THE MORITA 2-CATEGORY OF A STRICT MONOIDAL CATEGORY

SAM SILVER

Contents

Introduction 1
1. Category theory background 3
1.1. Categories 3
1.2. Functors 4
1.3. Isomorphisms 5
1.4. Natural transformations 5
1.5. Coequalizers and Fubini's theorem 7
2. 2-categories 9
2.1. Definition and examples 10
2.2. Pasting diagrams 11
2.3. 2-functors 11
3. Monoidal categories 11
3.1. Definition and examples 11
3.2. The Eckmann-Hilton argument 13
3.3. Monoids 14
3.4. Modules over monoids 15
3.5. Balanced tensor product of modules over monoids 17
4. The Morita (2-)category 21
4.1. The "upper" categories 21
4.2. Functoriality of composition 22
4.3. Unitality and associativity 23
4.4. Examples 24
5. Appendix 25
5.1. The strictification theorem for monoidal categories 25
5.2. The path 2-groupoid of a space X 25
References 28

Introduction

To some, mathematics is the study of numbers, and relationships between them. To others, it's the study of variation in numbers; to others still, it's the study of shapes and spaces, and how they fit together. The list goes on, but one aspect all these things have in common is *structure*. One might venture to characterize mathematics as *the study of structured things*. But then this begs a natural question: what *is* structure? What's the structure of structure itself?

Reflecting on this, one observation one might make is that we often have *things*, as well as *relationships* between them. We have equations and inequalities of numbers, flows from point to point in vector fields, maps of spaces and paths within them. Such things are what lend structure to our objects of study, and indeed they themselves often become our main objects of study. These are now notions that lend themselves to attempts at formalization, which brings us back into mathematics, in a relatively new

Date: Fall 2025.

subject known as *category theory*. A *category* is, roughly speaking, a collection of objects and arrows (or *morphisms*) between them, with a rule allowing you to *compose* consecutive arrows, in a manner in which bracketing of composition is irrelevant, and there is a "do-nothing" arrow for each object. These objects and arrows can be nearly anything: sets and functions, numbers and inequalities, spaces and maps, groups and homomorphisms, points and paths... the list goes on.

Category theory has been used to study and streamline a wide variety of fields of mathematics, including many aspects of geometry and topology, and certainly most aspects of abstract algebra. One notion that has seen widespread use and has blossomed into its own field of study is that of a *monoidal category*. Roughly speaking, this is a category equipped with an associative and unital multiplication operation which acts on objects and arrows in a manner compatible with the structure. These have broad applications to geometry, mathematical physics, commutative algebra, and more, but one thing that's special about them is that they are the perfect setting in which to define *monoids* and several associated types of objects.

Monoids are the general form of sets or other objects equipped with an associative and unital multiplication. They can act on things in a multiplicative way, which we call *modules*. The quintessential (and motivating) example of this comes from commutative algebra (and noncommutative), where rings are monoids in a certain monoidal category (abelian groups with tensor product), and they can act on modules by multiplication. Rings and modules form some of the most fundamental objects of study in algebraic geometry, subsuming notions of spaces, covers, bundles and projections, etc.

In the 1950s, the mathematician Kiiti Morita began studying rings using a new kind of equivalence relation, now known as *Morita equivalence* [10, 9]. Two rings R, S are Morita equivalent if the category of modules over R is (additively) equivalent to the category of modules over S. It has proven to be a useful tool and, more generally, philosophy in a number of geometric applications (see e.g. [8]). This equivalence has a natural categorical interpretation as the typical notion of categorical equivalence in a certain bicategory (a structure analogous to a category with a notion of "arrow between arrows", or 2-morphisms) called the *Morita bicategory* associated to a monoidal category.

The details of the construction of this category are folklore, however. The goal of this thesis is to make them fully precise, in complete explicit detail, for the case in which our monoidal category is *strict*, meaning there are no extra "coherence" data to deal with. In this case, we can avoid having to choose a specific model of higher categories by showing the stronger statement that for a strict monoidal category satisfying some mild constraints, we obtain a strict 2-category (a very tame type of bicategory) as the Morita bicategory. To be precise, we will prove the following theorem (Theorem 4.1):

Theorem 0.1. Let $\mathsf{C}^{\otimes}$ be a strict monoidal category admitting a calculus of balanced tensor products. There exists a 2-category $\mathfrak{M}(\mathsf{C}^{\otimes})$, called the *Morita 2-category of* $\mathsf{C}^{\otimes}$, in which

- the objects are monoids $A, B, C, \ldots$,
- the morphisms $A \to B$ are (A, B)-bimodules $M, N, P, \ldots$, and
- the 2-morphisms between (A, B)-bimodules $M \to N$ are (A, B)-bimodule homomorphisms.

with composition given by

- for 1-morphisms: balanced tensor product of bimodules, and
- for 2-morphisms: composition in C.

This thesis accompanies a paper [5] written in collaboration with my advisor Martina Rovelli and her graduate students Arghan Dutta and Stefano Luneia, in which we deal with the case of a general monoidal category (with the same mild constraints) by explicitly constructing a Morita $(\infty, 2)$-category in the complicial sets model of (∞, n)-categories. Some of the content overlaps and much of it is similar in spirit, but the techniques involved often differ substantially. We also offer herein a discussion of the Eckmann-Hilton argument for monoids and a detailed, explicit construction and treatment of a strict form of path 2-groupoid associated to a topological space.

1. Category theory background

We begin by introducing the necessary basic background for our construction. Most of the definitions in Sections 1 to 3 can be found in the excellent textbook [7] by Saunders Mac Lane, and for this section, much appears in the newer and also excellent textbook [11] by Emily Riehl.

1.1. Categories.

Definition 1.1. A *category* C consists of

- a collection $\operatorname{Ob}\mathsf{C}$ whose elements we call *objects of* C, as well as
- for any objects $X, Y \in \operatorname{Ob}\mathsf{C}$, a collection $\mathsf{C}(X,Y)$ called the *hom-set from X to Y*, with elements $f \in \mathsf{C}(X,Y)$ known as *morphisms* or *maps from X to Y*, usually denoted $f\colon X \to Y$, and
- for any $X, Y, Z \in \operatorname{Ob}\mathsf{C}$, a function $\mathsf{C}(X,Y) \times \mathsf{C}(Y,Z) \to \mathsf{C}(X,Z)$ called *composition*, usually denoted $(f,g) \mapsto g \circ f\colon X \to Z$ for $f\colon X \to Y$, $g\colon Y \to Z$

such that

- for each $X \in \operatorname{Ob}\mathsf{C}$, there exists some morphism $1_X\colon X \to X$, called the *identity* on X, such that for all $Y \in \operatorname{Ob}\mathsf{C}$ and any $f\colon X \to Y$ or $g\colon Y \to X$, we have $f \circ 1_X = f$ and $1_X \circ g = g$ (that is, composition is "unital"), and
- for any $W, X, Y, Z \in \operatorname{Ob}\mathsf{C}$ and $f\colon W \to X$, $g\colon X \to Y$, $h\colon Y \to Z$, we have $(h \circ g) \circ f = h \circ (g \circ f)$ (that is, composition is "associative").

This definition is just trying to capture our basic intuition about a few familiar examples of "structures".

Example 1.2. Let's take a look. We'll use $\{*\}$ to denote a set with one element.

- There is a category Set whose objects are sets, and whose morphisms $\mathsf{Set}(X,Y)$ consist of all functions $f\colon X \to Y$, with composition being the usual composition of functions. As we know, function composition is associative, and every set X has an identity function $1_X\colon X \to X$. This setting (no pun intended) is where the $\circ$ notation comes from.
- Similarly, there's a category Top whose objects are topologial spaces and whose morphisms are continuous maps.
- There is a category $\mathsf{Vect}_{\mathbb{R}}$ whose objects are real vector spaces and whose morphisms $\mathsf{Vect}_{\mathbb{R}}(V,W)$ are linear maps $T\colon V \to W$. Relatedly, there's a category $\mathsf{Mat}_{\mathbb{R}}$ whose objects are natural numbers and whose morphisms $m \to n$ are $m \times n$ matrices, with composition given by matrix multiplication.
- There is a category $\mathbb{Z}_{\leq}$ with $\operatorname{Ob}\mathbb{Z}_{\leq} = \mathbb{Z}$, and hom-sets given by
$$\mathbb{Z}_{\leq}(m,n) = \begin{cases} \{*\} & \text{if } m \leq n \\ \emptyset & \text{if } m > n \end{cases}$$
That is, the morphisms are just (weak) inequalities $m \leq n$ of integers. Identities come from the reflexive property $m \leq m$, and composition is unique if it exists, with its existence guaranteed by the transitive property $k \leq m, m \leq n \implies k \leq n$. The same construction can be made for any preordered set.
- For any topological space X, there is a category $\Pi_1(X)$, called the *fundamental groupoid* of X, whose objects are points $p \in X$, and whose morphisms $p \to q$ are path-homotopy classes of paths $\gamma\colon [0,1] \to X$ with $\gamma(0) = p$ and $\gamma(1) = q$. Composition is given by joining the paths end-to-end, and is unital and associative up to path-homotopy, with identities given by the classes of constant paths.
- For any set X, there is a category $\operatorname{disc} X$, with objects $\operatorname{Ob}\operatorname{disc} X = X$, and morphisms
$$(\operatorname{disc} X)(x,y) = \begin{cases} \{*\} & \text{if } x = y \\ \emptyset & \text{if } x \neq y \end{cases}$$
That is, there are no morphisms between different objects, and only one from any object to itself, which is trivially an identity under the unique choice for composition, which is trivially associative.

- For any group G, there is a category BG with just one object $\operatorname{Ob} BG = \{*\}$, and morphisms $BG(*,*) = G$, with composition given by the group multiplication. Since by definition, group multiplication is associative and unital, those properties hold for BG as well. This works more generally for any monoid.
- For any category C, we can form a category C^{op}, called the *opposite* of C, with the same objects, and where we set $\mathsf{C}^{\mathrm{op}}(X,Y) = \mathsf{C}(Y,X)$ and reverse the order of composition: $f \circ^{\mathrm{op}} g = g \circ f$, where $\circ^{\mathrm{op}}$ is composition in C^{op} and $\circ$ is composition in C. We can also take two categories C, D and form their product category $\mathsf{C} \times \mathsf{D}$, where objects and morphisms are ordered pairs of ones in C, D, and composition is component-wise.

So we see that categories can codify sets (with structure) and (structure-preserving) maps between them, but also might represent other kinds of structure altogether. What they all have in common is a notion of objects and various relationships between them, such that relationships can be composed in a coherent way.

1.2. **Functors.** In category theory, there are a host of informal "tenets" which philosophically guide the practitioner. The first of these we come upon is this: all meaningful structure is contained in the morphisms; even if we care primarily about the objects, we can and should focus on the relationships between them, with the objects as secondary data. For example: in Set, instead of using object-specific language like $x \in X$ or $Y \subseteq X$, we can phrase this and more in terms of maps, such as $x\colon \{*\} \to X$ with $* \mapsto x$ which picks out $x \in X$, or the inclusion $\imath\colon Y \to X$.

Under this philosophy, the first thing we should do when we come across a type of object is to ask: what are the appropriate morphisms between such objects? And indeed, we have just defined a type of object, *categories*. Here's the commonly accepted notion of "morphism of categories".

Definition 1.3. Let C, D be categories. A *functor* $F\colon \mathsf{C} \to \mathsf{D}$ from C to D consists of a map $\operatorname{Ob}\mathsf{C} \to \operatorname{Ob}\mathsf{D}$ and for all $X, Y \in \operatorname{Ob}\mathsf{C}$, maps $\mathsf{C}(X,Y) \to \mathsf{D}(FX, FY)$ (all denoted by F, by abuse of notation) such that

- For all $X \in \operatorname{Ob}\mathsf{C}$, we have $F(1_X) = 1_{F(X)}\colon F(X) \to F(X)$ and
- For all $X, Y, Z \in \operatorname{Ob}\mathsf{C}$ and morphisms $f\colon X \to Y$, $g\colon Y \to Z$, we have $F(g \circ f) = F(g) \circ F(f)\colon F(X) \to F(Z)$.

Example 1.4. Let's see some examples:

- For any category C, the identity maps on $\operatorname{Ob}\mathsf{C}$ and each $\mathsf{C}(X,Y)$ define a functor 1_{C} called the *identity functor*.
- Given any three categories $\mathsf{C}_1, \mathsf{C}_2, \mathsf{C}_3$ and functors $F\colon \mathsf{C}_1 \to \mathsf{C}_2$ and $G\colon \mathsf{C}_2 \to \mathsf{C}_3$, we obtain a functor denoted $G \circ F\colon \mathsf{C}_1 \to \mathsf{C}_3$ defined by composing the maps on objects and those on hom-sets. It's easy to check that this and the previous example satisfy the definition above.
- There is a functor $F\colon \mathsf{Set} \to \mathsf{Group}$ which maps a set X to the *free group* $F(X)$ of words in the elements of X and their formal inverses under concatenation. A map of sets $X \to Y$ induces a group homomorphism $F(X) \to F(Y)$, since it tells us where the generators go.
- There is a functor $U\colon \mathsf{Group} \to \mathsf{Set}$ given by "forgetting" the group structure: $U(G)$ is the underlying set of the group. Since a group homomorphism $G \to H$ is a map of sets with an extra condition, we can forget the condition to obtain a map of sets $U(G) \to U(H)$, too.
- There is a functor $(-)^*\colon \mathsf{Vect}_{\mathbb{R}}^{\mathrm{op}} \to \mathsf{Vect}_{\mathbb{R}}$ which takes a vector space V to its dual V^* and a linear map $T\colon V \to W$ to its transpose $T^*\colon W^* \to V^*$. We needed to put in the $^{\mathrm{op}}$ since transposing a map swaps its source and target. This is usually called a *contravariant functor* $\mathsf{Vect}_{\mathbb{R}} \to \mathsf{Vect}_{\mathbb{R}}$ (the typical ones we've seen are called *covariant*).

After having seen the first two examples above, one can verify that our categorical hopes are realized:

Proposition 1.5. There is a category Cat whose objects are categories and whose morphisms are functors, with functor composition as above.

And now we have three more examples of functors: check that in the examples of categories above, $\Pi_1\colon \mathsf{Top} \to \mathsf{Cat}$, $\mathrm{disc}\colon \mathsf{Set} \to \mathsf{Cat}$, and $B\colon \mathsf{Group} \to \mathsf{Cat}$ are all Cat-valued functors.

Remark 1.6. Technically, there is a slight issue with our definition here. In our definition of "category", we said there are "collections" of objects and morphisms, without specifying what kind of collections: sets, classes, etc. In most reasonable axiom systems, if we choose one of these, there cannot be a category whose "collection" of objects is "all categories", for the same reason that there cannot be a set of all sets. To make this precise, we can define a hierarchy of categories, each of whose objects are all categories with their notion of "collection" bounded by some notion, and go up a level whenever necessary (such as above). But this level of technical precision will not be useful for us, so we will ignore it altogether, as it does not affect any of the content we will discuss.

