many interesting theoretical questions related to both algebra and topology and also has applications to biology, physics, and cryptography.

Despite the range of questions and applications that arise in knot theory, some relatively simple questions remain unanswered. For example, how does one tell whether or not two knots are "the same"? We consider two knots to be the same if there exists an **ambient isotopy** between them; i.e. there is a homotopy of self-diffeomorphisms from $\mathbb{R}^3 \to \mathbb{R}^3$ that transforms one knot into the other. This means that two knots are the same if there exists a continuous deformation of space that takes the first knot to the second (this definition is consistent with the intuitive idea of equivalent knots). Kurt Reidemeister showed that two knots are connected by an ambient isotopy if and only if they are connected by a finite chain of Reidemeister moves [Cr], which are defined on knot diagrams, a special type of projection of the knots into $\mathbb{R}^2$ (see Figure 4.2).

Despite the lack of a complete answer, some progress has been made towards determining when knots are equivalent. One important although frequently overlooked step was the development of the knot quandle. In 1982, David Joyce [Jo] and Sergei Matveev [Mat] independently introduced a **knot invariant** (an object that is invariant under different representations of equivalent knots) that would aid in attempts to determine knot isomorphisms. They called this invariant the **knot quandle** and the **distributive groupoid**, respectively. Although this invariant is not perfect (it fails to distinguish between the right and left handed trefoils, for example), it proves to be complete up to orientation. Moreover, this invariant serves as a generalization of both the knot group and colorability, suggesting that the knot quandle is, perhaps, the most complete and the most fundamental invariant known today. Unfortunately, the difficulty inherent in proving that two quandles are isomorphic has severely limited the utility of this otherwise powerful invariant.

†Eleanor Birrell, Harvard '09, is a mathematics concentrator living in Pforzheimer House. She is originally from Los Altos, California where she graduated from Los Altos High School. Her academic interests include algebraic topology and computational complexity.

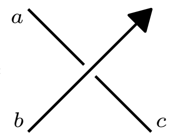
$P_i =$ 

Figure 4.1: A Crossing in a Knot Diagram.

4.2 The Knot Quandle

The knot quandle is defined by associating a quandle structure with a knot K.

Definition 1. A **quandle** is a set S equipped with a binary operation $\triangleright : S \times S \to S$ satisfying the following three axioms:

A1. $x \triangleright x = x$ for all $x \in S$
A2. For all $y, z \in S$, there exists a unique $x \in S$ such that $x \triangleright y = z$
A3. $(x \triangleright y) \triangleright z = (x \triangleright z) \triangleright (y \triangleright z)$ for all $x, y, z \in S$

Note that as the inverse operation $\triangleleft$ defined by $b \triangleleft a = x$ if and only if $x \triangleright a = b$ also fulfills the quandle axioms, a quandle is sometimes defined as a set with *two* binary operations: $\triangleright$ and $\triangleleft$. Distributive properties between the two operations, $(x \triangleright y) \triangleleft z = (x \triangleright z) \triangleleft (y \triangleright z)$ and $(x \triangleleft y) \triangleright z = (x \triangleleft z) \triangleright (y \triangleleft z)$, follow from the quandle axioms. For example, any group G is a quandle with the quandle product $\triangleright$ defined by conjugation, i.e. $a \triangleright b = bab^{-1}$.

The quandle structure can be associated with a knot in two distinct ways. The first method, which we shall refer to as the algebraic definition of the knot quandle, was first introduced by Joyce in [Jo]. In order to give the algebraic definition of the knot quandle, it is important to first understand the concept of a diagram of a knot.

Definition 2. A **diagram** D of a knot K is a projection of K onto a plane such that at most two strands of the knot intersect at any point and such that there are finitely many such intersections. By convention, at each crossing of two strands, one removes a segment of the projected image of the lower strand to convey relative height information. These breaks make the diagram a set of disjoint arcs. The orientation of the knot is also indicated on the diagram. Recall, from the introduction, that two diagrams represent the same knot if and only if they are connected by a finite chain of Reidemeister moves (see Figure 4.2).

Using this definition, it is possible to assign a quandle structure to any diagram of a knot.

Definition 3. Let D be a diagram of an oriented knot K. Let A_D = {arcs of D}. For all crossings P_i (as in Figure 4.1), define the relation $a \triangleright_A b = c$. Note that this definition depends on the orientation of the overcrossing arc but not the undercrossing arc. The **algebraic knot quandle** $\Gamma_A(K, D)$ of a given knot diagram D is then defined as $\langle A_D \mid \triangleright_A \rangle$. That is, it is the quandle with generators A_D subject to the relations $\triangleright_A$.

The algebraic knot quandle of a diagram does in fact prove to be a well-defined quandle. As we will show, it is independent of the chosen knot diagram; that is, it is a knot invariant.

Theorem 4. *The algebraic quandle, $\Gamma_A(K, D)$, is independent of D up to isomorphism.*

Proof. Let D_1, D_2 be two diagrams of a knot K. Since D_1 and D_2 are diagrams of the same knot, they are connected by a finite number of Reidemeister moves.

Consider the effect of each of these moves on the knot quandle (given an arbitrary choice of orientation), as in Figure 4.2. (Note that we show 8 such moves as opposed to the usual three, because we must have that the knot quandle is unchanged by Reidemeister moves regardless of orientation.) For each move, we can verify that the quandles obtained from the diagrams before

