Could ∞-Category Theory Be Taught to Undergraduates?

Emily Riehl

1. The Algebra of Paths

It is natural to probe a suitably nice topological space X by means of its *paths*, the continuous functions from the standard unit interval $I := [0, 1] \subset \mathbb{R}$ to X. But what structure do the paths in X form?

To start, the paths form the edges of a directed graph whose vertices are the points of X: a path $p : I \to X$ defines an arrow from the point $p(0)$ to the point $p(1)$. Moreover,

Emily Riehl is a professor of mathematics at Johns Hopkins University. Her email address is eriehl@jhu.edu.

this graph is *reflexive*, with the constant path refl_x at each point $x \in X$ defining a distinguished endoarrow.

Can this reflexive directed graph be given the structure of a category? To do so, it is natural to define the composite of a path p from x to y and a path q from y to z by gluing together these continuous maps—i.e., by concatenating the paths—and then by reparametrizing via the homeomorphism $I \cong I \cup_{1=0} I$ that traverses each path at double speed:

$$I \xrightarrow{\cong} I \cup_{1=0} I \xrightarrow{p \cup q} X \qquad (1.1)$$
$$\underset{p * q}{\longrightarrow}$$

But the composition operation $*$ fails to be associative or unital. In general, given a path r from z to w, the

composites $(p * q) * r$ and $p * (q * r)$ are not equal: while the have the same image in X, their parametrizations differ. However, they are *based homotopic*, in the sense that there exists a continuous function $h : I \times I \to X$ so that

$$h(-, 0) := (p * q) * r, \quad h(-, 1) := p * (q * r),$$
$$h(0, -) := \mathrm{refl}_x, \quad \text{and} \quad h(1, -) := \mathrm{refl}_w,$$

a situation we summarize by writing $(p * q) * r \simeq p * (q * r)$. Similarly, the paths $p * \mathrm{refl}_y \simeq \mathrm{refl}_x * p \simeq p$ are all based homotopic, though not equal (unless p is the constant path).

Paths are also invertible up to based homotopy: for any path $p : I \to X$ from x to y, its reversal $p^{-1} : I \to X$, defined by precomposing with the flipping automorphism $I \cong I$, defines an inverse up to based homotopy: $p * p^{-1} \simeq \mathrm{refl}_x$ and $p^{-1} * p \simeq \mathrm{refl}_y$. These observations motivate the following definition:

Definition 1.2. For a space X, the **fundamental groupoid** $\pi_1 X$ is the category whose:

- objects are the points of X and
- arrows are based homotopy classes of paths of X

with composition defined by concatenation, identity arrows defined by the constant paths, and inverses defined by reversing paths.

The fundamental groupoid of a space X answers a slightly different question than originally posed, describing the structure formed by the *based homotopy classes of paths* in X. The paths themselves form something like a weak groupoid where composition is not uniquely defined. Indeed, given paths p from x to y and q from y to z, the composite in $\pi_1 X$ is represented by *any* path s so that there is a based homotopy h witnessing $p * q \simeq s$. Here, the based homotopy h defines a continuous function from the solid 2-simplex into X that restricts along the boundary triangle to the map defined by gluing the three paths:

While multivalued composition operations are not well-behaved in general, this one has a certain homotopical uniqueness property: the witnessing homotopies for two such composites can be glued together—by a higher-dimensional analogue of the construction (1.1)—to define a homotopy $s \simeq p * q \simeq t$. This suggests that the witnessing homotopies that fill the triangles formed by the paths between x and y and z might be productively regarded as part of the composition data. Thus, the moduli space of composites of p and q is defined as follows:

Definition 1.3. Given composable paths p and q in a space X, the **space of composites** of p and q is the subspace

$\mathrm{Comp}(p, q) \subset \mathrm{Map}(\Delta^2, X)$ of continuous maps $h : \Delta^2 \to X$ that restrict on the boundary horn to $p \cup q$:

(1.4)

The homotopical uniqueness of path composition can be strengthened by the following observation:

Theorem 1.5. *For any composable paths p and q in a space X, the space of composites $\mathrm{Comp}(p, q)$ is contractible.*

So what do the paths in X form? We have seen that they form a weak groupoid with a multivalued but homotopically unique composition law. In fact, this weak category is an infinite-dimensional category. The based homotopies $h : I \times I \to X$ between parallel paths s and t themselves might be regarded as paths $h : I \to \mathrm{Map}(I, X)$ between points s and t in the space $\mathrm{Map}(I, X)$ to which Theorem 1.5 equally applies, and these observations extend iteratively to higher-dimensional homotopies. The points, paths, and higher paths in a space X assemble into a weak infinite-dimensional category—with interacting weak composition and identities for paths at all levels[1]—in which all morphisms are weakly invertible. Such a structure is known as an ∞-*groupoid*.

The homotopy hypothesis. The *fundamental ∞-groupoid* $\pi_\infty X$ of paths in X captures the data of all of the higher homotopy groups of the space, information which is referred to as the *homotopy type* of the space. In a famous letter to Quillen [G], Grothendieck formulated his *homotopy hypothesis*, positing that the fundamental ∞-groupoid construction defines an equivalence between homotopy types and ∞-groupoids. Grothendieck's vision was that this result should be provable for various models of ∞-groupoids that were then under development, though some instead use this thesis to *define* an ∞-groupoid to be a homotopy type.

There is something unsatisfying, though, about the naïve interpretation of the homotopy hypothesis as the assertion of a bijection between the collection of homotopy types and the collection of ∞-groupoids, or even as an equivalence between the homotopy category of spaces[2] and the homotopy category of ∞-groupoids. The disappointment lies in the fact that both spaces and ∞-groupoids live most naturally as the objects in weak infinite-dimensional categories, with the

---

[1]*Famously, the definition of a weak 3-dimensional category by Gordon, Power, and Street takes six pages to state.*

[2]*Homotopy types can be understood as isomorphism classes of objects in the* homotopy category of spaces, *the category obtained by localizing the category of spaces and continuous functions by the* weak homotopy equivalences, *those functions that induce isomorphisms on all homotopy groups.*

standard morphisms between them supplemented by higher-dimensional weakly invertible morphisms. Thus, a more robust expression of the homotopy hypothesis is the assertion that the ∞-*categories* of spaces and of ∞-groupoids are equivalent,[3] so now we must explain what ∞-categories are.