1.3. **Isomorphisms.** Now, we'll begin to talk about *invertibility*. Each category comes with an intrinsic notion of what it means for a morphism to be "invertible".

Definition 1.7. Let C be a category, and $f\colon X \to Y$ be a morphism. We say f is an *isomorphism* if there exists a morphism $g\colon Y \to X$ such that $g \circ f = 1_X$ and $f \circ g = 1_Y$. In this circumstance, one often denotes g by f^{-1} and calls it the *inverse* of f. If there exists such an f for some objects X and Y, then we say $X \cong Y$ (X is *isomorphic* to Y).

We briefly verify the following:

Proposition 1.8. Inverses are unique (when they exist). That is, given $f\colon X \to Y$ in some category C, if $g_1, g_2\colon Y \to X$ both satisfy $g_i \circ f = 1_X$ and $f \circ g_i = 1_Y$, then $g_1 = g_2$.

Proof. We have

$$g_1 = g_1 \circ 1_Y = g_1 \circ (f \circ g_2) = (g_1 \circ f) \circ g_2 = 1_X \circ g_2 = g_2$$

as desired. □

Thus, our terminology of "the" inverse of f is justified. However, we should be wary of calling two isomorphic objects "the same", since there may be multiple isomorphisms between any two objects, and the statement "$X \cong Y$" tells us nothing about this. If there are multiple dictionaries between two objects, and I choose a different one from you and we try to translate back and forth, the result after one cycle may differ from the initial state. However, if two objects are *uniquely* isomorphic, we have better reason to regard them as "the same".

Example 1.9. We already use the term "isomorphism" in other contexts, like "group isomorphism" or "linear isomorphism". We see that this notion recovers those, and gives us the right analogue in other settings:

- A group homomorphism $\varphi\colon G \to H$ is an isomorphism in Group if and only if it is a group isomorphism, that is, if it's a bijective group homomorphism.
- Similarly, the isomorphisms in $\mathsf{Vect}_{\mathbb{R}}$ are the bijective linear maps.
- Isomorphisms in Top are homeomorphisms, and they have precisely the same definition as above. Notably, they are not just bijective continuous maps: the map $\exp(i-)\colon [0,1) \to S^1$ is a counterexample.
- We obtain a notion of isomorphism in Cat, called *isomorphism of categories*.
- For any topological space X, every morphism in $\Pi_1(X)$ is an isomorphism, since all paths can be inverted. We call such categories *groupoids*.
- BG is another groupoid, for any group G. So is $\operatorname{disc} X$ for any set X.

We note that the composition of any two isomorphisms is an isomorphism, and the identity is an isomorphism, so we obtain:

Proposition 1.10. For any category C, we can form a category (in fact a groupoid) called the *core* of C, whose objects are the same and whose morphisms are just the isomorphisms in C. Furthermore, this construction is a functor $\mathsf{Cat} \to \mathsf{Cat}$.

1.4. **Natural transformations.** Now, one might think this is a nice place to stop, but there's so much more, and the next notion captures another vital piece of categorical philosophy: that of *higher-dimensional structure*. Why not treat morphisms as objects themselves, and try to put morphisms between them? Just as we can deform one map into another in topology, it turns out we can do the same to functors.

Definition 1.11. Let C, D be categories and $F, G\colon \mathsf{C} \to \mathsf{D}$ be functors. A *natural transformation* $\eta\colon F \Rightarrow G$ consists of, for each object $X \in \mathsf{C}$, a morphism $\eta_X\colon F(X) \to G(X)$ in D, such that for any morphism $f\colon X \to Y$ in C, we have $\eta_Y \circ F(f) = G(f) \circ \eta_X$. This is usually depicted by saying that "the following square commutes":

$$\begin{array}{ccc} F(X) & \xrightarrow{F(f)} & F(Y) \\ {\scriptstyle \eta_X}\downarrow & & \downarrow{\scriptstyle \eta_Y} \\ G(X) & \xrightarrow[G(f)]{} & G(Y) \end{array}$$

meaning that no matter which path one takes from the upper left to the lower right, the composition of morphisms along the path yields the same thing.

Example 1.12. Let's see some examples:

- For any functor $F\colon \mathsf{C} \to \mathsf{D}$, taking $\eta_X = 1_{F(X)}$ yields an "identity natural transformation" $\eta\colon F \Rightarrow F$.
- For any functors $F, G, H\colon \mathsf{C} \to \mathsf{D}$ and natural transformations $\eta\colon F \Rightarrow G$, $\nu\colon G \Rightarrow H$, we obtain a natural transformation $\nu \circ \eta\colon F \Rightarrow H$ with components $(\nu \circ \eta)_X = \nu_X \circ \eta_X$ (note that the commutative squares "glue together" to form the one we want).
- Recall the functors $F\colon \mathsf{Set} \to \mathsf{Group}$ and $U\colon \mathsf{Group} \to \mathsf{Set}$ from earlier. There is a natural transformation $\eta\colon 1_{\mathsf{Set}} \Rightarrow U \circ F$ in which η_X maps each element of X to itself in the free group (viewed as a set). Similarly, there is a natural transformation $\nu\colon F \circ U \Rightarrow 1_{\mathsf{Group}}$ in which $\nu_G\colon F(U(G)) \to G$ is defined on generators, which correspond to elements of G, by mapping them to the corresponding elements of G.
- There is a natural isomorphism $((-)^*)^* \Rightarrow 1_{\mathsf{Vect}_{\mathbb{R}}}$ which, for any vector space V, takes a linear functional $\psi \in (V^*)^*$ on linear functionals to the unique vector $v \in V$ such that $\psi(f) = f(v)$ for any linear functional $f \in V^*$. However, there is no natural isomorphism $(-)^* \Rightarrow 1_{\mathsf{Vect}_{\mathbb{R}}}$, even if we restrict to finite-dimensional vector spaces, despite the fact that $V^* \cong V$ for every finite-dimensional vector space V.

Natural transformations are sometimes called "morphisms of functors", and for good reason. By the first two examples above, we can verify:

Proposition 1.13. For any two categories C, D, there is a category D^{C} whose objects are functors $F\colon \mathsf{C} \to \mathsf{D}$ and whose morphisms are natural transformations, with composition as above.

This is strange - before the functors were our morphisms, and now they're our objects! We make sense of this by interpreting our data in a geometric way: the objects (categories) are like points, with morphisms (functors) like paths between them. Then morphisms between morphisms (natural transformations) are like paths between paths - that is, they are path homotopies, which we think of as 2-dimensional objects! An important difference is that unlike paths, morphisms are not always invertible, but the analogy is still incredibly useful for thinking about higher category theory.

Another (related) geometric perspective sees categories like topological spaces, functors as continuous maps, and natural transformations as homotopies of maps. The commutative squares in the definition are like "2-dimensional patches" that fit together to form the homotopy, and this analogy can be made precise. Taking this one step further to encompass the notion of homotopy equivalence from topology, we rediscover another important notion in category theory:

Definition 1.14. Let C, D be categories, and $F\colon \mathsf{C} \to \mathsf{D}$ be a functor. We say F is an *equivalence of categories* if there exists a functor $G\colon \mathsf{D} \to \mathsf{C}$ and natural isomorphisms $\epsilon\colon F \circ G \Rightarrow 1_{\mathsf{D}}$ and $\eta\colon 1_{\mathsf{C}} \Rightarrow G \circ F$. If there exists such an F, we say C and D are *equivalent*.

Example 1.15. Examples of equivalences abound—it's a very useful notion. Here are a few:

- Any category is equivalent to itself via the identity functor, and the composition of equivalences is an equivalence (it's a quick exercise to verify this). Furthermore, any isomorphism is an equivalence of categories, via the identity natural transformations.

- There's a functor $I\colon \mathsf{Mat}_{\mathbb{R}} \to \mathsf{FinVect}_{\mathbb{R}}$ (the category of finite-dimensional real vector spaces with linear maps) given on objects by $n \mapsto \mathbb{R}^n$ and on morphisms by taking an $n \times m$ matrix to the corresponding linear map $\mathbb{R}^n \to \mathbb{R}^m$ in the standard basis. This functor is an equivalence of categories, by choosing a basis for every finite-dimensional real vector space and using this to turn all linear maps into matrices (yielding a functor $\mathsf{FinVect}_{\mathbb{R}} \to \mathsf{Mat}_{\mathbb{R}}$ which one can check satisfies the above definition).
- For any category C and any object X in it, we obtain a group $\operatorname{Aut} X$ (called the *automorphism group of* X) given by the set of isomorphisms $X \to X$ under composition. For any connected groupoid G (meaning all objects are isomorphic) and any object X in G, the inclusion $B \operatorname{Aut} X \hookrightarrow G$ is an equivalence of categories, since picking an isomorphism $X \cong Y$ for every object Y lets us turn maps $Y \to Z$ into maps $X \to X$ via the compositions $X \cong Y \to Z \cong X$ (yielding a functor $\mathsf{G} \to B \operatorname{Aut} X$ which one can check satisfies the above definition). In particular, $\Pi_1(X)$ is equivalent to $B\pi_1(X)$ for any space X.
- Hilbert's *Nullstellensatz* establishes an equivalence of categories between (the opposite of) the category of reduced, finitely-generated k-algebras and the category of affine varieties over k, for any algebraically closed field k. Since the functors involved are contravariant, we call this a "duality".

As one explores more and more of the world of mathematics, one quickly finds that often, when we want to say two kinds of things are "the same", we cannot do so easily, because e.g. we have some non-canonical representation we're using, or somehow our correspondence only holds "up to isomorphism", and doesn't yield a true bijection. Consequently, we frequently find ourselves reaching for equivalences of categories, rather than isomorphisms thereof, to describe a dictionary between one setting and another.

1.5. **Coequalizers and Fubini's theorem.** The following is a construction on which we'll rely heavily. We will use it to define the so-called "balanced tensor product of modules" (Definition 3.19). It's a special case of a more general construction called a "colimit", but we won't go into that here.

Definition 1.16. Let C be a category, $X, Y \in \operatorname{Ob} \mathsf{C}$, and let $f, g\colon X \rightrightarrows Y$ be two morphisms. We say a morphism $\pi\colon Y \to C$ for some object C is a *coequalizer* of f and g if $\pi \circ f = \pi \circ g$ and for any $h\colon Y \to Z$ (for any object Z) with $h \circ f = h \circ g$, there is a *unique* morphism $h'\colon C \to Z$ such that $h = h' \circ \pi$. We can depict this using the following diagram:

$$\begin{array}{ccccc} X & \overset{f}{\underset{g}{\rightrightarrows}} & Y & \xrightarrow{\pi} & C \\ & & & \searrow^{h} & \downarrow{\scriptstyle \exists! h'} \\ & & & & Z \end{array}$$

The object C is denoted $\operatorname{coeq}(f, g)$. Often, one calls C a coequalizer and leaves the morphism implicit.

For a given parallel pair of morphisms in some category, a coequalizer may not exist. But if one does, it is "essentially unique". That is, there may be multiple coequalizers, but:

Proposition 1.17. If $\pi\colon Y \to C$, $\pi'\colon Y \to C'$ are two coequalizers of a parallel pair $f, g\colon X \rightrightarrows Y$, then there is a unique isomorphism $\varphi\colon C \to C'$ such that $\pi' = \varphi \circ \pi$.

Proof. By definition, since C is a coequalizer and $\pi' \circ f = \pi' \circ g$, there exists a unique $\varphi\colon C \to C'$ such that $\pi' = \varphi \circ \pi$. But similarly, since C' is a coequalizer and $\pi \circ f = \pi \circ g$, there exists a unique $\varphi'\colon C' \to C$ such that $\pi = \varphi' \circ \pi'$. Then, we compute

$$\pi = \varphi' \circ \pi' = \varphi' \circ \varphi \circ \pi$$

So we see that 1_C and $\varphi' \circ \varphi$ are both morphisms $a\colon C \to C$ such that $\pi = a \circ \pi$. By the uniqueness of such morphisms in the definition, we have $\varphi' \circ \varphi = 1_C$. A symmetric argument yields $\varphi \circ \varphi' = 1_C$, so we conclude that φ is an isomorphism, as desired. □

Thus, we're justified in speaking of "the" coequalizer of a parallel pair, since there's no ambiguity in translating between representatives. In the above definition and proof, we're seeing the first hints of another guiding principle in category theory: *universality*. Coequalizers are something defined by a

"universal property": in this case, $\mathrm{coeq}(f,g)$ is the *universal* morphism which sees f and g as equal, in the sense that it "comes before" any other such morphism in a unique way. Universally-defined things are always unique up to unique isomorphism (but may not always exist).

In the spirit of the categorical principles we've been seeing, it turns out coequalizers are, in a certain sense, functorial; and this is something we will make extensive use of.

Definition 1.18. A *map of coequalizer diagrams* from $f,g\colon X \rightrightarrows Y$ to $f',g'\colon X' \rightrightarrows Y'$ consists of a map $h\colon X \to X'$ and a map $k\colon Y \to Y$ such that $k\circ f = f'\circ h$ and $k\circ g = g'\circ h$, that is, the "upper" and "lower" squares in the following diagram commute:

$$\begin{array}{ccc} X & \overset{f}{\underset{g}{\rightrightarrows}} & Y \\ {\scriptstyle h}\downarrow & & \downarrow{\scriptstyle k} \\ X' & \overset{f'}{\underset{g'}{\rightrightarrows}} & Y' \end{array}$$

It is not hard to check that for any category C there is a category $\mathrm{coeq}(\mathsf{C})$ whose objects are coequalizer diagrams and whose morphisms are maps of coequalizer diagrams. We can also restrict to the category $\mathrm{coeq}_0(\mathsf{C})$ of coequalizer diagrams for which the coequalizer exists in C.

Proposition 1.19. The coequalizer construction extends to a functor $\mathrm{coeq}\colon \mathrm{coeq}_0(\mathsf{C}) \to \mathsf{C}$.

Proof. First, we must verify that given a map of coequalizer diagrams, there is an induced map of coequalizers. Suppose we have the following setup:

$$\begin{array}{ccc} X & \overset{f}{\underset{g}{\rightrightarrows}} & Y \\ {\scriptstyle h}\downarrow & & \downarrow{\scriptstyle k} \\ X' & \overset{f'}{\underset{g'}{\rightrightarrows}} & Y' \end{array}$$

as in Definition 1.18. Let $\pi'\colon Y' \to \mathrm{coeq}(f',g')$ be the coequalizer map. Note that $\pi'\circ k$ coequalizes f and g, since

$$\begin{aligned} \pi'\circ k\circ f &= \pi'\circ f'\circ h \\ &= \pi'\circ g'\circ h \\ &= \pi'\circ k\circ g \end{aligned}$$

by the coequalizer property and one of the commuting squares. Thus $\pi'\circ k$ factors uniquely through the coequalizer map $\pi\colon Y \to \mathrm{coeq}(f,g)$, giving us our desired induced map $\mathrm{coeq}(f,g) \to \mathrm{coeq}(f',g')$.