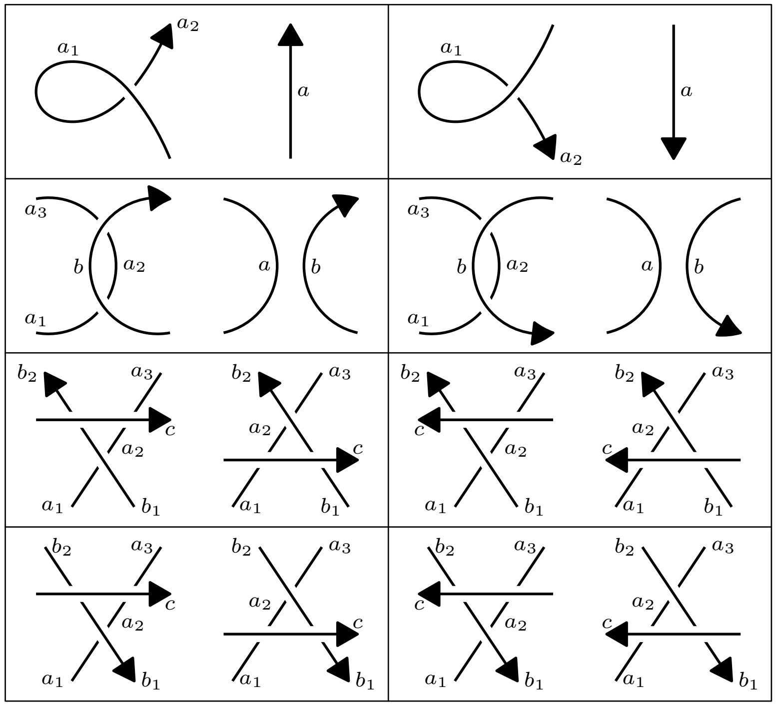


Figure 4.2: Reidemeister Moves.

and after the move are isomorphic. Consider the move depicted in the upper left-hand corner of Figure 4.2. Let Γ the the algebraic quandle corresponding to a knot including a strand a, as in the right-hand side of the diagram illustrating that move. Then Γ is generated by a and some collection of arcs $\{b, c, d, \ldots\}$, and is subject to some collection of relations. Let Γ' be the algebraic quandle corresponding to the same knot after the Reidemeister move has been performed, so that the knot diagram contains strands a_1 and a_2 as depicted on the left-hand side of the illustration. Then Γ' is generated by a_1, a_2 and $\{b, c, d, \ldots\}$, subject to the same relations as Γ (but with a_1 or a_2 substituted for occurrences of a in those relations, as appropriate) plus an additional relation $a_1 \triangleright a_1 = a_2$. But by quandle axiom A1, $a_1 \triangleright a_1 = a_1$. Hence $a_1 = a_2$, so the two quandles are manifestly isomorphic, by mapping $a_1 = a_2$ to a and each of the other generators to itself.

We leave it to the reader to check in a similar manner that the other Reidemeister moves also leave the algebraic quandle unchanged, up to canonical isomorphism. (One finds that each type of Reidemeister move corresponds to one of the quandle axioms.) Since D_1 and D_2 are related by a finite sequence of Reidemeister moves, it follows that $\Gamma_A(K, D_1) \cong \Gamma_A(K, D_2)$, as desired. $\square$

There is a second way of associating a quandle structure with a knot. This geometric definition was proposed by Matveev in his paper on distributive groupoids [Mat].

Definition 5. Let K be an oriented knot in $\mathbb{R}^3$. Let $N(K)$ be a small tubular neighborhood of K. Let $E(K) = \overline{\mathbb{R}^3 \setminus N(K)}$. Fix a base point $x \in E(K)$. Let $B_K = \{$homotopy* classes of paths in $E(K)$ with fixed initial point x and an endpoint on $\partial N(K)\}$. Here homotopy* means a homotopy which keeps the base point x fixed, such that the trajectory of the other endpoint is contained in $\partial N(K)$. The reader can check that being homotopic* is in fact an equivalence relation, which we henceforth denote by $\backsimeq$. Given a path b from x to some point y on $\partial N(K)$, let γ be the

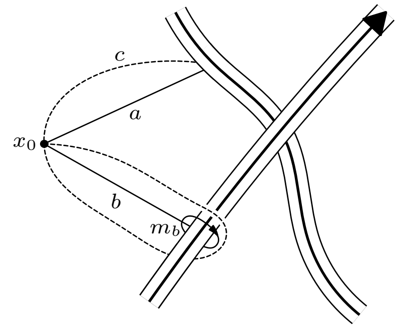


Figure 4.3: The loop c is a representative of $[a] \triangleright [b]$.

shortest line segment from y to the knot, and let the **oriented meridian** m_b be the loop in $\partial N(K)$ based at y, in the plane containing γ and perpendicular to the knot at $\gamma \cap K$.[1] Define an operation $\triangleright : B_K \times B_K \to B_K$ by $[a] \triangleright [b] \mapsto [bm_b b^{-1} a]$, as in Figure 4.3, where b^{-1} denotes the path b traversed backwards, and concatenation of paths denotes the standard composition law for paths from homotopy theory. The **geometric knot quandle** is defined as $\Gamma_B(K, x) = \langle B_D \mid \triangleright_B \rangle$. That is, it is the quandle generated by B_D with relations $\triangleright_B$.

The knot quandle, as defined by Matveev, satisfies the quandle axioms given in Definition 1. It also proves to be a knot invariant.

Theorem 6. *The geometric definition of the knot quandle is independent of the chosen base point, up to isomorphism.*

Proof. Choose two base points $x_1, x_2 \in \mathbb{R}^3 \setminus K$, for a fixed knot K. Let γ be a fixed path from x_2 to x_1; such a path exists because the complement of a knot is path-connected. Define $\phi : \Gamma_B(K, x_2) \to \Gamma_B(K, x_1)$ by $[\beta] \mapsto [\gamma\beta]$. There is a homotopy* $\beta_i \simeq \beta_j$ if and only if there is a homotopy* $\gamma\beta_i \simeq \gamma\beta_j$, so ϕ is well-defined. By the same token, there is also a well-defined map $\phi^{-1} : \Gamma_B(K, x_1) \to \Gamma_B(K, x_2)$ given by $\phi^{-1}[\beta] = [\gamma^{-1}\beta]$; it is easy to see that ϕ^{-1} inverts ϕ. If m_j denotes an oriented meridian of $N(K)$ at the endpoint of β_j, we can compute as follows, using the definition of $\triangleright$ in the geometric knot quandle:

$$\begin{aligned}
\phi[\beta_i] \triangleright \phi[\beta_j] &= [\gamma\beta_i] \triangleright [\gamma\beta_j] \\
&= [\gamma\beta_j m_j (\gamma\beta_j)^{-1} \gamma\beta_i] \\
&= [\gamma\beta_j m_j \beta_j^{-1} \beta_i] \\
&= \phi[\beta_j m_j \beta_j^{-1} \beta_i] \\
&= \phi([\beta_i] \triangleright [\beta_j]).
\end{aligned}$$

Therefore ϕ is a quandle homomorphism, and by the same reasoning so is ϕ^{-1}. Consequently ϕ is a quandle isomorphism. (Note that this isomorphism is not unique; it depends on the path chosen between x_1 and x_2.) □

There are therefore two distinct ways to associate a quandle structure with a knot. Both definitions give rise to a knot invariant (up to isomorphism). In fact, they give rise to the same invariant: the knot quandle.