2. ∞-Categories in Mathematics

Over the past few decades, ∞-*categories*—weak infinite-dimensional categories with weakly invertible morphisms above dimension one—have been invading certain areas of mathematics. In derived algebraic geometry, the derived category of a ring is now understood as the 1-categorical quotient of the ∞-category of chain complexes, and an ∞-categorical property called "stability" explains the triangulated structure borne by the derived category. In mathematical physics, Atiyah's topological quantum field theories have been "extended up" to define functors between ∞-categories.[4] ∞-categories have also made appearances in the Langlands program in representation theory and in symplectic geometry among other areas. Quillen's model categories [Q] from abstract homotopy theory are now understood as presentations of ∞-categories.

Ordinary categories "frame a template for a mathematical theory," with the objects providing the "nouns" and the morphisms the "verbs," in a metaphor suggested by Mazur. As the objects mathematicians study increase in complexity, a more robust linguistic template may be required to adequately describe their natural habitats—with adjectives, adverbs, pronouns, prepositions, conjunctions, interjections, and so on—leading to the idea of an ∞-category. Like an ordinary 1-category, an ∞-category has objects and morphisms, now thought of as "1-dimensional" transformations. The extra linguistic color is provided by higher-dimensional invertible morphisms between morphisms—such as chain homotopies or diffeomorphisms—and higher morphisms between these morphisms, continuing all the way up.

How might a researcher in one of these areas go about learning this new technology of ∞-categories? And how might ∞-category theory ultimately be distilled down into something that we could reasonably teach advanced undergraduates of the future?

Curiosities from the literature. ∞-categories were first introduced by Boardman and Vogt to describe the composition structure of homotopy coherent natural transformations between homotopy coherent diagrams [BV]. Joyal was the first to assert that "most concepts and results of category theory can be extended to [∞-categories]" and pioneered the development of ∞-categorical analogues of standard categorical notions [J]. Lurie then developed various aspects of ∞-category theory that were needed for his thesis on derived algebraic geometry [L1]. His books [L2, L3] and the online textbook *Kerodon* are primary references for uses of this technology, while texts written by Cisinski, Groth, Hinich, Rezk, and others provide parallel introductions to the field.

If one delves further into the ∞-categories literature, some curiosities soon become apparent:

(i) Particularly in talks or lecture series introducing the subject, the definition of ∞-category is frequently delayed, and when definitions are given, they don't always agree.

These competing definitions are referred to as *models* of ∞-categories, which are Bourbaki-style mathematical structures defined in terms of sets and functions that represent infinite-dimensional categories with a weak composition law in which all morphisms above dimension one are weakly invertible. In order of appearance, these include *simplicial categories*, *quasi-categories*, *relative categories*, *Segal categories*, *complete Segal spaces*, and *1-complicial sets*, each of which comes with an associated array of naturally occurring examples [Be].

(ii) Considerable work has gone into defining the key notions for and proving the fundamental results about ∞-categories, but sometimes this work is later redeveloped starting from a different model.

For instance, [KV] begins:

> In recent years ∞-categories or, more formally, (∞, 1)-categories appear in various areas of mathematics. For example, they became a necessary ingredient in the geometric Langlands problem. In his books [L2, L3] Lurie developed a theory of ∞-categories in the language of quasi-categories and extended many results of the ordinary category theory to this setting.
>
> In his work [R1] Rezk introduced another model of ∞-categories, which he called complete Segal spaces. This model has certain advantages. For example, it has a generalization to (∞, n)-categories (see [R2]).
>
> It is natural to extend results of the ordinary category theory to the setting of complete Segal spaces. In this note we do this for the Yoneda lemma.

(iii) Alternatively, authors decline to pick a model at all and instead work "model-independently."

---

[3] *In fact, a version of this result had been proven already by Quillen for the Kan complex model of ∞-groupoids* [Q], *but this was not so clearly understood at the time.*

[4] *Lurie's "fully extended" topological quantum field theories are also "extended down" so that they might be understood as functors between* (∞, n)-*categories, with non-invertible morphisms up to and including the dimension n of the indexing cobordisms. Here we reserve the term "∞-categories" for "*(∞, 1)-*categories," which have non-invertible morphisms only in the bottom dimension.*

One instance of this appears in the precursor [L1] to [L2], which avoids selecting a model of ∞-categories at all:

> We will begin in §1 with an informal review of the theory of ∞-categories. There are many approaches to the foundation of this subject, each having its own particular merits and demerits. Rather than single out one of those foundations here, we shall attempt to explain the ideas involved and how to work with them. The hope is that this will render this paper readable to a wider audience, while experts will be able to fill in the details missing from our exposition in whatever framework they happen to prefer.

The fundamental obstacle to giving a uniform definition of an ∞-category is that our traditional set-based foundations for mathematics are not really suitable for reasoning about ∞-categories: sets do not feature prominently in ∞-categorical data, especially when ∞-categories are only well-defined up to equivalence, as they must be when different models are involved. When considered up to equivalence, ∞-categories, like ordinary categories, do not have a well-defined set of objects. Essentially, ∞-categories are 1-categories in which all the sets have been replaced by ∞-groupoids. Where a category has a set of elements, an ∞-category has an ∞-groupoid of elements, and where a category has hom-sets of morphisms, ∞-categories have ∞-groupoidal mapping spaces.[5] The axioms that turn a directed graph into a category are expressed in the language of set theory: a category has a composition function satisfying axioms expressed in first-order logic with equality. By analogy, composition in an ∞-category can be understood as a morphism between ∞-groupoids, but such morphisms no longer define functions since homotopy types do not have underlying sets of points.[6] This is why there is no canonical model of ∞-categories.

Reimagining the foundations of ∞-category theory. Despite these subtleties, it is possible to reason rigorously and model-independently about ∞-categories without getting bogged down in the combinatorial scaffolding of a particular model. The framework introduced in §3 considerably streamlines the basic core theory of ∞-categories, though its scope is currently more limited than the corpus of results that have been proven using a model.

