Categorical Structures in Physics

Contents

Introduction

This thesis concerns the use of categorical structures in physics. It is largely self-contained, and as such introduces most of the necessary prerequisites in chapters 1 and 2, which concern mathematics and physics, respectively. There are a multitude of successful and interesting attempts by physicists to use category theory and its descendants to organize physical data, but we will only focus on a few. Namely, we will study the following:

- *Synthetic differential geometry*, the axiomatic ("synthetic") development of differential geometry in certain categories with enough structure to admit an internal logic capable of supporting notions of smoothness and infinitesimality, i.e. smooth topoi. It has been applied to physics in the form of both smaller attempts to reformulate certain physical models, such as general relativity [Guts and Grinkevich, 1996, Lawvere, 2002], and larger attempts to understand the ambient geometry in which physics occurs [Schreiber, 2013, Lawvere, 1997]. (F. W. Lawvere in particular has worked both on the general theory and its applications to continuum mechanics). We will study both categories of applications.
- *Topos quantum theory*, a separate attempt to apply topoi to physics, makes use of the fact that quantum physics derives largely from a quantum logic which, while radically non-classical, is a system of logic nonetheless. While classical (Newtonian) mechanics can be formalized in the topos of sets and its logic shown to be the internal logic of sets, it is hoped that we may find a topos whose internal logic resembles quantum logic, and thereby reformulate quantum mechanics in a consistent manner. This hope is realized by the work of Isham, Döring, Flori, and others [Döring and Isham, 2008a, Flori, 2013a, Flori, 2013b], who in particular use the category $\mathrm{Set}^{\mathcal{V}(\mathcal{H})^{\mathrm{op}}}$ of presheaves of abelian von Neumann algebras over a given Hilbert space $\mathcal{H}$.
- A *topological quantum field theory* (TQFT) is a quantum field theory defined on a manifold M whose Lagrangian density is independent of the metric, and hence depends only on the

> topology of M. TQFTs have vanishing Hamiltonians, and hence no actual dynamics, but coupling their Lagrangians to those of non-topological field theories provides topological constraints on the latter, making them interesting from a physical point of view. It was realized by Atiyah [Atiyah, 1988] that TQFTs can in general be realized as functors from "geometric" categories to "algebraic" categories, associating elements of e.g. modules to spaces in a functorial manner. In the usual formulation, as in e.g. [Lurie, 2009b], a TQFT is a symmetric monoidal functor from the category of (smooth, compact, oriented, closed) $(n-1)$-manifolds and their cobordisms to the category of vector spaces over a given field k, which in turn assigns an *element* of k, i.e. a number, to every n-manifold. This formulation of TQFTs admits a natural extension to higher categories: we assign numbers to n-manifolds, vector spaces to $(n-1)$-manifolds, categories of vector spaces to $(n-2)$-manifolds, and so on.

While it is too intricate to cover in the current work, Schreiber's project to formalize physics in modal homotopy type theory [Schreiber, 2016], which seeks to map the internal logic of an ∞-topos onto Hegel's logic of concepts and sublations as well as the logic of physics, is worth mentioning as well.

Typesetting and Acknowledgements

All figures were created by the author in Inkscape, including the diagrammatic calculus of categorical quantum mechanics. Commutative diagrams were created using tikzcd; Yichuan Shen's tikzcd editor (available at https://tikzcd.yichuanshen.de/) in particular cut down on several days' worth of tedious typesetting. Feynman diagrams were drawn with the software JaxoDraw [Binosi and Theussl, 2004]. Aside from the creators of these tools, I would like to thank Prof. Mirroslav Yotov for answering silly questions and providing a fruitful bird's-eye view of many subjects, especially algebraic geometry.

Chapter 1

Mathematics

This chapter, which introduces category theory and covers the study of spaces from many categorically oriented points of view, is a blend of many sources. Our sources for category theory include [Mac Lane, 2013, Riehl, 2017, Aluffi, 2009]. The section on homotopy theory borrows from [May, 1999, Hatcher, 2005, Munkres, 2018], in roughly that order, whereas the discussion on vector bundles is inspired by [Hatcher, 2003, Weibel, 2013]. The section on algebraic geometry is indebted to [Hartshorne, 2013, Vakil, 2017, Authors, 2018].

1.1 Category Theory

1.1.1 Categories

A **category** C is a class $\text{Ob}(\mathsf{C})$ of **objects** and, for every two objects $X, Y \in \text{Ob}(\mathsf{C})$, a class of **morphisms** denoted variously as $\mathsf{C}(X, Y)$ or $\text{Hom}_{\mathsf{C}}(X, Y)$. (We will have occasion to use both notations – while $\mathsf{C}(X, Y)$ is more concise and easier on the eyes, the Hom notation is sometimes more enlightening). For every triple of objects X, Y, Z, there is a composition function $\mathsf{C}(Y, Z) \times \mathsf{C}(X, Y) \to \mathsf{C}(X, Z)$ sending g, f to the **composition morphism** $g \circ f$, often abbreviated to gf, whose existence we require. We also require that composition is associative, in the sense that $(h \circ g) \circ f = h \circ (g \circ f)$, as well as the existence of **identity morphisms** id_X for each $X \in \text{Ob}(\mathsf{C})$ such that $g \circ \text{id}_X = g$ and $\text{id}_X \circ f = f$.

If $\text{Ob}(\mathsf{C})$ is a set rather than a proper class, C is said to be *small*. If $\mathsf{C}(X, Y)$ is a set for all $X, Y \in \mathsf{C}$, then C is *locally small*, and we often refer to $\mathsf{C}(X, Y)$ as a *hom-set*[1].

[1]"Hom" is an abbreviation of homomorphism, a relic from category theory's origins in algebraic topology.

Many common "types" of mathematical objects can be assembled into categories:

- There is a category Set whose objects are sets and whose morphisms are functions (a function $f : X \to Y$ being a selection of an element in Y for every element of X). Composition of functions is defined in the usual sense, and there is an obvious identity morphism $\mathrm{id}_X : X \to X, x \mapsto x$.
- The category Top consists of topological spaces and continuous functions.
- The category Ab consists of abelian groups and group homomorphisms.
- The category CRing consists of commutative rings and ring homomorphisms.
- The category $R\text{-}\mathsf{Mod}$ consists of modules over a commutative ring R and their homomorphisms[2].
- The category Man^p consists of C^p manifolds and maps. For instance, $\mathsf{Diff} := \mathsf{Man}^\infty$ consists of smooth manifolds and maps.

Set is a locally small category, as are all categories whose objects and morphisms can be thought of as sets and set functions, including all of the above examples.

Monomorphisms and Epimorphisms In Set, we can classify morphisms into injective, surjective, and bijective maps. This generalizes in the following manner: A morphism $f : X \to Y$ in a category C is an **epimorphism** if, for all $g, h : Y \to Z$, we have $gf = hf$ if and only if $g = h$. f is a **monomorphism** if, for all $g, h : W \to X$, we have $fg = fh$ if and only $g = h$. f is an **isomorphism** if there is an inverse morphism $g : Y \to X$ such that $fg = \mathrm{id}_Y$ and $gf = \mathrm{id}_X$. Two objects in C are **isomorphic** if there is an isomorphism between them. The isomorphisms in Grp, Set, Top, and Diff, for instance, are the group isomorphisms, bijections, homeomorphisms, and diffeomorphisms, respectively; for nearly all intents and purposes, isomorphic objects are to be regarded as equivalent. Note: we often shorten epimorphism to *epi*, or in its adjectival form, an *epic* morphism, whereas monomorphism is shortened to *mono*, or a *monic* morphism.

In Set, (i) epimorphisms are equivalent to surjections, (ii) monomorphisms are equivalent to injections, and (iii) isomorphisms are equivalent to bijections. To prove this, take a map of sets $f : X \to Y$.

(i) Suppose that there is some $y \in Y$ not contained in $f(X)$. Let $Z = \{0, 1\}$, and let $g, h : Y \to Z$ send $Y - y$ to 0 and y to 0 or 1, respectively. $gf = hf$, but $g \neq h$. So if f is an epimorphism, it must be a surjection. Conversely, suppose that f is a surjection, and let $g, h : Y \to Z$ satisfy $gf = hf$.

[2] We often write $R(X, Y)$ and $\mathrm{Hom}_R(X, Y)$ instead of $R\text{-}\mathsf{Mod}(X, Y)$ and $\mathrm{Hom}_{R\text{-}\mathsf{Mod}}(X, Y)$.

For every $y \in Y$ there is an x_y such that $f(x_y) = y$, so $g(y) = gf(x_y) = hf(x_y) = h(y)$, and $g = h$. Obviously, if $g = h$ then $gf = hf$ as well, so surjections are epimorphisms.

(ii) Similar to (i).

(iii) Bijections obviously have inverses. Conversely, let $f : X \to Y$ admit an inverse $g : Y \to X$ such that $g(f(x)) = x$ and $f(g(y)) = y$. If f is not surjective, then there is some $y \in Y$ mapped to by no $f(x)$, so we cannot have $f(g(y)) = y$, and if f is not injective, then there are $x \neq x' \in X$ with $f(x) = f(x')$ and therefore $x = g(f(x)) = g(f(x')) = x'$, a contradiction. So isomorphisms are injective and surjective, and hence bijective. Importantly, this proof hinges on the fact that injective surjections are bijections; in an arbitrary category, it is *not* necessarily true that a morphism which is both monic and epic is an isomorphism. A category where this is true is known as a **balanced category**.

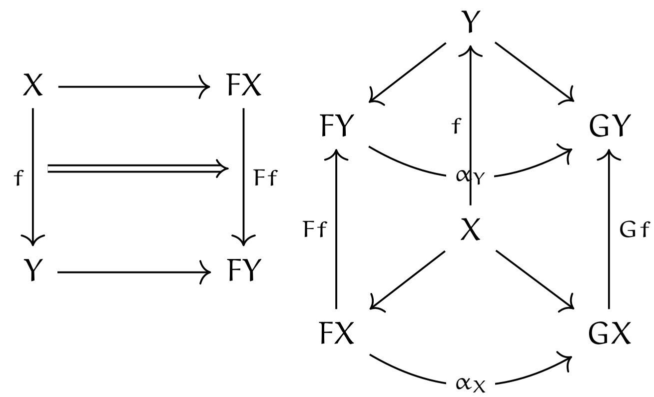
Most of our example categories are balanced, but CRing is not. To see this, take the inclusion $i : \mathbb{Z} \to \mathbb{Q}$. First, let $f, g : R \to \mathbb{Z}$ be such that $if = ig$. Since i is an inclusion, $f(r) = g(r)$ for all r, making i monic. Now let $h, k : \mathbb{Q} \to S$ be such that $hi = ki$. $h(p/q) = h(p)h(q^{-1}) = h(p)h(q)^{-1}$, so h and likewise k are completely determined by where they send the integers, and hence $hi = ki$ implies $h = k$. Despite being monic and epic, i fails to be an isomorphism.

Naturality In general, the vast majority of types of mathematical objects assemble into categories, the main concern being what the morphisms between objects of a certain type should be; generally, there is a natural notion of morphism between such objects (as in the above examples) which, when equipped to their category, allow that category to "encapsulate" the nature of that type of object. This natural notion is generally one that preserves precisely the structure associated to that type of object; given enough information about what is needed to define an object of that type, the structure we want morphisms to preserve generally becomes obvious.

For instance, we may define a natural notion of a morphism between categories: a morphism $F : C \to D$ should map objects $X \in C$ to objects $FX \in D$ and morphisms $f : X \to Y$ to morphisms $Ff : FX \to FY$ in a manner that preserves composition, identity, and associativity. Such a map has a special name: Given two categories C, D, a **functor** $F : C \to D$ consists of a map $\mathrm{Ob}(C) \to \mathrm{Ob}(D)$, as well as, for every $X, Y \in C$, a map $C(X, Y) \to D(FX, FY)$. We require $F(g \circ f) = (Fg) \circ (Ff)$ and $F\mathrm{id}_X = \mathrm{id}_{FX}$. (Associativity is trivial).

Given two functors $F, G : C \to D$, a **natural transformation** $\alpha : F \Rightarrow G$ is a family $\{\alpha_X : FX \to GX\}_{X \in C}$ of maps in D such that, for any $f : X \to Y$, we have $(Gf) \circ \alpha_X = \alpha_Y \circ (Ff)$. If each α_X is an isomorphism, α is known as a **natural isomorphism**.

We can define two new categories: the category Cat of small categories and functors, and, for any $\mathsf{C}, \mathsf{D} \in \mathsf{Cat}$, a category D^{C} whose objects are functors $\mathsf{C} \to \mathsf{D}$ and whose morphisms are natural transformations between functors. Both of these are subject to set-theoretic issues[3]. We will handwave these issues away, though especially curious/bored readers may see Appendix A for a discussion on the problems this can lead to, and the mechanisms for fixing them.



The data associated to a functor and natural transformation

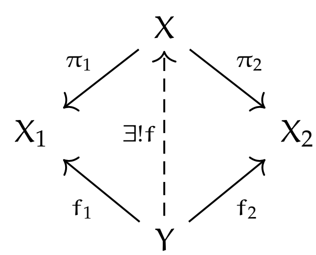
All of our example categories are locally small, and their objects are sets equipped with extra structure. Such locally small categories which can be modeled on sets are called **concrete**, and they admit functors $\mathsf{C} \to \mathsf{Set}$ which "forget" the structure on their objects, conveniently known as **forgetful functors**. For instance, the forgetful functor $\mathsf{Ab} \to \mathsf{Set}$ just maps abelian groups to their underlying sets, and group homomorphisms to their underlying set functions. In general, for a category to be concrete we require the existence of a forgetful functor which is injective on hom-sets, as otherwise two different maps in C will be sent to the same set map, so we cannot speak of their "underlying" set maps.

A functor F for which each map $\mathsf{C}(X, Y) \to \mathsf{C}(FX, FY)$ is injective is known as a **faithful functor**; in contrast, functors which are surjective on hom-sets are called **full**. Faithfully full functors are bijections on hom-sets. On objects, F is **essentially surjective** if every object $Y \in \mathsf{D}$ is isomorphic to some FX, $X \in \mathsf{C}$.

