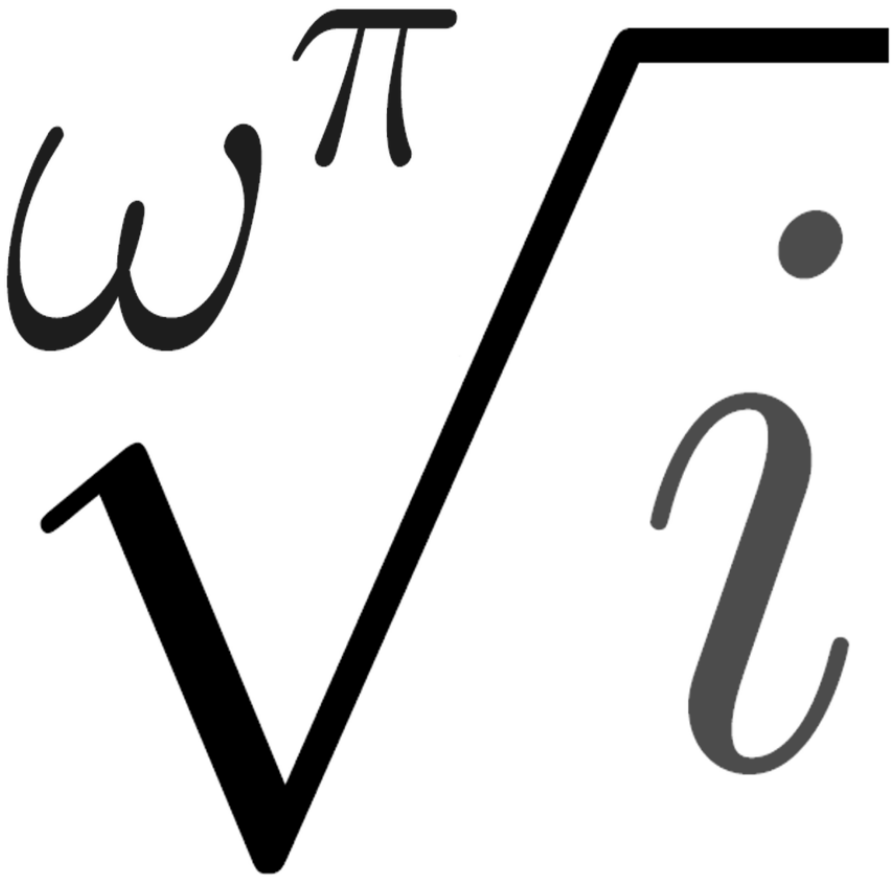
# THE invariant



*Oxford's first maths magazine*

Michaelmas Term 2017

Fellow Mathematicians,
Welcome (back) to a year of mathematical excitement. It has been over a year now since the latest version of the Invariant has been published. As far as I can tell, the magazine was first published at the very least in 1966, and very likely much earlier than that. Over the years the magazine has displayed the gigantic variety of ideas in Mathematics and its applications and a is great place to have a glimpse of the many realms of Maths that one hasn't explored. It has also been a voice for the members of the Invariants to share intriguing pieces of Mathematics that they have encountered, and I invite you to do the same.
In this edition of the Invariant you will find an appreciation for seeking multiple proofs of the same result and the various reasons people have done so in exploring mathematical structures. Phil Tootill, a former editor of this magazine, gives an interesting view of board game design and how it compares to mathematical rigor; after all, some formalists claim that Mathematics is nothing but a game where we follow a set of axioms. As a step away from the classical set-theoretic formalism, our colleague from UCL, introduces us to the world of Category Theory. Finally, you might enjoy reading about Rational Tangles and how interesting Mathematics arises even when we're just thinking about how a pair of strings can tangle up.

Wishing you reward in your mathematical explorations, and a successful term,
Hazem Hassan
Editor-in-chief

## Contents



## Term Card

week 1 - Tuesday 10 October - Common Room - 7:30PM
Invariants Social & Magazine Party

week 2 - Tuesday 17 October - L2 8:00PM
Lecture by Steve Roberts: *Finding Earth v2*

week 2 - Thursday 19 October - Common Room - 7:30PM
Morgan Stanley workshop

week 3 - Tuesday 24 October - L2 - 8:00PM
Lecture by Vicky Neal: *Reinventing the Wheel*

week 4 - Wednesday 1 November - Common Room - 7:00PM
TPP logic puzzle session with free snacks and pizza

week 5 - To be confirmed

week 6 - Tuesday 14 November - Common Room - 7:30PM
Puzzle competition

week 7 - Tuesday 21 November - L2 - 8:00PM
Lecture by Doyne Farmer

week 8 - Christmas Dinner- Details to be confirmed

# Reinventing the Wheel

Vicky Neale

## What is a proof for?

Of course proofs are to tell us that theorems are true. But they're about much more than that too, which explains why it is worth having more than one proof of a theorem. If a proof is only to demonstrate truth, then giving a second proof of a theorem is like being the second person to invent the wheel. The multitude of proofs of some theorems in mathematics reveals much about those results, but also about what proofs can do for us as mathematicians.

In this article, I'd like to reflect on a few other purposes of proofs, and to give some examples. These examples are inevitably biased towards the areas of mathematics that I have thought about most, they are not supposed to be a representative selection! My suggestions of purposes of proofs are not a complete list. Rather, I hope that the article will encourage you to think about the purposes of proofs for yourself, and to select your own favourite examples.

If you look online, it is not too hard to find collections of multiple proofs of the same result: of Pythagoras's Theorem, or the irrationality of 2, for example. Such collections are fascinating, and often contain little gems of proofs, with beautiful and unexpected ideas. However, there are probably good reasons why only a handful of these proofs are well known. Here, I want to focus on a slightly different type of situation, where the second proof of a theorem is at least as well known as the first.

### Generalising

Sometimes a merit of a proof is its applicability to other problems. Perhaps it opens up new avenues of generalisation. Perhaps it shows that more is true.

For example, one of the most famous proofs in mathematics is surely Euclid's proof that there are infinitely many primes. (You probably know the one: suppose there are finitely many primes, multiply them all together and 1 to obtain a number that is not divisible by any of the primes on the initial list, contradiction.) I think it is rightly famous: its elegant economy suffices to prove something important about the primes.

