Knots and Numbers:
18.095, January 5, 2022
Haynes Miller

A little history.

It all began with smoke rings. You will recall that Gandalf could make smoke rings chase each other around and end up clustered above his head. Long before Tolkien, in 1858 to be precise, the renowned German mathematical physicist Hermann von Helmholtz (1821–1894) worked out the theory of "vortex rings," and explained why they persisted in time. The smoke rotates around a circle in the core of the ring. Helmholtz showed that this dynamical system offers great resistance to getting broken.

This paper was read by the young Scottish physicist Peter Guthrie Tait (1831–1901). He was a school chum of James Clark Maxwell (and got a professorship at Edinburgh in 1859 in preference to Maxwell). In January, 1867, he succeeded in building an experimental apparatus that demonstrated Helmholtz's claims about the persistence of vortex rings. They rotate through each other, as Gandalf's did, one becoming small and squeezing through the other, only to grow and have the other squeeze through it, leapfrog style. They bounced off each other and vibrated. They acted like particles.

Tait was no slouch. He had just finished writing a textbook with William Thomson. Commoners were obliged to call Thomson Lord Kelvin now, because in 1866 he had been knighted for his role in the design of the first transatlantic telegraph cable; and of course he is known today also for his absolute scale of temperature. The book "T & T" set out the concept of energy and its conservation, clearly, for the first time. "The work was epoch-making, and created a revolution in scientific development," wrote one commentator.

Thomson was in the audience when Tait demonstrated his apparatus. He was fascinated by Tait's demonstration of Helmholtz's ideas, and saw in these smoke rings the makings of a theory of the atom. Thomson's vision was that perhaps atoms consisted of tiny vortex rings, persistent topological singularities in the aether. Different elements might correspond to different knot types!—so perhaps the hydrogen atom was a simple loop (the "unknot"), helium was the trefoil knot, ... (Here was something even Gandalf or the best modern vape artists couldn't do.) These rings vibrated, putting them in different energy states. And perhaps the mystery of the formation of molecules could be explained by the linking of several knots to form links.

Now, you may think that Thomson was smoking something himself ... and his vision might be explained by the fact that the smoke in Tait's apparatus was created by combining ammonia and sulfuric acid. But you have to remember that Rutherford's image of a tiny nucleus surrounded by orbiting electrons appeared only in 1909, 42 years later; Ernest Rutherford wasn't even born till 1871.

Kelvin's idea became the subject of scorn and ridicule when it failed to account successfully for any facts of chemistry whatsoever. But I submit that the basic idea is very deep, and is the same as the idea underlying the current use of loops in a theory of elementary particles, string theory: knots give a model for extended topological structures, and the combinatorial nature of the topology is what accounts for their persistence. Thomson's proposal can be thought of as reflecting a very early insight into the geometric notions underlying quantum mechanics. Today's ideas of string theory share this underlying idea, and it has re-entered chemistry in the study of polymer entanglement!

Some years later, starting in 1876, Tait began a serious attempt to list different knot types. Let's be clearer about what a knot is. Mathematically, a "knot" is a closed loop sitting in space in some way. Two knots are "the same" if one can be smoothly deformed to the other without ever crossing through itself. This operation is called a "regular isotopy." A "link" is a finite collection of nonintersecting embedded loops, so a knot is a link, and there is also the *empty* link.

Tait studied knots by drawing projections onto a plane. You can always arrange that this projection is free of kinks, tangencies, and points where three strands meet; this is then called a "knot projection." Tait published a correct and complete list of knots which can be presented using 7 or fewer crossings. The list on the back of your homework assignment was begun by him.

By the way, being Scottish, Tait was interested in golf, and in fact wrote a definitive account of the physics of golf. In most of the world he is known as the father of Freddie Tait, two time winner of the Scottish Open Golf Championship.

Knots.

Tait and his contemporaries had no idea how to actually prove that the knots they wrote down couldn't be deformed into each other, and it was only well into this century that enough algebraic topology was developed to address these questions.

A cat has just come into my life, and in consequence I have some very tangled bits of yarn around the house. I know that they were originally "unknots" (topologists' yarn always comes in closed loops) but I'll be damned if I can return them to that state now. How do I know, then that my evil neighbor didn't slip in and substitute a knotted bit of blue yarn? If I suspect that someone did, how could I prove, in a court of law, or to my colleagues here at MIT, that this thing could NEVER be returned to a simple loop? You need some "invariant," which takes on one value for the unknot and another for this one.