Now we must verify that this process respects composition. Looking at the definition above, it is clear that if h and k are identities 1_X and 1_Y, then $\pi'\circ k = \pi' = \pi$ and the unique factorization of one through the other is through the identity $1_{\mathrm{coeq}(f,g)}$. If we have a further map of coequalizer diagrams, the composite of the induced maps is the map induced by the composites, since both maps will factor the same morphism in C through the same coequalizer map, and this factorization must be unique by definition of coequalizer. Therefore, we conclude that coeq is a functor. □

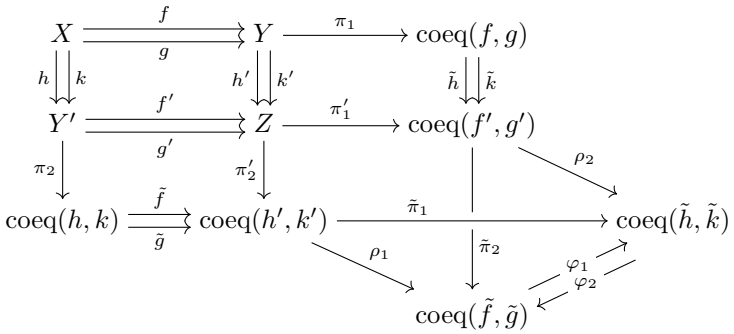
There's a certain statement about coequalizers which will be quite important for us later, as it corresponds to a notion of "associativity" for the balanced tensor product of modules (Proposition 4.5). It's called "Fubini's theorem", and though it may not look like it, it is indeed "the same" as the Fubini's theorem you may know from real analysis, in the sense that they are both instances of a common, vastly more general statement (as most things are in this field).

Theorem 1.20. Suppose we have the following setup:

$$\begin{array}{ccc} X & \overset{f}{\underset{g}{\rightrightarrows}} & Y \\ {\scriptstyle h}\downarrow\downarrow{\scriptstyle k} & & {\scriptstyle h'}\downarrow\downarrow{\scriptstyle k'} \\ Y' & \overset{f'}{\underset{g'}{\rightrightarrows}} & Z \end{array}$$

By this we mean, we're given the depicted maps, such that four squares commute: $h' \circ f = f' \circ h$, $h' \circ g = g' \circ h$, $k' \circ f = f' \circ k$, and $k' \circ g = g' \circ k$. Then, assuming all relevant coequalizers exist, there are induced maps $\tilde{h}, \tilde{k} \colon \operatorname{coeq}(f,g) \to \operatorname{coeq}(f',g')$ and $\tilde{f}, \tilde{g} \colon \operatorname{coeq}(h,k) \to \operatorname{coeq}(h',k')$, and there is a canonical isomorphism $\operatorname{coeq}(\tilde{f},\tilde{g}) \cong \operatorname{coeq}(\tilde{h},\tilde{k})$. That is, it doesn't matter in which order we compute the coequalizers: we obtain the same answer when computing the coequalizer across the rows, then the columns, as when we compute it across the columns, then the rows.

Proof. As a visual aid, this diagram may help one follow the proof.



The existence of $\tilde{f}, \tilde{g}, \tilde{h}, \tilde{k}$ was established above in Proposition 1.19. Furthermore, for the same reason, π_1 and π_1' induce a map $\tilde{\pi}_1 \colon \operatorname{coeq}(h',k') \to \operatorname{coeq}(\tilde{h},\tilde{k})$. Doing the same thing with f, g, h, k swapped yields a map $\tilde{\pi}_2 \colon \operatorname{coeq}(f',g') \to \operatorname{coeq}(\tilde{f},\tilde{g})$.

Next, let $\pi_2 \colon Y' \to \operatorname{coeq}(h,k)$ be the coequalizer map. Observe that

$$\begin{aligned}
\tilde{\pi}_1 \circ \tilde{f} \circ \pi_2 &= \tilde{\pi}_1 \circ \pi_2' \circ f' \\
&= \rho_2 \circ \pi_1' \circ f' \\
&= \rho_2 \circ \pi_1' \circ g' \\
&= \tilde{\pi}_1 \circ \pi_2' \circ g' \\
&= \tilde{\pi}_1 \circ \tilde{g} \circ \pi_2
\end{aligned}$$

by definition of $\tilde{\pi}_1$ and $\tilde{f}, \tilde{g}$. Since the map $\tilde{\pi}_1 \circ \tilde{f} \circ \pi_2$ clearly coequalizes h, k, it must factor uniquely through π_2, so we obtain $\tilde{\pi}_1 \circ \tilde{f} = \tilde{\pi}_1 \circ \tilde{g}$. Thus $\tilde{\pi}_1$ coequalizes $\tilde{f}, \tilde{g}$, and hence induces a map $\varphi_1 \colon \operatorname{coeq}(\tilde{f},\tilde{g}) \to \operatorname{coeq}(\tilde{h},\tilde{k})$. Doing the same thing again with f, g, h, k swapped yields a map $\varphi_2 \colon \operatorname{coeq}(\tilde{h},\tilde{k}) \to \operatorname{coeq}(\tilde{f},\tilde{g})$.

Finally, we check that

$$\begin{aligned}
\varphi_2 \circ \varphi_1 \circ \rho_1 \circ \pi_2' &= \varphi_2 \circ \tilde{\pi_1} \circ \pi_2' \\
&= \varphi_2 \circ \rho_2 \circ \pi_1' \\
&= \tilde{\pi}_2 \circ \pi_1' \\
&= \rho_1 \circ \pi_2'
\end{aligned}$$

so by uniqueness of factorization through π_2', we have $\varphi_2 \circ \varphi_1 \circ \rho_1 = \rho_1$. Then, by uniqueness of factorization through ρ_1, we have $\varphi_2 \circ \varphi_1 = 1_{\operatorname{coeq}(\tilde{f},\tilde{g})}$. By an analogous argument, we obtain $\varphi_1 \circ \varphi_2 = 1_{\operatorname{coeq}(\tilde{h},\tilde{k})}$. Therefore, φ_1 is the desired isomorphism $\operatorname{coeq}(\tilde{f},\tilde{g}) \cong \operatorname{coeq}(\tilde{h},\tilde{k})$. □

Remark 1.21. Since coequalizers are only defined up to unique isomorphism anyway, we might as well interpret Theorem 1.20 as saying that $\operatorname{coeq}(\tilde{f},\tilde{g}) = \operatorname{coeq}(\tilde{h},\tilde{k})$.

2. 2-CATEGORIES

Here, we return to the geometric perspective, as promised. Recall that we said natural transformations are "morphisms between morphisms" in Cat, and thus analogous to path-homotopies, which are 2-dimensional. How can we take this notion seriously in the categorical setting? Here's the first (but by no means last) attempt at answering that.

2.1. **Definition and examples.**

Definition 2.1. A 2-*category* C consists of

- a collection $\mathrm{Ob}\,\mathsf{C}$ whose elements we call "objects of C", as well as
- for any objects $X, Y \in \mathrm{Ob}\,\mathsf{C}$, a *category* $\mathsf{C}(X,Y)$ called the "hom-category from X to Y", with objects $f \in \mathsf{C}(X,Y)$ known as "1-morphisms from X to Y", usually denoted $f\colon X \to Y$, and morphisms $\eta \in \mathsf{C}(X,Y)(f,g)$ known as "2-morphisms from f to g", denoted $\eta\colon f \Rightarrow g$, and
- for any $X, Y, Z \in \mathrm{Ob}\,\mathsf{C}$, a *functor* $\mathsf{C}(X,Y) \times \mathsf{C}(Y,Z) \to \mathsf{C}(X,Z)$ called "horizontal composition", usually denoted $(f,g) \mapsto g \circ f\colon X \to Z$ for $f\colon X \to Y$, $g\colon Y \to Z$

such that

- for each $X \in \mathrm{Ob}\,\mathsf{C}$, there exists some morphism $1\colon X \to X$ such that for all $Y \in \mathrm{Ob}\,\mathsf{C}$ and any $f\colon X \to Y$ or $g\colon Y \to X$, we have $f \circ 1 = f$ and $1 \circ g = g$ (that is, composition is "unital"), and
- for any $W, X, Y, Z \in \mathrm{Ob}\,\mathsf{C}$ and $f\colon W \to X$, $g\colon X \to Y$, $h\colon Y \to Z$, we have $(h \circ g) \circ f = h \circ (g \circ f)$ (that is, composition is "associative").

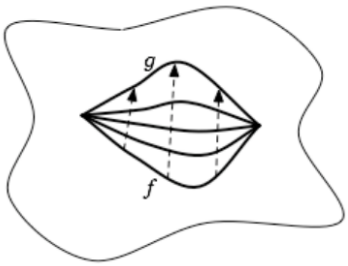
The composition of 2-morphisms in the hom-categories is typically referred to as "vertical composition", in contrast to the "horizontal composition" of 1-morphisms. We'll denote vertical composition by $\circ$, and horizontal composition will now be denoted by $*$.

We see that the definition is basically identical to the one we first gave for categories, except we've replaced the word "set" with "category", and the word "function" with "functor". This results in extra structure, literally of "morphisms between morphisms", which we call 2-morphisms, in contrast to the "1-dimensional" ones between objects. To emphasize the distinction, we often call morphisms between objects "1-morphisms", and even call objects "0-morphisms". Similarly, the categories and functors we defined at the very beginning will sometimes be called "1-categories" and "1-functors" for clarity. Carrying this analogy further, a "0-category" is just a set: it's a collection of objects with no nontrivial 1-morphisms.

In more general terms, we noticed that in the definition of category, we were already using objects of a category Set for hom-*sets*, and morphisms from it (functions) for composition; then we just replaced Set with another category, namely Cat. (It's important that Cat had a notion of product $\times$, in order for us to form the composition map.) This notion is known as "enriched category theory" - we'd say that a 2-category is a category "enriched in Cat". But we won't have much more to say about this.

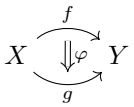
Example 2.2. Let's see some examples.

- Any category C is a 2-category $\mathrm{disc}\,\mathsf{C}$ in a trivial way: we can view each hom-set $\mathsf{C}(X,Y)$ as the discrete category $\mathrm{disc}\,\mathsf{C}(X,Y)$.
- It is immediate from our earlier remarks and the above definition that Cat is a 2-category in an interesting way, with hom-categories $\mathsf{Cat}(\mathsf{C},\mathsf{D}) = \mathsf{D}^{\mathsf{C}}$.
- For a more geometric example, one construction we can give is the *path 2-groupoid* $P_{\leq 2}(X)$ of a space X. The objects $P_{\leq 2}(X)$ are the points $x \in X$, the morphisms $x \to y$ are equivalence classes of paths $f\colon [0,1] \to X$ with $f(0) = x$, $f(1) = y$ under the equivalence relation of reparametrization (this is necessary in order for composition to be well-defined), and the 2-morphisms $[f] \to [g]$ are equivalence classes of path homotopies F from f to g under the equivalence relation of homotopy of path homotopies (suitably refined to account for reparametrizations of source and target paths). We can picture the objects, morphisms, and 2-morphisms something like this:

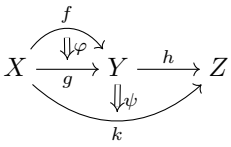


Intuitively, distinct 2-morphisms in $P_{\leq 2}(X)$ lie on different sides of a 2-dimensional hole. For a more explicit and detailed treatment of this example, see 5.2.

2.2. **Pasting diagrams.** Inspired by our geometric vision for 2-categories (see the last example above), it is frequently useful to represent 2-morphisms in a 2-category using 2-dimensional pictures called "pasting diagrams". Given objects X, Y and morphisms $f, g\colon X \to Y$, the pasting diagram for a 2-morphism $\varphi\colon f \Rightarrow g$ is



We can represent more complicated compositions this way. For example, given objects X, Y, Z, morphisms $f, g\colon X \to Y$, $h\colon Y \to Z$, $k\colon X \to Z$, and 2-morphisms $\varphi\colon f \Rightarrow g$, $\psi\colon g \circ h \Rightarrow k$, we could represent all this data and the composite $(\varphi * h) \circ \psi\colon f * h \Rightarrow k$ in the following concise pasting diagram:



Such diagrams make the structure of a 2-morphism much easier to see.

2.3. **2-functors.** As usual, we ought to define morphisms between 2-categories, known as 2-functors.

Definition 2.3. A 2-*functor* $F\colon \mathsf{C} \to \mathsf{D}$ between 2-categories C and D consists of a map $F\colon \operatorname{Ob}\mathsf{C} \to \operatorname{Ob}\mathsf{D}$ and for each $X, Y \in \operatorname{Ob}\mathsf{C}$, a functor $F\colon \mathsf{C}(X,Y) \to \mathsf{D}(F(X), F(Y))$ such that for any 2-morphisms $\varphi\colon f_1 \Rightarrow g_1$ in $\mathsf{C}(X,Y)$ and $\psi\colon f_2 \Rightarrow g_2$ in $\mathsf{C}(Y,Z)$ with the target of f_1 and g_1 equal to the source of f_2 and g_2, we have $F(\psi * \varphi) = F(\psi) * F(\varphi)$.

It's just a map of the data that "respects both levels of composition".

Example 2.4. Since we don't have many examples of 2-categories, it's hard to give examples of 2-functors, but we could say at least the following:

- There's an identity 2-functor on any 2-category and a composition of any two 2-functors with matching source and target, and indeed, 2-categories and 2-functors form a category $2\mathsf{Cat}$.
- The inclusion $\operatorname{disc}\mathsf{Cat} \hookrightarrow \mathsf{Cat}$ is a 2-functor - we're just adding in the nontrivial natural transformations.
- For any space X, there is a forgetful 2-functor $P_{\leq 2}(X) \to \operatorname{disc}\Pi_1(X)$ which sends any reparametrization class of paths to the path homotopy class containing it, and sends any path homotopy to the identity 2-morphism on the path homotopy class of its source and target.

3. Monoidal categories

There's a special case of 2-categories that's important enough to get its own name. It sits at the intersection of numerous areas of mathematics.

3.1. **Definition and examples.**

Definition 3.1. A *(strict) monoidal category* is a 2-category with one object.

Remark 3.2. Let $\mathsf{C}^{\otimes}$ be a monoidal category. Now all of the data is contained in one category $\mathsf{C}^{\otimes}(*,*)$ (where $*$ is the unique object in $\mathsf{C}^{\otimes}$), equipped with a functor $\otimes\colon \mathsf{C}^{\otimes}(*,*) \times \mathsf{C}^{\otimes}(*,*) \to \mathsf{C}^{\otimes}(*,*)$ coming from composition (let's call it $\otimes$ now, since that's what it's usually denoted by in this setting). Recall that this must be unital, that is, there's some 1-morphism $1_*\colon * \to *$ such that $1_* \otimes f = f = f \otimes 1_*$ for any morphism $f\colon * \to *$, and it's also associative, meaning $(f \otimes g) \otimes h = f \otimes (g \otimes h)$. This is the definition usually given for a monoidal category, and the terminology most commonly used involves a shift in dimension.