Perhaps less well known is the argument attributed to Euler that proceeds by considering the sum of the reciprocals of the primes. It turns out that this sum diverges (I shan't go into details here), and this demonstrates that in particular there are infinitely many primes.

If your goal is just to prove that there are infinitely many primes, then Euclid's argument is enough, and as a bonus it is beautiful. Euler's approach needs more work,

but reveals more about the distribution of the primes, and also extends nicely. It's possible to generalise Euclid's argument to show that there are infinitely many primes of certain forms. For example, a classic problem for undergraduates is to prove that there are infinitely many primes that are one less than a multiple of 4, and a suitable adaptation of Euclid's argument will do this very nicely. But it turns out to be rather hard to extend the argument to the whole family of such problems (even using it to prove that there are infinitely many primes that are one more than a multiple of 4 takes more thought).

As we would hope, a more general result is true, and this is known as Dirichlet's theorem: if $a$ and $d$ are coprime (have highest common factor 1), then there are infinitely many primes that are $a$ more than a multiple of $d$. It turns out to be possible to prove that in this case the sum of the reciprocals of the primes that are $a$ more than a multiple of $d$ diverges, and this proves that there are infinitely many such primes.

### Additional insights

In 1909, David Hilbert solved Waring's problem. Hilbert showed that for each $k$, there is some $s$ such that every positive integer is a sum of $s$ $k^{\text{th}}$ powers. This is a big generalisation of a famous theorem of Lagrange, which states that every positive integer is a sum of four squares. Edward Waring conjectured the generalisation in the eighteenth century, and Hilbert was able to prove the result.

In the 1920s, G.H. Hardy and J.E. Littlewood gave another proof of Waring's conjecture. This proof has gone on to be more significant and more well known—but why, when the theorem had already been proved?

Hardy and Littlewood took a very different approach from that of Hilbert. Their argument gives more information: it gives an asymptotic formula for the number of ways to write a large integer $N$ as a sum of $s$ $k^{\text{th}}$ powers (where $s$ is large enough in terms of $k$). This formula is an estimate for the number of representations, but it becomes a better estimate for larger values of $N$. By showing that the number of representations is positive for suitable $s$ and for large enough $N$, Hardy and Littlewood proved the conjecture of Waring, but their argument gives a more detailed insight. In addition, the Hardy–Littlewood approach gives a framework for solutions to a range of other similar problems in additive number theory.

One striking example is due to Ivan Vinogradov, who streamlined and adapted the approach of Hardy and Littlewood to prove that every sufficiently large odd number is a sum of three odd primes. The Hardy-Littewood circle method has since been used to resolve a range of problems with this general flavour, and there are books entirely about the circle method.

### A Rosetta stone

Terence Tao has described the various proofs of Szemerédi's Theorem as a Rosetta stone for mathematics.

The theorem, which Endre Szemerédi proved in 1975, states that for any $k \geq 1$ and any $\delta > 0$, there is $N$ such that if $A$ is a subset of $\{1, 2, \ldots, N\}$ of size at least $\delta N$, then $A$ contains an arithmetic progression of length $k$. (For example, if $N$ is large enough and we choose 1% of the numbers from 1 to $N$, then our chosen set must contain some structure in the form of an arithmetic progression of length 100.) This resolved a conjecture of Erdős and Turán from the 1930s.

So what is so special about this theorem? Here is a quick summary of its history.

In 1953, Klaus Roth gave a proof of the special case of arithmetic progressions of length 3, using a version of the Hardy-Littlewood circle method that today is usually phrased in terms of Fourier analysis.

In 1969, Szemerédi proved the theorem for progressions of length 4, and in 1975 he was able to deal with the general case of progressions of arbitrary length. He used ingenious combinatorial arguments, including the Regularity Lemma that he proved along the way (and that has gone on to have a huge number of applications to other problems).

In 1977, Hillel Furstenberg gave another proof of Szemerédi's Theorem, unexpectedly using ergodic theory.

As if these proofs were not enough, in 1998 and 2001 Timothy Gowers gave a new proof (for progressions of length 4 in 1998 and for arbitrary lengths in 2001), extending Roth's Fourier analytic ideas and introducing several new ingredients.

Since then, there have been several additional proofs, refining previous arguments but also introducing new ideas such as hypergraph regularity.

Each proof has its own intrinsic interest, of course, but what is so special, and what prompted Tao to use the Rosetta stone analogy, is that exploring the connections between the arguments has been a rich source of insights. It seems that the approaches have a number of common features, and these give ways to connect seemingly disparate areas of mathematics.

### An anticlimax

Sometimes the second proof of a theorem comes as something of an anticlimax. One of the highlights of nineteenth century mathematics was the proof in 1896 by Jacques Hadamard and Charles de la Vallée Poussin of the Prime Number Theorem. (They came up with their proofs independently but using essentially the same ideas at the same time.) This theorem says that the number of primes up to any value $x$ is asymptotically $x/\log x$. Perhaps surprisingly, their proof relied crucially on ideas from complex analysis: their argument proceeded by showing that the Riemann zeta function never takes the value zero for a certain set of complex values, and it turns out that this is sufficient to prove the theorem. Mathematicians wondered whether there was an 'elementary' proof, one that does not use complex analysis, since after all this is a result about counting prime numbers and it is not clear that such a result should rely on complex analysis.

In 1948, Erdős and Selberg gave an elementary proof of the Prime Number Theorem, and thereby resolved the question. It was not, however, quite the triumph it might have been, because their arguments are rather sophisticated and subtle. While the argument

is technically 'elementary', it is definitely not easy. The elementary argument has its fans, and has led to other new ideas, but undergraduates meeting the Prime Number Theorem still usually learn about a proof using complex analysis before they meet the elementary method—the second proof has not, as one might have expected, rendered the first one obsolete.

### Conclusion

I hope that these examples will prompt you to reflect on your own experiences of knowing multiple proofs of the same theorem, and in particular I encourage you not to settle for knowing just one proof if knowing another would deepen your understanding. This is as true of undergraduate mathematics as it is of cutting-edge research. Sometimes being the second person to invent the wheel can be a good thing!

## References

G.H. Hardy and E.M. Wright, *An Introduction to the Theory of Numbers*, sixth edition, Oxford University Press, 2009.