To say it differently, what if I give a trefoil of yarn to my cat for a few days and then simplify the resulting tangle. How do I KNOW that it won't come out the unknot, or some other knot?

The trefoil and its mirror image present another problem. If we put a mirror behind our knot and look at the image of the knot in it, we see what might be a different knot. The shape is the same but the crossings are reversed. The mirror image of the trefoil is different from the original trefoil: the trefoil is "chiral," in contrast with the figure 8 knot, which is "amphicheiral"—a good spelling bee word introduced in 1904 by William Thomson. Here's a property of knots which is clear and distinct, but very hard to guess in advance. The mirror image of K is written $\overline{K}$.

Given two knots I can cut each and splice the open ends together. This is an example of a general geometric operation called "surgery," and the result is the "sum" of the two knots. It's not quite well-defined yet though: I need to decide which ends to splice. This is resolved by giving the knots an "orientation." Once again, reversing the orientation might or might not result in an isotopic oriented knot (though the first examples of this were only given in 1964). 8_{17} is the first non-reversible knot, and the knots you get by tying two 8_{17}'s together in the two possible ways are distinct. The reverse of K is written $-K$.

A knot is "prime" if it can't be written as a nontrivial sum. (To make the analogy with numbers better we should speak of the "product" of knots rather than their "sum," but it's too late now.) A feature of numbers is preserved: an oriented knot is a sum of prime oriented knots in a unique way (up to order of course). So people (starting with Tait) make tables of prime knots. The table on the back of this handout lists all the prime knots of 8 or fewer crossings. It does not carry complete information, though; it does not indicate whether K is isotopic to $-K$, $\overline{K}$, or $-\overline{K}$, or all or none of them, and it does not display $\overline{K}$ even if K is chiral.

The number of distinct prime knots goes up very rapidly with the number of crossings in its most efficient projection. In this tabulation again K is considered equivalent to $-K$, $\overline{K}$, and $-\overline{K}$.

Crossings	3	4	5	6	7	8	9	10	11	12	13	14	15
Prime knots	1	1	2	3	7	21	49	165	552	2176	9988	46872	253293

Knot projections, Reidemeister moves, and coloration.

If you look at a knot from the right point, you can make what you see – its projection – avoid cusps, tangencies, and triple points.

If you watch a deformation, you won't be able to avoid these three bad things, but you will be able to avoid triple tangencies, quadruple points, and other pathologies.

So two knot projections represent the same knot exactly when they differ by a sequence of "Reidemeister moves." Type I avoids kinks; Type II avoids tangencies; Type III avoids triple points. You need both Type I moves, but only one – any one – of the six possible Type III moves. Knot invariants are constructed by assigning some algebraic object – a number or a polynomial, for example – to knot projections in such a way that the value is unchanged by Reidemeister moves. For example:

Definition. A knot projection is 3-colorable if you can assign an element of $\{R, G, Y\}$ to each arc in such a way that (1) at every crossing all three incoming arcs are either all the same color or all different colors, and (2) all three colors are used.

Observe that the standard trefoil (either of them) is 3-colorable, but that the standard unknot is not.

Lemma. The property of being 3-colorable is preserved by Reidemeister moves.

Corollary. The trefoil is nontrivial.

Rational tangles.

Many mathematicians have tried their hand at the problem of classifying knots. Among them is John Conway, a great group theorist and combinatorialist. When I first met him he was at the University of Cambridge, but he moved to Princeton, and died in 2020 (of COVID-19). He proposed to first isolate a ball containing a simple part of a knot, or link, that he wished to study. He demanded that the knot pierce the surface of the ball in exactly four spots. Fix those points and call them NE, SE, SW, and NW. What lies inside the ball is a "tangle." (There might be loops entirely inside the ball.) Two tangles are equivalent if you can move one to the other through tangles (and without detaching them from the sphere). For example the "zero tangle" has two horizontal lines. Perhaps it's called "zero" because the slope is zero. That makes two vertical lines "infinity." "1" will be a line of slope 1 with one of slope -1 passing under it, and "-1" is the other way around.

Here's a way to generate new tangles from old. Take a tangle and release the anchor points. Let the ends of the strands drift around on the sphere, till they come to rest again at the same anchor points, and re-attach them. Call this a "surface transformation."