However, this current state of the art employs proof techniques that are unfamiliar to non-category theorists and thus is not feasible to integrate into the undergraduate curriculum. The concluding §4 describes a more speculative dream for the future where enhancements to the foundations of mathematics would allow us to interpret

uniqueness as contractibility and automatically ensure that all constructions are invariant under equivalence.

3. The Formal Theory of ∞-Categories

In the paper "General theory of natural equivalences" that marked the birth of category theory, Eilenberg and Mac Lane observed that categories, functors, and natural transformations assemble into a 2-category CAT. The essence of this result is the observation that natural transformations can be composed in 2-dimensions' worth of ways: "vertically" along a boundary functor or "horizontally" along a boundary category.

More generally, Power proves that any *pasting diagram* of compatibly oriented functors and natural transformations has a unique composite.
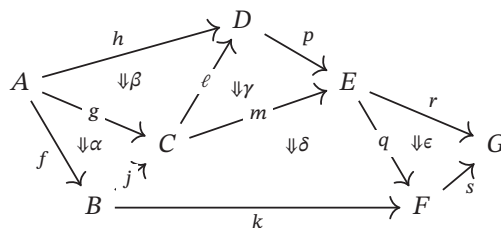
The pasting composite, which in the example above defines a natural transformation between the categories A and G from the functor rph to the functor skf, can be decomposed as a vertical composite of whiskerings of the atomic natural transformations $\alpha, \beta, \gamma, \delta, \epsilon$: for instance, the composite factors as $rp\beta$ followed by $r\gamma g$ followed by ϵmg followed by $sqm\alpha$ followed by $s\delta f$, among eight total possibilities. Power's theorem is that pasting composition is well-defined.

Similarly, in well-behaved models of ∞-categories:

Proposition 3.1 (Joyal, Riehl–Verity). ∞-*categories,* ∞-*functors, and* ∞-*natural transformations assemble into a cartesian closed 2-category* ∞-CAT.

This result was first observed in the quasi-category model by Joyal [J]. Since the category of quasi-categories is cartesian closed, it is enriched over itself, defining an $(\infty, 2)$-category of quasi-categories. The 2-category of quasi-categories is obtained by a quotienting process, that maps each hom-quasi-category to its homotopy category, by applying the left adjoint to the (nerve) inclusion of categories into quasi-categories. The other "well-behaved" models—including complete Segal spaces, Segal categories, and 1-complicial sets among others—are also cartesian closed, so a similar construction defines a 2-category of complete Segal spaces, and so on.[7] To adopt

---

[5]*"Large"* ∞-*categories also exist and behave like large 1-categories.*

[6]*Similar considerations have motivated Scholze et al to use the term "anima"—referring to the "soul" of a "space"—as a synonym for* ∞-*groupoids.*

[7]*While the meaning of the terms "*∞*-categories" and "*∞*-functors" in a given model is typically clear, the "*∞*-natural transformations," which can be understood as equivalence classes of 2-cells in the ambient* $(\infty, 2)$*-category, are less evident.*

a "model-independent" point of view, note that these 2-categories are all *biequivalent* or "the same" in the sense appropriate to 2-category theory [RV2, §E.2].

The good news, which is surprising to many experts, is that a fair portion of the basic theory of ∞-categories can be developed in the 2-category ∞-CAT. These aspects might be described as *formal category theory*, as they involve definitions of categorical notions such as equivalence, adjunction, limit, and colimit that can be defined internally to any 2-category. In the 2-category CAT, these recover the classical notions from 1-category theory, while in the 2-category ∞-CAT these specialize to the correct notions in ∞-category theory.[8] Thus, for the core basic theory involving these notions, ordinary category theory extends to ∞-categories simply by appending the prefix "∞−" [RV2, §2-4].

Equivalences and adjunctions. The following definitions make sense in an arbitrary cartesian closed 2-category, such as ∞-CAT.

Definition 3.2. An **equivalence** between ∞-categories is given by:

- a pair of ∞-categories A and B,
- a pair of ∞-functors $g : A \to B$ and $h : B \to A$, and
- a pair of invertible ∞-natural transformations $\alpha : \mathrm{id}_A \cong hg$, and $\beta : gh \cong \mathrm{id}_B$.

Definition 3.3. An **adjunction** between ∞-categories is given by:

- a pair of ∞-categories A and B,
- a pair of ∞-functors $u : A \to B$ and $f : B \to A$, and
- a pair of ∞-natural transformations $\eta : \mathrm{id}_B \Rightarrow uf$ and $\epsilon : fu \Rightarrow \mathrm{id}_A$

so that the following pasting identities hold:

$$
\begin{array}{c}
B \xmapsto{\Downarrow\eta} B \quad f \Downarrow\epsilon \quad = f\left(\right)f \quad u \Downarrow\epsilon \quad f \Downarrow\eta \, u \quad = u\left(\right)u.\\
A \quad A \quad A \quad A \quad A \quad A
\end{array}
$$

One commonly writes $A \underset{u}{\overset{f}{\rightleftarrows}} \perp B$ to indicate that f is **left adjoint** to u, its **right adjoint**, with the data of the **unit** η and **counit** ϵ being left implicit.

Among the many advantages of using definitions that are taken "off the shelf" from the 2-categories literature is that the standard 2-categorical proofs then specialize to

prove the following facts about adjunctions between ∞-categories.[9]

Proposition 3.4 ([RV2, 2.1.9]). *Adjunctions between ∞-categories compose:*

$$
A \underset{u}{\overset{f}{\rightleftarrows}} \perp B \underset{v}{\overset{g}{\rightleftarrows}} \perp C \qquad \rightsquigarrow \qquad A \underset{v\circ u}{\overset{f\circ g}{\rightleftarrows}} \perp C
$$

Proposition 3.5 ([RV2, 2.1.10]). *Given an adjunction*

$$A \underset{u}{\overset{f}{\rightleftarrows}} \perp B$$ *between ∞-categories and a functor* $f' : B \to A$, f' *is left adjoint to* u *if and only if there exists a natural isomorphism* $f \cong f'$.