# From Maths to Meeples

Phil Tootill

Leaving mathematics behind was difficult. At university, it was the core of my identity. My friends were mathematicians, my shelves were piled with maths books, and my coffee was drunk from mathematical mugs. Four years on, I'm happy to say it still has a huge impact on how I think: in my work, as a software developer, but also in my hobby as a board game designer.

To many people, board games are a symbol of humdrum afternoons on a disappointing holiday, but that perception is changing rapidly. I started playing board games in my final year of university, and suddenly found myself lacking time for anything else. Soon after, I attended the UK Games Expo in Birmingham, and was inspired to try designing my own games. Three years in, I have a range of promising prototypes, and one game due to be published later this year.

In this article, I'll be discussing some of the ways studying maths influenced my approach to game design.

## 1. The Need For Rigour

In both fields, there's a huge focus on seemingly trivial details. This was a huge cause of frustration in my first year of studying. I'd covered all the interesting cases- why would it matter if n was 0? I eventually came to appreciate the details, and this has been incredibly valuable when designing board games. It's easy to overlook special cases, such as a player needing to play a card when they have no cards left. Eventually these cases will all happen: in fact, players often deliberately seek out these cases and try to exploit them. While they're not the most exciting part of designing a game, they're a crucial one, and my mathematical training made them a lot easier to identify.

## 2. Choosing Axioms

In many ways, the rules of a game are like the axioms of a mathematical theory. An ideal set of rules needs to be consistent and cover everything your game should do, while being as simple as possible. If a rule is too complex, one solution is to replace it with an equivalent but more intuitive one. The ideal solution, however, is to remove the rule entirely.

There are some good examples of mathematicians trying to streamline a set of axioms. One example is the axiom of choice from set theory. Roughly speaking, this axiom is as follows:

> *Given any collection of sets, a set exists which takes one element from each set in the collection.*

At first glance, the axiom seems harmless. At worst, it seems redundant, and should follow from the other axioms. However, it has some hugely unintuitive consequences, such as the Banach-Tarski paradox.

Another example can be found in Euclidean geometry. Euclid put forward five axioms, from which all geometry should follow. The first four are totally intuitive, corresponding naturally to using a ruler and a compass to do geometry on paper. The fifth axiom, however, caused centuries of controversy. There are various formulations of this axiom, the most illustrative being the triangle postulate:

The angles in a triangle must add up to 180 degrees.

Why is this axiom needed? Including this statement as a basic rule of geometry is incredibly dissatisfying. Mathematicians searched for a way to derive this axiom from the other four for two thousand years. The conclusion was that the axiom is indeed necessary- but only in Euclid's geometry. The are other alternate geometries, where the axiom may not be true. To illustrate this, consider a triangle drawn on the surface of the earth. The base runs along the equator, and reaches a quarter of the way around. The apex lies at the north pole. The sides make a right angle at each vertex, showing the triangle postulate is false on the surface of a sphere.

Several times, I've seen an analogous situation in game design. I was working on a game for a year, and had a core rule which just wouldn't work. The rule had been there from the start, and seemed important, but consistently received negative feedback. I spent weeks looking for ways to fix this rule. Out of curiosity, I tried removing the rule entirely, and it fixed the whole problem. Like the new geometries which were discovered by removing the triangle postulate, games can take surprising and interesting directions when designers are brave enough to relax rules which seem unquestionable.

## 3. Structuring An Argument

When analysing a game, it's often useful to break it down into three stages, known as the MDA framework:

- Mechanics: The core rules of the game.
- Dynamics: How the rules interact.
- Aesthetics: The emotional responses to the game.

To illustrate, consider the game of poker. The core mechanics provide a series of moments where each player can choose to bet or to fold, and the ways hands are scored. Bluffing is an example of a dynamic. The rules never explicitly state that players should act as if their hand is better than it is, but its a natural consequence of the betting mechanics. The tension, which makes the game fun, is the aesthetic that is creating by the bluffing dynamic.

This structure is reflected in mathematics by the relationship between axioms, lemmas and theorems. Starting with axioms, the aim is to create theorems, and lemmas are used as a stepping stone in the process. A good lemma can be used in many situations. Similarly, a good dynamic, such as bluffing, can re-used in multiple games, just by implementing a small number of mechanics.

It's also useful to consider these two structures when considering how to explain a game. With the right collection of lemmas, a theorem should follow with little difficulty. Likewise, a game explanation becomes awkward if too much emphasis is placed on why the game will be fun. A large part of the challenge of game design isn't just trying to make rules that work- it's in making a set of rules which are exciting and interesting to players, even during the explanation.

# Category Theory

Lukas Kofler

What do the lowest common multiple of two natural numbers, the $\vee$ in logic and the disjoint union $\bigsqcup$ of two sets have in common?
At first glance, nothing. At second glance – still nothing, except maybe that all three take two variables and output another one of the same kind. It turns out though that they are examples of the very same concept – the so-called coproduct. This surprising result showcases the power of category theory. To quote Tom Leinster[1]:

> *Category theory takes a birds eye view of mathematics. From high in the sky, details become invisible, but we can spot patterns that were impossible to detect from ground level.*

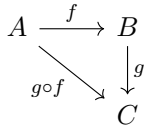
Category theory was invented by Saunders Mac Lane and Samuel Eilenberg in their 1945 paper "General theory of natural equivalences". They were motivated by their research in algebraic topology. The theory soon found application in homological algebra and algebraic geometry. Today category theory is an active area of research in its own right and is also being applied in theoretical physics, computer science and even linguistics.
But first things first. The starting point of category theory is the general observation that in many areas of mathematics we essentially work with just two different types of things which form a *category* when considered together: some (static) objects and some means of getting from one object to another.
There are plenty of examples: to "move" from one set $A$ to another one $B$, we can use a function $f : A \to B$. Groups in abstract algebra are linked by group homomorphisms, which are special functions that "preserve" the group structure. The same holds true for rings and ring homomorphisms, vector spaces and linear functions between them. If we work with topological spaces the functions we want to use are just continuous ones[2]. These are different instances of structures with "nice" structure-preserving functions between them! This is an important notion and we shall try to capture it.

Consequently, we will look at objects and function-y maps between them which we will call arrows (or maps, or morphisms as we please). Which properties do we want these arrows to have?

[1] A category theorist at the University of Edinburgh.

[2] If you haven't studied some of these concepts yet, don't worry. In fact, the only things we need to know about are sets and functions. You can ignore the other ones, they are just for extra illustration.