There are a couple of simple surface transformations. One can *Twist*: pass SE over NE. One can *Twist-inverse*: pass NE under SE. And one can *Rotate* (in the positive direction, by 90°). Call these operations T, T^{-1}, and R. For example, here's a picture of $T^n =$. $RT^n =$ is a barber pole.

It's a fact that T, T^{-1}, and R generate all surface transformations

Here's a little fact you can check directly: $(TR)^3 = I$. Students don't appreciate the letter "F" as much as they might: it's the first letter of the alphabet to look 8 different ways under the symmetries of a square. Let's label a tangle by F, to keep track how it is moved around. We can check this.

Clearly $R^4 = I$, the identity transformation. But in fact if you rotate a *rational* tangle by 180° you get the same tangle back: that is, $R^2 = I$. I'll let you check this for homework. It's obviously false for general tangles.

We assigned a rational number to some tangles. How far can this go? Here's the theorem, due to some combination of Schubert, Conway, Kauffman, and Goldman. Write Tang for the set of isotopy classes of tangles. Actually, we allow *links* here too, consisting of several non-intersecting components. We'll distinguish the tangles obtained from $=$ by our operations by calling them "rational tangles." They have no closed loops.

Theorem. There is a function $Q : \mathrm{Tang} \to \mathbf{C} \cup \{\infty\}$ such that
- $Q_= = 0$,
- $Q_{TF} = 1 + Q_F$,
- $Q_{RF} = -1/Q_F$,
- $Q_{F \sqcup O} = 0$ if O is a separated unknotted circle.

Q establishes a bijection from rational tangles to $\mathbf{Q} \cup \{\infty\}$.

So T not only twists, it also *translates*, and R not only rotates, it also *reciprocates*. For example, $Q_{||} = \infty$, $Q_{T^n=} = n$, and the value of the barber pole is $-1/n$.

We will prove this theorem, and then check that it's actually true by a simulation.

First, let's check that the two identities we discovered for tangles hold for translation and reciprocation. First R^2:

$$q \mapsto -1/q \mapsto -1/(-1/q)) = q$$

Then TR:

$$q \mapsto 1 - \frac{1}{q} = \frac{q-1}{q} \mapsto 1 - \frac{q}{q-1} = \frac{-1}{q-1} \mapsto 1 + (q-1) = q\,.$$

We'll assign a complex number to each "tangle projection," and then check that it is invariant under Reidemeister moves.

The idea is to assign to a tangle diagram a number that is constructed so as to be invariant under Reidemeister moves, and then check the desired properties. To do this, we "resolve" each crossing by breaking both strands and then tying them together in the each of the two possible ways to avoid the crossing. Since there are two ways, we'll be talking about formal *sums* of tangles.

We'll also want to keep track of which way by assigning a symbol to each: A if you turn left as you go down, B if you turn right as you go down. You don't need an orientation for these instructions to make sense! So the "value" of a tangle will be the same as the sum of A times the value of the same tangle with the crossing removed and the strands tied so that you turn left as you go down; and B if you turn right as you go down. These are "local" operations. So we'll be talking about polynomials in A and B whose coefficients are tangles! This "value" won't be quite our function Q, but we'll see how to get to Q from it.

The axioms suggest that we drop out terms that include a separated circle. If you apply this operation to every crossing and drop out terms that contain separated circles, you will ultimately wind up with a linear combination of the two basic tangles $=$ and $||$: say

$$N_F(A,B)|| + D_F(A,B) = \,.$$

where N_F and D_F are polynomials in A and B with integer coefficients. For example $N_= = 0, D_= = 1, N_{||} = 1, D_{||} = 0$, and for the $T=$ crossing $N = A, D = B$.

We'll want to think about how the Reidemeister moves alter the values of N and D. Type I is easy. Removing one loop produces a term with a separated loop, which we drop out, and another term that is either A or B times the tangle with the loop removed: both N and D get multiplied by the same thing, either A or B.