Notation 3.3. In a monoidal category $\mathsf{C}^{\otimes}$, one usually refers to the "underlying category" $\mathsf{C} = \mathsf{C}^{\otimes}(*,*)$, referring to morphisms $* \to *$ in the 2-category $\mathsf{C}^{\otimes}$ as objects in the monoidal category C, and likewise 2-morphisms become 1-morphisms. Composition of 1-morphisms in C (2-morphisms in $\mathsf{C}^{\otimes}$) is still denoted by $\circ$, and now the composition map $\mathsf{C}^{\otimes}(*,*) \times \mathsf{C}^{\otimes}(*,*) \to \mathsf{C}^{\otimes}(*,*)$ is denoted by $\otimes \colon \mathsf{C} \times \mathsf{C} \to \mathsf{C}$ and called the "monoidal product" or sometimes "tensor product". The unit 1_* is now just denoted by 1 and referred to as the "monoidal unit". Usually, one specifies a monoidal category by giving a category C, a functor $\otimes \colon \mathsf{C} \times \mathsf{C} \to \mathsf{C}$, and an object $1 \in \mathsf{C}$ such that $\otimes$ is strictly associative with strict unit 1.

Remark 3.4. As an exercise, the interested reader should try to make the above discussion precise using the language of equivalence of categories.(Find the suitable notion of "monoidal functor" between categories equipped with an associative and unital product $\otimes$, show that these form a category, and construct an equivalence of categories between this and the category of 2-categories with one object, with 2-functors as morphisms.)

Remark 3.5. We say "strict" because there is a slightly more general notion which goes by the name of "monoidal category", in which the monoidal product is associative/unital merely up to natural isomorphism, and there are some extra axioms to ensure this is coherent. These "non-strict" kinds of structures appear far more frequently "in the mathematical wild", but for simplicity, we won't deal with them here, since much of the substance is communicated through the strict theory.

Due to the above remark, it is not so easy to list meaningful examples of strict monoidal categories. For this reason, we'll start by listing a few strict examples, then give some examples of weak monoidal categories, noting that these do not fit any definition we've given yet.

Example 3.6. First, recall the classical definition of a *monoid*: a set A together with a binary operation $A \times A \to A$ denoted by $(x, y) \mapsto x \cdot y$ such that $(x \cdot y) \cdot z = x \cdot (y \cdot z)$ for all $x, y, z \in A$, and there is an element $1 \in A$ such that $1 \cdot x = x = x \cdot 1$ for all $x \in A$.

- Given a monoid A, there is a strict monoidal category $\operatorname{disc} A^{\cdot}$ whose underlying category is $\operatorname{disc} A$, with multiplication given on objects by $\cdot$. Clearly 1 serves as the unit.
- Given a *commutative* monoid A (one in which $x \cdot y = y \cdot x$ for all $x, y \in A$), there is a strict monoidal category $BA^{\cdot}$ with underlying category BA, with multiplication given on morphisms by $\cdot$ (which also serves as composition). Clearly the unique object $*$ serves as the unit.
- Let $\mathbb{R}^* = \mathbb{R} \cup \{\pm\infty\}$ be the poset of extended real numbers, viewed as a category. The functions max and min each endow $\mathbb{R}^*$ with the structure of a strict monoidal category.

The following are non-strict examples:

- Set with the Cartesian product $\times$ and a singleton set $\{*\}$ as unit. More generally, any complete category with the categorical product and terminal object as unit.
- Set with disjoint union $\amalg$ and the empty set $\emptyset$ as unit. More generally, any cocomplete category with the categorical coproduct and inital object as unit.
- Ab or Vect_k with tensor product over $\mathbb{Z}$ or k, and $\mathbb{Z}$ or k as unit. More generally, R-modules with tensor product $\otimes_R$ over R and R as unit. We will soon see how to go even more general.

Remark 3.7. It's not actually so bad to give these non-strict examples, since as it turns out, a so-called "strictification" theorem holds for monoidal categories:

Proposition 3.8. Every monoidal category is (strictly monoidally) equivalent to a strict monoidal category.

Proof. See Section 5.1. □

Without defining it precisely, "strictly monoidally equivalent" means "equivalent via functors respecting the monoidal structure, up to equality rather than just isomorphism". We won't set up the machinery to prove it here, but it's not hard with the right definitions, and it means that we are somewhat justified in thinking of monoidal categories as strict, even if they're not. The upshot for us is that general monoidal categories provide familiar examples on which it is valid to build our intuition. From here on, every time we say "monoidal category" we mean "strict monoidal category", and when we want to refer to a general one, we will say "non-strict".

3.2. **The Eckmann-Hilton argument.** One might wonder why we need the "commutative" hypothesis in the second strict example BA above. Here's a counterexample without it: consider the monoid $A = \mathsf{Set}(\{1,2\},\{1,2\})$ under composition. A has four elements: the identity function e, the constant map c_1 to 1, the constant map c_2 to 2, and the swap map s which maps $1 \mapsto 2$ and $2 \mapsto 1$. A is not commutative: $c_1 \circ s \neq s \circ c_1$ since the LHS has image $\{1\}$ and the RHS has image $\{2\}$. We can form the one-object category BA, but when we demand that the multiplication (and composition) $\circ$ defines a monoidal product, we fail the bifunctoriality condition:

$$(e \circ c_1) \circ (s \circ e) = c_1 \circ s \neq s \circ c_1 = (e \circ s) \circ (c_1 \circ e)$$

To see how commutativity rectifies this issue, we'll sketch the proof of the example here.

Proposition 3.9. Given a commutative monoid A, there is a strict monoidal category $BA^{\cdot}$ with underlying category BA, with multiplication given on morphisms by $\cdot$.

Proof. Keep in mind throughout that "composition" and "product" are the same thing. For the product $(-)\cdot(-)\colon BA \times BA \to BA$ to be a bifunctor, we must check that $(w \cdot x) \cdot (y \cdot z) = (w \cdot y) \cdot (x \cdot z)$. Since $\cdot$ is associative, we can drop the brackets, and since it's commutative, the condition holds, and hence $\cdot$ is a bifunctor. It's clearly associative and unital, too, so we conclude that $BA^{\cdot}$ is a monoidal category. □

But it turns out this strong assumption is necessary.

Proposition 3.10. Suppose $\mathsf{C}^{\otimes}$ is a monoidal category with one object. Then $\mathsf{C}^{\otimes} \cong BA^{\cdot}$ where A is a commutative monoid.

Proof. Let $A = \mathsf{C}(*,*)$ be the monoid of morphisms $* \to *$ under composition. Let $x, y \in A$. Using unitality and bifunctoriality, we observe that

$$\begin{aligned} x \otimes y &= (x \circ 1_*) \otimes (y \circ 1_*) \\ &= (x \otimes 1_*) \circ (y \otimes 1_*) \\ &= x \circ y \end{aligned}$$

Thus, $\otimes$ and $\circ$ are actually the same operation on the morphisms of C. Furthermore,

$$\begin{aligned} x \circ y &= (1_* \otimes x) \circ (y \otimes 1_*) \\ &= (1_* \circ y) \otimes (x \circ 1_*) \\ &= y \otimes x \\ &= y \circ x \end{aligned}$$

so indeed this operation on A is commutative. Now we can see that since $\otimes = \circ$, $\mathsf{C}^{\otimes}$ is given by our construction for BA above. □

Corollary 3.11. There is an equivalence of categories between the category of monoidal categories with one object and the category of commutative monoids.

Proof. For an equivalence of categories, we need functors in both directions. We have them: given a monoidal category $\mathsf{C}^{\otimes}$ with one object $*$, we obtain a monoid $\mathsf{C}(*,*)$ with the operation of composition (commutative by the above reasoning), and given a commutative monoid A, we obtain a monoidal category $BA^{\cdot}$ with one object (and one can check that both of these constructions are compatible with morphisms).

We also need a natural isomorphism between each composition and the relevant identity functor. On one side, we must check: $\mathsf{C}^{\otimes} \cong B\mathsf{C}(*,*)^{\circ}$, and this isomorphism is natural. We produced this isomorphism above (our choice of A was $\mathsf{C}(*,*)$ with composition) and it remains to check that for a monoidal functor $F\colon \mathsf{C}^{\otimes} \to \mathsf{D}^{\boxtimes}$, the square

$$\begin{array}{ccc} \mathsf{C}^{\otimes} & \xrightarrow{F} & \mathsf{D}^{\boxtimes} \\ \downarrow{\scriptstyle\cong} & & \downarrow{\scriptstyle\cong} \\ B\mathsf{C}(*,*)^{\circ} & \xrightarrow{BF_*} & B\mathsf{D}(*,*)^{\circ} \end{array}$$

commutes (where $F_*\colon \mathsf{C}(*,*) \to \mathsf{D}(*,*)$ denotes the map of monoids induced by F). Since each monoidal category has just one object, the diagram certainly commutes on objects, and on morphisms, we see that

it commutes, since the vertical isomorphisms are given essentially by reinterpretation of the data, and no arbitrary choices are made in the construction of $BA^{\cdot}$. Thus, the square commutes, and hence this isomorphism is natural.

On the other side, we must check that we have a natural isomorphism of monoids $A \cong BA^{\cdot}(*,*)$. The existence of the isomorphism, and its naturality, are clear, since the morphisms of $BA^{\cdot}$ are precisely the elements of A, so this isomorphism is essentially given by the identity on the underlying set of A, which is indeed a monoid isomorphism (and certainly natural with respect to A). Therefore, since we have functors in both directions whose compositions are naturally isomorphic to the identity, we conclude that there is an equivalence of categories between the category of monoidal categories with one object and the category of commutative monoids. □

This result is the first and most fundamental instance of what's known as the *Eckmann-Hilton principle*. Named for Beno Eckmann and Peter Hilton, who discovered and used it in a 1962 paper [6], the principle was actually known quite a bit sooner. In fact, if you've heard of the higher homotopy groups $\pi_n(X)$ ($n \geq 2$) of a space X, it is essentially the reason why they are all abelian: higher loops are basically maps out of an n-cube $[0,1]^n$, and once $n \geq 2$, we have two distinct faces along which to compose loops by gluing the cubes together, and these two operations are compatible in the same way we've been discussing. This initially caused doubt about the definition of π_n:

"In 1932, E. Čech submitted to the ICM at Zürich a paper on higher homotopy groups; however, Alexandroff and Hopf, the kings of topology at the time, objected to the fact that they were abelian for $n \geq 2$, and on these grounds persuaded Čech to withdraw his paper, so that only a small paragraph appeared in the ICM Proceedings." [4]

Eventually, people came to understand that the commutativity of the higher homotopy groups was a feature, not a bug, and that that the definition was correct. Further, with the development of monoidal category theory in the 60s and onward, people began to see that this was but the first example of a general principle: *given many compatible product operations on a (higher) category, they are all equivalent, and the resulting product is more commutative than each of the individual ones.* We shall see some more interesting examples of this further on.

3.3. **Monoids.** There are plenty of fun things to do with monoidal categories (see e.g. [3]), but for us, their primary purpose will be as the most general setting in which to construct monoids and some associated structures. Monoids are meant to capture the notion of an "abstract structure of multiplication", where we have a binary operation which is associative and has an identity element.

Definition 3.12. Let $\mathsf{C}^{\otimes}$ be a monoidal category. A *monoid* (A, μ, η) in $\mathsf{C}^{\otimes}$ (usually just A for short, with μ, η left implicit) consists of an object $A \in \mathsf{C}$, a morphism $\mu\colon A \otimes A \to A$ called "multiplication", and a morphism $\eta\colon 1 \to A$ called the "unit", such that the following diagrams commute:

$$\begin{array}{ccc} 1 \otimes A & \xrightarrow{\eta \otimes 1_A} & A \otimes A \\ & {\scriptstyle 1_A}\searrow & \downarrow{\scriptstyle \mu} \\ & & A \end{array} \qquad \begin{array}{ccc} A \otimes A & \xleftarrow{1_A \otimes \eta} & A \otimes 1 \\ {\scriptstyle \mu}\downarrow & \swarrow{\scriptstyle 1_A} & \\ A & & \end{array}$$

i.e. multiplication is unital, and

$$\begin{array}{ccc} A \otimes A \otimes A & \xrightarrow{1_A \otimes \mu} & A \otimes A \\ {\scriptstyle \mu \otimes 1_A}\downarrow & & \downarrow{\scriptstyle \mu} \\ A \otimes A & \xrightarrow[\mu]{} & A \end{array}$$

i.e. multiplication is associative. A *monoid homomorphism* between monoids (A, μ_A, η_A) and (B, μ_B, η_B) in $\mathsf{C}^{\otimes}$ is a morphism $\varphi\colon A \to B$ such that the following diagrams commute:

$$\begin{array}{ccc} A \otimes A & \xrightarrow{\varphi \otimes \varphi} & B \otimes B \\ {\scriptstyle \mu_A}\downarrow & & \downarrow{\scriptstyle \mu_B} \\ A & \xrightarrow{\varphi} & B \end{array} \qquad \begin{array}{ccc} A & \xrightarrow{\varphi} & B \\ & {\scriptstyle \eta_A}\nwarrow \quad \nearrow{\scriptstyle \eta_B} & \\ & 1 & \end{array}$$

i.e. φ commutes with the monoid multiplications and units.

Example 3.13. As always, let's see some (strict and non-strict) examples. In our strict examples:

- In any monoidal category $\mathsf{C}^{\otimes}$, the monoidal unit 1 has a canonical monoid structure, with unit $1 \to 1$ and multiplication $1 \otimes 1 = 1 \to 1$ both given by the identity 1_1. The required diagrams evidently commute, as everything's the identity.
- In $\operatorname{disc} A^{\cdot}$ for an ordinary monoid A, a monoid X must have a morphism $1 \to X$, which is only possible if $X = 1$, and then that morphism and $\mu\colon 1 \otimes 1 = 1 \to 1$ are both just the identity, since that's all there is in $(\operatorname{disc} A)(1,1)$. Therefore, there's just one trivial monoid.
- In $BA^{\cdot}$ for a commutative ordinary monoid A, there's just one object $*$, and we must pick morphisms $\mu\colon * \otimes * = * \to *$ and $\eta\colon 1 = * \to *$ (elements $\mu, \eta \in A$) such that $\mu\eta = \eta\mu = 1$ (unitality) and $\mu^2 = \mu^2$ (associativity, a tautology). Thus $\eta = \mu^{-1}$. Recall that inverses in an ordinary monoid are unique. Thus, monoids in $BA^{\cdot}$ are in bijection with the invertible elements of A.
- In $\mathbb{R}^*$ with min as monoidal product, the only object with a morphism from the unit ∞ is ∞ itself, so there's just the trivial monoid structure on ∞.
- In $\mathbb{R}^*$ with max as monoidal product, we have a unique choice for every object x (real number or $\pm\infty$): η is the unique morphism $-\infty \to x$ and $\mu = 1_x\colon \max(x,x) = x \to x$. Unitality and associativity are immediate since there's only one morphism in each hom-set. Hence, there's one uninteresting monoid structure on every real number x and on $\pm\infty$.