If we have a map $f$ between two objects $A$ and $B$ and another one called $g$ going from $B$ to $C$, we would like to be able to chain them together to get a new map from $A$ to $C$. Let's call it $g \circ f$, just like we would with functions. As a diagram:




It doesn't matter whether we go from $A$ to $C$ directly or whether we take a detour through $B$ – we get the same result. Whenever this is the case, we say that the diagram *commutes*.
Let's assume we have three arrows: $f$ from $A$ to $B$, $g$ from $B$ to $C$ and $h$ from $C$ to $D$. When forming the composite arrow $h \circ g \circ f$ from $A$ to $D$ it shouldn't matter whether we compose f and g first and apply h later or whether we compose g and h first and apply it to f afterwards: we want $h \circ (g \circ f) = (h \circ g) \circ f$ – like standard function composition. We also note that every set $S$ has an identity function $1_S : S \to S$ which maps every element of the set to itself. We shall require every object to have such a do-nothing-arrow, too. We arrive at the following:

## Categories

**Definition 1.** A category **C** consists of objects (typically denoted $A, B, C, ...$) and arrows (typically denoted $f, g, h, ...$). These objects and arrows obey some rules:

- Every arrow $f$ has a *domain* $dom(f)$ and a *codomain* $cod(f)$. We write $f : A \to B$ whenever we want to say that $A = dom(f)$ and $B = cod(f)$.
- For two arrows $f : A \to B$, $g : B \to C$ (i.e. $cod(f) = dom(g)$) there exists an arrow $g \circ f : A \to C$ called the *composite* of f with g.
- For every object $A$, there is an arrow $1_A : A \to A$ which we call the *identity arrow* of A.

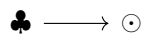
Lastly, there are two more rules governing the behaviour of the arrows:

- *Associativity of composition.* For any three arrows $f : A \to B$, $g : B \to C$ and $h : C \to D$ we have $h \circ (g \circ f) = (h \circ g) \circ f$.
- *Identity arrows act as a unit for composition.* For any arrow $f : A \to B$ we have $f \circ 1_A = f = 1_B \circ f$.

*Anything* that fits the above definition is a category. We immediately see that if we take all the sets as objects and all the functions as arrows, we can form the category

**Set**. Along the same lines we can speak of **Grp**, the category of groups and group homomorphisms and of **Top**, the category of topological spaces and continuous functions between them.
Can we make a category out of the set of integers $\mathbb{Z}$? Let the objects be the integers. Let there be an arrow between two integers $m$ and $n$ if and only if $m \leq n$. Since $n \leq n$ for every $n$, we have the required identity arrows. And because $n \leq m$ and $m \leq p$ implies that $n \leq p$, composition works too, so we see (after quickly checking the rest of the rules for arrows) that $\mathbb{Z}$ together with the operation $\leq$, denoted $(\mathbb{Z}, \leq)$, does indeed form a category!
Now for an example from logic: let's consider as objects logical statements $\psi, \phi, ...$ and arrows from some $\phi$ to some $\psi$ iff[3] we can deduce $\psi$ from $\phi$, usually denoted $\phi \vdash \psi$. For those familiar with formal theories, it should be easy to check that this, too, is a category.

A simple way of getting a new category, let's call it $\mathbf{C}^{op}$, from an old one $\mathbf{C}$ is to simply reverse all the arrows. To elaborate: the objects of $\mathbf{C}^{op}$ are the same as those of $\mathbf{C}$, while any arrow swaps its domain and codomain: $f : A \to B$ becomes $f^{op} : B \to A$. Consequently, identity arrows don't change. Composition of arrows is defined in a natural way: $f \circ^{op} g = g \circ f$ where $\circ^{op}$ denotes the composition operation in $\mathbf{C}^{op}$. This new category is called the *opposite* or *dual* of $\mathbf{C}$.
What is the dual of $(\mathbb{Z}, \leq)$? Reversing every arrow, we must get $(\mathbb{Z}, \geq)$! This formalizes the intuition that $\geq$ and $\leq$ are some kind of opposites of each other in a neat way.
Sadly, the dual of **Set** isn't quite as nice – when we reverse the direction of a function, we generally don't get another function. So the arrows of $\mathbf{Set}^{op}$ are just some relations which would be functions read "the other way around". This doesn't make for a very exciting category but it still exists.
All the categories we've considered so far have had infinitely many objects and arrows. But there are very simple ones too. For example, the category **1** which consists of a single object, let's call it $\bullet$, and one single identity arrow from $\bullet$ to itself. The category **2** looks like this (identity arrows not drawn):



We can call the objects and arrows whatever we want – the category stays the same for all intents and purposes. This brings up an important feature of category theory: the only things we are interested in from this new point of view are how objects relate to each other through arrows, i.e. how many, if any, arrows go from one object to another and how they might compose. Whatever's going on *within* objects is not of our concern. We can call this approach structural, or external.

[3] this is shorthand for if and only if.

## Isomorphisms

Which familiar concepts can we rephrase in such a way? Let's look at bijective functions and try to find a categorical analogue in **Set**.
Usually, part of the definition of a bijection is injectivity, i.e. that $f(x) = f(y) \implies x = y$. Category theory is too coarse for statements like this – the smallest thing we can see looking at the category of sets is a set. Since there is no categorical notion of an element of a set, we can't translate the former statement into category-speak.
We'll have to try from another angle: a function $f : A \to B$ is bijective iff there's an inverse function $f^{-1} : B \to A$ so that for every $x \in A$ we have $f^{-1} \circ f(x) = x$ and for all $y$ in $B$, $f \circ f^{-1}(y) = y$. This is the right way to look at it!

**Definition 2.** An *isomorphism*, sometimes simply called an *iso*, is an arrow $f : A \to B$ with an *inverse* $g : B \to A$ such that $g \circ f = 1_A$ and $f \circ g = 1_B$. We say that A and B are isomorphic and write $A \cong B$.