How about the effect of a move of type II? For example $\supset$ over $\subset$ goes to $AB|| + (A^2+B^2) =$. If this pattern appears in a bigger tangle, resolving both of the crossings will produce produce two sub-sums in the value. Here we can actually guarantee that the value is unchanged by requiring that

$$AB = 1 \quad , \quad A^2 + B^2 = 0\,.$$

These relations are equivalent to

$$B = A^{-1} \quad , \quad A^4 = -1\,.$$

So we can think of N_F and D_F as complex numbers by sending A to a fourth root of -1; or what's the same, to an 8th root of 1, say

$$A \mapsto \omega = \frac{1+i}{\sqrt{2}}\,.$$

I'll let you check the amazingly convenient fact that the type III move invariance is automatic once we arrange that type II holds.

Type I is more of a problem: even if we enforce our new values for A and B, removing a loop will change the value. As we saw, eliminating a loop multiplies the value by either A or B; or, using our new labels, by ω or ω^{-1}. Since it multiplies the entire value, it multiplies both N and D by the same number. We can eliminate this effect by taking the quotient, N/D. (This accounts for our choice of notation.) It's possible that both N and D vanish - when all terms include separated circles. Declare $0/0$ to be 0.

Let's check values on our basic tangles. $N_= = 0, D_= = 1$, with quotient 0. $N_{||} = 1, D_{||} = 0$, with quotient $\ldots$– well, this is where ∞ comes in: $N_{||}/D_{||} = \infty$. And the overcrossing T= the quotient is $A/B = A^2$, i.e. $\omega^2 = i$. This isn't quite what we wanted, but we can get exactly the right thing by dividing by i or, what's the same, multiplying by $-i$.

So, finally, define

$$Q_F = -i\frac{N_F}{D_F}\,.$$

This gives an invariant with values in $\mathbf{Q}[i] \cup \{\infty\}$.

We have to check the axioms. It has the right value on$=$. If there's a separated circle, every term in the splitting will include it, so we get $Q_F = 0/0 = 0$.

If we rotate a tangle by 90 degrees, numerator and denominator get swapped, and

$$Q_{RF} = -i\frac{N_{RF}}{D_{RF}} = -i\frac{D_F}{N_F} = \frac{1}{iN_F/D_F} = -\frac{1}{Q_F}\,.$$

To check the remaining axiom, we make the following observation. If we have two tangles, F and G, we can put them next to each other and tie NE of F to NW of G, and SE of F to SW of G. Write $F+G$ for the result. Then

$$=+= \;=\; =\,, \quad =+|| \;=\; || \;=\; ||+=\,, \quad ||+|| \;=\; |O|$$

so (since the loop O kills)

$$D_{F+G} \;=\; D_F D_G \quad , \quad N_{F+G} \;=\; N_F D_G + D_F N_G\,.$$

This is the rule for adding fractions! so

$$Q_{F+G} = Q_F + Q_G\,.$$

Q is an "additive" invariant – an unexpected bonus! In particular, since TF is obtained from F by adding a tangle with value 1,

$$Q_{TF} = Q_F + 1\,.$$

We have now constructed an invariant on all tangles that has all the required properties.

It clearly sends rational links to rational numbers (or ∞). Now we have to see why Q gives a bijection from rational tangles to $\mathbf{Q} \sqcup \{\infty\}$. Here's how I think of it. There's a group action involved. The relations $R^2 = I, (TR)^3 = I$ look very familiar. The matrices

$$T = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix} \quad , \quad R = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}$$

almost satisfy these identities; in fact for these matrices,

$$R^2 = -I = (TR)^3 \,.$$

It turns out that any 2×2 matrix with integer entries and determinant 1 is a product of these (and T^{-1}); these two matrices generate

$$\mathrm{SL}_2(\mathbf{Z}) = \left\{ A = \begin{bmatrix} a & b \\ c & d \end{bmatrix} : a, b, c, d \in \mathbf{Z}\,, \det A = 1 \right\} .$$

There is a famous action of this group on $\mathbf{C} \cup \{\infty\}$:

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix} q = \frac{aq + b}{cq + d}$$

so that

$$Tq = 1 + q \,, \quad Rq = -1/q \,.$$

The subset $\mathbf{Q} \cup \{\infty\}$ is a single orbit for this action. To see this, write the rational number as b/d with b and d relatively prime integers. Then there exist integers a, c such that $ad - bc = 1$, and then $A0 = q$.

This group also acts on the set of tangles, with R rotating and T twisting. The theorem asserts that the map Q is *equivariant*:

$$Q_{AF} = AQ_F \,.$$

This implies surjectivity.