[1]While it will not be important for our purposes, it is worth pointing out that the meridians m_b must be chosen with coherent orientations. One way to do this is to require that relative to a fixed orientation of the knot K, the meridians are oriented using the "right hand rule."

Theorem 7. *The algebraic and geometric definitions of the knot quandle are equivalent up to orientation.*

Sketch of proof. Let $\Gamma_A(K)$ be the algebraic knot quandle for a diagram D of K, and let $\Gamma_B(K)$ be the geometric knot quandle. Define a map $\phi : \Gamma_A(K) \to \Gamma_B(K)$ by mapping an arc $a \in \Gamma_A(K)$ to a homotopy class $[s_a]$ in $\Gamma_B(K)$ such that the following conditions hold: (1) The path s_a connects the base point x_0 to a point on the section of $\partial N(K)$ whose distance to the strand of the knot that projects to arc a is minimized; (2) at all points where the projection of s_a intersects D, the path s_a is *above* the strands of the knot it crosses, as in Figure 4.4.

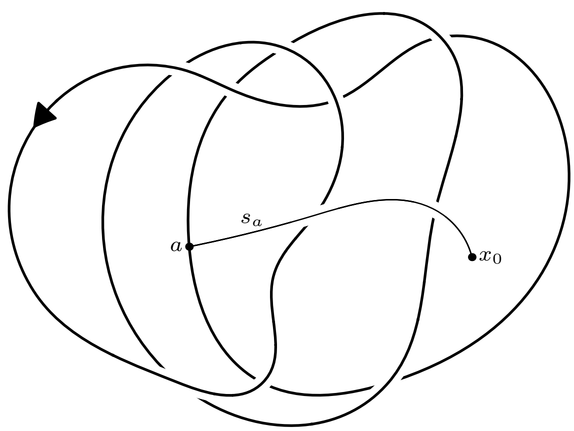


Figure 4.4: Choosing the Path s_a.

Having defined ϕ on the generators of $\Gamma_A(K)$, we extend it to a map on the quandle. It is possible to show that in fact ϕ is a quandle homomorphism.

Define $\psi : \Gamma_B(K) \to \Gamma_A(K)$ as follows. Given $[s] \in \Gamma_B(K)$ choose a representative s such that its projection onto D intersects the diagram in a finite number of points. Let $\{a_1, ..., a_n\}$ be the set of arcs in $\Gamma_A(K)$ that are above s in the diagram D, and let a_0 be the arc containing the endpoint of $[s]$; see Figure 4.5. Set

$$\psi([s]) = ((...((a_0\epsilon_1 a_1)\epsilon_2 a_2)\epsilon_3 ...)\epsilon_{n-1} a_{n-1})\epsilon_n a_n,$$

where ϵ_i is $\triangleright$ if the crossing between s and a_i is positive in D and ϵ_i is $\triangleleft$ if the crossing is negative. The sign of a crossing is defined as in Figure 4.6.

It is possible to show that ψ also defines a quandle homomorphism, and that ϕ and ψ are mutually inverse. So ϕ is an isomorphism between $\Gamma_A(K)$ and $\Gamma_B(K)$, so $\Gamma_A(K) \cong \Gamma_B(K)$, as desired. □

In order to gain a better understanding of the knot quandle, let us consider the example of the knot 5_2.

Example. Deriving the knot quandle $\Gamma(5_2)$.

Consider the standard diagram of the oriented knot 5_2 (see Figure 4.7a). We assign a generator to each of the arcs of the diagram. Using the algebraic definition of the knot quandle, we assign the relation $x_i \triangleright x_j = x_k$ to each of the crossings. This gives us $\Gamma_A(5_2) = \langle a, b, c, d, e \mid d \triangleright a = e, b \triangleright d = a, a \triangleright b = c, c \triangleright e = d, e \triangleright c = b\rangle$.

Now consider $E(K) = \overline{S^3 \setminus N(K)}$ for a small tubular neighborhood $N(K)$ of K. Fix a base point $x_0 \in E(K)$. The generators of the quandle $\Gamma_B(5_2)$ are the set of equivalence classes of

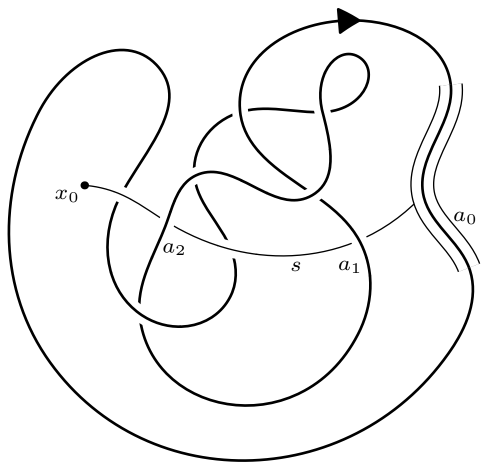


Figure 4.5: Choosing the Representative s.