Proposition 3.6 ([RV2, 2.1.12]). *Any equivalence can be promoted to an adjoint equivalence at the cost of replacing either of the natural isomorphisms.*

Relative adjunctions. The unit and counit in an adjunction satisfy a universal property in the 2-category ∞-CAT that we now explore in the case of the counit $\epsilon : fu \Rightarrow \mathrm{id}_A$. Given any ∞-functors $a : X \to A$ and $b : X \to B$, pasting with ϵ defines a bijection between ∞-natural transformations $\alpha : fb \Rightarrow a$ and $\beta : b \Rightarrow ua$. Any $\alpha : fa \Rightarrow b$ factors through a unique $\beta : b \Rightarrow ua$ as displayed below:

$$
\begin{array}{ccc}
X \xrightarrow{b} B & & X \xrightarrow{b} B \\
a \downarrow \;\; \forall\Downarrow\alpha \;\; \downarrow f & = & a \downarrow \;\; \exists\Downarrow\beta \nearrow u \;\; \downarrow f \\
A = A & & A = A \quad \Downarrow\epsilon
\end{array}
\tag{3.7}
$$

which is to say that the pair (u, ϵ) defines an *absolute right lifting* of id_A through f. Indeed, f is left adjoint to u with counit ϵ if and only if (u, ϵ) defines an absolute right lifting of id_A through f [RV2, 2.3.7].

Any component ϵa of the counit ϵ of an adjunction satisfies a universal property analogous to (3.7):

$$
\begin{array}{ccc}
Y \xrightarrow{b} B & & Y \xrightarrow{b} B \\
x \downarrow \;\; \forall\Downarrow\alpha \;\; \downarrow f & = & x \downarrow \;\; \exists\Downarrow\beta \nearrow u \;\; \downarrow f \\
X \xrightarrow{a} A & & X \xrightarrow{a} A \quad \Downarrow\epsilon \;\; A
\end{array}
$$

which asserts that $(ua, \epsilon a)$ defines an absolute right lifting of a through f. Motivated by examples such as these, absolute right liftings (r, ρ) of a generic ∞-functor $g : C \to A$ through $f : B \to A$ can be thought of as exhibiting r as *right adjoint to f relative to g with partial counit ρ.*

---

[8]*The technical part of this story involves proofs that the "synthetic" or "formal" notions introduced here agree with the previously-defined "analytic" notions in the quasi-categories model [RV2, §F]. But we encourage those not already acquainted with the analytic theory of some model of ∞-categories to learn these definitions first.*

[9]*What is less obvious is that these definitions are "correct" for ∞-category theory. In the case of equivalences, this is easily seen in any of the models. In the case of adjunctions, this is quite subtle, and involves the fact that the low-dimensional data enumerated in Definition 3.3 suffices to determine a full "homotopy coherent adjunction," uniquely up to a contractible space of choices [RV1].*

Definition 3.8. Given ∞-functors $g : C \to A$ and $f : B \to A$ with common codomain, a functor $r : C \to B$ and ∞-natural transformation $\rho : fr \Rightarrow g$ define an **absolute right lifting** of g through f or a **right adjoint of f relative to** g if pasting with ρ induces a bijection between ∞-natural transformations as displayed below:

$$
\begin{array}{ccc}
X \xrightarrow{\ b\ } B & & X \xrightarrow{\ b\ } B \\
{\scriptstyle c}\downarrow \quad \forall\Downarrow\alpha \quad \downarrow{\scriptstyle f} & = & {\scriptstyle c}\downarrow \quad {}^{\exists\Downarrow\beta}\nearrow^{r} \downarrow{\scriptstyle f} \\
C \xrightarrow[g]{} A & & C \xrightarrow[g]{} A
\end{array}
$$

The following lemmas about relative right adjoints admit straightforward 2-categorical proofs:

Lemma 3.9 ([RV2, 2.3.6]). *If (r,ρ) is right adjoint to $f : B \to A$ relative to $g : C \to A$ then for any $c : X \to C$, $(rc, \rho c)$ is right adjoint to f relative to gc.*

$$
X \xrightarrow{\ c\ } C \xrightarrow[g]{\ \ } A \qquad
\begin{array}{c}
B \\
{}^{r}\nearrow \ \downarrow{\scriptstyle f} \\
\Downarrow\rho
\end{array}
$$

Lemma 3.10 ([RV2, 2.4.1]). *Suppose that (r,ρ) is right adjoint to $f : B \to A$ relative to $g : C \to A$. Then in any diagram of ∞-functors and ∞-natural transformations*

$$
\begin{array}{c}
D \\
{}^{s}\nearrow \ \Downarrow\sigma \ \downarrow{\scriptstyle k} \\
B \\
{}^{r}\nearrow \ \Downarrow\rho \ \downarrow{\scriptstyle f} \\
C \xrightarrow[g]{\ \ } A
\end{array}
$$

(s,σ) is right adjoint to k relative to r if and only if $(s, \rho \cdot f\sigma)$ is right adjoint to fk relative to g.

Limits and colimits. Relative adjunctions can be used to define limits and colimits of diagrams valued inside an ∞-category A and indexed by another ∞-category J. As a cartesian closed 2-category, ∞-CAT contains a terminal ∞-category 1, which admits a unique ∞-functor $! : J \to 1$ from any other ∞-category J. Maps $a : 1 \to A$ from the terminal ∞-category into another ∞-category define *elements* in the ∞-category A.[10] Using the cartesian closed structure, any diagram $d : J \to A$ defines an element $d : 1 \to A^J$ in the ∞-*category of J-shaped diagrams in A*. Exponentiation with the unique functor $! : J \to 1$ defines the *constant diagrams functor* $\Delta : A \to A^J$, which carries an element of A to the constant J-shaped diagram at that element.