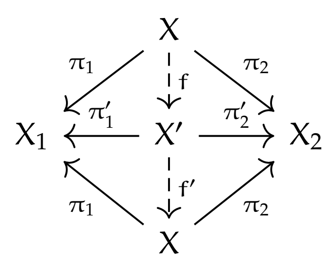
[3]It is for this reason that Cat consists of *small* categories; the set-theoretically problematic CAT is defined as the category of all categories.

1.1.2 Limits and Colimits

To see how categorical thinking can encapsulate the nature of certain types of mathematical objects, consider the product of topological spaces: given a pair of topological spaces X_1, X_2, we define their product to be a space X equipped with canonical projection maps $\pi_i : X \to X_i$, and give X the smallest topology that makes the π_i continuous. Every open set in this initial topology is required for continuity, making this the "most efficient" space with continuous morphisms into X_1 and X_2. This can be made rigorous by the following observation: any space Y equipped with a pair of functions $(f_1 : Y \to X_1, f_2 : Y \to X_2)$ admits a continuous map $f : Y \to X, y \mapsto (f_1(y), f_2(y))$ such that $\pi_1 f = f_1$ and $\pi_2 f = f_2$; in fact, this f is *uniquely* determined by f_1 and f_2. Pictorially, there is a unique arrow $f : Y \to X$ such that the triangles in the following diagram commute:



In particular, if we set $Y = X$, we get $f = \mathrm{id}_X$. We see that $X = X_1 \times X_2$ encodes pairs of morphisms $(f_1 : Y \to X_1, f_2 : Y \to X_2)$ in the most efficient possible way; in fact, if any other space X' with morphisms $(\pi'_1 : X' \to X_1, \pi'_2 : X' \to X_2)$ satisfies this property, then the diagram

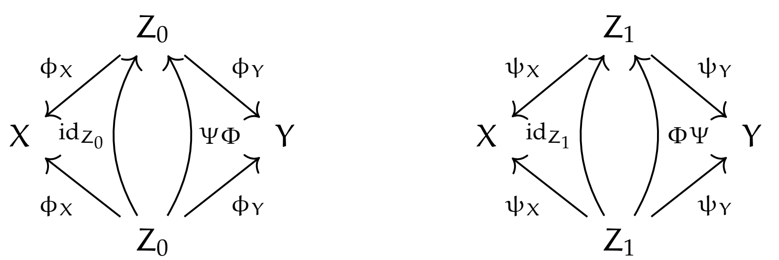


demonstrates that the unique morphism $f'f : X \to X$ satisfies $\pi_1 = \pi_1 f'f$ and $\pi_2 = \pi_2 f'f$; since id_X also satisfies this property, we have $f'f = \mathrm{id}_X$, and by the same reasoning $ff' = \mathrm{id}_{X'}$, making X' and X homeomorphic to one another. It follows that the product of topological spaces can be *defined* (up to homeomorphism) by this category-theoretic requirement, which takes place abstractly in Top. We can generalize this to an arbitrary category C:

The product of two objects X, Y is, if it exists, an object $X \times Y$ equipped with morphisms $\pi_X : X \times Y \to X$ and $\pi_Y : X \times Y \to Y$ such that for every Z equipped with a pair of morphisms $f : Z \to X$ and $g : Z \to Y$, there is a unique morphism $h : Z \to X \times Y$ with $\pi_X h = f$ and $\pi_Y h = g$.

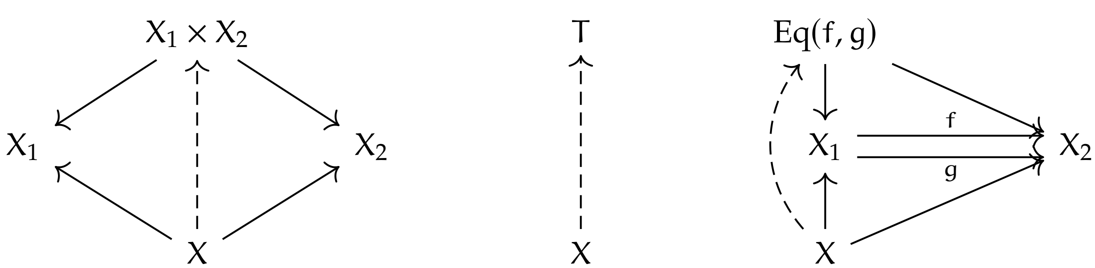
The product in Top is the topological product, as we've seen; in Ab, Set, and CRing, it's the product of abelian groups, cartesian product of sets, and product of rings, respectively. All of these share the same property of being unique up to isomorphism. In general, suppose two objects X, Y in a category C have two products, Z_0 and Z_1. Then Z_0 and Z_1 are isomorphic.

Proof. Let ϕ_X, ϕ_Y be the canonical projections from Z_0 and ψ_X, ψ_Y the canonical projections from Z_1. By the universal property of the product, Z_1 has an arrow $\Psi : Z_1 \to Z_0$ such that $\phi_X \circ \Psi = \psi_X$ and $\phi_Y \circ \Psi = \psi_Y$, and Z_0 has an arrow $\Phi : Z_0 \to Z_1$ such that $\psi_X \circ \Phi = \phi_X$ and $\psi_Y \circ \Phi = \phi_Y$. It follows that $\phi_X \circ \Psi \circ \Phi = \psi_X \circ \Phi = \phi_X$, and $\phi_Y \circ \Psi \circ \Phi = \phi_Y$. Likewise, $\psi_X \circ \Phi \circ \Psi = \psi_X$ and $\psi_Y \circ \Phi \circ \Psi = \psi_Y$. It follows that both the morphisms $\Psi \circ \Phi$ and id_{Z_0} satisfy the required factorization identities in the product diagram for Z_0, and likewise for Z_1, as indicated in the following diagrams:



So $\mathrm{id}_{Z_0} = \Psi\Phi$ and $\mathrm{id}_{Z_1} = \Phi\Psi$, making Φ and Ψ isomorphisms between Z_0 and Z_1. ∎

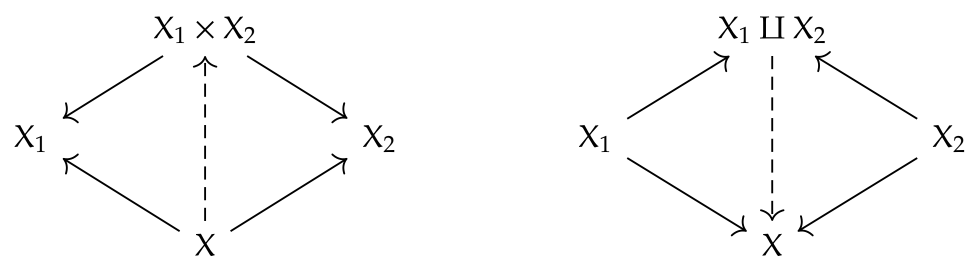
This manner of thinking about categorical constructions can be vastly generalized: for instance, we may ask for an object that classifies morphisms into *no* objects, i.e. an object T that has a unique morphism $f : X \to T$ for all $X \in C$. Such an object is known as a **terminal object**. We may even throw morphisms into the mix: given a diagram $f, g : X_1 \rightrightarrows X_2$, we may ask for an object Y equipped with morphisms $i : Y \to X_1, j : Y \to X_2$ such that $fi = gi = j$ any other object equipped with commuting morphisms to X_1 and X_2 bears a unique morphism to Y; such a Y, when it exists, is known as the **equalizer** of f and g, $\mathrm{Eq}(f, g)$.



Diagrams for products, terminal objects, and equalizers; dashed arrows are unique

This process is generalized in the obvious way to arbitrary diagrams; the object corresponding to a certain diagram is known as the **limit** of that diagram. For instance, the limit of the empty diagram is the terminal object, the limit of the diagram $X_1\ X_2$ is the product $X_1 \times X_2$, and the limit of the diagram $f, g : X_1 \rightrightarrows X_2$ is the equalizer $\mathrm{Eq}(f, g)$. The proof of the uniqueness of products up to isomorphism generalizes easily to the uniqueness of any kind of limit. In particular, any category can have at most one terminal object up to isomorphism. In Set, all singletons are terminal objects – for X an arbitrary set, there's only a single function $f : X \to \{*\}$ sending all $x \in X$ to the single object $*$ – and all singletons are isomorphic, allowing us to just speak of "a" terminal set; if we need a specific one, we'll use the ordinal $1 := \{\varnothing\}$.

Duality Given any category C, we can flip all the arrows, obtaining the opposite category C^{op}. For instance, a morphism $X \to Y$ in Set^{op} is given by a function $f : Y \to X$. In general, every arrow-theoretic statement and construction has a dual, given by flipping all the arrows and attaching the prefix 'co'; this is known as the principle of duality. For instance, the *co*product of two objects $X_1, X_2 \in C$ is an object $X_1 \amalg X_2$ equipped with two morphisms $i_1 : X_1 \to X_1 \amalg X_2$, $i_2 : X_2 \to X_1 \amalg X_2$ such that any X also equipped with such morphisms has a unique morphism *from* $X_1 \amalg X_2$ making everything commute.



Comparing product and coproduct diagrams

We similarly have **coequalizers**, *co*terminal (**initial**) objects, and in general, **colimits**.

An especially ubiquitous notion is given by that of a cofunctor, or a contravariant functor: A **contravariant functor** $F : C \to D$ is a functor $C^{op} \to D$. Specifically, each arrow $f : X \to Y$ in C is sent to an arrow $Ff : FY \to FX$, and composition works backwards, sending $g \circ f : X \to Y \to Z$ to $F(gf) = (Ff)(Fg) : FZ \to FY \to FX$. Normal functors are often called **covariant** when specification is required.

Example. For every object X in a category C, there is a covariant functor $C(X, -) : C \to \mathrm{Set}$ sending $Y \in C$ to the set $C(X, Y)$, and a morphism $f : Y \to Z$ to the set map $f_* : C(X, Y) \to C(X, Z)$ sending a $g : X \to Y$ to $f_*(g) = f \circ g$. The dual, contravariant functor is given by $C(-, X)$, which sends an object Y to $C(Y, X)$ and a map $f : Y \to Z$ to $f^* : C(Z, X) \to C(Y, X)$, $g \mapsto g \circ f$. $C(X, -)$ and $C(-, X)$ are known as the covariant and contravariant **representable functors** for X.

Example. A **lattice** is a poset which, as a category, has all binary products and coproducts. The coproduct is to be interpreted as the join (or sup, logical OR) $x \vee y$ and the product as the meet (or inf, logical AND) $x \wedge y$. Since the categorical structure on an arbitrary poset is given by writing an arrow $x \to y$ whenever $x \leq y$, the join of two elements x, y is an element $x \vee y$ satisfying $x, y \leq x \vee y$, and such that any object z satisfying $x, y \leq z$ also satisfies $x \vee y \leq z$. In this way, $x \vee y$ is the least upper bound of x, y, while $x \wedge y$ is the greatest lower bound.

If L has elements 0 and 1 such that $0 \leq x \leq 1$ for all $x \in L$, then 0 and 1 are the initial and terminal objects of L as a category. Equalizers and coequalizers are trivial in lattices, so a lattice with 0 and 1 is a poset which, as a category, has all finite limits and colimits.

We may also define lattices with 0 and 1 equationally: a lattice is a set with two distinguished elements 0 and 1, and two associative, commutative binary operations $\vee$ and $\wedge$ such that $x \wedge x = x \vee x = x$, $1 \wedge x = 0 \vee x = x$, and $x \wedge (y \vee x) = (x \wedge y) \vee x = x$. The partial order is recovered by the relation $x \leq y \iff x = x \wedge y \iff y = x \vee y$. If also $x \wedge (y \vee z) = (x \wedge y) \vee (x \wedge z)$, or, *equivalently*, $x \vee (y \wedge z) = (x \vee y) \wedge (x \vee z)$, we say that the lattice is distributive. If L has for each x an element $\neg x$ such that $x \wedge \neg x = 0$ and $x \vee \neg x = 1$, then such a $\neg x$ is unique, and is known as the **complement** of x. A **Boolean algebra** is a distributive lattice with 0 and 1 in which every element x has a complement. In such a lattice, the DeMorgan laws hold: $\neg(x \vee y) = \neg x \wedge \neg y$, $\neg(x \wedge y) = (\neg x) \vee (\neg y)$, and $\neg\neg x = x$. For instance, every poset of subsets of a given set is a Boolean algebra under the operations of union, intersection, and complement; in fact, every Boolean algebra can be constructed up to isomorphism in this manner.

Equivalence and Universality As indicated earlier, the notion of naturality plays a large role in category theory; categories and their morphisms serve as a method of organizing objects of a certain type, and basic constructions on categories (taking limits, opposites, etc.) yield natural constructions on the corresponding objects. The key ingredient in all of these constructions is universality, which can be thought of as selecting the "most general" or "best" way of doing something: for instance, the product $X \times Y$ of two objects is the most general object that bears morphisms to both X and Y, in the sense that all other objects with morphisms to X and Y see those morphisms factor *uniquely* through those of $X \times Y$ [4]. Even without the use of category theory, universal properties show up throughout mathematics: for instance, the tensor product $M \otimes N$ of R-modules M and N satisfies the universal property that any bilinear morphism $M \oplus N \to P$ factors uniquely through $M \otimes N$; informally, it is the *most general way* to turn bilinear homomorphisms into linear morphisms. The localization of a ring A at a multiplicatively closed subset $S \subset A$ satisfies the universal property that every ring homomorphism $A \to B$ which sends A to an invertible element of B factors uniquely through $S^{-1}A$; it is the most general way to add inverses to A.

Category theory also allows us to weaken the notion of equivalence from strict equality ($=$) to isomorphism ($\cong$). Many categories have a natural notion of a "morphism between morphism", or a **2-morphism**: e.g., natural transformations serve as the 2-morphisms in Cat. In a category with 2-morphisms, known as a **2-category**, we can further weaken the notion of equivalence: let X, Y be objects of a 2-category C with morphisms $F : X \to Y$ and $G : Y \to X$ such that FG admits a 2-isomorphism $\alpha : FG \cong id_Y$ and GF a 2-isomorphism $\beta : GF \cong id_X$. In $C = CAT$, this concept bears a special name: An **equivalence of categories** $C \cong D$ is a pair of functors $F : C \to D$, $G : C \to D$ equipped with natural isomorphisms $\alpha : FG \cong id_Y$ and $\beta : GF \cong id_X$.