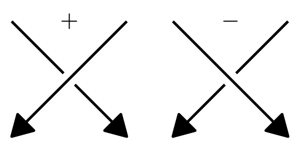


Figure 4.6: The Sign of a Crossing.

paths in $E(K)$, as shown in Figure 4.7b. Equivalently, the generators are the set of paths that end at a distinct arc of the diagram and go over all other arcs by the definition of ϕ in the proof of Theorem 7. The relations of the geometric quandle are given by $[x] \triangleright [y] = [y m_y y^{-1} x]$; therefore, the geometric definition generates the quandle

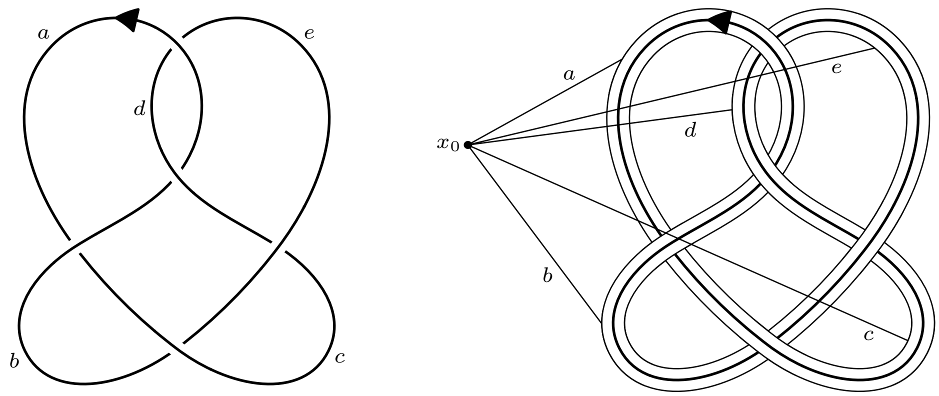
$$\Gamma_B(5_2) = \langle [a], [b], [c], [d], [e] \mid [d] \triangleright [a] = [a m_a a^{-1} d] = [e], [b] \triangleright [d] = [d m_d d^{-1} b] = [a],$$
$$[a] \triangleright [b] = [b m_b b^{-1} a] = [c], [c] \triangleright [e] = [e m_e e^{-1} c] = [d], [e] \triangleright [c] = [c m_c c^{-1} e] = [b] \rangle.$$

The equalities of the form $[a m_a a^{-1} d] = [e]$ are all geometrically self-evident; given a parameterization of the knot and its tubular neighborhood, it would be straightforward to write down the corresponding homotopies.

The quandles $\Gamma_A(5_2)$ and $\Gamma_B(5_2)$ are trivially isomorphic, so in the case of 5_2 we can see explicitly that the algebraic and geometric quandles are isomorphic, as required by Theorem 7.

4.3 A Complete Invariant

Having established that the knot quandle is actually a well-defined knot invariant, we are left to wonder how useful this new invariant is. In order to answer this, we must look at two independent questions. First, how good is the knot quandle at distinguishing between knots? Second, how easy is it to show whether or not two knots have the same knot quandle? The answer to the first question turns out to be that the knot quandle is extremely good at distinguishing between knots. In fact, it is a complete invariant up to orientation.



(a) (b)

Figure 4.7: (a) Constructing $\Gamma_A(5_2)$; (b) Constructing $\Gamma_B(5_2)$.

In order to prove this, we need to first establish some definitions and background theorems.

Definition 8. A 2-surface F in a 3-manifold M is **compressible** if one of the following conditions holds:

1. F is an embedding of a 2-sphere and bounds the embedding of a 3-ball in M,
2. F is the embedding of a disk in ∂M,
3. F is the embedding of a disc in M and there is an embedded 3-ball in M whose boundary is contained in $F \cup \partial M$,
4. F is not the embedding of a 2-sphere or a disc and there exists an embedded disk $\Delta \subset M$ such that $\Delta \cap F = \partial\Delta$ and such that $\partial\Delta$ is not **nullhomotopic** in F (that is, it is a nontrivial element of the fundamental group $\pi_1(F)$).

Definition 9. If every embedding of a 2-sphere in M is compressible, then M is **irreducible**.

Definition 10. A 3-manifold with boundary M is **boundary irreducible** if its boundary ∂M is not compressible.

Definition 11. A **handlebody** in $\mathbb{R}^n$ is a closed, regular neighborhood of a finite graph in $\mathbb{R}^n$.

Definition 12. A manifold M is called **sufficiently large** if one can embed a handlebody $H \neq S^2$ in M such that the map of fundamental groups induced by the inclusion $H \hookrightarrow M$ is injective.

The proof of the completeness of the knot quandle also relies on four important theorems.

Theorem 13 (Dehn's Lemma)**.** *Let M be a 3-manifold with boundary and let γ be a closed curve on its boundary ∂M. Then if there exists an immersed 2-disc $D \to M$ such that $\partial D = \gamma$, then there exists an embedded disk $D' \subset M$ with the same boundary $\partial D' = \gamma$.*

This theorem was first proven in 1910 by Max Dehn, a German mathematician, but his proof was later discovered to contain holes. It was finally rigorously proven by Christos Papakyriakopoulos in 1956. One of the important consequences of Dehn's lemma (from a knot theoretic point of view) is that it can be used to prove Dehn's Theorem.

Theorem 14 (Dehn's Theorem). *A knot K is the unknot if and only if $\pi_1(\mathbb{R}^3 \setminus K)$ is isomorphic to $\mathbb{Z}$.*

Proof. One direction of Dehn's theorem is not too bad; using the Wirtinger presentation defined in Definition 26 (or see [GP]), the fundamental group of the complement of the unknot has one generator and no relations, and is therefore isomorphic to $\mathbb{Z}$.

The other direction follows from Dehn's lemma. Let K be a knot such that $\pi(\mathbb{R}^3 \setminus K) = \mathbb{Z} = \langle x \rangle$ where x is a loop on $\partial N(K)$. Let μ be a meridian of $N(K)$ and let λ be a **longitude** (a loop in $N(K)$ not in a homotopy class generated by a meridian). The fundamental group $\pi_1(\mathbb{R}^3 \setminus K)$ clearly contains $\langle \mu \rangle = \mathbb{Z}$, so the curve λ must be homotopic to a constant, which implies that there is an immersed disc Δ with $\partial\Delta = \lambda$. By Dehn's lemma, there is a disc embedded in $\mathbb{R}^3 \setminus N(K)$ bounded by λ. By contracting $N(K)$ to K along the radial chords, we obtain a disc embedded in $\mathbb{R}^3 \setminus K$ bounded by K, so K is the unknot. □

Definition 15. Given two knots K_1, K_2 with associated meridians and longitudes m_1, m_2, l_1, l_2 we say a homomorphism $\phi : \pi_1(\mathbb{R}^3 \setminus K_1) \to \pi_1(\mathbb{R}^3 \setminus K_2)$ **preserves peripheral structure** if the image of $\langle [m_1], [l_1] \rangle \subset \pi_1(\mathbb{R}^3 \setminus K_1)$ through ϕ is conjugate to a subgroup of $\langle [m_2], [l_2] \rangle \subset \pi_1(\mathbb{R}^3 \setminus K_2)$. We say that $\langle [m_1], [l_1] \rangle \subset \pi_1(\mathbb{R}^3 \setminus K_1)$ is the **peripheral structure** of the knot K_1, and similarly with K_2; note that we may also view this as simply the normalizer of $[m_1]$, i.e. $N([m_1])$. This definition can be generalized to maps between manifolds—the Waldhausen theorem below holds for this more general definition.