Definition 3.11. An ∞-category A **admits J-shaped limits** if the constant diagrams functor $\Delta : A \to A^J$ admits a

right adjoint and **admits J-shaped colimits** if the constant diagrams functor admits a left adjoint:

$$
A \underset{\underset{\text{lim}}{\longleftarrow}}{\overset{\overset{\text{colim}}{\longleftarrow}}{\underset{\perp}{\overset{\perp}{\underset{\Delta}{\longrightarrow}}}}} A^J
$$

The counit λ of the adjunction $\Delta \dashv \lim$ and the unit γ of the adjunction $\mathrm{colim} \dashv \Delta$ encode the limit and colimit cones, as is most easily seen when considering their components at a diagram $d : 1 \to A^J$. By Lemma 3.9, the right and left adjoints restrict to define relative adjunctions:

$$
\begin{array}{cc}
\begin{array}{c}
A \\
{}^{\lim}\nearrow \ \Downarrow\lambda \ \downarrow{\scriptstyle\Delta} \\
1 \xrightarrow[d]{} A^J = A^J
\end{array}
&
\begin{array}{c}
A \\
{}^{\mathrm{colim}}\nearrow \ \Uparrow\gamma \ \downarrow{\scriptstyle\Delta} \\
1 \xrightarrow[d]{} A^J = A^J
\end{array}
\end{array}
$$

These observations allow us to generalize Definition 3.11 to express the universal properties of the limit or colimit of a single diagram in settings where the limit and colimit functors do not exist.

Definition 3.12. A diagram $d : J \to A$ between ∞-categories **admits a limit** just when Δ admits a relative right adjoint at d and **admits a colimit** just when Δ admits a relative left adjoint at d, as encoded by absolute lifting diagrams:

$$
\begin{array}{cc}
\begin{array}{c}
A \\
{}^{\lim d}\nearrow \ \Downarrow\lambda_d \ \downarrow{\scriptstyle\Delta} \\
1 \xrightarrow[d]{} A^J
\end{array}
&
\begin{array}{c}
A \\
{}^{\mathrm{colim}d}\nearrow \ \Uparrow\gamma_d \ \downarrow{\scriptstyle\Delta} \\
1 \xrightarrow[d]{} A^J
\end{array}
\end{array}
$$

There is an easy formal proof of the ∞-categorical version of a classical theorem:

Theorem 3.13. *Right adjoints preserve limits and left adjoints preserve colimits.*

Proof. Consider an adjunction as in Definition 3.3 and a limit of a diagram as in Definition 3.12. To show that the limit is preserved by the right adjoint, we must show that

$$
\begin{array}{c}
A \xrightarrow{\ u\ } B \\
{}^{\lim d}\nearrow \ \Downarrow\lambda_d \ \downarrow{\scriptstyle\Delta} \quad \downarrow{\scriptstyle\Delta} \\
1 \xrightarrow[d]{} A^J \xrightarrow[u^J]{} B^J
\end{array}
$$

defines a relative right adjoint. By Lemma 3.10, it suffices to demonstrate this after pasting with the relative right adjoint encoded by the component at d of the counit ϵ^J of the adjunction $f^J \dashv u^J$, as below-left:

$$
\begin{array}{ccc}
\begin{array}{c}
A \xrightarrow{u} B \\
{}^{\lim d}\nearrow \Downarrow\lambda_d \downarrow\Delta \downarrow\Delta \\
1 \xrightarrow[d]{} A^J \xrightarrow[u^J]{} B^J \\
\Downarrow\epsilon^J \quad \downarrow f^J \\
A^J
\end{array}
& = &
\begin{array}{c}
B \\
{}^{u}\nearrow \Downarrow\epsilon \downarrow f \\
A = A \\
{}^{\lim d}\nearrow \Downarrow\lambda_d \downarrow\Delta \downarrow\Delta \\
1 \xrightarrow[d]{} A^J = A^J
\end{array}
& = &
\begin{array}{c}
B \\
{}^{u\lim d}\nearrow \Downarrow\epsilon_{\lim d} \downarrow f \\
A \\
{}^{\lim d}\nearrow \Downarrow\lambda_d \downarrow\Delta \\
1 \xrightarrow[d]{} A^J
\end{array}
\end{array}
$$

From the 2-functoriality of the exponential in a cartesian closed 2-category, $f^J\Delta = \Delta f$ and $\epsilon^J\Delta = \Delta\epsilon$. Hence,

---

[10]*We find this terminology less confusing than referring to the "objects" in an ∞-category A, which is itself an object of ∞-CAT.*

the pasting diagram displayed above-left equals the one displayed above-center, which equals the diagram above-right. This latter diagram is a pasted composite of relative right adjoints, and is then a relative right adjoint in its own right by Lemma 3.10. □

Universal properties. Definitions 3.3 and 3.12 do not express the full universal properties of adjunctions between and co/limits in ∞-categories. These require additional structure borne by the 2-category ∞-CAT, namely the existence of *comma ∞-categories* for any pair of ∞-functors $f : B \to A$ and $g : C \to A$. The comma ∞-category is an ∞-category $Hom_A(f, g)$ equipped with canonical functors and an ∞-natural transformation as below-left, formed by the pullback of ∞-categories below-right:

$$\begin{array}{ccc} Hom_A(f,g) \xrightarrow{\mathrm{dom}} B & & Hom_A(f,g) \xrightarrow{\phi} A^2 \\ \mathrm{cod}\downarrow \quad \Downarrow\phi \quad \downarrow f & & (\mathrm{cod,dom})\downarrow \quad \lrcorner \quad \downarrow (\mathrm{cod,dom}) \\ C \xrightarrow{g} A & & C \times B \xrightarrow{g \times f} A \times A \end{array} \quad (3.14)$$

While the universal property of $Hom_A(f, g)$ as an object of ∞-CAT is weaker than the standard notion of comma object in 2-category theory, nevertheless:

Theorem 3.15 ([RV2, 3.5.8]). *An ∞-functor $f : B \to A$ admits a right adjoint $r : C \to B$ relative to $g : C \to A$ if and only if there exists an equivalence over $C \times B$:*

$$Hom_A(f, g) \simeq_{C \times B} Hom_B(B, r).$$

When Theorem 3.15 is applied to $f : B \to A$ and id_A, it specializes to:

Corollary 3.16 ([RV2, 4.1.1]). *An ∞-functor $f : B \to A$ admits a right adjoint $u : A \to B$ if and only if $Hom_A(f, A) \simeq_{A \times B} Hom_B(B, u)$.*