Yoneda's Lemma For a category C, we will denote the functor category $\mathrm{Set}^{C^{op}}$ of contravariant functors $C \to \mathrm{Set}$ by $\widehat{C}$; its elements are known as **presheaves**. Yoneda's lemma states that C admits a full and faithful embedding into its category of presheaves $\widehat{C}$.

For a covariant functor $F : C \to \mathrm{Set}$, the set $\widehat{C}(C(X, -), F)$ of natural transformations from $C(X, -)$ to F is isomorphic to FX. For a contravariant $F : C^{op} \to \mathrm{Set}$, $\widehat{C}(C(-, X), F) \cong FX$.

Proof. For F covariant, take an arbitrary $a \in FX$. Letting $\alpha_X(id_X) = a$ defines a unique natural transformation in which any $f : X \to Y$ must be mapped to $(Ff)(a)$. Conversely, any $a \in FX$ defines a unique natural transformation $\alpha_Y(f) = (Ff)(a)$. For F contravariant, flip the direction of f. ■

Note that when $F = C(-, Y)$, the contravariant version yields $\widehat{C}(C(-, X), C(-, Y)) \cong C(X, Y)$. We may use this to define an embedding of C in $\widehat{C}$: the **Yoneda embedding** is the functor よ $: C \to \widehat{C}$ sending X to $C(-, X)$ and $f : X \to Y$ to the natural transformation $C(-, X) \Rightarrow C(-, Y)$ corresponding to f. Since the sets of natural transformations between two functors *are* the hom-sets in the functor category $\widehat{C}$, よ is a full and faithful functor, and hence a proper embedding.

[4] The name "universality" derives from the fact that this property is expressed via *universal properties*, as $\forall \ldots \exists! \ldots$.

Furthermore, $\widehat{\mathsf{C}}$ also contains all colimits in a natural way: (Co-Yoneda lemma) Every element of $\widehat{\mathsf{C}}$ is a colimit of a diagram of contravariant representable functors in a canonical manner. For further details and a proof, see [MacLane and Moerdijk, 2012], pgs. 41-43.

1.1.3 Adjunctions

The "best" relation two functors $F : C \to D$ and $G : D \to C$ can have is their forming an equivalence of categories $C \cong D$. Then, morphisms in C can be mapped to morphisms in D in a natural and reversible manner (up to isomorphism). The *next* best relation F and G can have is a failure of equivalence on objects, but an equivalence on morphisms, in the sense that $D(FX, Y)$ is in bijection with $C(X, GY)$ for all $X \in C$, $Y \in D$. If this happens in a natural manner, we say that F and G are adjoint functors. Adjunctions show up everywhere, as we will demonstrate.

Given locally small categories C and D, along with functors $F : C \to D$ and $G : D \to C$, we call F and G **adjoint** functors if there's a natural isomorphism Φ between the following functors from $C^{op} \times D$ to Set:

$$\Phi : D(F-, -) \cong C(-, G-)$$

Then, F is said to be left adjoint to G, and G is said to be right adjoint to F. This relation is written as $F \dashv G$, with the $\dashv$ symbol pointing towards the *left* adjoint (we could also write $G \vdash F$).

The name "adjoint" comes from linear algebra, where the adjoint of an operator A on an inner product space V is another operator $A^{\dagger}$ satisfying $\langle Av, w\rangle = \langle v, A^{\dagger}w\rangle$: we "move" the operator to the other side by taking its adjoint.

Example. The free abelian group on a set S, is defined to be an abelian group $F(S)$ along with an inclusion set map $i_S : S \to F(S)$ such that every set map $u : S \to A$, where A is an abelian group, factors as $u = \varphi \circ i$ for a unique homomorphism φ. A set map $f : S \to T$ generates by composition a map $i_T \circ f : S \to F(T)$, and hence a unique homomorphism $F(S) \to F(T)$; it can be verified that when $f = id_S$, this homomorphism is $id_{F(S)}$, and furthermore that composition of these induced maps is associative. This evidences F as a functor Set $\to$ Ab, known as a **free functor**. If we let J be the forgetful functor Ab $\to$ Set, then we see that $\mathrm{Set}(S, JA)$ is in bijection with $\mathrm{Ab}(FS, A)$: the map from set maps to group homomorphisms is given by the definition of the free group, and the map from group homomorphisms to set maps is given by taking $\varphi : F(S) \to A$ to the set map $\varphi \circ i : S \to JA$. This bijection is natural in both S and A, rendering F the left adjoint to J. Free-forgetful adjunctions of this nature are extremely common: in fact, we may define free functors as left adjoints to forgetful functors.

Example. In Set, maps $X \times Y \to Z$ can be identified with maps $X \to \mathsf{Set}(Y, Z)$ by **currying**: in lambda notation, we send $\lambda x, y.f(x, y)$ to $\lambda x.(\lambda y.f(x, y))$. This yields an adjunction with $- \times Y$ on the left and $\mathsf{Set}(Y, -)$ on the right. As we'll see later, this is the defining feature of a **cartesian closed category**.

Example. A **Heyting algebra** is a lattice H with 0 and 1 which has an right adjoint known as exponentiation associated to the functor $- \wedge y$. That is, there is for every x, y an object, generally written as $x \Rightarrow y$, such that $z \leq (x \Rightarrow y)$ iff $x \wedge x \leq y$, i.e. $x \Rightarrow y$ is a least upper bound for all elements z with $z \wedge x \leq y$. In particular, $y \leq (x \Rightarrow y)$.

The unit and counit of the exponential adjunction give us inclusions $x \leq (y \Rightarrow (x \wedge y))$ and $y \wedge (y \Rightarrow x) \leq x$. The properties $1^X \cong 1$ and $X^1 \cong X$, valid in any category with a right adjoint to its product functor, become $(x \Rightarrow 1) = 1$ and $(1 \Rightarrow x) = x$, and the properties $(y \times z)^x \cong y^x \times z^x$ and $x^{y \times z} \cong (x^y)^z$ become $(x \Rightarrow (y \wedge z)) = ((x \Rightarrow y) \wedge (x \Rightarrow z))$ and $((y \wedge z) \Rightarrow x) = (z \Rightarrow (y \Rightarrow x))$. Heyting algebras are distributive due to the fact that $- \wedge y$ is a left adjoint, and hence preserves coproducts: $((x \vee z) \wedge y) = ((x \wedge y) \vee (z \wedge y))$.

In a Heyting algebra, we may define the negation of x as $\neg x := (x \Rightarrow 0)$, the idea being that "not x" means "x implies falsity". This is not a strict negation: while $x \wedge \neg x = 0$, as evidenced by the identity $x \wedge (x \Rightarrow y) \leq y$, $x \vee \neg x$ isn't necessarily equal to 1. If x does have a strict negation, though, it is $\neg x$. So while $x \leq \neg\neg x$, this isn't a strict equality as in a Boolean algebra. However, $\neg x = \neg\neg\neg x$, and $x \leq y$ implies that $\neg y \leq \neg x$, so we're not totally lost. These features tell us that the logic of a Heyting algebra doesn't necessarily satisfy the law of double negation $x = \neg\neg x$, and as such is an *intuitionistic* logic rather than a classical one.

Given a predicate $S(x, y)$, where $x \in X$ and $y \in Y$ are elements of sets, we may regard S as the subset $S \subseteq X \times Y$ of those pairs for which $S(x, y)$ is true. The statement $(\forall x)S(x, y)$ then picks out a subset $T \subseteq Y$ consisting of all those y such that $X \times y \subseteq S$. Letting p denote the projection $X \times Y \to Y$, we may denote this subset as $\forall_p S$. The statement $(\exists x)S(x, y)$ is equivalent to $y \in p(S)$, and we will denote the corresponding subset by $\exists_p S$. Let $\mathcal{P}Y$ be the Boolean algebra of all subsets $T \subseteq Y$ and $\mathcal{P}(X \times Y)$ the Boolean algebra of all predicates S. Viewing these as categories, we have a pair of functors $\forall_p, \exists_p : \mathcal{P}(X \times Y) \rightrightarrows \mathcal{P}(Y)$. There is a third functor, $p^* : \mathcal{P}(Y) \to \mathcal{P}(X \times Y)$ which sends each subset $T \subseteq Y$ to its inverse image $p^*T = X \times T$. Then, there is an adjoint triple $\exists_p \dashv p^* \dashv \forall_p$. This follows from the fact that $p^*T \subseteq S \iff T \subseteq \forall_p S$ and $S \subseteq p^*T \iff \exists_p S \subseteq T$.

Example. Ab is naturally a subcategory of Grp, so we can define an inclusion functor $i : \mathsf{Ab} \to \mathsf{Grp}$ which just drops the 'abelian' prefix. The left adjoint of this functor is given by abelianization, sending a group G to $G/[G,G]$ and a group homomorphism $\varphi : G \to H$ to the map $\varphi^* : G \to H \to H/[H,H]$, which satisfies $\varphi(xy) = \varphi(x)\varphi(y) = \varphi(y)\varphi(x) = \varphi(yx)$ and hence extends to a morphism $G/[G,G] \to H/[H,H]$. In general, a subcategory C_0 of a category C is **reflective** when its inclusion functor has a left adjoint, and **coreflective** when the inclusion functor has a right adjoint.

A paramount feature of adjoints which we will state but not prove is their ability to preserve limits and colimits. Let $F : C \to D$ be left adjoint to $G : C \to D$, let Γ be a diagram in C, and let Δ be a diagram in D. Then, $\operatorname{colim} F\Gamma = F(\operatorname{colim}\Gamma)$ and $\lim G\Delta = G(\lim \Delta)$. Succinctly, *left adjoints preserve colimits and right adjoints preserve limits*.

Units and Counits Given an adjunction $\Phi : C(X, GY) \cong D(FX, Y)$, suppose we set $Y = FX$, giving us a bijection $C(X, GFX) \cong D(FX, FX)$. Plugging the identity 1_{F_X} in on the right side gives us a unique $\eta_X : X \to GFX$. Doing this for all X gives us a natural transformation $\mathrm{id}_C \to GF$, since an $h : X' \to X$ is translated to a $GFh : GFX' \to GFX$ such that $GFh \circ \eta_{X'} = \eta_X \circ h$ (proof: $GFh \circ \eta_{X'} = GFh \circ \Phi(\mathrm{id}_{FX'}) = \Phi(Fh \circ \mathrm{id}_{FX'}) = \Phi(\mathrm{id}_{FX} \circ Fh) = \Phi(\mathrm{id}_{FX}) \circ h = \eta_X \circ h$). Dually, we can let $X = GY$, so that plugging in id_{GY} into the right hand side of the bijection $C(GY, GY) \cong D(FGY, Y)$ gives us a natural transformation $\varepsilon : FG \to \mathrm{id}_D$. Both the composites $G \xrightarrow{\eta G} GFG \xrightarrow{G\varepsilon} G$ and $F \xrightarrow{F\eta} FGF \xrightarrow{\varepsilon F} F$ reduce to the identities 1_G and 1_F; from this, we obtain the adjunction's **zig-zag identities**

$$(\varepsilon F)(F\eta) = 1_F \qquad (\eta G)(G\varepsilon) = 1_G$$

We call η the **unit** of the adjunction and ε the **counit**.

Monads Consider the iterated composites of an endofunctor $T : C \to C$, i.e. $T^2 = TT, T^3, \ldots$. If $\mu : T^2 \to T$ is a natural transformation, with μ_X a morphism $T^2X \to TX$, then $T\mu = \{T\mu_X\}_{X\in C}$ is a natural transformation from T^3 to T^2, defined by $(T\mu)_X$ to $T(\mu_X)$. μT is another natural transformation between T^3 and T^2, defined by $(\mu T)_X := \mu_{TX}$.

A **monad** in a category C consists of an endofunctor T on C and two natural transformations $\eta : \mathrm{id}_C \to T$ and $\mu : T^2 \to T$ known as the **unit** and **multiplication** such that the following diagrams commute:

$$\begin{array}{ccc} T^3 & \xrightarrow{T\mu} & T^2 \\ {\scriptstyle \mu T}\downarrow & & \downarrow{\scriptstyle \mu} \\ T^2 & \xrightarrow{\mu} & T \end{array} \qquad \begin{array}{ccccc} 1T & \xrightarrow{\eta T} & T^2 & \xleftarrow{T\eta} & T1 \\ & \searrow\!\!\!\!\!= & \downarrow{\scriptstyle \mu} & =\!\!\!\!\!\swarrow & \\ & & T & & \end{array}$$

where 1 is the natural transformation $\{\mathrm{id}_X\}_{X\in C}$.

The structure is meant to resemble that of a monoid (identity, associative composition), with η the **unit** of T and μ the multiplication of T. In this sense, the left diagram just expresses the associativity of multiplication, and the right diagram expresses the left and right unit laws.

Example. As an example, the powerset functor $\mathcal{P} : \mathsf{Set} \to \mathsf{Set}, X \mapsto \mathcal{P}X$, $(\mathcal{P}f)(S) = f(S)$ forms a monad. The unit sends $X \in \mathsf{Set}$ to the map $\eta_X : \mathrm{id}_{\mathsf{Set}}(X) \to \mathcal{P}X, x \mapsto \{x\}$, and the multiplication sends X to the map $\mu_X : \mathcal{P}\mathcal{P}X \to \mathcal{P}X$, $\{S_\lambda\} \mapsto \bigcup_\lambda S_\lambda$.