To see that Q maps rational tangles injectively, it's enough to show that for some rational tangle F, if $AQ_F = Q_F$ then $AF = F$. Well, $Q_{||} = \infty$, and $A\infty = \infty$ exactly when

$$A = \pm \begin{bmatrix} 1 & b \\ 0 & 1 \end{bmatrix} = \pm T^b \,.$$

And, indeed, $T|| = ||$.

Square dance.

Let's "prove" the rational part of this theorem by a square dance!

Some things to think about.

1. There are six variants of the Type III Reidemeister move. Show that (in the presence of Type II moves) they are all equivalent; you only need one.

2. Tangles are divided into three distinct classes: those connecting NW to NE, to SE, and to SW. Are these classes separated by our invariant? What is the corresponding decomposition of rational numbers? How should ∞ be handled?

3. If K is a rational tangle with value q, what is the value of the mirror image of K?

4. Prove that $R^2 = I$ on rational tangles. (This must be the case if the Schubert-Conway theorem is true, since $R^2 q = q$ for any rational number.) This operation rotates by 180° around an axis poking out of the page. How about 180° rotations around the other two principal axes - the x-axis and the y-axis? In fact, you may find it convenient to prove simultaneously that all three operations fix all rational tangles.

Some more things to think about.

5. A "two-bridge link" is a link that can be presented with just two highest points and two lowest points. Explain why such links are the same as links obtained by tying NE to NW and SE to SW in a rational tangle. Draw pictures of the link corresponding to $T^n =$ and some other examples. It turns out that all prime knots with fewer than 8 crossings are 2-bridge knots. Can you draw pictures of them showing that? It turns out that exactly 12 of the 21 8-crossing prime knots are 2-bridge knots. I think it's rather hard to pick them out from the pictures! Different rational tangles may produce the same link by this process. Come up with a list of modifications of a rational tangle that don't change the corresponding link. A complete list will lead to a complete invariant for two-bridge links.

6. Find out about "continued fractions" and explain how they provide an algorithmic and optimal way to pass from $=$ to a rational tangle F with invariant $Q_F = q$, for any $q \in \mathbf{Q}$.

7. Work out Q_F for some simple non-rational tangles. Longer term question: What is the image of $Q : \mathrm{Tang} \to \mathbf{C} \sqcup \{\infty\}$?

8. The equality $R^2 F = F$ only has a chance of being true because we do not color the strings. Can you develop an analogous invariant of "colored tangles" in which the strands are colored and the equivalence respects the colors? More questions: How about if you impose orientations on the strands?

Two books on knot theory:

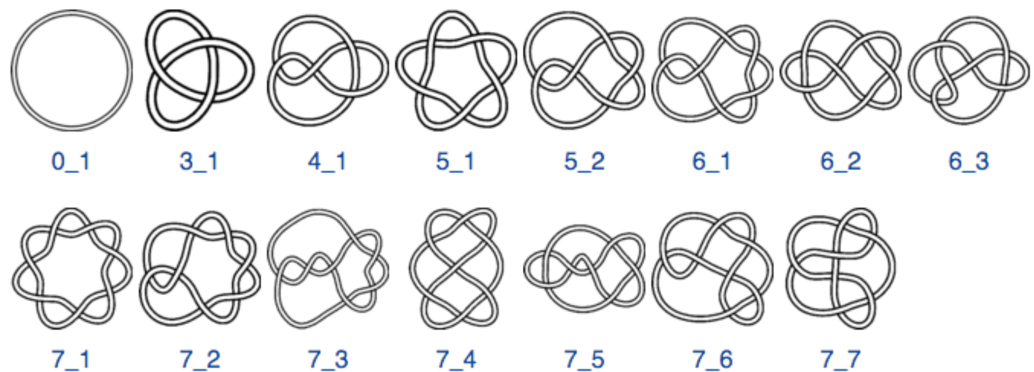
Colin Adams, *The Knot Book: An Elementary Introduction to the Mathematical Theory of Knots*, W. H. Freeman and Co., 1994.

Charles Livingston, *Knot Theory*, Math. Assoc. of America, 1993.

Continued fractions occur in any elementary number theory book, e.g.

Harold Davenport, *The Higher Arithmetic*, Dover, 1983.

Knots with 7 or fewer crossings



Knots with 8 crossings