Theorem 16 (Waldhausen Theorem). *Let M, N be irreducible, boundary-irreducible 3-manifolds. Let M be sufficiently large and let $\phi : \pi_1(N) \to \pi_1(M)$ be an isomorphism preserving peripheral structure. Then there exits a homeomorphism $f : N \to M$ that induces ϕ.*

The final theorem was first suggested in 1908, although it remained unproven until 1989.

Theorem 17 (Gordon-Luecke Theorem). *If K_1 and K_2 are unoriented knots in S^3 and there is an orientation preserving homeomorphism between their complements, then K_1 and K_2 are equivalent as unoriented knots.*

Now, for convenience, we will establish some useful lemmas, some of whose proofs we omit, before we proceed to the proof that the knot quandle is a complete invariant.

Lemma 18. *Let K be a non-trivial knot. Let $N(K)$ be a small tubular neighborhood of K, and let $E(K)$ be the closure of $S^3 \setminus N(K)$. Then $E(K)$ is irreducible, boundary irreducible, and sufficiently large.*

Lemma 19. *If K is non-trivial, then there is an injective homomorphism from the fundamental group $\pi_1(\partial N)$ into $\pi_1(\mathbb{R}^3 \setminus K)$.*

Lemma 20. *The peripheral structure of a knot is determined by the knot quandle.*

Proof. The group $\pi_1(\mathbb{R}^3 \setminus K)$ is determined by the knot quandle, as is shown in Section 4.5. We will show that the homotopy class of a meridian of $N(K)$ can be constructed from elements of the geometric knot quandle. Choose any $[x], [y] \in \Gamma(K)$, not necessarily distinct; we can do this because the geometric quandle is nonempty by definition. Now computing in the fundamental groupoid of $E(K)$, define a homotopy class by $[x^{-1}]([y] \triangleright [x])[y^{-1}][x]$. Using the definition of $\triangleright$ in the geometric quandle, we see that this is equal to the homotopy class $[m_x]$ of the meridian m_x associated to $[x]$. Since m_x is a loop, this homotopy class is an element of the fundamental group $\pi_1(\mathbb{R}^3 \setminus K)$. As the knot quandle determines the fundamental group of the knot complement (see Section 4.5) and the homotopy class of a meridian, it determines the normalizer of $[m_x]$ in $\pi_1(\mathbb{R}^3 \setminus K)$, which is precisely the peripheral structure. □

Using this background, it is possible to prove that the knot quandle is a complete invariant up to orientation.

Theorem 21. *The knot quandle $\Gamma(K)$ is a complete invariant up to orientation.*

Proof. Let K_1, K_2 be two knots such that there exists an isomorphism $\psi : \Gamma(K_1) \to \Gamma(K_2)$.

- If K_1 is trivial: The fundamental group of the complement, $\pi_1(\mathbb{R}^3 \setminus K)$ can be derived from the knot quandle $\Gamma(K)$. Therefore, since $\Gamma(K_1) \cong \Gamma(K_2)$, we know that $\pi_1(\mathbb{R}^3 \setminus K_1) \cong \pi_1(\mathbb{R}^3 \setminus K_2)$. The knot K_1 is trivial, so by Dehn's theorem, $\pi_1(\mathbb{R}^3 \setminus K_1) \cong \mathbb{Z}$, which implies that $\pi_1(\mathbb{R}^3 \setminus K_2) \cong \mathbb{Z}$. So by Dehn's theorem K_2 is also trivial.
- If K_1, K_2 are nontrivial: Since $\Gamma(K_1) \cong \Gamma(K_2)$ and the quandle determines the peripheral structure (Lemma 20), K_1 and K_2 have the same peripheral structure. Therefore the isomorphism $\phi : \pi_1(\mathbb{R}^3 \setminus K_1) \to \pi_1(\mathbb{R}^3 \setminus K_2)$ induced by the isomorphism $\psi : \Gamma(K_1) \to \Gamma(K_2)$ preserves peripheral structure. Also, E_1, E_2 are boundary-irreducible, sufficiently large, irreducible 3-manifolds (Lemma 18).

  Therefore, by the Waldhausen theorem, there exists a homeomorphism $f : E(K_1) \to E(K_2)$ that induces ϕ. It follows immediately that there is a homeomorphism between $\mathbb{R}^3 \setminus K_1$ and $\mathbb{R}^3 \setminus K_2$ to which we may apply the Gordon-Luecke theorem, to conclude that K_1 and K_2 are equivalent up to orientation. □

It is again important to note that the knot quandle is not a truly complete invariant; it is only complete up to the orientations of the knot and the ambient space.

Example. Consider both the right- and left-handed trefoils (see Figure 4.8). We know that these are not equivalent knots because their signatures (an invariant defined in [Cr]) are different ($\sigma(3_1) = -2$ and $\sigma(3_1^*) = 2$). However, these two knots have isomorphic knot quandles.

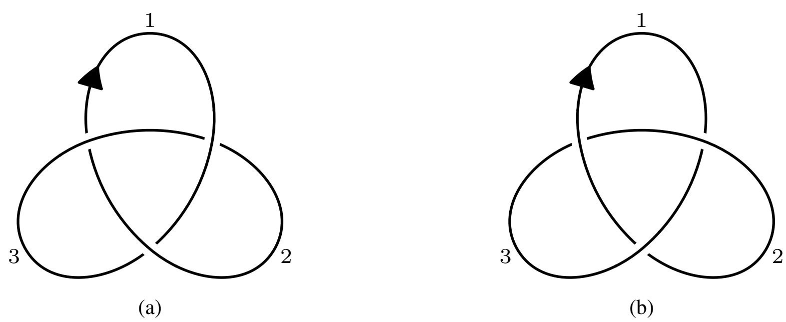


Figure 4.8: (a) The right-handed trefoil knot 3_1; (b) the left-handed trefoil knot 3_1^*.