This is our first real category theoretic definition of an important concept. We don't have to "look inside" the objects and arrows at hand; the definition is framed purely in terms of arrows and how they relate to each other.
In **Set**, the isos turn out to be precisely the bijective functions[4]. Furthermore, we can see that two sets are isomorphic iff they have the same number of objects. In **Grp**, we get the group isomorphisms (hence the name) and in $\mathbf{Vect}_K$, the category of vector spaces and linear transformations/matrices over some field $K$, we get the invertible matrices, just as we would expect. In $(\mathbb{Z}, \leq)$ we only get isos whenever $n \leq m$ and $m \leq n$ at the same time, so $n \cong m \iff n = m$.
We can ask what's the dual concept of an isomorphism, i.e. an isomorphism in a dual category? Let's try it out: if we reverse the arrows in the definition we get $f^{op} : B \to A$ and $g^{op} : A \to B$ with the requirements that $g^{op} \circ^{op} f^{op} = 1_A$ and $f^{op} \circ^{op} g^{op} = 1_B$. Save for the distracting superscripts, the equations look just like before. This shows that the dual of an iso is, again, an iso.
Isomorphisms appear almost everywhere in mathematics in different disguises, so it feels natural that we are able to define them categorically, without having to consider concrete isomorphic objects.

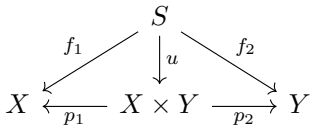
## Products

Another ubiquitous concept in maths is that of a product. We can form the product of two natural numbers, the cartesian product of two sets, the tensor product of two vector spaces, the direct product of two groups, and so on. Since the categorical notion of a

[4] You can work out the details yourself or look them up in any of the books referenced at the end.

product needs to be "arrow-theoretic", it is a bit more involved and doesn't really look like any of the products we know and love at first glance, but bear with me.

**Definition 3.** In a category $\mathbf{C}$ the *product* of objects $X$ and $Y$ is an object $X \times Y$ together with two so-called *projection arrows* $p_1 : X \times Y \to X$ and $p_2 : X \times Y \to Y$ such that for an arbitrary object $S$ coming with arrows $f_1 : S \to X$ and $f_2 : S \to Y$ there is a unique arrow $u : S \to X \times Y$ such that the following diagram commutes, which means that $f_1 = p_1 \circ u$ and $f_2 = p_2 \circ u$:




Given an arbitrary object $S$ with two arbitrary arrows into $X$ and $Y$ the product is the *only* object (up to isomorphism[5]) so that you can always find an arrow $u$ making our arbitrary arrows "factorise": $f_i = p_i \circ u$. So together with a product object we implicitly get a (bijective) function taking two functions as arguments ($f_1$ and $f_2$ in the definition) and returning another function $u$. This property is what makes the product such a special object.
Let's try some examples right away.

As always, looking at **Set** should clear things up a little. As mentioned above, we will consider the cartesian product $X \times Y = \{(x, y) \mid x \in X \text{ and } y \in Y\}$.
We want to show that, for example, a function $\mathbb{N} \to \mathbb{R} \times \mathbb{Q}$ given by $n \mapsto (\sqrt{n}, n^2)$ is essentially the same thing as two functions $\mathbb{N} \to \mathbb{R}$ and $\mathbb{N} \to \mathbb{Q}$, namely $n \mapsto \sqrt{n}$ and $n \mapsto n^2$.
We will use as projection arrows $p_1 : X \times Y \to X$ with $(x, y) \mapsto x$ and $p_2 : X \times Y \to Y$ with $(x, y) \mapsto y$. These functions forget about one of the components of $(x, y)$. If we consider a figure in the cartesian $X$-$Y$ plane, i.e. a subset of $X \times Y$, we can view the image of these functions as the shadows of the figure when shining a light onto the $x$-, respectively $y$-axis.
Now say we are given some set $S$ and functions $f_1 : S \to X$ and $f_2 : S \to Y$. We want to find a function $u$ from $S$ to the product $X \times Y$ such that $f_1$ has the same effect on any element of $S$ as $p_1 \circ u$ (analogously for $f_2$).
We can set $u = (f_1, f_2)$. This function sends $s \in S$ to $(f_1(s), f_2(s))$ in $X \times Y$. Applying the projection arrow for $X$ afterwards we get $p_1((f_1(s), f_2(s))) = f_1(s)$. The same with $p_2$ and we are done.
This shows that $u$ does indeed make the triangles in the above definition commute. Although it might not be immediately apparent, the way we defined $u$ makes it the unique such function too. Since $S$, $f_1$ and $f_2$ are arbitrarily chosen, this procedure will work for any set equipped with two such functions. This proves that the cartesian product really

[5]So we could prove that if there exists another product $X \times' Y$, then $X \times Y \cong X \times' Y$.

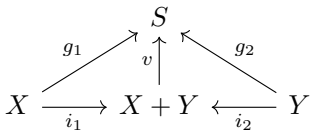
is the incarnation of the categorical product in **Set**.

Every pair of arrows into $X$ and $Y$ corresponds to exactly one arrow into $X \times Y$. This is an instance of a *universal mapping property*, UMP for short. Many concepts in category theory can be characterised by different UMPs.
Since category theory doesn't give us many tools for inspecting arrows, it's no big surprise that one of the most coveted properties an arrow can have is uniqueness.
Even though we didn't mention it before, even isomorphisms have such a UMP – namely that the inverse to a given iso is unique (prove it!).
A less obvious instance of a product can be found in the category consisting of the natural numbers with a single arrow from a number $n$ to another one $m$ iff n divides m. The "product" of natural numbers $m$ and $k$, let's call it $l$, has to satisfy the following conditions: it has to divide $m$ and $k$ (the projection arrows), and whenever a number $n$ divides both $m$ and $k$, $n$ needs to divide $l$ too (this condition equates to the unique arrow making everything commute).
If $m = 60$ and $k = 72$, then any number that divides both $A$ and $B$ – 1, 2, 3, 4, 6 or 12 – has to divide $l$, so $l$ must be 12. This might look familiar now. Indeed, the unique number for which this is true in general is the greatest common factor $gcf(m, k)$!
Let us look at the category of finite sets and functions between them for a second. This category's product works just like the one of **Set**. Consider just the number of elements in the sets, i.e. their cardinalities, which we denote $|A|$ for some finite set $A$. It is easy to show that $|A \times B| = |A| \cdot |B|$. So if $A$ has $n$ elements and $B$ has $m$, their product has $n \cdot m$ elements. This shows that the categorical product generalises even the standard multiplication of natural numbers we learned in primary school! Isn't that neat?
Continuing with our familiar examples, the product in **Grp** is the direct product $G \times H$ of groups $G$ and $H$, similarly to the product in **Set**. The product of vector spaces in $\mathbf{Vect}_K$ is the tensor product $V \otimes W$.
Now there's just one more thing to do: dualize.