To verify the coherence laws, let $\mathcal{S} = \{\{S_{\lambda_\xi}\}_{\xi\in\Xi}\}_{\lambda\in\Lambda}$, where each S_{λ_ξ} is a subset of X, be an arbitrary element of $\mathcal{P}\mathcal{P}\mathcal{P}X$. We want to verify that $(\mu_X\mu_{\mathcal{P}X})(\mathcal{S}) = (\mu_X\mathcal{P}\mu_X)(\mathcal{S})$. On one side, $(\mu_X\mu_{\mathcal{P}X})(\mathcal{S}) = \bigcup_{\lambda\in\Lambda}\left(\bigcup_{\xi\in\Xi} S_{\lambda_\xi}\right) = \bigcup_{\lambda,\xi} S_{\lambda_\xi}$. On the other side, note that $\mathcal{P}\mu_X$ is a map $\mathcal{P}\mathcal{P}\mathcal{P}X \to \mathcal{P}\mathcal{P}X$ sending $\mathcal{S}$ to $\{\bigcup_{\xi\in\Xi} S_{\lambda_\xi}\}_{\lambda\in\Lambda}$, so $(\mu_X\mathcal{P}\mu_X)(\mathcal{S}) = \bigcup_{\lambda\in\Lambda}\left(\bigcup_{\xi\in\Xi} S_{\lambda_\xi}\right) = \bigcup_{\lambda,\xi} S_{\lambda_\xi}$ as well. To verify the law for η, we must show that $\mu_X\eta_{\mathcal{P}X} = \mu_X\mathcal{P}\eta_X = \mathrm{id}_{\mathcal{P}X}$, which is evident from the trivial action of μ on singletons.

Every adjunction $F : C \to D \dashv G : D \to C$ gives rise to a monad in the category C. GF is the endofunctor on C, the unit $\eta : \mathrm{id}_C \to GF$ of the adjunction the unit of the monad, and, given the counit ε, the multiplication is given as $G\varepsilon F : GFGF \to GF$. The coherence laws then look like

$$\begin{array}{ccc} GFGFGF & \xrightarrow{GFG\varepsilon F} & GFGF \\ {\scriptstyle G\varepsilon FGF}\downarrow & & \downarrow{\scriptstyle G\varepsilon F} \\ GFGF & \xrightarrow[G\varepsilon F]{} & GF \end{array} \qquad \begin{array}{ccc} FGFG & \xrightarrow{FG\varepsilon} & FG \\ {\scriptstyle \varepsilon FG}\downarrow & & \downarrow{\scriptstyle \varepsilon} \\ FG & \xrightarrow[\varepsilon]{} & \mathrm{id}_D \end{array} \qquad \begin{array}{ccccc} \mathrm{id}_C\, GF & \xrightarrow{\eta GF} & GFGF & \xleftarrow{GF\eta} & GF\,\mathrm{id}_C \\ & \searrow\!\!\!\!\!= & \downarrow{\scriptstyle G\varepsilon F} & =\!\!\!\!\!\swarrow & \\ & & GF & & \end{array}$$

The middle diagram is just a restatement of the right, obtained by removing the G on the left and the F on the right; it must hold, since $\varepsilon\varepsilon = \varepsilon \cdot (FG\varepsilon) = \varepsilon \cdot (\varepsilon FG)$. The right diagram must hold since $1 = G\varepsilon \cdot \eta G = \varepsilon F \cdot F\eta$.

Example. Consider the free abelian group - forgetful functor adjunction $F \dashv U$. This yields a monad with unit $\eta : \mathrm{id}_{\mathsf{Set}} \to UF$ with $\eta_X : X \to UFX$ sending $x \in X$ to x considered as a basis element of FX and multiplication $U\varepsilon F : UFUF \to UF$, where $\varepsilon : FU \to 1_{\mathsf{Ab}}$ sends an abelian group A to a morphism $FUA \to A$ that takes the elements of an *element* of FUA (a collection of arbitrary un-concatenated elements of A) and multiplies them all together to get an element of A. This is conceptually similar to the power set monad, in that the unit "wraps" a set ($x \mapsto \{x\}$ vs. $x \mapsto \{\text{basis element } x\}$), whereas the multiplication gives us a way to reduce several elements at the same level (set of sets $\mapsto$ set of union of sets vs. set of elements of abelian group $\mapsto$ sum of elements in abelian group). This similarity comes from the fact that both monads involve Set as the base category.

Given a monad $T = (T, \mu, \eta)$ on C, an **algebra** over T, or a T-algebra, is an object $X \in C$ along with a morphism $f : TX \to X$ such that $f\eta_X = \mathrm{id}_X$ and $f(Tf) = f\mu_X$. In the power set monad on Set, for instance, an algebra is an assignment to each subset S of a given object X an element $f(S)$ such that $f(\{x\}) = x$ and $f(\{f(S_\lambda)\}) = f\left(\bigcup_\lambda S_\lambda\right)$. A **morphism** of T-algebras $(X, f) \to (Y, g)$ is a morphism $\alpha : X \to Y$ where the obvious square commutes: $g(T\alpha) = \alpha f$. Thus, any monad T on C gives us a category C^T of T-algebras, known as the **Eilenberg-Moore category** of T. While there is no natural choice of map $TX \to X$ (we have to choose a T-algebra structure), there is a natural map $\mu_X : T^2X \to TX$ giving TX a T-algebra structure. The functor $F^T : C \to C^T$ sending X to the algebra (TX, μ_X) is known as the **free algebra** functor, and the subcategory of C^T consisting of the free algebras is known as the **Kleisli category** C_T.

The free algebra functor $F^T : C^T \to C$ is left adjoint to the forgetful functor $C^T \to C, (X, f) \mapsto X$. The counit of this adjunction is the natural transformation $\mu : T^2 \to T$ and the unit is $\eta : 1 \to T$. In this way, not only does every adjunction generate a monad, but every monad comes from an adjunction.

1.2 Homotopy Theory

1.2.1 Homotopy Equivalence

Given two continuous functions $f, g : X \rightrightarrows Y$ between topological spaces, we may ask whether there is a "continuous transformation" of f into g. For instance, we may wonder whether two different loops on a torus (continuous functions $\gamma : [0, 1] \to T^2$ with $\gamma(0) = \gamma(1)$) can be morphed into one another continuously, i.e. without breaking one of the loops. Such a transformation between two paths, say f and g, would look like a *family* of paths $F_t(x)$, where $s \in [0, 1]$, such that $F_0(x) = f(x)$ and $F_1(x) = g(x)$. The right definition is as follows: A **homotopy** between two continuous maps $f, g : X \rightrightarrows Y$ is a continuous map $F : X \times [0, 1] \to Y$ such that $F(x, 0) = f(x)$ and $F(x, 1) = g(x)$. If there is a homotopy from f to g, the two maps are said to be **homotopic**, written as $f \simeq g$. We think of the second argument t as moving along the continuous family, and the first argument x as selecting a point in F_t.

Homotopy is an equivalence relation on the set of continuous maps $X \to Y$, and composition is compatible with this relation.

Proof. Every map f is homotopic to itself, by letting $F(x,t) = f(x)$. If F is a homotopy from f to g, then $F'(x,t) = F(x, 1-t)$ is a homotopy from g to f. Finally, if F is a homotopy from f to g and G a homotopy from g to h, defining $H(x,t) = F(x, 2t)$ for $0 \leq t \leq 1/2$ and $G(x, 2t-1)$ for $1/2 \leq t \leq 2$ yields a homotopy from f to h. So the relation whereby $f \sim g$ if $f \simeq g$ is reflexive, symmetric, and transitive, and hence an equivalence relation on $\mathsf{Top}(X,Y)$. Given two homotopies $f \simeq g : X \to Y$ and $h \simeq k : Y \to Z$, we may extend the homotopy $f \simeq g$ to a homotopy $h \cong f \simeq h \cong g$ that leaves h fixed but moves f to g, and likewise obtain a chain of homotopic maps $h \circ f \simeq h \circ g \simeq k \circ f \simeq k \circ g$. Therefore, we can define $[h] \circ [f]$ by taking the homotopy class of the composition of any representative of $[h]$ with any representative of $[f]$. ■

We may define a new category whose objects are those of Top, but whose morphisms are *homotopy classes* of morphisms in Top. This category, which is famously *not* concrete, is known as hTop.

We may sometimes want to restrict the set of homotopies between two maps $f, g : X \rightrightarrows Y$, requiring that all morphisms in our continuous family $F(x, -)$ preserve all points p in a subspace $X_0 \subseteq X$; such a homotopy is known as a homotopy relative to X_0. This is also an equivalence relation, the proof being more or less unchanged. The prototypical example is when $X = I, X_0 = \{0, 1\}$, and f, g are paths $I \to Y$; in this case, f is homotopic to g relative to the endpoints $\{0, 1\}$ when $F(x, 0) = f(x)$, $F(x, 1) = g(x)$, and $F(s,t) = f(s) = g(s)$ for all $s \in \{0, 1\}$.

A **pointed space** is a topological space X equipped with a specified element $x \in X$ known as the **basepoint**. A **basepoint-preserving map** f between pointed spaces (X, x) and (Y, y) is a continuous map $X \to Y$ sending x to y. When working in the category Top_* of pointed spaces and basepoint-preserving maps, we often denote all basepoints as $*$, lazily stating that $f(*) = *$ and so on. Homotopies in this category must necessarily be relative to the basepoint.

Quotienting the hom-sets in Top_* by the equivalence relation of basepoint-preserving homotopy yields the homotopy category hTop_* of pointed topological spaces. The product in Top_* is the product in Top, with the basepoint being the product of the two basepoints. The coproduct is not the disjoint union, however, since there would be no canonical basepoint; Top_* remedies this in the most obvious possible way, by identifying the basepoints of the two spaces with a single point. This forms the **wedge product** $X \vee Y$. Denoting the basepoints of X and Y by x_0 and y_0, there is a canonical inclusion $X \vee Y \hookrightarrow X \times Y$ sending $x \in X \subseteq X \vee Y$ to (x, y_0) and $y \in Y \subseteq X \vee Y$ to (x_0, y). Identifying this subspace of $X \times Y$ with a point yields the smash product $X \wedge Y = X \times Y / X \vee Y$.

1.2.2 Categories of Topological Spaces

Since $\mathsf{Top}_*(-,-)$ is a bifunctor, we can immediately form four important endofunctors on Top_*. Letting S_1 have an arbitrary basepoint 0, and defining I to be the interval $[0, 1]$ with the basepoint 0, these are:

- The loop space functor $\Omega = \mathsf{Top}_*(S^1, -)$
- The path space functor $P = \mathsf{Top}_*(I, -)$
- The reduced suspension functor $\Sigma = S^1 \wedge -$
- The reduced cylinder functor $C = I \wedge -$

The action of Ω and P on functions are canonically defined. The action of Σ and C on functions comes from the universal property of quotient spaces: if $A_0 \subseteq A$ and $B_0 \subseteq B$, then $f : A \to B$ extends to a unique map. Since $X \vee Y$ is sent to $X \vee f(Y) \subseteq X \vee Z$, this lets us define Σf and Cf for $X = S^1, I$. The action of the functor $X \wedge -$ on a map $f : Y \to Z$ is to send the image of (x, y) in $X \wedge Y$, which we can denote $x \wedge y$, to $x \wedge f(y) \in X \wedge Z$.

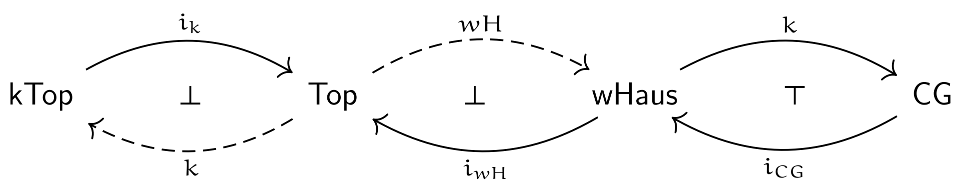
There are many nice properties of these functors which hold for most conceivable examples but fail to hold in general; for instance, the smash product is "usually" associative up to natural isomorphism, but fails to be so in general: as detailed in [May and Sigurdsson, 2006], $(\mathbb{Q} \wedge \mathbb{Q}) \wedge \mathbb{N}$ is not homeomorphic to $\mathbb{Q} \wedge (\mathbb{Q} \wedge \mathbb{N})$. As such, we may want to move to a more nicely behaved subcategory of Top_*, of which there are many. To specify certain subcategories, we need additional topological definitions. A space X is **weak Hausdorff** if, for all compact Hausdorff spaces Y and continuous functions $f : Y \to X$, the image of f is closed in X. X is a **k-space** if any subset $X_0 \subset X$ all of whose preimages are closed is itself closed. X is **compactly generated** if it is both weak Hausdorff and a k-space.

Topological manifolds, metric spaces, and compact Hausdorff spaces are all both compactly generated and Hausdorff, and are therefore contained in all of the following full subcategories of Top:

- kTop, the category of k-spaces

- wHaus, the category of weak Hausdorff spaces
- CG = kTop ∩ wHaus, the category of compactly generated spaces
- CGHaus, the category of compactly generated Hausdorff spaces

All of these have pointed, homotopy, and pointed homotopy variants. Letting i_k denote the inclusion functor kTop $\longrightarrow$ Top and i_{wH} the inclusion functor wHaus $\longrightarrow$ Top, we have a triplet of adjunctions:



The right adjoint k to i_k is known as k-ification, and the left adjoint wH to i_{wH} as weak Hausdorffification; k-ification turns a weak Hausdorff space into a compactly generated space, and, as a functor wHaus $\longrightarrow$ CG, is itself left adjoint to the inclusion functor CG $\longrightarrow$ wHaus. wHaus is complete, and right adjoints preserve limits, allowing us to construct limits in CG by constructing them in wHaus and then k-ifying. We will implicitly work in CG, letting $X \times Y$ denote the k-ification of the product in wHaus, and Y^X the k-ification of the space of maps from X to Y [5].