Using the algebraic definition of the knot quandle, we get that:

$$\Gamma(3_1) = \langle 1, 2, 3 \mid 1 \triangleright 3 = 2, 3 \triangleright 2 = 1, 2 \triangleright 1 = 3 \rangle.$$

Furthermore:

$$\Gamma(3_1^*) = \langle 1, 2, 3 \mid 2 \triangleright 3 = 1, 1 \triangleright 2 = 3, 3 \triangleright 1 = 2 \rangle.$$

Using the quandle axioms, it is easy to see that these give rise to the same quandle. Consider $\Gamma(3_1)$. The relation $1 \triangleright 2$ is not explicitly defined, but $2 \triangleright 2 = 2$ by A1, $3 \triangleright 2 = 1$, and there exists a unique x such that $x \triangleright y = z$ for any y, z (by A2), so we must have $1 \triangleright 2 = 3$. Similarly, we find that $2 \triangleright 3 = 1$ and $3 \triangleright 1 = 2$, and hence $\Gamma(3_1) = \Gamma(3_1^*)$ despite the face that 3_1 and 3_1^* are not equivalent knots.

Nonetheless, the knot quandle is an extremely powerful invariant. However, in order to be computationally useful, we would have to be able to efficiently determine whether or not two knots give rise to isomorphic knot quandles.

4.4 Computational Complexity

Unfortunately, despite the fact that the knot quandle is a complete invariant up to orientation, it does not turn out to be easy to use. The primary problem with the knot quandle is that it is difficult to determine whether or not two quandles are isomorphic. One approach to this problem has been suggested by Sam Nelson and Benita Ho [HN].

Definition 22. The **quandle matrix** associated with a finite quandle Q with n elements, M_Q, is the $n \times n$ matrix whose (i,j)-th entry is given by $x_i \triangleright x_j$,

$$M_Q = \begin{pmatrix} x_1 \triangleright x_1 & \dots & x_1 \triangleright x_n \\ \vdots & \ddots & \vdots \\ x_n \triangleright x_1 & \dots & x_n \triangleright x_n \end{pmatrix}.$$

We can define an equivalence relation on quandle matrices.

Definition 23. Let $\rho \in S_n$ be a permutation of $\{1, ..., n\}$. Set $\rho(M_Q) = A_\rho^{-1}(\rho a_{ij})A_\rho$ where A_ρ is the permutation matrix of ρ and ρa_{ij} is the image of the (i,j)-th entry of M_Q under the permutation ρ, acting on the elements $\{x_1, \dots, x_n\}$ of Q in the natural way. Then we say $\rho(M_Q)$ is **permutation equivalent** or **p-equivalent** to M_Q.

This equivalence allows us to determine whether or not two knot quandles are isomorphic.

Theorem 24. *Two $n \times n$ quandle matrices determine isomorphic quandles if and only if they are p-equivalent by a permutation $\rho \in S_n$.*

Proof. Let $Q = \langle x_1, \dots, x_n \mid \triangleright \rangle$ and $Q' = \langle y_1, \dots, y_n \mid \triangleright \rangle$ be finite quandles and let $M_Q, M_{Q'}$ be their respective matrices. Let $\psi : Q \to Q'$ be an isomorphism of finite quandles. We have that ψ, considered on subscripts, induces a bijection $\rho : \{1, ..., n\} \to \{1, ..., n\}$, so $\rho \in S_n$. The fact that ψ is an isomorphism gives us that $\psi(x_i \triangleright x_j) = \psi(x_i) \triangleright \psi(x_j) = x_{\rho(i)} \triangleright x_{\rho(j)}$, so we obtain a permutation of the quandle matrix for Q' by applying $\rho \in S_n$ to every element in M_Q. Conjugation by the permutation matrix of ρ puts the matrix back in standard form, yielding the equality $\rho(M_Q) = A_\rho^{-1}(\rho(a'_{ij}))A_\rho$ where a'_{ij} is an element in the matrix $M_{Q'}$, as desired.

The argument for the reverse implication is essentially identical. □

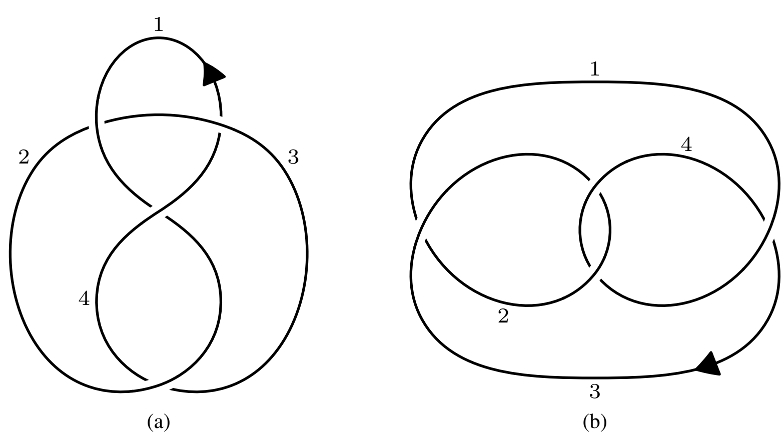


Figure 4.9: Two presentations of the knot 4_1: (a) the knot D_1; (b) the knot D_2.

Example. Consider the two representations of 4_1 in Figure 4.9. Using the algebraic definition of the knot quandle, we derive the following two quandles: $\Gamma(D_1) = \langle 1,2,3,4 \mid 3 \triangleright 1 = 2, 1 \triangleright 4 = 2, 3 \triangleright 2 = 4, 1 \triangleright 3 = 4\rangle$ and $\Gamma(D_2) = \langle 1,2,3,4 \mid 1 \triangleright 3 = 2, 1 \triangleright 2 = 4, 3 \triangleright 4 = 2, 3 \triangleright 1 = 4\rangle$.