## Coproducts

Dualizing just means turning the arrows around. If we do that to the definition of a product, we get its dual which is called a coproduct:

**Definition 4.** The *coproduct* of two objects $X$ and $Y$ in a category $\mathbf{C}$ is an object denoted $X + Y$with two *injection arrows* $i_1 : X \to X + Y$ and
$i_2 : Y \to X + Y$ such that for any object $S$ and arrows $g_1 : X \to S$ and $g_2 : Y \to S$ there is always a unique arrow $v : X + Y \to S$ such that the following diagram commutes:



So analogously to before some object is a coproduct iff, given $g_1$ and $g_2$ as above, we can find exactly one $v$ such that $g_1 = v \circ i_1$ and $g_2 = v \circ i_2$.

The coproduct of two sets $A$ and $B$ in **Set** (which is the same thing as the product in $\mathbf{Set}^{op}$, by the way) is the disjoint union $A \sqcup B$. Unlike the ordinary union, it keeps track of where its elements came from: if an element $x$ is a member of both $A$ and $B$, it shows up as two different elements in $A \sqcup B$: the one coming from $A$ is called $(x, 1)$, the one from $B$ $(x, 2)$.
We get intuitive injection functions $i_1 : X \to X \sqcup Y$ sending $x$ to $(x, 1)$ and $i_2 : Y \to X \sqcup Y$ sending $y$ to $(y, 2)$. Suppose now we are given some functions $g_1 : X \to S$ and $g_2 : Y \to S$. Then the function $v : X \sqcup Y \to S$ sending $(x, 1)$ to $g_1(x)$ and $(y, 2)$ to $g_2(y)$ makes the diagram commute.
Moreover, if $v'$ makes the diagram commute too, then $v'((x, 1)) = v' \circ i_1(x) = g_1(x) = v((x, 1))$ for any $x \in X$ and likewise $v'((y, 2)) = v((y, 2))$. Thus, we must have $v = v'$, so $v$ is unique as required. This shows that $X \sqcup Y$ is indeed the coproduct in **Set**!
If we again consider only finite sets and the functions between them, the coproduct generalises addition just like the product generalises multiplication since $|A| + |B| = |A \sqcup B|$. Why isn't the ordinary union $A \cup B$ the coproduct though? Consider the union $\{1\} \cup \{1\} = \{1\}$. Let $S = \{a, b\}$ and let the functions into $S$ be given by $g_1$ sending 1 to $a$ and $g_2$ sending 1 to $b$. What could our special function $v$ from the union $\{1\}$ to $S$ be? We can send the union's only element either to $a$ or to $b$, but not to both! Consequently, one of the triangles will not commute and the UMP of the coproduct is not satisfied.

Let's look for some more coproducts in familiar categories. Consider again the category of natural numbers with an arrow between two numbers $n$ and $m$ iff $n$ divides $m$. We found that the product in this category is the greatest common factor. Intuitively we might suspect that the lowest common multiple is the dual concept – and we would be right: two naturals $n$ and $m$ need to divide their coproduct by the existence of the injection arrows. And if $n$ and $m$ both divide some other number $k$, their coproduct – appealing to the UMP of the coproduct – needs to divide $k$ too. Thus, $n+m = lcm(n, m)$ where $+$ stands for the coproduct operation.
For those familiar with these concepts: in $\mathbf{Vect}_K$ the coproduct is the direct sum $V \oplus W$ of vector spaces together with the evident inclusion maps from $V$ and $W$. In **Grp** it is the free product $G * H$ equipped with injective homomorphisms from $G$ and $H$.
The abstract category consisting of logical statements as objects and proofs as arrows (p. 3) has an interesting coproduct too. We must be able to deduce the coproduct $\phi + \psi$ from $\phi$ as well as from $\psi$, this is granted by the injection arrows. And whenever we are able to infer some other statement $\xi$ from both $\phi$ and $\psi$ (the diagonal arrows in the diagram of the coproduct), we can infer it from $\phi + \psi$ too by the existence of this unique arrow making the triangles commute. We can check that $\phi \vee \psi$ satisfies these conditions! The co-coproduct, i.e. the product, in this category is $\phi \wedge \psi$ as you might want to verify.

## Conclusion

We have shown that the product of natural numbers, the greatest common factor, the product of sets and the logical conjunction $\wedge$ are, deep down, the same concept. In the natural environment of the objects they act on, they sit in the same spot. This demonstrates how category theory sheds light on the similarities of very different areas in mathematics, or categories, on one hand. On the other hand, it can grant us additional information about the structures within a category – duality makes rigorous the apparent connection between $\wedge$ and $\vee$, between $gcf(n, m)$ and $lcm(n, m)$, between addition and multiplication.
What might be next? We can generalise other concepts like the evaluation of a function $f(x)$ when $x$ equals some element of the domain, the quotient of groups $G/K$ or the empty set $\varnothing$. A natural way to abstract even further would be to do to categories what we did to sets – consider some or all of them together as a huge object, a *2-category*, whose arrows are called *functors*.
This train of thought uncovers many more beautiful results we will learn about in the years to come.
Category theory might seem to make things difficult rather than easy right now. But the more mathematics we study and the more complex the objects we work with become, the more we will appreciate its unifying aspects and powerful generalisations.

> *I hope most mathematicians continue to fear and despise category theory, so I can continue to maintain a certain advantage over them.*
>
> - John Baez

## Further reading

Category Theory: A Gentle Introduction - Peter Smith (2016)

Category Theory - Steve Awodey (2006)

The n-Category Caf - golem.ph.utexas.edu/category

nLab - ncatlab.org

# Rational Tangles

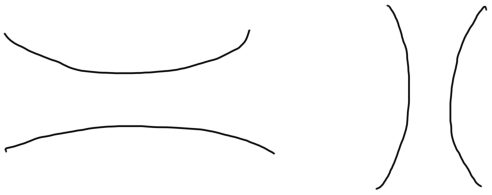
Hazem Hassan

*Very recently John Conway presented to me a result of his which I found to be extremely surprising. I find it astonishing that this result is not very well known (at least in my own experience). It is about a certain class of tangles, rational tangles, and in particular it turns out that they have a very interesting and well behaved structure to them.*

## 1 Introduction

**Note:** *a lot of the ideas in this article are best understood by playing with a pair of strings and transforming them, so I would highly recommend removing the shoelace of a pair of shoes to follow along with.*

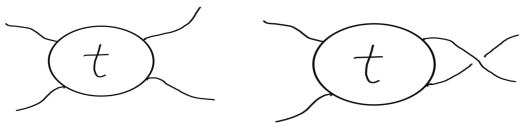
I'll start with a very non-rigorous introduction to the result, and will later make the proper definitions needed to prove it. We start by considering the following relatively arbitrary looking construction. Label the following tangles 0 and $\infty$, respectively.