In CG, there is an adjunction $- \times Z \dashv (-)^Z$ for all Z, such that maps $X \times Z \longrightarrow Y$ can be identified in a natural way with maps $X \longrightarrow Y^Z$. In particular, $\mathsf{CG}(X \times I, Y) \cong \mathsf{CG}(X, PY)$ and $\mathsf{CG}(X \times S^1, Y) \cong \mathsf{CG}(X, \Omega Y)$. In the based version, CG_*, this becomes $- \wedge Z \dashv (-)^Z$, yielding the adjunctions $C \dashv P$ and $\Sigma \dashv \Omega$. (The exponential here ranges over basepoint-preserving maps, and its basepoint is the map that sends all points in the domain to the basepoint of the codomain). These adjunctions are preserved upon passing to homotopy classes. We will write $[X, Y]$ for $\mathsf{hCG}_*(X, Y)$, leaving the basepoints implicit.

1.2.3 Homotopy Groups

Given two loops $\gamma_0, \gamma_1 : (S^1, *) \longrightarrow (X, *)$, the **composite loop** $\gamma_0 * \gamma_1$ is defined by $(\gamma_0 * \gamma_1)(t) = \gamma_0(2t)$ if $0 \le t \le 1/2$, and $\gamma_1(2t - 1)$ if $1/2 \le t \le 1$. Under the operation of composition of loops, $[S^1, X]$ has the structure of a group.

[5] This space is equipped with the compact-open topology, whose subbase contains, for all $X_0 \subseteq X$, $Y_0 \subseteq Y$, the set of all functions $f : X \longrightarrow Y$ with $f(X_0) \subseteq Y_0$.

Proof. The proof that $*$ respects homotopy equivalence is similar to that of $\circ$ respecting homotopy equivalence. We define the identity element on $[S^1, X]$ to be the constant loop $e(t) = *$, and define the inverse of a loop $\gamma : S^1 \to X$ by the loop $\gamma^{-1}(t) = \gamma(1-t)$. To see that $[\gamma^{-1} * \gamma] = [e]$, use the homotopy $F(s,t) = \gamma_s(t) * \gamma_s(t)^{-1}$, where $\gamma_s(t) = \gamma(t)$ for $t \leq s$ and $\gamma(s)$ for $t \geq s$. This implies that $[\gamma * \gamma^{-1}] = [(\gamma^{-1})^{-1} * \gamma^{-1}] = [e]$ as well, so $([S_1, X], *)$ has a multiplication, inverses, and a two-sided identity. ∎

The **fundamental group** of a pointed space $(X, *)$ is defined as $\pi_1(X, *) := [S^1, X]$, with the group structure defined above. We will generally omit the $*$, just writing $\pi_1(X)$. The **higher homotopy groups** of a pointed space X are defined as $\pi_n(X) := [S^1, \Omega^{n-1}X] = \pi_1(\Omega^{n-1}X)$, $n \geq 1$. Since $S^n = \Sigma S^{n-1}$, we have $\pi_n(X) = [S^1, \Omega^n X] \cong [\Sigma^n S^1, X] \cong [S^n, X]$. This alternative definition allows us to interpret the nth homotopy group of a space X as the homotopically distinct ways of mapping the n-sphere into X in a basepoint-preserving manner, as well as to clearly demonstrate the functoriality of π_n; Every based map $f : X \to Y$ induces a map $\pi_n(X) \to \pi_n(Y)$ given by sending a loop $\ell : S^1 \to X$ to the loop $f \circ \ell : S^1 \to Y$. We can also define a *zeroth* homotopy group $\pi_0(X)$; this is just the set of path-connected components of X, and doesn't necessarily have a group structure.

As a consequence of the functoriality of homotopy groups, homeomorphic spaces have isomorphic fundamental groups. In fact, the motivation behind the introduction algebraic topology was the development of algebraic tools to figure out when two groups are homeomorphic.

A based map $f : X \to Y$ that induces isomorphisms $\pi_n(X) \cong \pi_n(Y)$ is known as a **weak equivalence**. Two spaces X, Y are **weakly equivalent**, written as $X \simeq Y$, when there is a weak equivalence between them. Homeomorphisms are weak equivalences, but the converse is not true in general; this means that, while two spaces X, Y with differing homotopy groups cannot be homeomorphic, verifying that all homotopy groups are the same isn't enough to verify that X and Y are homeomorphic.

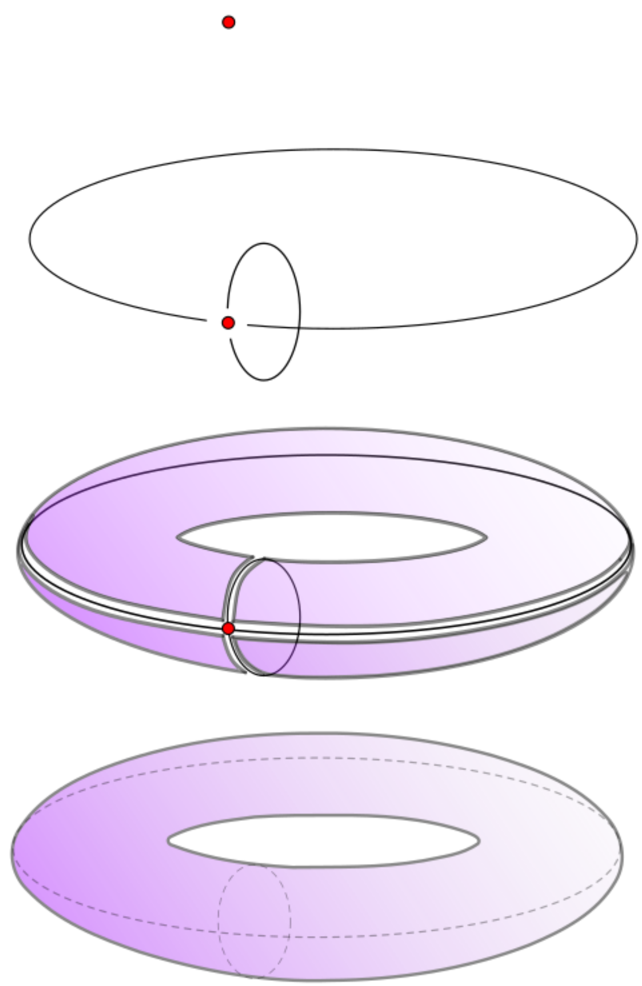
Homotopy groups will serve as one of our primary methods of classifying topological spaces, and weak equivalence will serve as an important notion of equality in this classification. Another notion of equivalence is similar to that of categories: two spaces X and Y are **homotopy equivalent** if there are continuous $f : X \to Y$ and $g : Y \to X$ such that fg and gf are homotopic to the identity maps on Y and X, respectively. This is also a weaker property than homeomorphism.

1.2.4 CW Complexes

The vast majority of spaces that come to mind when one thinks of a topological space all share a common trait: they can be pieced together using points and n-disks in a systematic manner. The circle S^1, for instance, is constructed by attaching $D^1 \cong [0,1]$, to a single point at both ends. Attaching two copies of D^2 to the circle along their boundaries yields a sphere. A torus can be constructed in a similar manner with one point, two 1-disks, and one 2-disk, as shown below.

We can make this construction pattern rigorous. The general process is as follows:

Figure 1.1: Construction of the torus T^2 as a CW complex.



1. Start with a set of points X^0.

2. Form an n-skeleton X^n from X^{n-1} by attaching a collection of open n-disks e^n_α via maps specifying where their boundary goes, $\varphi_n : S^{n-1} \to X^{n-1}$. We can say that X^n is the quotient space $X^{n-1} \coprod_\alpha D^n_\alpha$ of X^{n-1} under the identifications $x \sim \varphi_\alpha(x)$ for $x \in \partial D^n_\alpha$; as a set, $X^n = X^{n-1} \coprod_\alpha e^n_\alpha$.

3. Either stop at a finite stage (in which case X is finite-dimensional, and its dimension is n), or take the infinite union $X = \bigcup_n X^n$ and give it the weak topology, where A is open/closed in X iff $A \cap X^n$ is open/closed in X^n for all n.

Spaces constructed in this way are called **CW complexes**, a.k.a. cell complexes. Some examples:

- A 1-dimensional CW complex is a **graph**. (It's actually a multigraph, but we call it a graph).

- S^n is constructed with the cells e^0, a single point, and e^n, the disk D^n attached by the constant map $S^{n-1} \to e_0$. By part 2 of the construction, we can see that $S^n = D^n/\partial D^n$.

A **subcomplex** of a CW complex X is a closed subspace $A \subset X$ that's a union of cells in X; the closedness implies that the characteristic map of each of these cells has image contained in A, making A itself a CW complex. A pair (X, A) of a CW complex X and a subcomplex A is called a **CW pair**. Since each skeleton X^n of a subcomplex X is a closed subspace of X, (X, X^n) is a CW pair.

CW complexes are especially well behaved; they are all compactly generated Hausdorff, locally contractible, and paracompact; the full subcategory CW of Top consisting of the CW complexes is closed under topological products, wedge sums, and smash products. Homotopy equivalence between CW complexes is equivalent to weak equivalence, and every topological space is weakly equivalent to a CW complex.

1.3 Bundles

1.3.1 Vector Bundles

Let k denote any of the fields $\mathbb{R}, \mathbb{C}$, or $\mathbb{H}$. A **family of** (k-)**vector spaces** over a topological space X is a topological space E along with a continuous map $\pi : E \to X$ and a finite-dimensional k-vector space structure on each fiber $E_x := \pi^{-1}(x)$ such that the k-module addition and scalar multiplication $E \times_X E \to E$ and $k \times E \to E$ are continuous. A **homomorphism** between families

$\pi : E \to X$ and $\rho : F \to X$ is a continuous map which reduces to a linear transformation $E_x \to F_x$ on all x and satisfies $\rho f = \pi$.

For a family of vector spaces $\pi : E \to X$ and an open subset $Y \subseteq X$, write $E|_Y$ for the restriction $\pi^{-1}(Y) \subseteq E$. A **vector bundle** is a family of vector spaces $\eta : E \to X$ which is *locally trivial*: every $x \in X$ is contained in an open set $x \in U \subseteq X$ such that $\pi|_U : E|_U \to U$ is a **trivial bundle**, or a bundle which is isomorphic to one of the form $E = X \times k^n$. Given a vector bundle $E \to X$, the function $x \mapsto \dim E_x$ is a continuous function $X \to \mathbb{N}$, and therefore constant on connected components. If it is constant everywhere, then we can define the **dimension** $\dim E$ of the vector bundle. 1-dimensional bundles in particular are known as **line bundles**.

The category of k-vector bundles over X and their homomorphisms forms a category $\mathsf{VB}_k(X)$. A useful categorical property of $\mathsf{VB}_k(X)$ is its closure under pullbacks: given a continuous $f : X \to Y$ and a vector bundle $F \to Y$, there is an induced bundle $f^*F \to X$ given by the pullback $F \times_Y X$ and its projection maps:

$$\begin{array}{ccc} f^*F = F \times_Y X & \xrightarrow{\pi_F} & F \\ {\scriptstyle f^*\psi = \pi_X}\downarrow & & \downarrow{\scriptstyle \psi} \\ X & \xrightarrow[f]{} & Y \end{array}$$

So, in fact, we can interpret $\mathsf{VB}_k(-)$ as a contravariant functor from Top into some category of vector bundles in general. The **Whitney sum** $E \oplus F$ of two vector bundles over X is defined locally as $(E \oplus F)_x = E_x \oplus F_x$, and is given the subspace topology of $E \times F$; this is the finite product in $\mathsf{VB}_k(X)$, as well as the finite coproduct.

Given a vector bundle $\pi : E \to X$ and an open covering $\{U_i\}$ of X for which each E_{U_i} is locally trivial, with an isomorphism $h_i : U_i \times k^n \cong E_{U_i}$, we have on each intersection $U_i \cap U_j$ a map $h_i^{-1} h_j : (U_i \cap U_j) \times k^n \cong E_{U_i \cap U_j} \cong (U_i \cap U_j) \times k^n$ sending (x, v), considered as a point in E_{U_j}, to $(x, g_{ij} v)$ for some nonsingular n-dimensional matrix g_{ij}; the g_{ij} satisfy the *cocycle condition* $g_{ij} g_{jk} = g_{ik}$. While they must all lie in $GL(n; k)$, the set of all invertible $n \times n$ matrices over k, if they lie in any subgroup G of $GL(n; k)$, we call G the **structure group** of the vector bundle.

1.3.2 Principal Fiber Bundles

A **principal fiber bundle** (PFB) is a surjection of smooth manifolds $\pi : E \to X$ along with a choice of Lie group G which acts freely on E, with each $g \in G$ generating a diffeomorphism $R_g : E \to E, p \mapsto pg$. We require that for each $x \in X$, $\pi^{-1}(x)$ is diffeomorphic to G and there is a neighborhood $U \supset x$ such that $\pi^{-1}(U)$ is diffeomorphic to $U \times G$.

Examples of Lie Groups It will be useful to compile a list of Lie groups. Since most Lie groups have variants over $\mathbb{R}$ and $\mathbb{C}$, we will use k to denote a field which is either $\mathbb{R}$ or $\mathbb{C}$. Most (but not all) Lie groups encountered in nature are **matrix Lie groups**, or subgroups G of $GL(n;k)$, the set of automorphisms of k^n / $n \times n$ non-singular matrices with the elementwise convergence topology, which are closed in $GL(n;k)$. The Lie algebra associated to a matrix Lie group G can be calculated as the set of all $n \times n$ matrices X such that $e^{tX} \in G$ for all $t \in \mathbb{R}$, and equipped with the commutator $[X,Y] = XY - YX$. All groups below will be matrix Lie groups unless specified otherwise, and hence groups of invertible $n \times n$ matrices with product and inverse given by matrix multiplication and inversion.

The **special linear group** $SL(n;k)$ is the matrix Lie group consisting of $n \times n$ matrices over k with determinant 1, which has dimension $n^2 - 1$ (as a k-vector space). Since $\det(e^{tX}) = e^{\mathrm{Tr}tX} = 1$ if and only if $\mathrm{Tr}X = 0$, the corresponding Lie algebra $\mathfrak{sl}(n;k)$ consists of all $n \times n$ matrices over k with trace 0.