In our non-strict examples:

- Monoids in $\mathsf{Set}^{\times}$ are precisely ordinary monoids. In $\mathsf{Cat}^{\times}$, they are categories with an associative, unital product, which might sound familiar - they are strict monoidal categories.
- Monoids in $\mathsf{Set}^{\amalg}$ are in bijection with sets; more generally, in $\mathsf{C}^{\amalg}$, they are just objects. To see this, note that for any object A, there's a unique morphism from the initial object to A, so η contains no information, and a morphism $A \amalg A \to A$ which is associative must restrict to the identity on each copy of A (we call the resulting morphism the *fold map*). Thus each object has a unique monoid structure.
- Monoids in $\mathsf{Ab}^{\otimes_{\mathbb{Z}}}$ are just rings (abelian groups with a compatible product), monoids in $\mathsf{Vect}_k^{\otimes_k}$ are k-algebras, etc.

Remark 3.14. The above objects (particularly those in our non-strict examples) are ubiquitous throughout objects, and form the basic objects of study of a wide variety of fields. For example, rings are of fundamental interest in the field of algebraic geometry. Thus, we see the study of monoidal category theory begin to "subsume" a number of other subjects central to modern mathematics.

3.4. **Modules over monoids.** Monoids capture abstract multiplicative structures, but these would be of little use if they could not act on concrete objects in a multiplicative way. This is captured by the notion of a *module*.

Definition 3.15. Let (A, η, μ) be a monoid in a monoidal category $\mathsf{C}^{\otimes}$. A *(left) module* (M, ν) *over* A (usually just M or ${}_AM$ for short) consists of an object $M \in \mathsf{C}$, together with a morphism $\nu\colon A \otimes M \to M$ called the "action", such that the following diagrams commute:

$$\begin{array}{ccc} 1 \otimes M & \xrightarrow{\quad 1_M \quad} & M \\ & {\scriptstyle \eta \otimes 1_M} \searrow \quad \nearrow {\scriptstyle \nu} & \\ & A \otimes M & \end{array}$$

i.e. the action is "unital",and

$$\begin{array}{ccc} A \otimes A \otimes M & \xrightarrow{1_A \otimes \nu} & A \otimes M \\ {\scriptstyle \mu \otimes 1_M}\downarrow & & \downarrow{\scriptstyle \nu} \\ A \otimes M & \xrightarrow[\nu]{} & M \end{array}$$

i.e. the action is "associative". An *A-module homomorphism* between A-modules (M, ν_M) and (N, ν_N) is a morphism $f\colon M \to N$ such that the following diagram commutes:

$$\begin{array}{ccc} A \otimes M & \xrightarrow{1_A \otimes f} & A \otimes N \\ {\scriptstyle \nu_M}\downarrow & & \downarrow{\scriptstyle \nu_N} \\ M & \xrightarrow[f]{} & N \end{array}$$

that is, f commutes with the actions of A on M and N. Right A-modules and their morphisms are defined analogously.

Given monoids (A, η_A, μ_A) and (B, η_B, μ_B) in $\mathsf{C}^{\otimes}$, an *(A,B)-bimodule* is an object M with a left A-module structure (M, ℓ) and a right B-module structure (M, r), such that the two actions commute:

$$\begin{array}{ccc} A \otimes M \otimes B & \xrightarrow{1_A \otimes r} & A \otimes M \\ {\scriptstyle \ell \otimes 1_B}\downarrow & & \downarrow{\scriptstyle \ell} \\ M \otimes B & \xrightarrow[r]{} & M \end{array}$$

and an *(A,B)-bimodule homomorphism* is a morphism that's simultaneously a left A-module homomorphism and a right B-module homomorphism for the respective module structures. We will typically say "M is an (A,B)-bimodule" and if needed, denote the maps by ℓ and r as above.

Example 3.16. As always, let's see some examples. Strict monoidal categories first:

- In any monoidal category $C^{\otimes}$, any monoid A has a canonical (A,A)-bimodule structure where both the left and right actions are given by the multiplication $A \otimes A \to A$. The left and right actions commute with each other since multiplication is associative.
- In any monoidal category $\mathsf{C}^{\otimes}$, a module over the unit 1 with its canonical monoid structure is just an object of C. That is, every object has a unique left 1-module structure with action given by the identity $1_M\colon 1 \otimes M = M \to M$. Uniqueness is forced by the unitality diagram, and it clearly satisfies associativity too. Since 1 is the only monoid in $\operatorname{disc} A^{\cdot}$ for an ordinary monoid A, this determines all the (bi)modules in $\operatorname{disc} A^{\cdot}$, and the same is true for $(\mathbb{R}^*)^{\min}$.
- In $BA^{\cdot}$ for a commutative ordinary monoid A, a module over an invertible element $\mu \in A$ (recall the unit $\eta = \mu^{-1}$) must have the object $*$, and a morphism $\nu \in A$ such that $\nu\eta = 1\ (= \eta\nu)$ and $\nu^2 = \nu\mu$, so $\nu = \eta^{-1} = \mu$ and this clearly works. Thus, there's just the one canonical bimodule over every monoid.
- In $\mathbb{R}^*$ with max as monoidal product, over a monoid given by $x \in \mathbb{R}^*$, we have one module structure on each element $y \in \mathbb{R}^*$ with $y \geq x$, since what we need is a morphism $x \otimes y = \max(x,y) \to y$ and this exists iff $\max(x,y) \leq y$, which is equivalent to $x \leq y$. In this case, $\max(x,y) = y$ and the action is given by the identity. All diagrams commute since all hom-sets have one element. The same is true for right actions, so the (x_1, x_2)-bimodules are precisely the elements $y \geq \max(x_1, x_2)$.

Now non-strict:

- A module over an ordinary monoid A in $\mathsf{Set}^{\times}$ is just an A-set, that is, a set with an ordinary monoid action. (This is more commonly studied for groups specifically.)
- In $\mathsf{C}^{\amalg}$, modules Y over an object X are in bijection with morphisms $X \to Y$. To see this, note that we need a morphism $X \amalg Y \to Y$ which makes the above unitality and associativity diagrams involving the fold map $X \amalg X \to X$ commute. Equivalently, the data of a morphism $X \amalg Y \to Y$ is the same as the data of two morphisms: one $X \to Y$ and one $Y \to Y$. The unitality diagram forces the one $Y \to Y$ to be the identity, since the morphism $\emptyset \amalg Y = Y \to X \amalg Y$ is the inclusion. This just leaves us with a morphism $f\colon X \to Y$ as the remaining data, and for any $f\colon X \to Y$, we see that the unitality and associativity diagrams commute. An (X_1, X_2)-bimodule Y is then precisely a *cospan*, that is, a pair of maps $X_1 \to Y \leftarrow X_2$. It's easy to check that the actions always commute.
- Modules over monoids in $\mathsf{Ab}^{\otimes_{\mathbb{Z}}}$ are just modules over rings in the ordinary sense.

Remark 3.17. We continue to recover more objects of mathematical interest. Adding to the algebraic geometry story, we now have modules, which subsume both ideals (for a commutative ring or algebra R, these are just submodules of the (R,R)-bimodule R) and objects such as the space of sections of a vector bundle over an affine variety X (this is canonically an $\mathcal{O}_X$-module).

We further add one more proposition of note: modules also subsume monoid homomorphisms (though we haven't dealt with composition yet, we will soon).

Proposition 3.18. Let $\mathsf{C}^{\otimes}$ be a monoidal category, and let $\varphi\colon A \to B$ be a monoid homomorphism. Then we obtain an (A,B)-bimodule structure on B.

Proof. We define the right action of B on B by $r := \mu_B\colon B \otimes B \to B$, and we define the left action as the composite $\ell := \mu_B \circ (\varphi \otimes \mathrm{id}_B)\colon A \otimes B \to B \otimes B \to B$. We already know from the monoid axioms that r is unital and associative, so we must verify that the action of A is unital and associative. The unitality triangle breaks up as follows:

$$\begin{array}{ccccc} 1 \otimes B & & \xrightarrow{\quad 1_B \quad} & & B \\ {\scriptstyle \eta_A \otimes 1_B}\downarrow & \searrow^{\eta_B \otimes 1_B} & & \nearrow_{\mu_B} & \\ A \otimes B & & \xrightarrow[\varphi \otimes 1_B]{} & B \otimes B & \end{array}$$

The right triangle commutes since it expresses unitality for the monoid B, and the left triangle commutes since it is obtained by sending the unitality diagram for the monoid homomorphism φ through the functor $- \otimes B$. The associativity diagram breaks up like so:

$$\begin{array}{ccccc} A \otimes A \otimes B & \xrightarrow{1_A \otimes \varphi \otimes 1_B} & A \otimes B \otimes B & \xrightarrow{1_A \otimes \mu_B} & A \otimes B \\ \downarrow{\scriptstyle \mu_A \otimes 1_B} & \searrow^{\varphi \otimes \varphi \otimes 1_B} & \downarrow{\scriptstyle \varphi \otimes 1_B \otimes 1_B} & & \downarrow{\scriptstyle \varphi \otimes 1_B} \\ & & B \otimes B \otimes B & \xrightarrow[1_B \otimes \mu_B]{} & B \otimes B \\ & & \downarrow{\scriptstyle \mu_B \otimes 1_B} & & \downarrow{\scriptstyle \mu_B} \\ A \otimes B & \xrightarrow[\varphi \otimes 1_B]{} & B \otimes B & \xrightarrow[\mu_B]{} & B \end{array}$$

The "square" on the left commutes since it's obtained by pushing the square which says φ is a monoid homomorphism through the functor $- \otimes B$. The triangle above it commutes by bifunctoriality of the monoidal product. The top square on the right also commutes by bifunctoriality of $\otimes$, and the bottom square on the right commutes by associativity of the multiplication μ_B in B. Therefore, the action is associative too, and hence a well-defined right A-action on B. Finally, we must check that these actions commute. The relevant diagram breaks up like so:

$$\begin{array}{ccccc} A \otimes B \otimes B & \xrightarrow{\varphi \otimes 1_B \otimes 1_B} & B \otimes B \otimes B & \xrightarrow{\mu_B \otimes 1_B} & B \otimes B \\ \downarrow{\scriptstyle 1_A \otimes \mu_B} & & \downarrow{\scriptstyle 1_B \otimes \mu_B} & & \downarrow{\scriptstyle \mu_B} \\ A \otimes B & \xrightarrow[\varphi \otimes 1_B]{} & B \otimes B & \xrightarrow[\mu_B]{} & B \end{array}$$

The left square commutes by bifunctoriality of $\otimes$, and the right square commutes by associativity of multiplication in B. Therefore, we have obtained an (A,B)-bimodule structure on B. □

3.5. **Balanced tensor product of modules over monoids.** The following definition gives us a notion of "multiplying" (as we will see, really "composing") modules in a canonical way.

Definition 3.19. Let $\mathsf{C}^{\otimes}$ be a monoidal category. Let A be a monoid in $\mathsf{C}^{\otimes}$, and suppose we're given a right A-module M and a left A-module N. We define the *balanced tensor product* $M \otimes_A N$ *of* M *and* N *over* A to be the coequalizer of the following parallel pair:

$$M \otimes_A N := \mathrm{coeq}\left(M \otimes A \otimes N \underset{1_M \otimes \ell_N}{\overset{r_M \otimes 1_N}{\rightrightarrows}} M \otimes N \right)$$

whenever it exists. We will denote the coequalizer map by $\pi_{M,N}\colon M \otimes N \to M \otimes_A N$.

Definition 3.20. Let $\mathsf{C}^{\otimes}$ be a monoidal category. We say $\mathsf{C}^{\otimes}$ *admits a calculus of balanced tensor products* if for any monoids A, B, C, D and (A, B)-bimodule M, (B, C)-bimodule N, and (C, D)-bimodule P, we have that

(1) the coequalizer
$$M \otimes_B N := \operatorname{coeq}\left(M \otimes B \otimes N \underset{1_M \otimes \ell_N}{\overset{r_M \otimes 1_N}{\rightrightarrows}} M \otimes N \right)$$
exists in C,

(2) the functor $M \otimes (-) \colon \mathsf{C} \to \mathsf{C}$ preserves coequalizer diagrams of the form
$$N \otimes_C P := \operatorname{coeq}\left(N \otimes C \otimes P \underset{1_N \otimes \ell_P}{\overset{r_N \otimes 1_P}{\rightrightarrows}} N \otimes P \right)$$
and

(3) the functor $(-) \otimes P \colon \mathsf{C} \to \mathsf{C}$ preserves coequalizer diagrams of the form
$$M \otimes_B N := \operatorname{coeq}\left(M \otimes B \otimes N \underset{1_M \otimes \ell_N}{\overset{r_M \otimes 1_N}{\rightrightarrows}} M \otimes N \right)$$

That is, the coequalizers from Definition 3.19 all exist and commute with the monoidal product.

Proposition 3.21. Let $\mathsf{C}^{\otimes}$ be a monoidal category, and let A, B be monoids with M a right B-module and N a left B-module. There is a canonical right B-module structure on $A \otimes M$ with action given by the composite $r_{A \otimes M} := 1_A \otimes r_M \colon (A \otimes M) \otimes B = A \otimes (M \otimes B) \to A \otimes M$, and if $\mathsf{C}^{\otimes}$ admits a calculus of balanced tensor products, then $A \otimes (M \otimes_B N) = (A \otimes M) \otimes_B N$ (so we are justified in writing $A \otimes M \otimes_B N$ with no brackets) and $\pi_{A \otimes M, N} = A \otimes \pi_{M,N}$. Similarly, there is a left B-module structure on $N \otimes A$, defined analogously, and if $\mathsf{C}^{\otimes}$ admits a calculus of balanced tensor products, then $(M \otimes_B N) \otimes A = M \otimes_B (N \otimes A)$ and $\pi_{M, N \otimes A} = \pi_{M,N} \otimes A$.

Proof. We will just prove the statements with A on the left; the other proof is very similar. The above-defined right B-module structure is clearly well-defined, since the relevant unitality and associativity diagrams are given simply by sending those of M_B through the functor $A \otimes -$. The hypothesis that $\mathsf{C}^{\otimes}$ admits a calculus of balanced tensor products means that the top row of the following diagram is a coequalizer diagram:

$$\begin{array}{ccccc}
A \otimes (M \otimes B \otimes N) & \underset{1_A \otimes 1_M \otimes \ell_N}{\overset{1_A \otimes r_M \otimes 1_N}{\rightrightarrows}} & A \otimes (M \otimes N) & \longrightarrow & A \otimes (M \otimes_B N) \\
\| & & \| & & \\
(A \otimes M) \otimes B \otimes N & \underset{1_{A \otimes M} \otimes \ell_N}{\overset{r_{A \otimes M} \otimes 1_N}{\rightrightarrows}} & (A \otimes M) \otimes N & \longrightarrow & (A \otimes M) \otimes_B N
\end{array}$$

By our above definition and the definition of a monoidal category, the squares on the left commute, since their top and bottom rows are the same. Thus, their coequalizers and the relevant maps must be equal too, and the conclusion follows. □

Proposition 3.22. Suppose $\mathsf{C}^{\otimes}$ is a monoidal category which admits a calculus of balanced tensor products, and let M be an (A, B)-bimodule and N be a (B, C)-bimodule for some monoids A, B, C. Then we obtain a canonical (A, C)-bimodule structure on the balanced tensor product $M \otimes_B N$.