When Theorem 3.15 is applied to $\Delta : A \to A^J$ and $d : 1 \to A^J$ it specializes to:

Corollary 3.17 ([RV2, 4.3.1]). *A diagram $d : J \to A$ has a limit $\ell : 1 \to A$ if and only if $Hom_{A^J}(\Delta, d) \simeq_A Hom_A(A, \ell)$.*

Here the comma ∞-category $Hom_{A^J}(\Delta, d)$ defines the *∞-category of cones* over d. The ∞-category $Hom_A(A, \ell)$ admits a *terminal element* $id_\ell : 1 \to Hom_A(A, \ell)$, which defines a right adjoint to the unique functor $! : Hom_A(A, \ell) \to 1$. Via the equivalence $Hom_{A^J}(\Delta, d) \simeq_A Hom_A(A, \ell)$ we see that the limit cone is terminal in the ∞-category of cones. Indeed, d admits a limit if and only if the ∞-category of cones admits a terminal element [RV2, 4.3.2].

Could we teach this to undergraduates? If classical category theory were a standard course in the undergraduate mathematics curriculum, then perhaps students with a particular interest in the subject might go on to learn some 2-category theory, like students with a particular interest in algebra might go on to take more specialized courses in classical algebraic geometry, algebraic number theory, or

representation theory when offered. From this point, the proofs of the results stated here are not too difficult.

But these results encompass only a small portion of ∞-category theory, and for most of the rest there is a greater jump in complexity when extending from 1-categories to ∞-categories. This is best illustrated by considering the Yoneda lemma, which in an important special case is stated as follows:

Theorem 3.18 ([RV2, 5.7.1]). *Given an ∞-category A and elements $a, b : 1 \to A$, the ∞-groupoid $Hom_A(a, b)$ is equivalent to the ∞-groupoid of functors $Hom_A(A, a) \to Hom_A(A, b)$ over A.*

One direction of this equivalence is easy to describe: the map from right to left is defined by evaluation at the identity element $id_a : 1 \to Hom_A(a, a)$. The inverse equivalence defines the Yoneda embedding, which is notoriously difficult to construct in ∞-category theory as it involves equipping A with a homotopy coherent composition function [L2, §5.1.3].

To achieve further simplifications of ∞-category theory, one idea is to ask our foundation system to do more of the work.

4. A Synthetic Theory of ∞-Categories

To explain the desiderata for an alternative foundation, consider the default notions of "sameness" for ∞-categorical data. Two elements in an ∞-category are the same if and only if they are connected by a path in the underlying ∞-groupoid, while parallel morphisms are the same if and only if they are connected by an invertible 2-cell, defining a path in the appropriate mapping space. For an ∞-categorical construction to be "well-defined" it must:

(i) respect the notions of sameness encoded by paths in suitable spaces, and

(ii) respect equivalences between these spaces themselves, since they are only well-defined as homotopy types.

Axiomatizing sameness. In traditional foundations, there is a similar axiom that mathematical constructions or results must respect sameness as encoded by equality. The axioms that define the binary relation "=" in first-order logic are:

• reflexivity: for all x, $x = x$ is true.
• indiscernibility of identicals: for all x, y and for all predicates P, if $x = y$ then $P(x)$ holds if and only if $P(y)$ holds.

Analogous rules govern Martin-Löf's *identity types* in an alternative foundational framework for constructive mathematics known as *dependent type theory* [M-L]. Here the primitives include *types* like $\mathbb{N}$, $GL_n(\mathbb{R})$, and Group and *terms* like $17 : \mathbb{N}$, $I^n : GL_n(\mathbb{R})$, $S_n : $ Group, which may depend on an arbitrary *context* of variables drawn from previously defined types (e.g., the $n : \mathbb{N}$ appearing in three

examples above).[11] In traditional foundations, mathematical structures are defined in the language of set theory, while proofs obey the rules of first-order logic. In dependent type theory, mathematical constructions and mathematical proofs are unified: both are given by terms in appropriate types. The ambient type then describes the sort of object being constructed or the mathematical statement being proven, while the context describes the inputs to the construction or the hypotheses for the theorem.

The rules for Martin-Löf's identity types include:[12]

- identity-formation: in the context of two variables $x, y : A$ there is a type $x =_A y$.

Types encode mathematically meaningful statements or assertions, so this rule says that it is reasonable to inquire whether x and y might be identified, once x and y are terms in the same type. Certain identity types, such as $3 =_\mathbb{N} 4$, will be empty, since the terms 3 and 4 cannot be identified, but the type nevertheless exists.

- identity-introduction: for any $x : A$, there is a term $\text{refl}_x : x =_A x$.

Terms in types are witnesses to the truth of the statement encoded by the type, so this rule corresponds to the reflexivity axiom.

- identity-elimination: given any family of types $Q(x, y, p)$ depending on $x, y : A$ and $p : x =_A y$, to provide a family of terms $q_{x,y,p} : Q(x, y, p)$ for all x, y, p it suffices to provide a family of terms $d_x : Q(x, x, \text{refl}_x)$ for all x.

This rule implies Leibniz's indiscernibility of identicals and much more besides:

Theorem 4.1 ([BG, Lu]). *The iterated identity types provide any type with the structure of an ∞-groupoid, in which the terms are the points, the identifications are the paths, and the higher identifications encode homotopies.*

In low dimensions, the ∞-groupoid structure on the iterated identity types of a type A provides:

- functions $(-)^{-1} : (x =_A y) \to (y =_A x)$ that invert identifications,
- functions $- * - : (x =_A y) \times (y =_A z) \to (x =_A z)$ that compose identifications,
- higher identifications $\text{assoc} : (p*q)*r =_{x=_A w} p*(q*r)$ between composable triples of identifications,

and much more. All of these terms are constructed using the identity-elimination rule from the reflexivity terms provided by the identity-introduction rule. Since an

identification $p : x =_A y$ defines a path from x to y in the ∞-groupoid A, identifications in dependent type theory are often referred to colloquially as "paths."