The **orthogonal group** $O(n)$ consists of $n \times n$ *real* matrices which are orthogonal, or satisfy $X^T X = I_n$. Any orthogonal matrix X must satisfy $\det(X^T X) = \det(X)^2 = \det(I_n) = 1$, implying that $\det(X) = \pm 1$. Since the determinant is a continuous function $M_n(k) \to k$, this implies that $O(n)$ consists of two topological components: matrices with determinant 1, and matrices with determinant -1. Indeed, these form its two connected components. The component with determinant 1 is known as the **special orthogonal group** $SO(n)$. We can equivalently think of $O(n)$ as the $n(n-1)/2$-dimensional group of symmetries of S^n, and $SO(n)$ as the subgroup (with unchanged dimension) of those symmetries which preserve orientation, i.e. aren't reflections. $O(1)$, the set of 1×1 real numbers pretending to be matrices satisfying $x^2 = 1$, is clearly $\{-1, 1\}$, with $SO(1) = \{1\}$. $SO(2)$ is diffeomorphic to S^1, and $SO(3)$ to $\mathbb{RP}^3$. Since $\det e^{tX} = e^{\mathrm{Tr}tX} > 0$, the Lie algebras of $O(n)$ and $SO(n)$ are the same, and consist of antisymmetric $n \times n$ matrices, or matrices X satisfying $X^T = -X$.

The expansion of $O(n)$ to the complex numbers is known as the **unitary group** $U(n)$, which is the set of all $n \times n$ complex matrices X such that $X^\dagger X = I_n$ (known as unitary matrices), where $(X^\dagger)_{ij} = \overline{X_{ji}}$ is the adjoint. Since $\det(X^\dagger) = \overline{\det(X)}$, we must have $\overline{\det(X)}\det(X) = 1$, and hence the determinant of a unitary matrix must lie in the circle group S^1. For $n = 1$, $\det(X)$ is equivalent to X, making $U(1)$ is equivalent to S^1. In general, $U(n)$ is connected of dimension n^2, but not simply connected: its fundamental group is $\mathbb{Z}$, regardless of n. The subgroup of all $n \times n$ unitary matrices with determinant 1 is known as the **special unitary group** $SU(n)$, and has dimension $n^2 - 1$. $SU(1)$ is trivial, whereas $SU(2)$ is equivalent to the space of all versors (unit quaternions) and hence diffeomorphic to S^3. The Lie algebra $\mathfrak{u}(n)$ consists of anti-Hermitian matrices, whereas $\mathfrak{su}(n)$ consists of anti-Hermitian matrices with vanishing trace.

Another variant of $O(n)$ is given as follows: define an inner product $\langle\cdot,\cdot\rangle_{n,k}$ on $\mathbb{R}^{n+k}$ by the formula $\langle x,y\rangle_{n,k} = x_1y_1 + \ldots + x_ny_n - x_{n+1}y_{n+1} - \ldots - x_{n+k}y_{n+k}$, or equivalently by the matrix $I_{n,k}$, as $\langle x,y\rangle_{n,k} = x^T \mathrm{diag}(1,\ldots,1,-1,\ldots,-1)y$ where the inner product matrix $I_{n,k}$ has n 1's and k -1's. The **generalized orthogonal group** $Or(n,k)$ (also written as $O(n,k)$) is defined as the subgroup of $GL(n+k;\mathbb{R})$ satisfying $\langle Ax,Ay\rangle_{n,k} = \langle x,y\rangle_{n,k}$, or equivalently $A^T I_{n,k} A = I_{n,k}$. Of particular interest in physics is the **Lorentz group** $Or(1,3)$.

Principal Bundles Take a fiber bundle $E = E \xrightarrow{\pi} X$ with typical fiber G. A specific choice T_U of diffeomorphism $\pi^{-1}(U) \to U \times G$ for every open set U is known as a **local trivialization** of E, or, to physicists, a **choice of gauge**. If we can choose a *global* trivialization $\pi^{-1}(X) \to E \times G$, then E is known as a **trivial bundle**.

The pushforward $\pi_* : TE \to TX$ of the projection map $\pi : E \to X$ sends a $v \in T_pE$ to a vector $\pi^*(v) \in T_pX$ acting on functions $f \in C^\infty(E)$ as $\pi^*(v)(f) = v(f \circ \pi)$, and the pullback $\pi^* : T^*X \to T^*E$ sends a one-form $\omega \in T_pX$ to the one-form $\pi^*(\omega)(v) = \omega(\pi_*(v))$. Each $g \in G$ generates a diffeomorphism $R_g : E \to E$, and hence also has a pushforward $(R_g)_*(v)(f) = v(f \circ R_g)$ and pullback $R_g^*(v)(\omega) = \omega((R_g)_*(f))$. We define the **vertical subspace** V_p at a point p of the bundle E to be $\ker\pi_*$, or the space of vectors sending all functions of the form $f \circ \pi$ to zero. Since the only vector $v \in T_pE$ that sends all functions f to zero is zero itself, V_p measures the extent to which π "flattens" E.

The vertical subspace V_p of T_pE can be explicitly calculated, but in general there are many different "horizontal" subspaces H_p such that $T_pE = V_p \oplus H_p$. A smooth selection $p \mapsto H_p$, generated by a set of $h = \dim V - \dim E$ smooth vector fields that span H_p at each point, would allow us to split an arbitrary vector field on E into a horizontal part, coming from X, and a vertical part inherent to E. A **connection** is such a smooth assignment $p \mapsto H_p$ which is invariant under the action of G, in the sense that $(R_g)_*(H_p) = H_{pg}$. Such a connection defines a map $\omega_p : T_pE \to \mathfrak{g}$ given by setting $\omega_p(v) = 0$ for precisely all $v \in H_p$, and, for the vector field $X_p^* = \frac{d}{dt}\left(pe^{tX}\right)\big|_{t=0}$ which is necessarily in V_p, setting $\omega_p(X_p^*) = X$. Since V_p and H_p depend smoothly on p, so does ω_p, allowing us to consider it as a section of the trivial bundle $\mathfrak{g} \otimes E$ tensored with $\Omega^k(E)$, i.e. an element of $\Gamma\left((\mathfrak{g} \times E) \otimes \left(\Lambda^k T^*E\right)\right) \cong \Gamma(\mathfrak{g} \times E) \otimes_{C^\infty(E)} \Omega^k(E)$, which we write as $\Omega^k(E,\mathfrak{g})$ Such an object is known as a $\mathfrak{g}$**-valued one-form**. Given an $A \in \mathfrak{g}$ and an

$\mathbb{R}$-valued (normal) k-form $\omega \in \Omega(E)$, we define $A \otimes \omega$ to be the $\mathfrak{g}$-valued one-form given by tensoring the section $p \mapsto A$ with ω. For a basis $E_\alpha, \alpha \in \{1, \ldots, \dim \mathfrak{g}\}$ of $\mathfrak{g}$, we can clearly write $\omega = \sum_\alpha E_\alpha \otimes \omega^\alpha$, where the ω^α are a set of $\mathbb{R}$-valued k-forms thought of as the components of ω in the basis $\{E_\alpha\}$.

Given $\mathfrak{g}$-valued k and ℓ-differential forms ω, η on a manifold M, we define their bracket to be the $(k+\ell)$-form

$$[\omega, \eta](v_1, \ldots, v_{k+\ell}) = \frac{1}{k!\ell!} \sum_{\sigma \in S_n} (-1)^\sigma [\omega(v_{\sigma(1)}, \ldots, v_{\sigma(k)}), \eta(v_{\sigma(k+1)}, \ldots, v_{\sigma(k+\ell)})]$$

where $[-,-]$ is the bracket of $\mathfrak{g}$. For $\mathbb{R}$-valued forms ω, η, we have $[A \otimes \omega, B \otimes \eta] = [A, B] \otimes (\omega \wedge \eta)$, and for $\mathfrak{g}$-valued k, ℓ, m-forms ω, η, ρ we have $[\omega, \eta] = (-1)^{k\ell-1}[\eta, \omega]$ and $(-1)^{km}[[\omega, \eta], \rho] + (-1)^{\ell m}[[\rho, \omega], \eta] + (-1)^{k\ell}[[\eta, \rho], \omega] = 0$. (In particular, for odd-valued forms ω, η, we have $[\omega, \eta] = [\eta, \omega]$). We say that $\mathfrak{g}$-valued differential forms on M, with their bracket, form a **graded Lie algebra**. Since d is an antiderivation on the graded algebra of $\mathbb{R}$-valued differential forms, it follows that d is an antiderivation on the graded Lie algebra of $\mathfrak{g}$-valued differential forms.

The $\mathfrak{g}$-valued one-form ω derived from a connection is known as a **connection one-form**. The connection $p \mapsto H_p$ also allows us to assign to each choice of gauge $T_U : \pi^{-1}(U) \to U \times G$ a $\mathfrak{g}$-valued one-form on U. Take the pullback of the local section map $\sigma_U : U \to E, p \mapsto T_U^{-1}(p, e)$ to get a map $\sigma_U^* : T^*E \to T^*U$, $\sigma_U^*(\omega)(v) = \omega((\sigma_U)_*(v))$, and, after tensoring with $\Gamma(\mathfrak{g} \times E)$, let $\omega_U := \sigma_U^* \omega$ be the promised $\mathfrak{g}$-valued one-form on U. These one-forms on open subsets of X are known as **gauge potentials**. When G is abelian, these trivially piece together to form a well-defined one-form on the whole of X, but when G is nonabelian, we must take into account the choice of gauge on each open set.

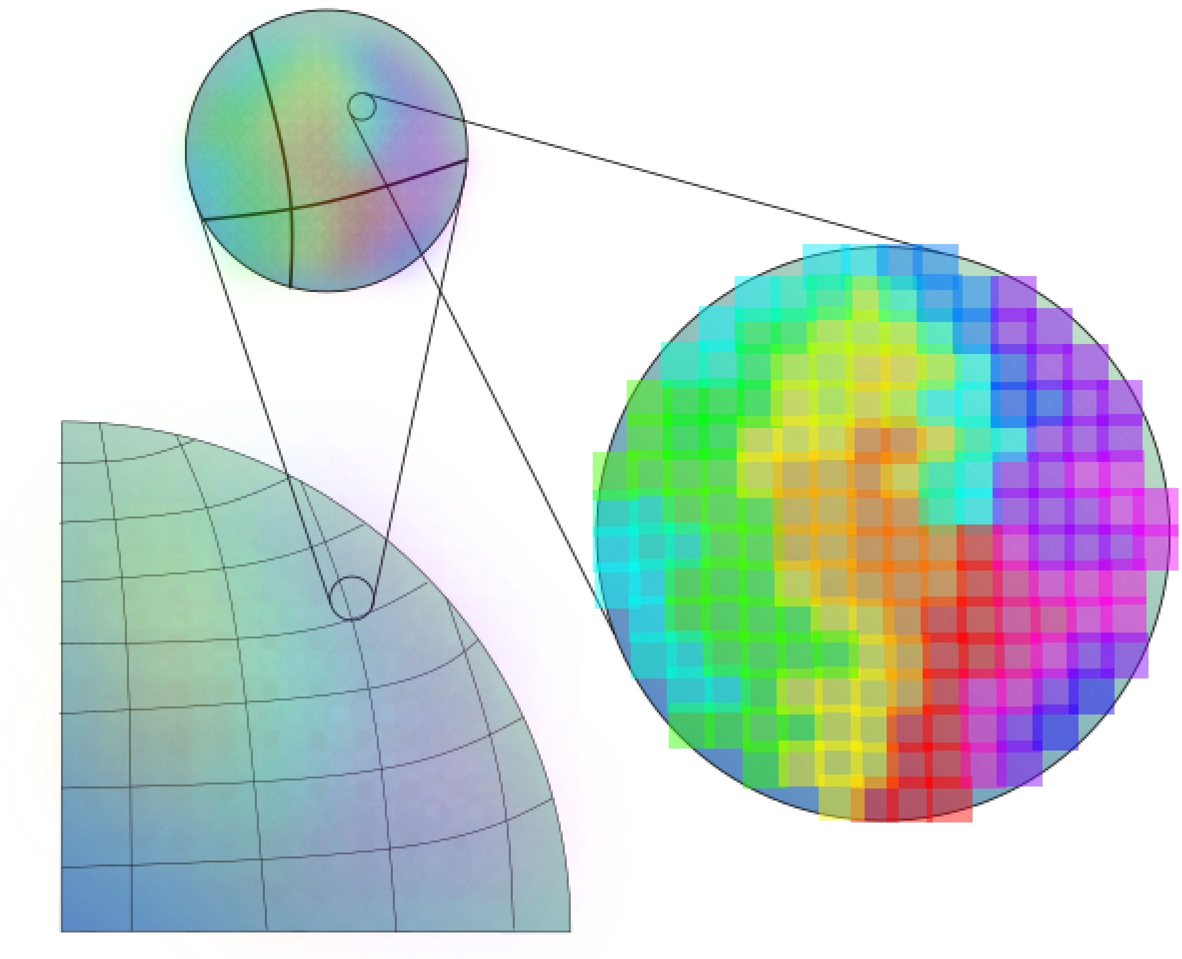
For instance, the Lie algebra $\mathfrak{u}(1)$ is the imaginary line $i\mathbb{R}$, so a connection one-form on a $U(1)$-bundle $E \xrightarrow{\pi} X$ – an S^1-bundle with an extra "rotating" action – looks like an ordinary one-form in the imaginary domain. Since $\mathfrak{u}(1)$ is abelian, a connection gives us a $\mathfrak{u}(1)$-valued one-form ω on X which we can write as $\omega = -iA$, where A is an ordinary one-form on X.

1.4 Enriched Categories

Many families of objects that naturally assemble into categories can be endowed with additional operations. Some motivating examples, some of which we have already seen:

- (Monoidal structure) Given two R-modules M and N, their tensor product is the module

Figure 1.2: A section of a principal $U(1)$-bundle.



$M \otimes N$, unique up to isomorphism, such that bilinear maps $\varphi : M \times N \to P$ are naturally in bijection with maps $M \otimes N \to P$. The operator $\otimes$ can be extended to a bifunctor $\mathsf{R\text{-}Mod} \times \mathsf{R\text{-}Mod} \to \mathsf{R\text{-}Mod}$, and equips $\mathsf{R\text{-}Mod}$ with the structure of a *monoid*.