We have the two quandle matrices

$$M_1 = \begin{pmatrix} 1&3&4&2 \\ 4&2&1&3 \\ 2&4&3&1 \\ 3&1&2&4 \end{pmatrix},\ M_2 = \begin{pmatrix} 1&4&2&3 \\ 3&2&4&1 \\ 4&1&3&2 \\ 2&3&1&4 \end{pmatrix}.$$

Define $\rho = (1)(2)(34) \in S_4$. Then

$$\begin{aligned} A_\rho^{-1}(\rho(M_1))A_\rho &= \begin{pmatrix} 1&0&0&0 \\ 0&1&0&0 \\ 0&0&0&1 \\ 0&0&1&0 \end{pmatrix}^{-1} \rho\left(\begin{pmatrix} 1&3&4&2 \\ 4&2&1&3 \\ 2&4&3&1 \\ 3&1&2&4 \end{pmatrix}\right)\begin{pmatrix} 1&0&0&0 \\ 0&1&0&0 \\ 0&0&0&1 \\ 0&0&1&0 \end{pmatrix} \\ &= \begin{pmatrix} 1&0&0&0 \\ 0&1&0&0 \\ 0&0&0&1 \\ 0&0&1&0 \end{pmatrix}\begin{pmatrix} 1&4&3&2 \\ 3&2&1&4 \\ 2&3&4&1 \\ 4&1&2&3 \end{pmatrix}\begin{pmatrix} 1&0&0&0 \\ 0&1&0&0 \\ 0&0&0&1 \\ 0&0&1&0 \end{pmatrix} = \begin{pmatrix} 1&4&2&3 \\ 3&2&4&1 \\ 4&1&3&2 \\ 2&3&1&4 \end{pmatrix} = M_2. \end{aligned}$$

So by Theorem 24, $\Gamma(D_1) \cong \Gamma(D_2)$, as M_1 is p-equivalent to M_2. As a result, $K(D_1) = K(D_2)$ up to orientations; they are both diagrams of the knot 4_1. (Note that the ambient isotopy between the two representations is easy to visualize explicitly: lift strand 3 in Figure 4.9(b) and pull it up to the top of the picture.)

This method is reasonably effective; it allows us to determine whether two knots are equivalent if they generate small, finite, easily understood quandles. The problem is that determining whether two matrices are permutation-equivalent is thought to be a computationally intractable problem.

There is no currently known algorithm for determining whether or not two matrices are permutation equivalent in polynomial time. Therefore, as the size of the knot quandle increases, determining whether two knot quandles are isomorphic becomes difficult. Since all small knots have already been completely classified, this renders the knot quandle relatively useless as a computational tool.

The difficulty in actually making use of the quandle, despite the fact that it is a complete invariant up to orientation, has caused mathematicians to resort to less complete, but more useful, invariants that can be derived from the knot quandle.

4.5 The Knot Group

The concept of the fundamental group gives rise to the first such knot invariant: the knot group.

Definition 25. The **knot group** of a knot K is defined as the fundamental group of its complement, $\pi_1(\mathbb{R}^3 \setminus K)$. For convenience, however, we shall denote the knot group by $\pi_1(K)$.

Just as there are two definitions for the knot quandle, there are two equivalent definitions for the knot group.

Definition 26. Let K be a knot. Let D be an oriented diagram of K. Let A_D be the set of arcs of D. For each crossing P_i as in Figure 4.1, define the relation R_i by $c = bab^{-1}$. The **knot group** of K with respect to D is the group $\langle A_D \mid R_i\rangle$. This representation of the knot group is called the **Wirtinger presentation**.

To prove that the Wirtinger presentation of a knot is independent of the choice of diagram, one can simply check that it is invariant under Reidemeister moves. The proof that the Wirtinger presentation gives rise to the same group as the fundamental group of the complement of a knot is more complex and is omitted.

Using the Wirtinger presentation, we can show that the knot quandle determines the knot group up to isomorphism. In this manner, the knot quandle gives rise to an invariant that is easier to calculate than the quandle itself, but which is unfortunately less exact; while the knot quandle is a complete invariant up to orientation, the knot group is not.

Theorem 27. *The knot group $\pi_1(K)$ is determined by the knot quandle $\Gamma(K)$.*

Proof. Let K be a knot. Choose a diagram D of K. We will define an association ψ from knot quandles to knot groups. Let $\Gamma(K) = \Gamma_A(K, D) = \langle A_D \mid R_i \rangle$. Define a group $\psi(\Gamma(K)) = G = \langle A_D \mid S_i \rangle$ where S_i is defined as follows: if R_i is the relation $a \triangleright b = c$, then S_i is the relation $bab^{-1} = c$. By definition, G is the Wirtinger presentation of the knot group, so the knot group can be derived from the knot quandle. □

This method of determining the knot group given a knot quandle can be more easily seen by working through the example of the knot 5_2.

Example. Consider the knot 5_2, as in Figure 4.7a. As seen in Section 4.2,

$$\Gamma(5_2) = \langle a, b, c, d, e \mid d \triangleright a = e, b \triangleright d = a, a \triangleright b = c, c \triangleright e = d, e \triangleright c = b \rangle.$$

The group $\psi(\Gamma(5_2))$ determined from $\Gamma(5_2)$ as in the proof of Theorem 27 is $\langle a, b, c, d, e \mid ada^{-1} = e, dbd^{-1} = a, bab^{-1} = c, ece^{-1} = d, cec^{-1} = b \rangle$. This is exactly the knot group $\pi_1(5_2)$, as can easily be checked using Wirtinger presentation.

However, because the information about the peripheral structure is not retained by the map ψ from the knot quandle to the knot group, the latter is not a complete invariant.

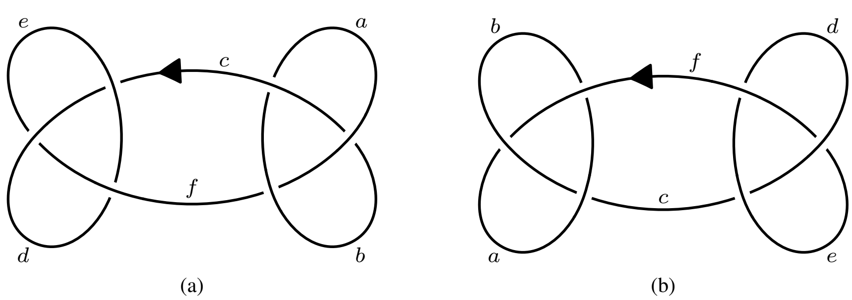


Figure 4.10: Two inequivalent knots with the same knot group: (a) $3_1\#3_1$, and (b) $3_1\#3_1^*$.