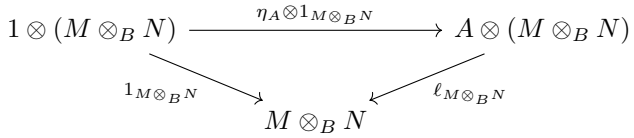
Proof. We'll construct the left A-module structure; the right C-module structure is analogous. Consider the following diagram:

$$\begin{array}{ccccc}
A \otimes M \otimes B \otimes N & \underset{1_A \otimes 1_M \otimes \ell_N}{\overset{1_A \otimes r_M \otimes 1_N}{\rightrightarrows}} & A \otimes M \otimes N & \xrightarrow{1_A \otimes \pi_{M,N}} & A \otimes M \otimes_B N \\
\downarrow{\scriptstyle \ell_M \otimes 1_B \otimes 1_N} & & \downarrow{\scriptstyle \ell_M \otimes 1_N} & & \vdots{\scriptstyle \ell_{M \otimes_B N}} \\
M \otimes B \otimes N & \underset{1_M \otimes \ell_N}{\overset{r_M \otimes 1_N}{\rightrightarrows}} & M \otimes N & \xrightarrow{\pi_{M,N}} & M \otimes_B N
\end{array}$$

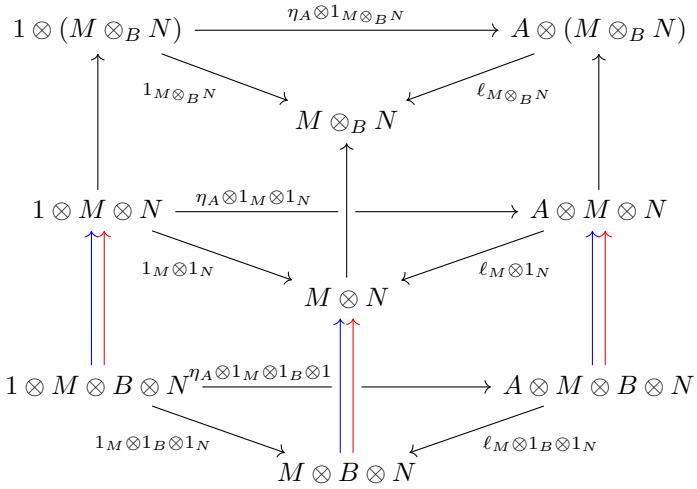
By the previous proposition, we know what we've written makes sense, and the top row is a coequalizer diagram. The two squares on the left both commute: the one with upper horizontal arrows commutes

since it's the bimodule square for M sent through the functor $- \otimes N$, and the one with lower horizontal arrows commutes by bifunctoriality of $\otimes$. Thus, the vertical maps induce a map $A \otimes M \otimes_B N \to M \otimes_B N$, which we call $\ell_{M \otimes_B N}$. We must check that this defines a left A-module structure on $M \otimes_B N$.

To verify unitality, we must establish the commutativity of the following diagram:

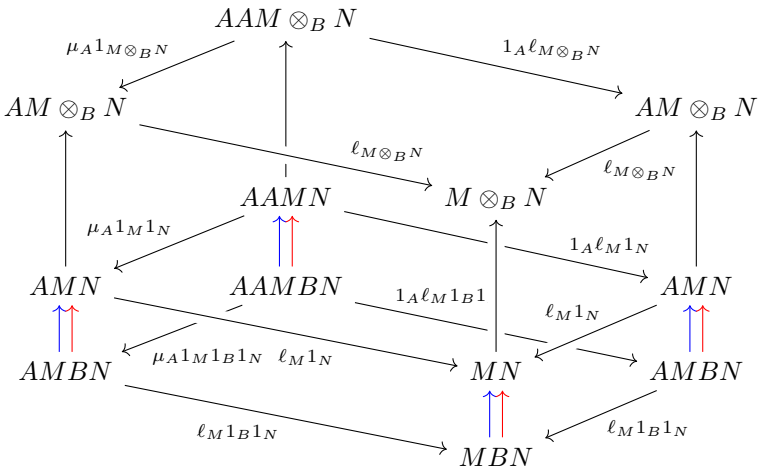


Since taking colimits is functorial, it is enough to establish a "commutative prism" in the pre-balanced tensor setting. That is, we must argue that "every" square and triangle in the following diagram, besides the top one, commutes:



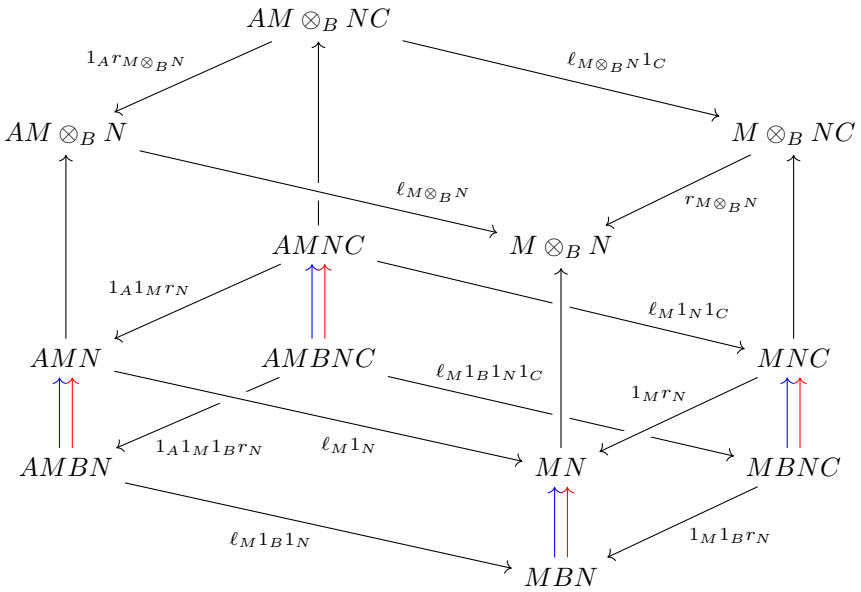
The blue arrows correspond to B acting from the right on M, and red to B acting from the left on N. The other vertical arrows are coequalizer morphisms. The front left face commutes by our definition of $\ell_{M \otimes_B N}$ above, and the preceding reasoning. The front right face commutes since the two coequalizer diagrams are equal. The map involving η on top is induced by those below simply by bifunctoriality of $\otimes$. The triangles commute since they're functorial images of the unitality triangle for the A-module M (the bottom one through $- \otimes (B \otimes N)$, and the middle through $- \otimes N$). Thus, we have a commutative diagram of parallel pairs of morphisms on the bottom, so by functoriality of coequalizers, the top triangle commutes.

We set up a similar argument for associativity:



(I've dropped the $\otimes$ for monoidal product since it was getting cumbersome; can add them later.) Again the vertical arrows form coequalizer diagrams and we want to establish that the top square is induced by the bottom prism, and that the bottom prism commutes, so that the top square commutes. As before, the left- and right-hand front faces (they're the same) both commute by our definition of $\ell_{M\otimes_B N}$. The right-hand back face is just the image of the front face under the functor $A\otimes -$, so it commutes too. The left-hand back face commutes by bifunctoriality of $\otimes$. The horizontal rectangles both commute as functorial images of the associativity diagram for ${}_AM$ (the middle under $-\otimes N$ and the bottom under $-\otimes B\otimes N$). Thus, by functoriality of coequalizers, the square on top commutes too, meaning $\ell_{M\otimes_B N}$ is associative, and hence defines a left A-action on $M\otimes_B N$.

The same arguments (mirrored) apply to show that $M\otimes_B N$ has an analogously-defined left C-module structure. Now to show that it's an (A,C)-bimodule, we must show that these actions commute, using a similar argument one more time:



As usual, both front faces commute by definition, and the back faces are functorial images of the front faces (under $A\otimes -$ and $-\otimes C$) so they commute as well. The horizontal squares both commute by bifunctoriality of $\otimes$. Therefore, the top square commutes, and hence we have obtained an (A,C)-bimodule structure on $M\otimes_B N$, as desired. □

Now we've seen that we can multiply bimodules together to get new ones, in a way that preserves their "outer actions". In fact, it looks an awful lot like morphism composition from our first definition, that of an ordinary category: an (A,B)-bimodule M is like a "morphism" $A\to B$, and the (A,C)-bimodule $M\otimes_B N$ is like the composite of M with a morphism $N\colon B\to C$. Spoiler alert: this is exactly where we're headed, and precisely the insight leading to the Morita bicategory! But let's not get ahead of ourselves. We haven't seen any concrete examples yet.

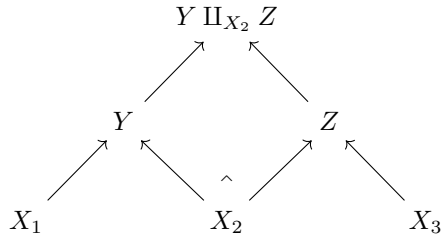
Example 3.23. Let's see what this looks like in our examples of strict monoidal categories.

- If A is a monoid in a monoidal category $\mathsf{C}^\otimes$ admitting a calculus of balanced tensor products, then with its canonical (A,A)-bimodule structure, we have $A\otimes_A A\cong A$ as (A,A)-bimodules. More generally, for a left A-module M, we have $A\otimes_A M\cong M$, and similarly for a right A-module. We will prove this in the next section. This takes care of all tensor products in $BA^\cdot$ for a commutative ordinary monoid A.
- In any monoidal category $\mathsf{C}^\otimes$, regarding any object M,N as right and left 1-modules, respectively, we have $M\otimes_1 N\cong M\otimes N$. This is easy to see from the relevant coequalizer diagram. This takes care of all tensor products in $\operatorname{disc} A^\cdot$ and $(\mathbb{R}^*)^{\min}$.

- In $(\mathbb{R}^*)^{\max}$, let $x_1, x_2, x_3 \in \mathbb{R}^*$ be monoids, and let $y \geq \max(x_1, x_2)$ be an (x_1, x_2)-bimodule, and $z \geq \max(x_2, x_3)$ be an (x_2, x_3)-bimodule. Then $y \otimes_{x_2} z$ is the coequalizer of two copies of the (trivial) inequality $\max(y, x_2, z) = \max(y, z)$, so it is the universal (minimal) object with a map from $\max(y, z)$, which is itself: hence $y \otimes_{x_2} z = \max(y, z) = y \otimes z$.

Now, in the non-strict examples:

- In $\mathsf{Set}^{\times}$, the balanced tensor product of A-sets doesn't have any special name; it just is what it is.
- In $\mathsf{C}^{\amalg}$, the balanced tensor product of cospans $X_1 \to Y \leftarrow X_2$ and $X_2 \to Z \leftarrow X_3$ is given by the *pushout*:



  that is, the universal ("minimal") object $Y \amalg_{X_2} Z$ and maps $Y, Z \to Y \amalg_{X_2} Z$ completing the two maps $X_2 \to Y, Z$ into a commutative square (the arrows in the cospan for the bimodule are given by composing these with those $X_1 \to Y$ and $X_3 \to Z$). Like the coequalizer, it is defined up to unique isomorphism.
- In $\mathsf{Ab}^{\otimes_{\mathbb{Z}}}$, balanced tensor product is just the ordinary tensor product of modules (in particular, of vector spaces, when the module is over a field).

Remark 3.24. The story of applications continues: tensor products of modules are ubiquitous throughout algebraic geometry and modern algebra. Things are even getting interesting in the "cocartesian" setting $\mathsf{C}^{\amalg}$: cospans and pushouts are frequently of interest, even in e.g. Set, where they carry some nontrivial combinatorial information, or in the study of cobordisms in differential topology.

4. The Morita (2-)category

We can now begin carrying out the construction which we've been building up to. For the duration of this section, fix a (strict) monoidal category $\mathsf{C}^{\otimes}$ admitting a calculus of balanced tensor products. The goal of this section is to prove the following theorem:

Theorem 4.1. There exists a 2-category $\mathfrak{M}(\mathsf{C}^{\otimes})$, called the *Morita 2-category of* $\mathsf{C}^{\otimes}$, in which

- the objects are monoids $A, B, C, \ldots$,
- the morphisms $A \to B$ are (A, B)-bimodules $M, N, P, \ldots$, and
- the 2-morphisms between (A, B)-bimodules $M \to N$ are (A, B)-bimodule homomorphisms.

with composition given by

- for 1-morphisms: balanced tensor product of bimodules, and
- for 2-morphisms: composition in C.

Proof. The proof will be broken up throughout the following subsections. First, Proposition 4.2 verifies that the hom-categories are well-defined. Next, Proposition 4.3 shows that horizontal composition is bifunctorial on these hom-categories. Finally, Proposition 4.4 and Proposition 4.5 verify that horizontal composition is both unital and associative. Therefore, $\mathfrak{M}(\mathsf{C}^{\otimes})$ is a well-defined 2-category. □

We'll start with the simplest part first: the 1-category structure between any two monoids.

4.1. The "upper" categories.

Proposition 4.2. Let A, B be monoids. Then there is a category $\mathfrak{M}(\mathsf{C}^{\otimes})(A, B)$ whose objects are (A, B)-bimodules, with (A, B)-bimodule homomorphisms as morphisms, and composition given by composition in C.

Proof. The only thing we have left to check is that the composition of bimodule homomorphisms is a bimodule homomorphism, and that the identity is a bimodule homomorphism, since we know composition in C is associative. The identity part is clear, since for any (A, B)-bimodule M, the diagram

$$\begin{array}{ccc} A \otimes M & \xrightarrow{1_A \otimes 1_M} & A \otimes M \\ {\scriptstyle \ell_M}\downarrow & & \downarrow{\scriptstyle \ell_M} \\ M & \xrightarrow[1_M]{} & M \end{array}$$

commutes, and similarly for the right B-action on M. The composition part is similarly verified, since for any morphisms $f\colon M \to N$ and $g\colon N \to P$ of (A, B)-bimodules, the outer square

$$\begin{array}{ccccc} A \otimes M & \xrightarrow{1_A \otimes f} & A \otimes N & \xrightarrow{1_A \otimes g} & A \otimes P \\ {\scriptstyle \ell_M}\downarrow & & {\scriptstyle \ell_N}\downarrow & & \downarrow{\scriptstyle \ell_P} \\ M & \xrightarrow[f]{} & N & \xrightarrow[g]{} & P \end{array}$$

commutes, since the two inner ones do. Hence, the category $\mathfrak{M}(\mathsf{C}^{\otimes})(A, B)$ is well-defined. □

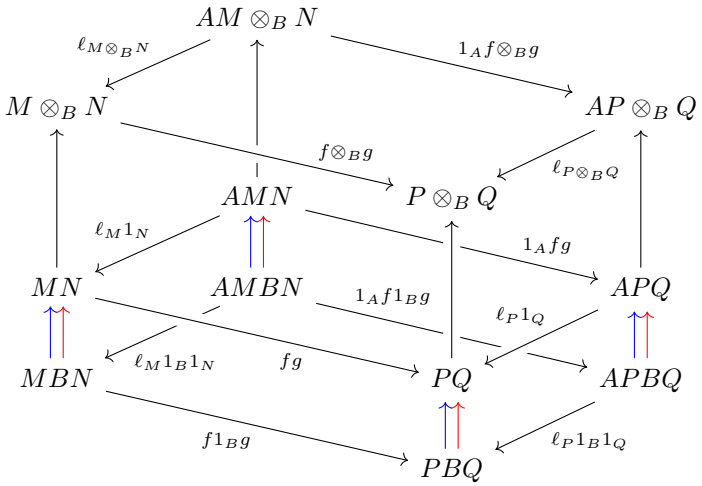
4.2. **Functoriality of composition.**

Proposition 4.3. Let A, B, C be monoids. Then the map $- \otimes_B -\colon \mathfrak{M}(\mathsf{C}^{\otimes})(A, B) \times \mathfrak{M}(\mathsf{C}^{\otimes})(B, C) \to \mathfrak{M}(\mathsf{C}^{\otimes})(A, C)$ is a functor, for the category structure introduced in Proposition 4.2.