- (Cartesian closed structure) Every function of the form $f : X \times Y \to Z$ in Set is equivalent to a function of the form $X \to \mathrm{Hom}_{\mathsf{Set}}(Y, Z)$ via currying. Similarly, in the category CGWH, the adjunction $- \times X \dashv -^X$ allows us to identify maps $Y \times X \to Z$ with maps $Y \to (X \to Z)$ in a manner entirely internal to CGWH.
- (Model structure) Every morphism in Top can be factored as a fibration followed by a cofibration [Riehl, 2014]. Any morphism which is both a fibration and a cofibration is a weak equivalence, inducing isomorphisms on all higher homotopy groups. The fibrations

and cofibrations on Top tell us what we need to know in order to do homotopy theory, and by defining fibrations and cofibrations in arbitrary categories, we may do homotopy theory in categories other than Top.

- (Enriched structure) Every hom-set in R-Mod is an abelian group in a natural way: the identity is the zero map $0(m) = 0$, and addition is given by $(\varphi + \psi)(m) = \varphi(m) + \psi(m)$. Composition is a bilinear map $\circ_{XYZ} : R(X, Y) \times R(Y, Z) \to R(X, Z)$ as well, so we say that R-Mod is *enriched* over Ab.

- (n-categorical structure) In Cat, morphisms are functors. The set D^C of functors $C \to D$ is itself a category, with natural transformations as morphisms; we can therefore say that Cat has not just hom-sets but hom-*categories*.

- (Abelian structure) In many categories enriched over Ab, such as R-Mod, morphisms have kernels, images, cokernels, and coimages; we can correspondingly find quotient objects and speak of the homology of chain complexes. [Weibel, 1995].

- (Topological structure) Diff admits a natural notion of a covering, in which a function family $\{M_i \to M\}$ covers the smooth manifold M if the images of all functions form an open cover of M [MacLane and Moerdijk, 2012]. It is possible to extend this notion of a covering to the notion of a topology on a category, known as a *Grothendieck topology*.

We will use these examples to construct a few hierarchies of structures that can be placed on (arbitrary) categories. Enriched categories, in particular, give us a way to replace the hom-sets of a category C with hom-*objects* in a category V with some additional structure necessary to define composition; n-categories are examples of enriched categories, and abelian categories are categories enriched over Ab with some additional niceness properties.

R-Mod is a very useful case study. Not only does the tensor product give it a monoidal structure, but every R-Mod is enriched over $\mathbb{Z}$-Mod = Ab in a manner compatible with the monoidal structure on Ab: the composition map $\circ_{XYZ} : R(X, Y) \times R(Y, Z) \to R(X, Z)$ is a bilinear map in Ab, and hence can be reduced to a single arrow $R(X, Y) \otimes_{\mathbb{Z}} R(Y, Z) \to R(X, Z)$. So, hom-sets in R-Mod are objects in Ab, and composition in R-Mod is described by morphisms in Ab in a manner compatible with Ab's monoidal structure. In general, any category C whose objects and morphisms can be described by a "monoidal category" V in a similar manner is said to be enriched over V.

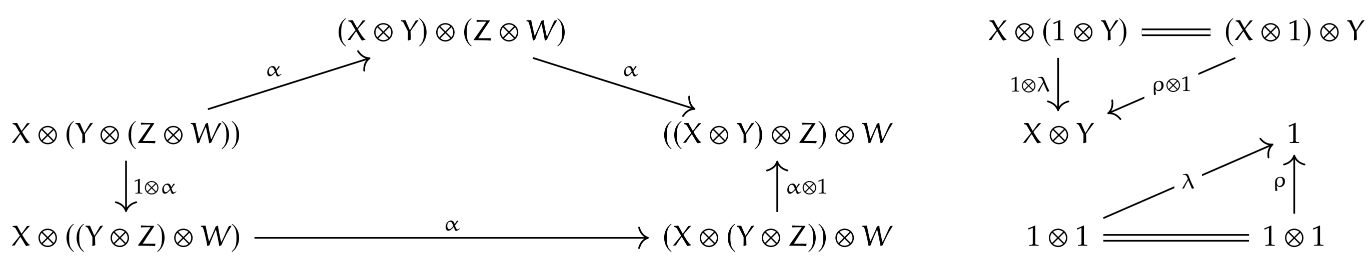
Our discussion of monoidal categories and enrichment is based largely off of [Mac Lane, 2013, Fong and Spivak, 2018, Riehl, 2014], with extra details pertaining to structures in monoidal categories based off of Coecke's articles [Coecke, 2010, Abramsky and Coecke, 2009].

1.4.1 Monoidal Categories

In many categories, there is a natural notion of a product of objects, which is functorial in nature: in Set, any object X gives rise to a functor $X \times - : \mathsf{Set} \to \mathsf{Set}$, sending Y to $X \times Y$ and a map $f : Y \to Z$ to $\mathrm{id}_X \times f : X \times Y \to X \times Z$. Due to the commutativity (up to isomorphism) of the cartesian product, this allows us to regard $- \times -$ as a functor $\mathsf{Set} \times \mathsf{Set} \to \mathsf{Set}$, also called a bifunctor (since it's functorial in both arguments). The same happens in R-Mod, with the tensor product $\otimes$ yielding a bifunctor $- \otimes - : R\text{-}\mathsf{Mod} \times R\text{-}\mathsf{Mod} \to R\text{-}\mathsf{Mod}$.

Monoidal categories generalize this kind of structure; we equip a category V with a bifunctor $\otimes : V \times V \to V$ which gives V the structure of a monoid. We generally don't require commutativity and associativity to hold exactly, but only up to natural isomorphism; this is sometimes called a weak monoidal structure, in contrast to a strong monoidal structure, but more often it is just called a monoidal structure.

Monoidal Categories A category V is a **monoidal category** when it is equipped with the following objects: a bifunctor $\otimes : V \times V \to V$, a selected object $1 \in V$ known as the **unit**, a natural isomorphism $\alpha : -_1 \otimes (-_2 \otimes -_3) \cong (-_1 \otimes -_2) \otimes -_3$ known as the **associator**, and a pair of natural isomorphisms $\lambda : 1 \otimes - \cong \mathrm{id}_V$ and $\rho : - \otimes 1 \cong \mathrm{id}_V$ known as the left and right **unitors**. We require that the following three diagrams commute:



In the special case that $\otimes = \times$, V is known as **cartesian monoidal**.

We may sum up all this information by defining a monoidal category as the tuple $(V, \otimes, 1, \alpha, \lambda, \rho)$, but we generally just write $(V, \otimes, 1)$, leaving the natural isomorphisms implicit (they are usually canonical). The canonical example is, again, $(R\text{-}\mathsf{Mod}, \otimes, R)$. If a functor F between monoidal categories satisfies $F(X \otimes Y) \cong (FX) \otimes (FY)$, F is known as **strong monoidal**.

A monoidal category $(V, \otimes, 1)$ is **symmetric** if there is an additional natural isomorphism $\gamma : -_1 \otimes -_2 \Rightarrow -_2 \otimes -_1$ such that $\gamma_{X,Y} \circ \gamma_{Y,X} = \mathrm{id}_{X \times Y}$ and $\rho_X = \lambda_X \circ \gamma_{X,1}$.

Again, γ is generally canonical, and so left implicit. For instance, in R-Mod, $\gamma_{XY} : X \otimes Y \to Y \otimes X$ sends $x \otimes y$ to $y \otimes x$, which is obviously natural, commuting with other morphisms. Since the categorical product is naturally commutative, cartesian monoidal categories are symmetric monoidal.

Cartesian Closed Categories A symmetric monoidal category $(V, \otimes, 1)$ is **closed** when the functor $- \otimes X$ has a right adjoint, denoted variously by $[X, -]$, $X \Rightarrow -$, or $-^X$. If $\otimes = \times$, V is known as a **cartesian closed category**. Explicitly, for every X, A, B there is an isomorphism $V(A \times X, B) \cong V(A, B^X)$, natural in all three variables; this isomorphism is known as currying[6]. The object B^X is known as the *internal* hom of B and X; it gives us a way to interpret hom-sets in V as actual objects in V.

The canonical example of a cartesian closed category is, as remarked above, $(\mathsf{Set}, \times, 1)$ (where the singleton $1 \cong \{\varnothing\}$ is a terminal object); here, $B^X = \mathsf{Set}(X, B)$. Another important example of a closed monoidal category is given by R-Mod; it is well known that $R(X, B)$ *can* be given the structure of an R-module as $(\varphi + \psi)(a) = \varphi(a) + \psi(a)$ and $(r\varphi)(a) = r(\varphi(a))$, and B^X is, up to isomorphism, precisely this R-module.

In an arbitrary cartesian closed category $(V, \otimes, 1)$, the counit of the $- \otimes X \dashv -^X$ adjunction is a morphism $B^X \otimes X \to B$ known as the *evaluation* morphism; in Set, this morphism takes a function $f : X \to B$ and an element $x \in X$ and simply returns $f(x)$, hence the name. The unit is known as the coevaluation morphism, and in Set sends an element $x \in X$ to the function that takes a $b \in B$ and yields the pair (x, b).

When C has terminal objects and binary products, the category of presheaves $\widehat{C}$ is cartesian closed: finite products are computed pointwise, and the exponential Q^P is given by $Q^P(X) = \widehat{C}(\text{よ}(X) \times P, Q)$, so we can write $Q^P = \widehat{C}(\text{よ}(-) \times P, Q)$. The evaluation counit $\mathrm{eval}_X : (Q^P \times P)(X) = \widehat{C}(\text{よ}(X) \times P, Q) \times P(X) \to Q(X)$ sends a natural transformation $\alpha : h_X \times P \to Q$ and an element $p \in P(X)$ to $\alpha(1_X, p) \in Q(X)$.

[6] In computer science, currying is the partial evaluation of functions, e.g. taking the binary function $f : X \times Y \to Z$ and plugging in a fixed x to get a unary function $f_{x_0} : Y \to Z, y \mapsto f(x, y)$; this operation is itself a function $\mathrm{Hom}(X \times Y, Z) \to \mathrm{Hom}(X, \mathrm{Hom}(Y, Z)), \lambda x, y.f(x, y) \mapsto \lambda x. (\lambda y.f(x, y))$.

1.4.2 Enriched Categories

A category V is closed monoidal when its hom-sets can be thought of as objects in V. If we can think of *another* category C's hom-sets as being objects in V, then C is said to be enriched over V, or a V-category. If V is not concrete, then a V-category C may not even be a category in the traditional sense (it would have hom-objects rather than hom-sets), so we must define V-categories in a more abstract manner.

Enriched Categories Given a symmetric monoidal category $(V, \otimes, 1)$, a V-**category**, or category **enriched** over V, is a collection $C = \{X_\lambda\}$ of objects, along with the following data: For each pair $X, Y \in C$, we have an object $C(X, Y) \in V$ known as the **hom-object**. For each $X \in C$ we have a morphism $\mathrm{id}_X : 1 \to C(X, X)$ representing the identity morphism, and, for each triplet $X, Y, Z \in C$, we have a morphism $\circ_{XYZ} : C(X, Y) \otimes C(Y, Z) \to C(X, Z)$. We require composition to be associative, in the sense that

$$\circ_{XYW} \circ (\mathrm{id}_{C(X,Y)} \otimes \circ_{YZW}) = \circ_{XZW} \circ (\circ_{XYZ} \otimes \mathrm{id}_{C(Z,W)}) : C(X, Y) \otimes C(Y, Z) \otimes C(Z, W) \to C(X, W)$$

for all X, Y, Z, W, and we require the identity to play nicely with composition in the usual sense, for which we require

$$\circ_{XXY} \circ (\mathrm{id}_{C(X,Y)} \otimes \mathrm{id}_X) = \rho_{C(X,Y)} : C(X, Y) \otimes 1 \to C(X, Y)$$

and

$$\circ_{YYX} \circ (\mathrm{id}_Y \otimes \mathrm{id}_{C(X,Y)}) = \lambda_{C(X,Y)} : 1 \otimes C(X, Y) \to C(X, Y)$$

where ρ and λ are the right and left unitors. When C is enriched over V, we call V the **base category**.

The idea behind using a morphism $1 \to C(X, X)$ to represent the identity morphisms in C is that in most naturally occurring base categories, such as $(\mathsf{Set}, \times, 1)$ and $(\mathsf{Ab}, \otimes_{\mathbb{Z}}, \mathbb{Z})$, morphisms from the unit to an arbitrary object X of the base category correspond to elements of X, so in the general case we think of a morphism $1 \to C(X, X)$ as corresponding to an "element" of $C(X, X)$.

In fact, this idea allows us to extract a **underlying category** C_0 from any V-category C. This has the same objects as C, but its hom-sets are given by $C_0(X, Y) := V(1, C(X, Y))$. The identity morphism id_X is the above specified morphism $1 \to C(X, X)$, and composition sends the pair $f : 1 \to C(X, Y)$, $g : 1 \to C(Y, Z)$ first to $f \otimes g : 1 \otimes 1 \cong 1 \to C(X, Y) \otimes C(Y, Z)$, and then to a morphism $1 \to C(X, Z)$ given by composing $\circ_{XYZ}$ with $f \otimes g$.

Enriched Functors Given two V-categories C, D, a V-**functor** $F : C \to D$ is a map on objects $X \mapsto FX$ along with, for every $C(X, Y) \in V$, a morphism in V, $F_{X,Y} : C(X, Y) \to D(FX, FY)$. We require that these morphisms commute with composition morphisms in V, in the sense that

$$\circ^D_{FX,FY,FZ} \circ (F_{Y,Z} \times F_{X,Y}) = F_{X,Z} \circ \circ^C_{X,Y,Z}$$

and we require that the identity map $1 \to C(X, X)$ composed with $F_{X,X} : C(X, X) \to D(FX, FX)$ be equal to the identity map $1 \to D(FX, FX)$.