Example. Consider the diagrams of the knots $3_1\#3_1$ and $3_1\#3_1^*$, in Figure 4.10.
Using Wirtinger presentations, we calculate that:

$$\pi_1(3_1\#3_1) = \langle a, b, c, d, e, f \mid aba^{-1} = c, faf^{-1} = b, bfb^{-1} = a, cdc^{-1} = e, ece^{-1} = e, ded^{-1} = f \rangle,$$
$$\pi_1(3_1\#3_1^*) = \langle a, b, c, d, e, f \mid aba^{-1} = c, faf^{-1} = b, bfb^{-1} = a, ede^{-1} = c, dfd^{-1} = e, fef^{-1} = d \rangle.$$

These knot groups simplify to:

$$\pi_1(3_1\#3_1) = \langle b, c, d \mid bcb = cbc, cdc = dcd \rangle,$$
$$\pi_1(3_1\#3_1^*) = \langle a, e, f \mid afa = faf, efe = fef \rangle.$$

These knot groups are clearly isomorphic, despite the fact that the knots $3_1\#3_1$ and $3_1\#3_1^*$ are inequivalent. Therefore, as claimed, the knot group (unlike the knot quandle from which it is derived) is not a complete invariant up to orientation.

4.6 Colorability

The second useful invariant that can be derived from the knot quandle is colorability.

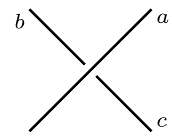


Figure 4.11: Each crossing in a 3-coloring must satisfy $2a - b - c \equiv 0 \bmod 3$.

Definition 28. A diagram, D, of a knot K is **3-colorable** if its arcs can be labeled with elements of the color set $\{0, 1, 2\}$ such that at each crossing P_i we have the relation $2a - b - c \equiv 0 \bmod 3$, as in Figure 4.11, and at least two distinct colors are used across the entire knot.

Note that because we are dealing with only three colors, we can simplify out understanding of 3-colorability by saying that a knot diagram D is 3-colorable if at each crossing P_i either the arcs a, b, c are all the same color or they are all different colors and such that at least two distinct colors are used across the entire knot. As you can check, these two definitions are equivalent.

Theorem 29. *3-colorability is a knot invariant.*

Proof. Consider each of the Reidemeister moves in Figure 4.2. With appropriate relabeling (in the case of the second move), you can check that none of these moves change the 3-colorability of the diagram. So 3-colorability is a knot invariant, as claimed. □

It turns out that it is possible to derive the 3-colorability of a knot K from the knot group $\pi_1(K)$.

Theorem 30. *A knot K is 3-colorable if and only if there exists a surjective homomorphism $\phi : \pi_1(K) \to D_3$ from the knot group to the dihedral group on three elements.*

We omit the proof, as it is technical and has little to do with the knot quandle.

This method of deriving the colorability of a knot from the knot group can be taken one step further to give us a way to derive the colorability directly from the knot quandle.

3-colorability is, however, a very weak invariant. For example, it fails to distinguish between the trefoil and the cinquefoil 5_1 (neither of them is 3-colorable). However, 3-colorability can be generalized to p-colorability for any prime p. This defines a labeling of the arcs of the diagram by the elements of $\{0, ..., p-1\}$. The crossing relations are defined again by $2x - y - z = 0 \bmod p$. As in the case of 3-coloring, a knot is p-colorable if and only if there exists a surjective homomorphism between the knot quandle and the dihedral group D_p. Therefore, by combining colorability by p colors for all prime p, one defines a much stronger invariant that is, in some sense, determined by the knot quandle (or the knot group).

4.7 Conclusion

A quandle structure can be naturally associated with any knot. This gives rise to the knot quandle, an invariant that is complete up to orientation. Although it is unable to distinguish between the right- and left-handed trefoils, the knot quandle can distinguish between all knots that are not related by a change in orientations, including mutants (for example $3_1\#3_1$ and $3_1\#3_1^*$).

Although the knot quandle is extremely powerful, it is not computationally useful because it is difficult to determine when two quandles are isomorphic. Therefore, mathematicians spend more time considering the weaker but more useful invariants that can be derived from the knot quandle. Two of these invariants, the knot group and p-colorability, have already been discussed. Other invariants can also be derived from the knot quandle, including the Alexander and Conway polynomials.

Despite its limited usefulness, the knot quandle proves to be a interesting invariant. Not only does it provide an almost complete invariant, but it also serves as a generalization for a collection of more familiar (and more useful) knot invariants.

Acknowledgments

I am very grateful to Professor Elizabeth Denne, under whose guidance this paper was written. I would also like to acknowledge the aid and advice of Zachary Abel, Scott Kominers, Sam Lichtenstein, Daniel Litt, Alison Miller, and Charles Nathanson. I am especially indebted to Daniel Litt for his invaluable assistance throughout the editorial process.

References

[Cr] Peter Cromwell: *Knots and Links*. Cambridge, UK: Cambridge University Press, 2004.

[GP] N.D. Gilbert and T. Porter: *Knots and Surfaces*. New York: Oxford University Press, 1994.

[HN] Benita Ho and Sam Nelson: Matrices and finite quandles, *Homology, Homotopy and Applications* **7**#1 (2005), 197–208.

[Jo] David Joyce: A classifying invariant of knots, the knot quandle, *Journal of Pure and Applied Algebra* **23**#1 (1982), 37–65.

[Lic] W.B. Raymond Lickorish: *An Introduction to Knot Theory*. New York: Spring-Verlag, 1997.

[Liv] Charles Livingston: *Knot Theory*. Washington D.C.: Mathematical Association of America, 1993.

[Man] Vassily Manturov: *Knot Theory*. Boca Raton, Florida: CRC Press, 2004.

[Mat] Sergei Matveev: Distributive Groupoids In Knot Theory, *Math. USSR Sb*. **47**#1 (1984), 73-83.

[Mu] James R. Munkres: *Topology*. Upper Saddle River, NJ: Prentice Hall, 2000.