(a) 0 tangle (b) $\infty$ tangle

While these may not look like tangles in the traditional sense, it is in the same spirit that we say 0 is a number. We also define a couple of actions on these tangles, namely *Twist* and *Turn*. Given a tangle labeled by some rational number $t$, a twist is accomplished by twisting the two ends on the right side of the tangle as shown below, the resulting tangle is then labeled $t + 1$.

A turn is defined to be a rotation of the entire tangle, including the ends, by 90° clockwise. For a tangle labeled $t$ the resulting tangle after turning is labeled $-\frac{1}{t}$.




(c) $t$ (d) $t+1$

Twist


(e) $t$ (f) $-1/t$

Turn

It is natural then to combine these actions to create numerous different tangles, and the astonishing result is that these labels turn out to be well defined, that is to say, if two tangles are equivalent[1] (in the sense that we can deform one into the other while keeping all four ends fixed, and without any cutting or gluing) then they will have the same label.


(g) 1 (h) 3 (i) $-1/3$

As a challenge I would recommend using twist and turn only to transform the $-\frac{1}{3}$ tangle shown above into a 0 tangle, which would result in a messy looking tangle, however since it has the same label as the 0 tangle defined above it must be the "same" tangle, and thus just pulling on all four ends, it will magically untangle itself into two separate strings.

## 2 Rational Tangles

Our goal now is to prove the above result, namely, if two tangles have the same label then they must be equivalent. We are only really interested in tangles that we can make using only twists and turns, and these are the so called *Rational Tangles*; these are the tangles that can be untangled through only twisting pairs of adjacent ends. In this

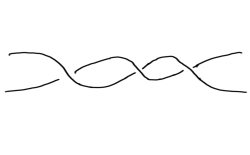
[1] The technical term for the equivalence here is ambient isotropy

section we will more formally define rational tangles and prove a very handy result on them.

We start of by considering a much simpler class of tangles, the tangles constructed by twisting the same pair of ends multiple times.

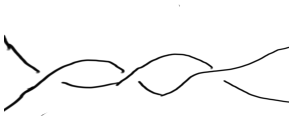
**Definition.** *A tangle t is a horizontal integer tangle if it is the result of applying twist$^a$ to the* 0 *tangle, and is denoted by $t_a$. A tangle is said to be a vertical integer tangle if it is the result of applying a turn to horizontal integer tangle, and is denoted by $t'_a$.*

Note that in the above definition $a$ is also allowed to be a negative integer, which simply corresponds to twisting in the other direction.



(j) $t_3$

(k) $t'_3$

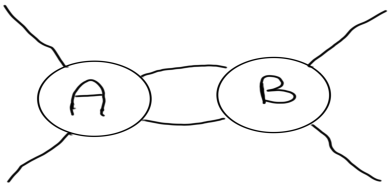


(l) $t_{-3}$

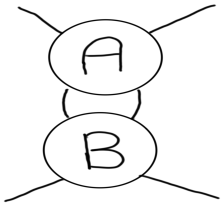
(m) $t'_{-3}$

We can then see that gluing $t_a$ to to the ends of a tangle on the right or left , is equivalent to twisting $a$ times, and similarly with $t'_a$ and the top or bottom ends, thus since rational tangles can be untangled via a sequence of twists, we can repeatedly "add" integer tangles to construct rational tangles, which motivates the following definition of addition.

**Definition.** *For any tangles $A$ and $B$ we define their sums $A + B$ and $A +' B$ as in the image below.*




$A + B$




$A +' B$

As one would expect $t_a + t_b = t_{a+b}$ and $t'_a +' t'_b = t'_{a+b}$.

**Definition.** *A tangle $t$ is said to be a rational tangle if one of the following conditions holds*

1. $t = t_a$ *or* $t = t'_a$
2. $t = t_a + B$, $t = B + t_a$, $t = t'_a +' B$ *or* $t = B +' t'_a$ *where $B$ is a rational tangle.*

This definition aligns with our intuition of a rational tangle being a finite amount of twists away from being two untangled strings.

**Theorem** (Flip Theorem)**.** *Let $t$ be a rational tangle, then a rotation of $t$ by* $180°$ *along the horizontal or vertical axis is equivalent to $t$.*

This theorem can be proved by an induction not too dissimilar to induction on natural numbers due to the recursive nature of our definition of rational tangles. To start with, it's easy to see that $t_a$ and $t'_a$ remain the same after a rotation. We assume then that $t = t_a + B$ where the $B$ is a rational tangle which is preserved by such rotations, then a rotation around the horizontal axis preserves both $t_a$ and $B$ so it's not too hard to see that $t_a + B$ is preserved by rotating around the horizontal axis.

The trickier part is the rotation around the vertical axis, which reduces to proving that $B + t_a$ is equivalent to $t_a + B$. We see that by rotating $B$ $a$ times around the horizontal axis while holding all the ends fixed that $B$ is equivalent to $t_{-a} + B + t_a$, as shown in the figure below


$$t_{-2} + t + t_2$$

Adding $t_a$ on the left of both tangles we see that $t_a + B$ is equivalent to $B + t_a$.

## 3 Basic Tangles

The above result that $B + t_a$ is equivalent to $t_a + B$ allows us to re-write tangles in multiples equivalent forms, for example

$$(t'_4 +' ((t'_2 +' t_1) + t_3)) + t_5 \sim (t'_4 +' ((t_1 +' t'_2) + t_3) + t_5 \sim (((t_1 +' t'_2) + t_3) +' t'_4) + t_5.$$

As we can see there are too many ways to write the same tangle as the sum of integer tangles and thus we would like to be able to have a standard way to write rational tangles.