Proof. First, we should verify that for any maps $f\colon M \to P$ of (A, B)-bimodules and $g\colon N \to Q$ of (B, C)-bimodules, there is an induced map $f \otimes_B g\colon M \otimes_B N \to P \otimes_B Q$ of (A, C)-bimodules. We define this map to be the one induced by the universal property of coequalizers in the following diagram:

$$\begin{array}{ccccc} M \otimes B \otimes N & \underset{1_M \otimes \ell_N}{\overset{r_M \otimes 1_N}{\rightrightarrows}} & M \otimes N & \longrightarrow & M \otimes_B N \\ {\scriptstyle f \otimes 1_B \otimes g}\downarrow & & \downarrow{\scriptstyle f \otimes g} & & \downarrow{\scriptstyle f \otimes_B g} \\ P \otimes B \otimes Q & \underset{1_P \otimes \ell_Q}{\overset{r_P \otimes 1_Q}{\rightrightarrows}} & P \otimes Q & \longrightarrow & P \otimes_B Q \end{array}$$

Both relevant squares on the left commute by bifunctoriality of $\otimes$ and the fact that f and g are bimodule maps, so $f \otimes_B g$ is well-defined as a morphism in C. Next, we'll verify that $f \otimes_B g$ is a left A-module homomorphism (the verification that it's a right C-module homomorphism is very similar). We set up the following diagram of coequalizers, omitting the $\otimes$ symbol for brevity:



We must check that all the squares besides the top one commute. We saw above that the front face commutes, and the back face is just the front face after applying the functor $A \otimes -$. The left and right faces commute by definition of the left actions of A on the balanced tensor products. The middle and

bottom horizontal squares commute by bifunctoriality of $\otimes$ and the fact that f is a left A-module map. Therefore, the whole diagram commutes, so the top face commutes, showing $f \otimes_B g$ is a map of left A-modules, as desired. Similarly, we see that $f \otimes_B g$ is a map of right C-modules, so we conclude that $\otimes_B$ extends to a map on pairs of (A,B)- and (B,C)-bimodule homomorphisms.

It remains to check that this map is functorial. First, for an (A,B)-bimodule M and a (B,C)-bimodule N, we must check that $1_M \otimes_B 1_N = 1_{M\otimes_B N}$, but looking at the definition above, this is true, by functoriality of coeq. Similarly, we must check that $(g_1 \otimes_B g_2)\circ(f_1 \otimes_B f_2) = (g_1 \circ f_1)\otimes_B (g_2 \circ f_2)$ for any maps $f_1 \colon M_1 \to N_1$, $g_1 \colon N_1 \to P_1$ of (A,B)-bimodules and maps $f_2 \colon M_2 \to N_2$, $g_2 \colon N_2 \to P_2$ of (B,C)-bimodules, but this is also immediate from functoriality of coequalizers. Therefore, we conclude that the composition map $\otimes_B \colon \mathfrak{M}(\mathsf{C}^{\otimes})(A,B) \times \mathfrak{M}(\mathsf{C}^{\otimes})(B,C) \to \mathfrak{M}(\mathsf{C}^{\otimes})(A,C)$ is a well-defined functor. □

4.3. **Unitality and associativity.**

Proposition 4.4. Let A be a monoid. There exists an (A,A)-bimodule, namely ${}_AA_A$, such that for any (A,B)-bimodule M, we have $A \otimes_A M = M$ as (A,B)-bimodules, and for any (C,A)-bimodule N, we have $N \otimes_A A = N$ as (C,A)-bimodules.

Proof. We will just prove the first statement; the second is proved similarly. By definition, $A \otimes_A M$ is the unique (up to unique isomorphism) object C equipped with a morphism $\pi \colon A \otimes M \to C$ which coequalizes the two natural maps $A \otimes A \otimes M \rightrightarrows A \otimes M$ (which are $\mu_A \otimes 1_M$ and $1_A \otimes \ell_M$), such that any other such morphism $A \otimes M \to \mathbb{Z}$ factors uniquely through π. Now consider the associativity diagram for the left action of A on M:

$$\begin{array}{ccc} A \otimes A \otimes M & \xrightarrow{1_A \otimes \ell_M} & A \otimes M \\ {\scriptstyle \mu_A \otimes 1_M}\downarrow & & \downarrow{\scriptstyle \ell_M} \\ A \otimes M & \xrightarrow[\ell_M]{} & M \end{array}$$

We see that M, in fact, is equipped with a morphism $\ell_M \colon A \otimes M \to M$, which coequalizes the two maps $\mu_A \otimes 1_M$ and $1_A \otimes \ell_M$ (since the above square commutes). We have a morphism $\eta_A \otimes 1_M \colon M = 1 \otimes M \to A \otimes M$, and we verify that it makes the following diagram commute:

$$\begin{array}{ccccc} A \otimes A \otimes M & \rightrightarrows & A \otimes M & \xrightarrow{\ell_M} & M \\ & & & {\scriptstyle 1_{A\otimes M}}\searrow & \downarrow{\scriptstyle \eta_A \otimes 1_M} \\ & & & & A \otimes M \end{array}$$

This is just the unitality diagram for ${}_AM$. Now for any $\psi \colon A \otimes M \to Z$ which coequalizes the two maps, we have a map $\psi \circ (\eta_A \otimes 1_M) \colon M \to Z$, and $\psi = (\psi \circ (\eta_A \otimes 1_M)) \circ \ell_M$ factors through ℓ_M by the above commutative diagram from unitality for ${}_AM$. Furthermore, this factorization is unique, since for any other $h \colon M \to Z$ such that $\psi = h \circ \ell_M$, we have

$$h = h \circ \ell_M \circ (\eta_A \otimes 1_M) = \psi \circ (\eta_A \otimes 1_M)$$

Thus, $\ell_M \colon A \otimes M \to M$ satisfies the definition of $\pi \colon A \otimes M \to A \otimes_A M$, so we conclude $A \otimes_A M = M$ as objects and $\pi = \ell_M$ as coequalizer maps.

It remains to verify that the (A,B)-bimodule structure on M is the same as the induced one on $A \otimes_A M$. We'll just verify it for the left A-action. Under our identifications from above, the defining diagram for $\ell_{A\otimes_A M}$ looks like

$$\begin{array}{ccccc} A \otimes A \otimes A \otimes M & \overset{1_A\otimes\mu_A\otimes 1_M}{\underset{1_A\otimes 1_M\otimes\ell_M}{\rightrightarrows}} & A \otimes A \otimes M & \xrightarrow{1_A\otimes\ell_M} & A \otimes M \\ {\scriptstyle \mu_A\otimes 1_A\otimes 1_M}\downarrow & & \downarrow{\scriptstyle \mu_A\otimes 1_M} & & \vdots{\scriptstyle \ell_{A\otimes_A M}=\ell_M} \\ A \otimes A \otimes M & \overset{\mu_A\otimes 1_M}{\underset{1_A\otimes\ell_M}{\rightrightarrows}} & A \otimes M & \xrightarrow[\ell_M]{} & M \end{array}$$

We know that there is a unique morphism $\ell_{A\otimes_A M} \colon A \otimes M \to M$ that makes the above diagram commute, and we know that ℓ_M does this since then the right-hand square is just the associativity diagram for ${}_AM$, so indeed the actions agree. A similar argument shows that the right B-actions agree. Therefore, we have $A \otimes_A M = M$ as (A,B)-bimodules. □

Proposition 4.5. Let A, B, C, D be monoids, and let M be an (A, B)-bimodule, N a (B, C)-bimodule, and P a (C, D)-bimodule. Then $M \otimes_B (N \otimes_C P) = (M \otimes_B N) \otimes_C P$ as (A, D)-bimodules.

Proof. This follows immediately from Fubini's theorem for coequalizers (Theorem 1.20). Indeed, we are in the following situation:

$$\begin{array}{ccc} M \otimes B \otimes N \otimes C \otimes P & \rightrightarrows & M \otimes N \otimes C \otimes P \\ \downarrow\downarrow & & \downarrow\downarrow \\ M \otimes B \otimes N \otimes P & \rightrightarrows & M \otimes N \otimes P \end{array}$$

The relevant squares all commute by bifunctoriality of $\otimes$, and the two coequalizers in question are by definition $M \otimes_B (N \otimes_C P)$ and $(M \otimes_B N) \otimes_C P$, so we obtain $M \otimes_B (N \otimes_C P) = (M \otimes_B N) \otimes_C P$, as desired. □

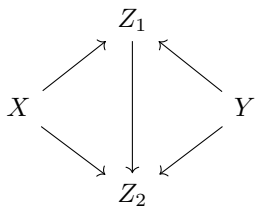
4.4. **Examples.** Now we have this 2-category, but what does it look like in our examples?

Example 4.6. Strict monoidal categories:

- For an ordinary commuatative monoid A, $\mathfrak{M}(BA^{\cdot})$ has as objects the invertible elements of A, as 1-morphisms only the identities, and a 2-morphism (bimodule map) from the identity bimodule on some invertible elemnt $\mu \in A$ to itself would be an element $f \in A$ which commutes with the action μ, that is, $f\mu = \mu f$, a tautology in a commutative monoid. Thus, the 2-morphisms $\mu \Rightarrow \mu$ are in bijection with elements of A.
- For a monoid A, $\mathfrak{M}(\operatorname{disc} A^{\otimes})$ has one object (1), and one 1-morphism $1 \to 1$ for each element of A. A bimodule homomorphism $a \Rightarrow b$ for $a, b \in A$ would be a morphism in $\operatorname{disc} A$, which can only be an identity, so we must have $b = a$ and there's just one 2-morphism $a \Rightarrow a$ for each a. We note that this is a monoidal category, and in fact $\mathfrak{M}(\operatorname{disc} A^{\otimes}) \cong \operatorname{disc} A^{\otimes}$.
- Similarly, $\mathfrak{M}((\mathbb{R}^*)^{\min})$ has just one object ∞, and one 1-morphism $\infty \to \infty$ for each element of $\mathbb{R}^*$. A 2-morphism $x \Rightarrow y$ would be in particular a morphism $x \to y$ in $\mathbb{R}^*$, which is unique and exists if $x \leq y$. Since hom-sets have at most one element here, all squares commute, so these are all bimodule maps: the 2-morphisms are precisely the morphisms in $\mathbb{R}^*$. In fact, we see that again, this is a monoidal category, and $\mathfrak{M}((\mathbb{R}^*)^{\min}) \cong (\mathbb{R}^*)^{\min}$.
- The objects of $\mathfrak{M}((\mathbb{R}^*)^{\max})$ are given by the elements of $\mathbb{R}^*$, and the 1-morphisms $x_1 \to x_2$ for $x_1, x_2 \in \mathbb{R}^*$ are given by the elements $y \in \mathbb{R}^*$ such that $y \geq \max(x_1, x_2)$. For $y, z \in \mathbb{R}^*$ such elements, an (x_1, x_2)-bimodule homomorphism would be a morphism $y \to z$, which exists and is a valid bimodule map iff $y \leq z$. Hence, the 2-morphisms are just given by inequalities of bimodules.

Non-strict monoidal categories:

- In $\mathfrak{M}(\mathsf{C}^{\amalg})$, the objects are precisely those of C, but now the 1-morphisms $X \to Y$ are the cospans $X \to Z \leftarrow Y$, and the 2-morphisms are *maps of cospans*, that is, given $X \to Z_i \leftarrow Y$ for $i = 1, 2$, maps $Z_1 \to Z_2$ such that the relevant triangles commute:



- In $\mathfrak{M}(\mathsf{Ab}^{\otimes})$, the objects are (not necessarily commutative) rings, the morphisms $R \to S$ are (R, S)-bimodules, and the 2-morphisms are bimodule maps. This was the setting in which this theory was originally developed and applied by Kiiti Morita. In the 1950s, he proved some structure theorems about an equivalence relation he defined between rings, known as *Morita equivalence*, which turns out to coincide with equivalence in the above-defined Morita 2-category. For more information, the reader might consult the notes [2] by Hyman Bass which first popularized the subject, or his exposition in Chapter 2 of [1]. One prototypical example (not proved here): for

any ring R and any integer $n > 0$, the ring $M_n(R)$ of $n \times n$ matrices over R is Morita equivalent to R.

5. Appendix

5.1. **The strictification theorem for monoidal categories.** So as not to derail the discussion, the proof of Proposition 3.8 has been relegated to the appendix. We will need several more definitions, whose statements we will omit here for brevity, but which can be found in [7]: the general definition of monoidal category, monoidal functors, and monoidal natural transformations (giving rise to a 2-category of monoidal categories, and thus to monoidal equivalence), as well as a skeleton of a category C and the fact that C is equivalent to its skeleton. With these in mind, we recall the statement we'd like to prove:

Proposition 5.1. Every monoidal category is (strictly monoidally) equivalent to a strict monoidal category.

Proof. Let $\mathsf{C}^{\otimes}$ be a monoidal category. Choose a skeleton S of C and an equivalence $F\colon \mathsf{C} \to S$ (relative to the inclusion $I\colon S \hookrightarrow \mathsf{C}$) with $FI = \mathrm{id}_S$ and a natural isomorphism $\eta\colon IF \Rightarrow \mathrm{id}_{\mathsf{C}}$.

Define a product $\boxtimes\colon S \times S \to S$ by $x \boxtimes y = F(x \otimes y)$ (on objects and morphisms), and a unit object for $S^{\boxtimes}$ given by $F(1_{\mathsf{C}})$. We also have an associator, also by composing the source and target functors with F. As for the unitors, let $x \in S$ and choose $y \in S$ with $F(y) = x$. Then

$$1_S \boxtimes x = F(1_{\mathsf{C}}) \boxtimes x = F(F(1_{\mathsf{C}}) \otimes x) = F(F(1_C \otimes y)) = F(1_C \otimes y) \xrightarrow{F(\lambda_y)} F(y) = x$$

so we can just define the left and right unitors by composing with F as well. All the axioms of a monoidal category now follow immediately from functoriality of F.