The homotopical interpretation of dependent type theory. The homotopical interpretation of dependent type theory was discovered in the early 21st century by Awodey–Warren and Voevodsky, building on earlier work of Hofmann and Streicher. This connection inspired Voevodsky to make the following definition:

Definition 4.2. A type A is **contractible** just when there is a term of type

$$\Sigma_{x:A}\Pi_{y:A} x =_A y,$$

which may be read as "there exists $x : A$, so that for all $y : A$, x is identifiable with y" with all quantifiers interpreted as requiring continuous specifications of data.[13]

Just as contractible spaces might contain uncountably infinitely many points, contractible types might contain more than one term. But since any two terms in a contractible type may be identified, and since identity-elimination implies that all of dependent type theory respects identifications, the theory behaves as if a contractible type had a unique term.

Univalent foundations. Martin-Löf's dependent type theory is a formal system in which all constructions are continuous in paths, our first desideratum. The second desideratum, of well-definedness under equivalences between types, follows once Voevodsky's *univalence axiom* is added to dependent type theory, resulting in a formal system called *homotopy type theory* or *univalent foundations*.

Voevodsky's univalence connects two notions of sameness for types. In traditional dependent type theory, types can be encoded as terms in a universe of types.[14] Thus, one notion of sameness between types A, B belonging to a universe $\mathcal{U}$ is given by paths $p : A =_\mathcal{U} B$ in the universe. Another notion of sameness is suggested by types as ∞-groupoids: for any A and B there is a type $A \simeq B$ of (homotopy) equivalences from A to B.[15] By identity-elimination, to define a natural map id-to-equiv : $A =_\mathcal{U} B \to A \simeq B$ it suffices to define the image of $\text{refl}_A : A =_\mathcal{U} A$, which we take to be the identity equivalence. The *univalence axiom* asserts that this map is an equivalence for all types A and B.

By univalence, an equivalence $e : A \simeq B$ gives rise to a path $\text{ua}(e) : A =_\mathcal{U} B$, which can then be used to transport

---

[11]*In a mathematical statement of the form "Let … be … then … " the stuff following the "let" likely declares the names of the variables in the context described after the "be," while the stuff after the "then" most likely describes a type or term in that context.*

[12]*One additional rule, not listed here, concerns the computational behavior of this system.*

[13]*The definition of contractible types makes use of the* dependent pair *and* dependent function *types in dependent type theory. A term of type $\Sigma_{x:A}\Pi_{y:A} x =_A y$ provides a term $c : A$, the "center of contraction," together with a family of paths $p_z : c =_A z$ for all $z : A$, the "contracting homotopy."*

[14]*Russell's paradox can be resolved by a cumulative hierarchy of universes.*

[15]*It is an interesting challenge to define a type whose data witnesses that a given map $f : A \to B$ is an equivalence in such a way that this type is contractible whenever it is inhabited.*

terms in any type involving A to terms in the corresponding type involving B. Thus, univalent mathematics is automatically invariant under equivalence between types.

∞-categories in univalent foundations. There is an experimental exploration of ∞-category theory [RS, BW, Ba] in an extension of homotopy type theory in which every type A has a family of arrows $\mathrm{Hom}_A(x, y)$ in addition to the family of paths $x =_A y$. These types are obtained from a new type-forming operation that produces *extension types*, whose terms are type-valued diagrams that strictly extend a given diagram along an inclusion of "shapes" (polytopes embedded in directed cubes constructed in the theory of a strict interval) [RS, §2-4]. Extension types include types analogous to the pullbacks of (1.4) and (3.14).

The formal system has semantics in complete Segal spaces [RS, W] so it provides a rigorous way to prove theorems about ∞-categories as understood in traditional foundations. At the same time, the experience of working with ∞-categories in this "univalent" foundational setting is much more akin to the experience of working with 1-categories in traditional foundations.

It takes work to set up this formal system and considerably more to describe its interpretation in complete Segal spaces; indeed, this is where the hard work of solving homotopy coherence problems goes. But once this is done, it is possible to define the notion of an ∞-category, something that would no doubt be reassuring to undergraduates hoping to learn about them.[16] The two axioms express the "Segal" and "completeness" conditions of Rezk's model [R1], respectively.

Definition 4.3 ([RS, 5.3,10.6]). An ∞-**category** is a type A in which:

- every composable pair of arrows has a unique composite, and
- for any pair of terms $x, y : A$ the natural map $x =_A y \to x \cong_A y$ from paths in A to isomorphisms in A is an equivalence.

The "uniqueness" in the first axiom is in the sense of Definition 4.2: what it asserts is that a suitable space of composites analogous to Definition 1.3 is contractible—the default meaning of uniqueness in this formal system. Like in the space defined by (1.4), a term in the space of composites provides a higher-dimensional witness to the composition relation. A composition function

$$\circ : \mathrm{Hom}_A(y, z) \times \mathrm{Hom}_A(x, y) \to \mathrm{Hom}_A(x, z),$$

is obtained by throwing away these witnesses. Like any construction in homotopy type theory, this function respects identifications and is therefore well-defined. It

follows from this axiom that composition is associative and unital, using canonical identity arrows $\mathrm{id}_x : \mathrm{Hom}_A(x, x)$.

The second axiom involves a type of isomorphisms in A, which may be defined in a similar manner to the type of equivalences between types [RS, §10].

The Yoneda lemma. The advantages of this synthetic framework for ∞-category theory are on display when comparing the proofs of the Yoneda lemma in [RV2, 5.7.1] and in [RS, 9.1]. We sketch the latter here in the same special case considered in Theorem 3.18.