**Definition.** *A horizontal basic tangle is a rational tangle that is constructed as follows*

1. *begin with some $t_a$*
2. *add some $t'_b$ to the bottom, then some $t_c$ to the right and repeat for a finite number of steps*

*A vertical basic tangle is a rational tangle that is constructed as follows*

1. *begin with some $t'_a$*
2. *add some $t_b$ to the left, then some $t'_c$ to the bottom and repeat for a finite number of steps*

**Example.**

$$(((t_1 +' t'_2) + t_3) +' t'_4) + t_5$$

*is a horizontal basic tangle,*

$$(t_1 + ((t_4 + t'_3) +' t'_5)) +' t'_2$$

*is a vertical basic tangle.*

It shouldn't come as a big surprise that every rational tangle is equivalent to some basic tangle, because they are constructed to be our standard way of writing rational tangles, by using the flip theorem.

# 4 Fractions

We have so far considered twists and the effect of adding them to rational tangles, we now switch our attention to turns.

**Definition.** *Let $t$ be a rational tangle then we define $-t$ to be the mirror image of $t$, and $-\frac{1}{t}$ to be the result of turning $t$.*

Note that $\frac{1}{t}$ is the result of a turn followed by taking the mirror image.

The following results are then not too hard to see by playing with a pair of strings and applying the flip theorem.

**Theorem.** *1. $b$ is a horizontal basic tangle if and only if $1/b$ is a vertical basic tangle*

*2. $t'_a = \frac{1}{t_a}$*

*3. if $t$ is a rational tangle then*

$$t +' t'_a = \frac{1}{t_a + \frac{1}{t}}$$

Using the above result we can now write basic tangles as fractions

**Example.**

$$(((t_1 +' t'_2) + t_3) +' t'_4) + t_5 \sim t_5 + \cfrac{1}{t_4 + \cfrac{1}{t_3 + \cfrac{1}{t_2 + \frac{1}{t_1}}}}.$$

$$(t_1 + ((t_4 + t'_3) +' t'_5)) +' t'_2 \sim \cfrac{1}{t_2 + \cfrac{1}{t_1 + \cfrac{1}{t_5 + \cfrac{1}{t_4 + \frac{1}{t_3}}}}}.$$

Fractions of this shape are called *continued fractions* which are a very well-studied subject in Number Theory, and for our purposes, they include both aspects of twists and turns from our original construction to represent basic tangles, and so it is natural to use them to define a tangle's label.

**Definition.** *let $t \sim t_{a_0} + \cfrac{1}{t_{a_1} + \cfrac{1}{t_{a_2} + \cfrac{1}{t_{a_3} + \cfrac{1}{t_{a_4} + \cdots}}}}$ be some basic tangle then, we define*

*its label* $F(t)$ *to be* $a_0 + \cfrac{1}{a_1 + \cfrac{1}{a_2 + \cfrac{1}{a_3 + \cfrac{1}{a_4 + \cdots}}}}$.

By construction this definition of a tangle's label matches with our earlier understanding of the label, thus we now have all the tools we need to prove the main result with the aid of a couple ideas from continued fractions.

**Theorem.** *Let* $t, s$ *be basic tangles. If* $F(t) = F(s)$ *then* $t$ *is equivalent to* $s$.

In order to note this we need to first use the fact that for every number there exists a unique continued fraction $a_0 + \cfrac{1}{a_1 + \cfrac{1}{a_2 + \cfrac{1}{a_3 + \cfrac{1}{a_4 + \cdots}}}}$ such that $a_i > 0$ for $i \geq 1$, and thus there is a unique basic tangle with $a_i > 0$ for $i \geq 1$, so our plan is to prove that every basic tangle is equivalent to such a basic tangle. To do so we are going to need to use the following formula,

$$a - \frac{1}{b} = (a-1) + \cfrac{1}{1 + \cfrac{1}{(b-1)}},$$

in fact we are going to need the integer tangle version, that is if $a$ and $b$ are rational tangles then

$$a + \frac{1}{-b} = (a + t_{-1}) + \cfrac{1}{t_1 + \cfrac{1}{b + t_{-1}}},$$

which isn't too hard to verify.

We use the above formula to move the negative numbers higher in the continued fractions.

**Example.** *let* $t \sim t_1 + \cfrac{1}{t_{-3} + \cfrac{1}{t_{-4} + \cfrac{1}{t_5}}} \sim t_1 - \cfrac{1}{t_3 + \cfrac{1}{t_4 + \cfrac{1}{t_{-5}}}}$. *Using the formula we get that*

$$t_4 + \frac{1}{t_{-5}} \sim t_3 + \cfrac{1}{t_1 + \cfrac{1}{t_4}}$$

*thus*

$$t \sim t_1 - \cfrac{1}{t_3 + \cfrac{1}{t_3 + \cfrac{1}{t_1 + \cfrac{1}{t_4}}}}.$$

*We can use the formula again with* $a = t_1$ *and* $b = t_3 + \cfrac{1}{t_3 + \cfrac{1}{t_1 + \cfrac{1}{t_4}}}$ *to get*

$$t \sim \cfrac{1}{t_1 + \cfrac{1}{t_2 + \cfrac{1}{t_3 + \cfrac{1}{t_1 + \cfrac{1}{t_4}}}}}.$$

We have thus demonstrated that starting out with a rational tangle, we can use the Flip Theorem to transform it to an equivalent basic tangle, which we can again transform to a basic tangle that is unique for every label, and so if $F(s) = F(t)$ then they are equivalent.

Enjoyed the read?
Now it's your turn!
Have you ever seen an interesting lemma that you want to share? Do you want to have a go at writing an article about Mathematics?
Write for us! We're waiting for your contribution.
Please contact
magazine@invariants.org.uk

The Invariants Society is Oxford University's student society for Mathematics. We're here to promote Maths and we host weekly informal lectures often given by leading mathematicians. To find out more, please see our website
http://www.invariants.org.uk/ or find us on Facebook at
https://www.facebook.com/oxford.invariants/.
Loved an article? Hated it? Do share! Send us your comments and we'll get back to you next term!

# Learning how to code?

Join **CodeSoc**: Oxford University Coding Society

- Free weekly coding classes
- Free social coding sessions
- A network of code-mentors to help you out

Open to everyone, from beginners to advanced coders.

Find out more on www.codesoc.co.uk or find us on Facebook!

$\omega^{\pi}\sqrt{i\text{nvariants}}$

The Invariants Society is sponsored by

Jane Street

tpp