For monoidality of F, the morphism $\epsilon\colon 1_S = F(1_{\mathsf{C}})$ is just the identity, and so is $\mu_{x,y}\colon F(x) \boxtimes F(y) = F(x \otimes y)$. The rest of the axioms for monoidality of F and η are immediate to verify by looking at the diagrams, by functoriality of F and our definitions. Therefore, $\mathsf{C}^{\otimes}$ is strictly monoidally equivalent to the strict monoidal category $S^{\boxtimes}$. □

5.2. **The path 2-groupoid of a space X.** The following is included as an explicit definition and verification of a toy example of a 2-category, in the hopes that it would be instructive to see the motions that one might have to carry out in such a definition/proof.

Claim 5.2. We would like to claim that given a space X, there is a 2-category with something like "points of X as objects, reparametrization classes of paths as morphisms, and homotopy classes of path homotopies as 2-morphisms". But we must make this precise. Fix a topological space X, and denote the interval $[0,1] \subseteq \mathbb{R}$ by I.

Definition 5.3. Let $x, y \in X$ be points, and let $f, f'\colon I \to X$, $f(0) = f'(0) = x$, $f(1) = f'(1) = y$ be paths from x to y. We say f' is a *reparametrization* of f if there exists a homeomorphism $r\colon I \to I$ from the interval to itself, relative to (fixing) its boundary, the endpoints $\partial I = \{0,1\}$, such that $f' = f \circ r$. We'll denote this by $f \sim_r f'$.

Proposition 5.4. The relation of reparametrization defines an equivalence relation on the set of paths from x to y, for any points $x, y \in X$.

Proof. Each path is a reparametrization of itself, via the identity, which fixes the endpoints: $f = f \circ 1_I$. If f' is a reparametrization of f via r, so $f' = f \circ r$, then f is a reparametrization of f' via r^{-1}: $f = f' \circ r^{-1}$, and r^{-1} fixes $\{0,1\}$ if r does. Finally, if $f' = f \circ r$ and $f'' = f' \circ r'$ are two reparametrizations, then $f'' = f \circ (r \circ r')$, and $r \circ r'$ is a homeomorphism $I \to I$ preserving the endpoints whenever r, r' are. Therefore, reparametrization is reflexive, symmetric, and transitive, and hence an equivalence relation. □

Definition 5.5. Let $x, y \in X$ be points, and let f, g be two paths x to y. We say a path homotopy F from f to g is *thin* if it's of the form $F = f \circ G$ for some path homotopy $G\colon I \times I \to I$ starting at the identity path $0 \to 1$ in I.

Lemma 5.6. Let $x, y \in X$, and let f, f' be paths $x \to y$. If f' is a reparametrization of f, then there is a thin path homotopy from f to f'.

Proof. Since $I = [0,1]$ is a convex subset of $\mathbb{R}$, we can take the straight-line homotopy from $1_{[0,1]}$ to r. Since both fix the endpoints, the homotopy will fix the endpoints too, since the line from each endpoint to itself is constant. Hence, composing this with f, we are done. □

Lemma 5.7. Let x, y, f, f' be as in Lemma 5.6. Then any two thin path homotopies from f to f' are homotopic as path homotopies.

Proof. Similarly, since $I \times I$ is a convex subspace of $\mathbb{R}^2$, we can take the straight-line homotopy from the first path homotopy to the second. Since they agree on the boundary $\partial(I \times I)$, this homotopy is constant on $\partial(I \times I)$, and thus its composition with f is a homotopy of path homotopies. □

Now we must set up the structure for our morphisms.

Definition 5.8. Let $x, y \in X$, let f, f', g, g' be paths $x \to y$, and let φ, ψ be path homotopies, from f to g and from f' to g' respectively. We say $\varphi \sim \psi$ if $f \sim_r f'$ and $g \sim_r g'$, and there exist thin path homotopies $F, G \colon I \times I \to X$ from f to f' and g to g' respectively such that the concatenation of path homotopies $(F \cdot \psi) \cdot \tilde{G}$ is homotopic (relative to $\partial(I \times I)$, that is, homotopic as a path homotopy) to φ. Here, $\tilde{G}$ denotes the reversal of G, a path homotopy from g' to g.

Proposition 5.9. The relation $\sim$ is an equivalence relation on the set of path homotopies between arbitrary paths $x \to y$. Furthermore, each equivalence class consists of path homotopies from paths in one reparametrization class to paths in another.

Proof. Each path homotopy is equivalent to itself, via the identity reparametrizations on its source and target, the constant path homotopies on them, and by the usual proof of unitality for path concatenation, applied to path homotopies. If $\varphi \sim \psi$ via the data $\varphi \colon f \to g$, $\psi \colon f' \to g'$, $F \colon f \to f'$ and $G \colon g \to g'$ such that $(F \cdot \psi) \cdot \tilde{G} \simeq \varphi$, then $\psi \sim \varphi$ via the data $f', g', f, g, \tilde{F}, \tilde{G}$, such that $(\tilde{F} \cdot \varphi) \cdot G \simeq \psi$. Finally, if $\varphi \sim \psi$ and $\psi \sim \chi$, via respective data

- $\varphi \colon f \to g$, $\psi \colon f' \to g'$, $F \colon f \to f'$ and $G \colon g \to g'$ such that $(F \cdot \psi) \cdot \tilde{G} \simeq \varphi$ and
- $\psi \colon f' \to g'$, $\chi \colon f'' \to g''$, $F' \colon f' \to f''$ and $G' \colon g' \to g''$ such that $(F' \cdot \chi) \cdot \tilde{G'} \simeq \psi$

then $\varphi \sim \chi$ via $\varphi \colon f \to g$, $\chi \colon f'' \to g''$, $F \cdot F' \colon f \to f''$ and $G \cdot G' \colon g \to g''$ such that

$$\begin{aligned}((F \cdot F') \cdot \chi) \cdot \widetilde{(G \cdot G')} &\simeq ((F \cdot F') \cdot \chi) \cdot (\tilde{G'} \cdot \tilde{G}) \\ &\simeq (F \cdot ((F' \cdot \chi) \cdot \tilde{G'})) \cdot \tilde{G} \\ &\simeq (F \cdot \psi) \cdot \tilde{G} \\ &\simeq \varphi\end{aligned}$$

Therefore, $\simeq$ is reflexive, symmetric, and transitive, so it defines an equivalence relation on arbitrary paths $x \to y$.

Furthermore, we note that the definition of $\varphi \sim \psi$ in particular implies that the sources and targests are reparametrizations of each other, so each equivalence class consists of path homotopies from paths in one reparametrization class to paths in another. □

We need a few lemmas for well-definedness of compositions.

Lemma 5.10. Let x, y, z be points in X, and let f, f' be paths $x \to y$, and g be a path $y \to z$. If f' is a reparametrization of f, then $f' * g$ is a reparametrization of $f * g$. The dual statement for $f \circ g, g'$ also holds, and together, these guarantee that path concatenation $*$ descends to a well-defined binary operation on reparametrization classes of paths with appropriate sources and targets.

Proof. We'll just prove the first version; the second is proved the same way. Let $r \colon I \to I$ be a homeomorphism relative to ∂I such that $f' = f \circ r$. Define $\bar{r} \colon I \to I$ by

$$\bar{r}(t) = \begin{cases} r(t/2) & \text{if } 0 \leq t \leq 1/2 \\ t & \text{otherwise} \end{cases}$$

Then by construction (recalling the definition of $f * g$), $\bar{r}$ is a self-homeomorphism of I, still fixing the endpoints, such that $f' * g = (f * g) \circ \bar{r}$. □

Lemma 5.11. Let x, y, z be points in X, and let $\varphi \sim \varphi'$ be equivalent path homotopies between some paths $x \to y$, call them f, g and f', g' respectively, and ψ be a path homotopy between paths h, h' from y to z. Then $\varphi' * \psi \sim \varphi * \psi$. The dual statement for $\varphi * \psi, \psi'$ also holds, and together, these guarantee that horizontal path homotopy concatenation $*$ descends to a well-defined binary operation on equivalence classes under $\sim$ of path homotopies with appropriate sources and targets.

Proof. Again, we prove only one of the statements. The construction given above extends verbatim to the entirety of φ, φ', and the chosen path homotopy between them (compatible with some thin path homotopies on the sources and targets), evaluated at a fixed time, and is clearly continuous with respect to this time. Unitality and associativity of path concatenation then yield the desired equivalence $\varphi' * \psi \sim \varphi * \psi$. □

Definition 5.12. Let $x, y \in X$ be points, let f, g, g', h be paths from x to y, with g' a reparametrization of g, and let φ, ψ be path homotopies, from f to g and from g' to h, respectively. Fix a thin path homotopy G from g to g'. We define $\varphi \circ_G \psi$ as the vertical path homotopy concatenation $(\varphi \cdot G) \cdot \psi$.

Lemma 5.13. Let $x, y \in X$ be points, let f, g, g', h be paths from x to y, with g' a reparametrization of g, and let φ, ψ be path homotopies, from f to g and from g' to h, respectively. For any thin path homotopies G, G' from g to g', we have $\varphi \circ_G \psi \sim \varphi \circ_{G'} \psi$.

Proof. This is an immediate consequence of Lemma 5.7: we take any homotopy of path homotopies from G to G' and whisker it on the left and right with the constant homotopies of path homotopies on φ and ψ. Taking the constant (thin) path homotopies on f and h completes the equivalence data for $\varphi \circ_G \psi \sim \varphi \circ_{G'} \psi$, as desired. □

Lemma 5.14. Let x, y be points in X, let f, f', g, g', h be paths $x \to y$, let $\varphi \sim \varphi'$ be equivalent path homotopies $f \to g$ and $f' \to g'$ respectively, and let ψ be a path homotopy from g to h. Fixing a thin path homotopy G from g' to g, we have $\varphi \circ_1 \psi \sim \varphi' \circ_G \psi$.

Proof. As usual, we just prove one of the statements. We have a thin path homotopy F from f to f', and we check that

$$\begin{aligned}(F \cdot ((\varphi' \cdot G) \cdot \psi)) \cdot 1_h &\simeq ((F \cdot \varphi') \cdot G) \cdot \psi \\ &\simeq \varphi \cdot \psi \\ &\simeq (\varphi \cdot 1) \cdot \psi\end{aligned}$$

Thus, we have our equivalence data for $\varphi \circ_G \psi \sim \varphi \circ_{G'} \psi$, as desired. □

Corollary 5.15. Note that Lemmas 5.13 and 5.14 together imply that we obtain a well-defined binary operation $\circ$ (which we call "concatenation") at the level of $\sim$-equivalence classes of path homotopies whose appropriate sources and targets lie in the same reparametrization class.

We are now ready to state the definition of our 2-category.

Theorem 5.16. There is a 2-category $P_{\leq 2}(X)$ whose objects are points $x, y, \dots \in X$, in which the hom-category $P_{\leq 2}(x, y)$ has

- as objects (1-morphisms), reparametrization classes of paths $[f], [g], \dots : x \to y$,
- as morphisms (2-morphisms) $[f] \to [g]$, equivalence classes under $\simeq$ of path homotopies such that (by Proposition 5.9) their representatives have sources in $[f]$ and targets in $[g]$,

with vertical composition given by concatenation of path homotopy equivalence classes (in the sense of Corollary 5.15), and horizontal composition given on 1-morphisms by concatenation of reparametrization equivalence classes of paths, and on 2-morphisms by horizontal concatenation of path homotopy equivalence classes.

Proof. The proof will be broken up into Propositions 5.17 to 5.20 below. □

Proposition 5.17. Vertical composition of 2-morphisms is unital and associative.

Proof. For unitality, we just take the class of the constant path homotopy at the source or target. Associativity follows from the associativity of usual path homotopy concatenation applied to $\varphi \cdot G \cdot \psi \cdot H \cdot \chi$, where G and H are some thin path homotopies in between the appropriate sources and targets. □

Proposition 5.18. Horizontal composition is functorial.

Proof. Again, this follows from regular interchange of path homotopies applied to the setting where we insert the appropriate reparametrizations. □

Proposition 5.19. Horizontal composition is unital and associative on 1-morphisms.

Proof. Note that, in the usual proof of unitality and associativity up to homotopy of path concatenation, the homotopies are actually through reparametrizations, so in fact path concatenation is unital and associative up to reparametrization. □

Proposition 5.20. Horizontal composition is unital and associative on 2-morphisms.

Proof. The same exact proof as above applies, extending the usual path homotopy to a homotopy of path homotopies. □

References

[1] H. Bass. *Algebraic K-theory*. Hyman Bass. W. A. Benjamin, 1968. ISBN: 9780805306606. URL: https://books.google.com/books?id=zOaEAAAAIAAJ.

[2] Hyman Bass. *The Morita Theorems*. 1962. URL: https://catalog.hathitrust.org/Record/009372348/Home.

[3] Tai-Danae Bradley. *What is Applied Category Theory?* 2018. arXiv: 1809.05923 [math.CT]. URL: https://arxiv.org/abs/1809.05923.

[4] Ronnie Brown (https://math.stackexchange.com/users/28586/ronnie brown). *Eckmann-Hilton and higher homotopy groups*. Mathematics Stack Exchange. URL:https://math.stackexchange.com/q/162329 (version: 2021-01-03). eprint: https://math.stackexchange.com/q/162329. URL: https://math.stackexchange.com/q/162329.

[5] Arghan Dutta, Stefano Luneia, Martina Rovelli, and Sam Silver. *The Morita* $(\infty, 2)$*-category of a monoidal category as a* 2*-complicial set*. 2025. arXiv: 2509.21472 [math.CT]. URL: https://arxiv.org/abs/2509.21472.

[6] B. Eckmann and P. J. Hilton. "Group-like structures in general categories I multiplications and comultiplications". In: *Mathematische Annalen* 145.3 (June 1962), 227–255. ISSN: 1432-1807. DOI: 10.1007/bf01451367. URL: http://dx.doi.org/10.1007/bf01451367.

[7] Saunders Mac Lane. *Categories for the working mathematician*. Second. Vol. 5. Graduate Texts in Mathematics. Springer-Verlag, New York, 1998, pp. xii+314. ISBN: 0-387-98403-8.

[8] Ralf Meyer. *Morita Equivalence In Algebra And Geometry*. 1997. URL: https://ncatlab.org/nlab/files/MeyerMoritaEquivalence-2.pdf.

[9] Kiiti Morita. "Adjoint pairs of functors and Frobenius extensions". In: *Science Reports of the Tokyo Kyoiku Daigaku (Tokyo University of Education)* Section A 9 (1965), 40–71.

[10] Kiiti Morita. "Duality for modules and its applications to the theory of rings with minimum condition". In: *Science Reports of the Tokyo Kyoiku Daigaku (Tokyo University of Education)* Section A 150 (1958), 83–142.

[11] Emily Riehl. *Category theory in context*. Dover Publications, 2016. ISBN: 9780486809038.