Theorem 4.4 ([RS, 9.1]). *Given an ∞-category A and elements $a, b : 1 \to A$, the type $\mathrm{Hom}_A(a, b)$ is equivalent to the type $\Pi_{x:A}\mathrm{Hom}_A(x, a) \to \mathrm{Hom}_A(x, b)$ of fiberwise functions.*

A key difference between Theorem 3.18 and Theorem 4.4 is that in the present framework it is straightforward to define the inverse equivalence to the evaluation at identity map. The inverse equivalence takes an arrow $f : \mathrm{Hom}_A(a, b)$ to the natural map $x \mapsto g \mapsto f \circ g : \Pi_{x:A}\mathrm{Hom}_A(x, a) \to \mathrm{Hom}_A(x, b)$. The usual proof then demonstrates that these maps are inverse equivalences. In one direction, this is given by an identification $f \circ \mathrm{id}_a =_{\mathrm{Hom}_A(a,b)} f$, which expresses the fact that composition in an ∞-category is unital. In the other direction, we must show that a fiberwise map ϕ agrees with the map whose component at x is given by $g \mapsto \phi_a(\mathrm{id}_a) \circ g : \mathrm{Hom}_A(x, a) \to \mathrm{Hom}_A(x, b)$. By a consequence of the univalence axiom called "function extensionality," it suffices to show that $\phi_a(\mathrm{id}_a) \circ g =_{\mathrm{Hom}_A(x,b)} \phi_x(g)$. This follows from the fact that fiberwise maps are automatically "natural," providing an identification $\phi_a(\mathrm{id}_a) \circ g =_{\mathrm{Hom}_A(x,b)} \phi_x(\mathrm{id}_a \circ g)$. Since the function ϕ_x respects the identification $\mathrm{id}_a \circ g =_{\mathrm{Hom}_A(x,a)} g$, we obtain the required identification.

In fact, [RS, 9.5] proves a generalization of the Yoneda lemma that was first discovered with this formal system and later proven in traditional foundations [RV2, 5.7.2]

Conclusion. Significant technical problems remain to make it feasible to teach ∞-category theory to undergraduates, but I have hope that in the future the subject will not seem as forbidding as it does today.

---

[16]*Recall that in the formal ∞-category theory developed in the 2-category ∞-CAT, a concrete definition of ∞-categories is not used.*

References

[Ba] C. Bardomiano Martínez, *Limits and colimits of synthetic ∞-categories*, 2022, arXiv:2202.12386.

[BG] Benno van den Berg and Richard Garner, *Types are weak ω-groupoids*, Proc. Lond. Math. Soc. (3) **102** (2011), no. 2, 370–394, DOI 10.1112/plms/pdq026. MR2769118

[Be] Julia E. Bergner, *The homotopy theory of (∞, 1)-categories*, London Mathematical Society Student Texts, vol. 90, Cambridge University Press, Cambridge, 2018. MR3791455

[BV] J. M. Boardman and R. M. Vogt, *Homotopy invariant algebraic structures on topological spaces*, Lecture Notes in Mathematics, Vol. 347, Springer-Verlag, Berlin-New York, 1973. MR0420609

[BW] U. Buchholtz and J. Weinberger, *Synthetic fibered (∞, 1)-category theory*, 2021, arXiv:2105.01724.

[G] A. Grothendieck, *Pursuing stacks*, 1983, arXiv:2111.01000

[J] A. Joyal, *The theory of quasi-categories and its applications*, Advanced Course on Simplicial Methods in Higher Categories, vol. 2, Centre de Recerca Matemática, 2008.

[KV] Ya. Varshavskiĭ and D. Kazhdan, *The Yoneda lemma for complete Segal spaces* (Russian, with Russian summary), Funktsional. Anal. i Prilozhen. **48** (2014), no. 2, 3–38, DOI 10.1007/s10688-014-0050-3; English transl., Funct. Anal. Appl. **48** (2014), no. 2, 81–106. MR3288174

[L1] J. Lurie, *On ∞-topoi*, 2003, arXiv:0306109.

[L2] Jacob Lurie, *Higher topos theory*, Annals of Mathematics Studies, vol. 170, Princeton University Press, Princeton, NJ, 2009, DOI 10.1515/9781400830558. MR2522659

[L3] J. Lurie, *Higher algebra*, 2017, www.math.ias.edu/~lurie/papers/HA.pdf.

[Lu] Peter LeFanu Lumsdaine, *Weak ω-categories from intensional type theory*, Typed lambda calculi and applications, Lecture Notes in Comput. Sci., vol. 5608, Springer, Berlin, 2009, pp. 172–187, DOI 10.1007/978-3-642-02273-9_14. MR2550047

[M-L] Per Martin-Löf, *An intuitionistic theory of types: predicative part*, Logic Colloquium '73 (Bristol, 1973), Studies in Logic and the Foundations of Mathematics, Vol. 80, North-Holland, Amsterdam, 1975, pp. 73–118. MR0387009

[Q] Daniel G. Quillen, *Homotopical algebra*, Lecture Notes in Mathematics, No. 43, Springer-Verlag, Berlin-New York, 1967. MR0223432

[R1] Charles Rezk, *A model for the homotopy theory of homotopy theory*, Trans. Amer. Math. Soc. **353** (2001), no. 3, 973–1007, DOI 10.1090/S0002-9947-00-02653-2. MR1804411

[R2] Charles Rezk, *A Cartesian presentation of weak n-categories*, Geom. Topol. **14** (2010), no. 1, 521–571, DOI 10.2140/gt.2010.14.521. MR2578310

[RV1] Emily Riehl and Dominic Verity, *Homotopy coherent adjunctions and the formal theory of monads*, Adv. Math. **286** (2016), 802–888, DOI 10.1016/j.aim.2015.09.011. MR3415698

[RV2] Emily Riehl and Dominic Verity, *Elements of ∞-category theory*, Cambridge Studies in Advanced Mathematics, vol. 194, Cambridge University Press, Cambridge, 2022, DOI 10.1017/9781108936880. MR4354541

[RS] Emily Riehl and Michael Shulman, *A type theory for synthetic ∞-categories*, High. Struct. **1** (2017), no. 1, 147–224, DOI 10.1007/s42001-017-0005-6. MR3912054

[W] J. Weinberger, *Strict stability of extension types*, 2022, arXiv:2203.07194.
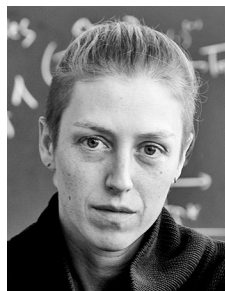
Emily Riehl

Credits

Opening image is courtesy of nevarpp via Getty.

Photo of Emily Riehl is courtesy of Marshall Clarke.