Given a V-category D with an arrow $g : 1 \to D(Y, Z)$ in V, we can for any $X \in D$ define a $g_* : D(X, Y) \cong 1 \times D(X, Y) \to D(Y, Z) \times D(X, Y) \to C(X, Z)$; this is the equivalent of postcomposition by g. Equivalently, we can define a $g^* : D(Z, W) \cong D(Z, W) \times 1 \to D(Z, W) \times D(Y, Z) \to D(Y, W)$ which is equivalent to precomposition by g.

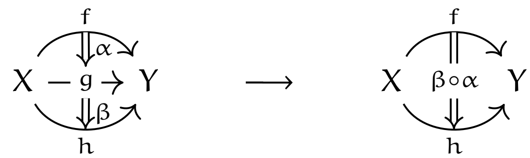
A V-**natural transformation** $\alpha : F \to G$ between V-enriched $F, G : C \to D$ is defined in the usual way, as a family of morphisms $\alpha_X : 1 \to D(FX, GX)$, but we require $(\alpha_Y)_* \circ F_{X,Y} = (\alpha_X)^* \circ G_{X,Y}$.

The set of V-categories along with V-functors forms a category of V-enriched categories, which we will call V-Cat. The notion of an equivalence of V-categories is roughly the same as in ordinary categories: we want an essentially surjective V-functor $F : C \to D$ that is V-fully faithful, in the sense that each $F_{X,Y} : C(X, Y) \to D(X, Y)$ is, as a morphism in V, an isomorphism. Similarly, a V-adjunction $F : C \to D \dashv G : D \to C$ is a natural isomorphism $D(F-, -) \cong C(-, G-)$, or equivalently V-natural transformations $\eta : 1 \to GF$ and $\varepsilon : FG \to 1$ satisfying the zig-zag identities.

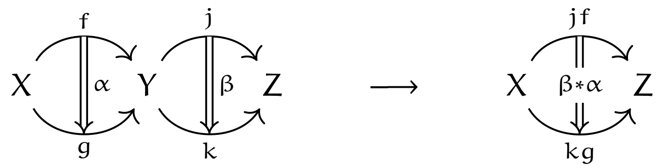
1.4.3 2-Categories

The category Cat of (small) categories has a cartesian product $\times$ and a terminal object consisting of the one-object category $*$ whose only morphism is the identity. The formation of the functor category D^C gives Cat an exponential, and thus makes it cartesian closed. In particular, it is symmetric monoidal, and we can therefore enrich over it. A category enriched over Cat, or a Cat-category, is known as a **2-category**.

More concretely, a 2-category C consists of objects $X, Y, \ldots$, and for every $X, Y \in C$ a category $C(X, Y)$. The objects of this category correspond to typical morphisms $X \to Y$, and are known as **1-morphisms** or **1-cells**. The morphisms of $C(X, Y)$ correspond to morphisms between functions, and are known as **2-morphisms** or **2-cells**. The categorical structure of $C(X, Y)$ allows us to **vertically compose** 2-morphisms as



and horizontally compose pairs of 2-cells as



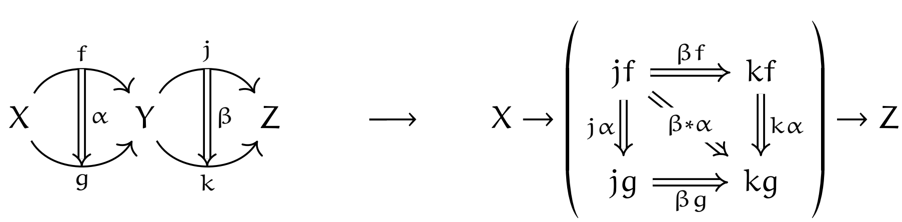
Horizontal composition comes from the fact that a 2-category C, being enriched over Cat, has a composition rule $\circ_{XYZ} : \mathsf{C}(Y,Z) \times \mathsf{C}(X,Y) \to \mathsf{C}(X,Z)$ which is an arrow in Cat, a.k.a. a *functor*: if on objects $\circ_{XYZ}(j,f) = jf$ and $\circ_{XYZ}(k,g) = kg$, then the morphism $(\beta,\alpha) : (j,f) \Rightarrow (k,g)$ must be sent by $\circ_{XYZ}$ to a 2-morphism $jf \Rightarrow kg$, which we denote as $\beta * \alpha$. The identity 2-cells $\mathrm{id}_{\mathrm{id}_X} : \mathrm{id}_X \Rightarrow \mathrm{id}_X$ are the identites for horizontal composition, and id_f is the identity for vertical composition on f. The horizontal composition of vertical composites is equal to the vertical composition of horizontal composites, in the sense that

$$(\delta\gamma) * (\beta\alpha) = (\delta * \beta) \cdot (\gamma * \alpha)$$

This is known as **middle-four interchange**.

The trivial horizontal composition

is known as the **whiskered composite** $b \cdot \alpha \cdot a : bfa \Rightarrow bga : X \to Y$. Of course, we can only whisker on one side if we want, letting the other side silently denote the identity morphism. Whiskering is natural, in the sense that every horizontally composable pair of 2-cells gives rise to a commutative square as follows:



We can also prove this by middle-four interchange:

$$(\beta g)(j\alpha) = (\beta * \mathrm{id}_g)(\mathrm{id}_j *\alpha) = (\beta\, \mathrm{id}_j) * (\mathrm{id}_g *\alpha) = \beta * \alpha = \ldots = (k\alpha)(\beta f)$$

Two objects X, Y in a 2-category C are **equivalent** if there are 1-morphisms $f : X \to Y$ and $g : Y \to X$ along with isomorphisms $\alpha : fg \Rightarrow \mathrm{id}_Y$, $\beta : gf \to \mathrm{id}_X$. For instance, in the 2-category Cat, this reduces to the notion of equivalence between categories. If fg and gf are not just isomorphic but strictly equal to id_Y and id_X, X and Y are isomorphic; in this way, equivalence is a "loosening" of isomorphism.

Given a category C and a 2-category D, we may define functors $C \to D$ by simply ignoring the 2-morphisms of D. A **pseudo-functor** $F : C \to D$ is, however, a looser notion of a functor which takes advantage of D's 2-categorical structure. Specifically, a pseudofunctor $F : C \to D$ is an assignment of objects $FX \in D$ to objects $X \in C$, 1-morphisms $Ff : FX \to FY$ in D to morphisms $f : X \to Y$ in C, along with for every $X \in C$ a 2-isomorphism $\alpha_X : \mathrm{id}_{FX} \Rightarrow F(\mathrm{id}_X)$ and, for every $X \xrightarrow{f} Y \xrightarrow{g} Z$ in C, a 2-isomorphism $\alpha_{g,f} : F(gf) \Rightarrow (Fg)(Ff)$. We require that for every $f : X \to Y$ in C we have $\alpha_{\mathrm{id}_Y,f} = \alpha_Y * \mathrm{id}_{Ff}$ and $\alpha_{f,\mathrm{id}_X} = \mathrm{id}_{Ff} * \alpha_X$. Furthermore, for any $W \xrightarrow{f} X \xrightarrow{g} Y \xrightarrow{h} Z$, we require the identity

$$(\mathrm{id}_{Fh} * \alpha_{g,f})\alpha_{h,gf} = (\alpha_{h,g} * \mathrm{id}_{Ff})\alpha_{hg,f}$$

In most cases, D is a 2-category of categories, with 1-morphisms being functors and 2-morphisms natural transformations, and F is a contravariant functor. In such a case, we can simplify the above definition. A **pseudo-functor** F on C is an assignment of a category FX to each $X \in C$, along with a functor $Ff : FY \to FX$ for each $f : X \to Y$, generally denoted by f^*. We require natural isomorphisms $\alpha_X : \mathrm{id}_X^* \cong \mathrm{id}_{FX}$ and $\alpha_{g,f} : f^*g^* \cong (gf)^*$ satisfying the following identities, where f, g, h are as above:

$$\alpha_{\mathrm{id}_X,f} = \alpha_X f^* \qquad \alpha_{f,\mathrm{id}_Y} = f^*\alpha_Y \qquad \alpha_{gf,h}\alpha_{f,g}h^* = \alpha_{f,hg}f^*\alpha_{g,h}$$

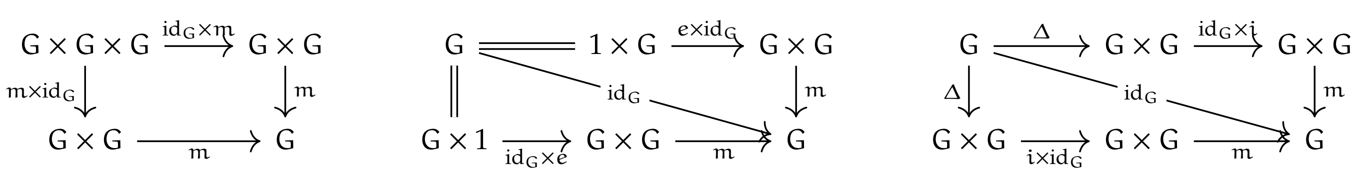
1.4.4 Internalization

Internal Category Theory Every small category C is in particular a pair of sets ($C_0 = \mathrm{Obj}(C)$, $C_1 = \mathrm{Mor}(C)$) equipped with the appropriate codomain, domain, and composition morphisms. In this manner, small categories are categories "internal" to set. In an arbitrary category C, we may define an **internal category** to be an object of objects C_0, an object of morphisms C_1, a domain morphism $d_0 : C_1 \to C_0$, a codomain morphism $d_1 : C_1 \to C_0$, an identity morphism $e : C_0 \to C_1$, and a composition morphism $m : C_2 \coloneqq C_1 \times_{C_0} C_1 \to C_1$, where the pullback is taken over the morphisms $d_0, d_1 : C_1 \to C_0$, expressing the fact that in order to compose two morphisms we require the codomain of the first to be the domain of the next. We require the obvious diagrammatic versions of the composition and identity laws. Writing $C = (C_0, C_1, d_0, d_1, e, m)$, a **internal functor** $F : C \to D$ between internal categories in C is a pair of morphisms $(F_0 : C_0 \to D_0, F_1 : C_1 \to D_1)$ in C which commute with the morphisms d_0, d_1, e, and m of each internal category. An **internal natural transformation** $\alpha : F \Rightarrow G$ is a morphism $\alpha : C_0 \to D_1$ such that $d_0\alpha = F_0$, $d_1\alpha = G_0$, and $m \circ ((\alpha \circ d_0) \times_{D_0} F_1) = m \circ (G_1 \times_{D_0} (\alpha \circ d_1))$. With internal functors and internal natural transformations, the collection of internal categories in C forms a 2-category $\mathsf{Cat}(C)$.

Example. For instance, a category internal to Vect is known as a 2-**vector space**: to be precise, a 2-vector space is a pair (V_0, V_1) of vector spaces along with linear maps $d_0, d_1 : V_1 \to V_0$, $e : V_0 \to V_1$, and a linear $m : V_1 \times_{V_0} V_1 \to V_0$ [Baez and Crans, 2003]. Since $d_0 \circ e = d_1 \circ e = \mathrm{id}_{V_0}$, e must be a surjection $V_0 \to V_1$.

Internalization The process of finding categories *internal* to a given category C is a specific example of the idea of **internalization**, the transportation of mathematical objects from Set to arbitrary categories by an analysis of the arrows involved.

For instance, we may define a group internal to a category C with all finite products, also known as a **group object** as an object G along with an identity map $e : 1 \to G$, a multiplication map $m : G \times G \to G$, and an inversion map $i : G \to G$ such that the following diagrams commute:



These diagrams tell us that multiplication is associative, the identity is two-sided, and inverses are two-sided. A group object homomomorphism is a morphism $\varphi : G \to H$ between group objects such that $\varphi \circ e_G = e_H$ and $m_H \circ (\varphi \times \varphi) = \varphi \circ m_G : G \times G \to H \times H \to H$; group objects and their homomorphisms yield a category $\mathsf{Grp}(C)$ of group objects in a category C.

Group objects in Set are clearly groups, but group objects in other categories behave more interestingly: in Top and Diff, the group objects are the topological and Lie groups, respectively.

A group object in Sch_S is known as an S**-group scheme**, and a group object in Cat is known as a **strict** 2**-group**. We may also internalize rings, R-modules, and other algebraic theories.

1.5 Homological Algebra

In this section, we'll add an increasing amount of structures to an arbitrary Ab-category, culminating in the definition of an *abelian category*. Such categories allow us to define homology and cohomology, and are very useful in the study of algebraic topology. $R\text{-}\mathsf{Mod}$ is the prototypical example of an abelian category, and in a sense is the universal example: the *Freyd-Mitchell embedding theorem* allows us to embed any category C, by means of a full and faithful functor, into some $R\text{-}\mathsf{Mod}$. As such, we'll think of the elements of abelian categories as being R-modules, allowing us to work with elements rather than arrow-theoretic language.

1.5.1 Abelian Categories

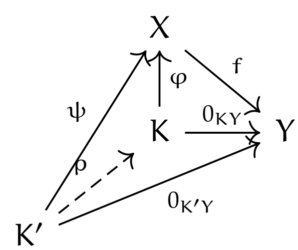
In an Ab-category C, every hom-set is an abelian group, and composition is a bilinear operation $\circ_{XYZ} : \mathsf{C}(X, Y) \times \mathsf{C}(Y, Z) \to \mathsf{C}(X, Z)$. An Ab-functor $F : \mathsf{C} \to \mathsf{D}$ between Ab-categories is a functor such that each mapping $\mathsf{C}(X, Y) \to \mathsf{D}(FX, FY)$ is a morphism in Ab, i.e. an abelian group homomorphism. Since Ab is a concrete category whose morphisms $1 = \mathbb{Z} \to G$ are in bijection with elements of G, the definition of an Ab-natural transformation simplifies to a family of homomorphisms $FX \to GX$ satisfying the usual commutativity condition.

Additive Categories In an Ab-category C, the finite product is, if it exists, equivalent to the coproduct. To see this, suppose for objects $X, Y \in \mathsf{C}$ we have a product $X \times Y$ with projections p_X and p_Y. Then, the pair of maps $(\mathrm{id}_X, 0_{XY})$ induces a morphism $i_X : X \to X \times Y$ such that $p_X i_X = \mathrm{id}_X$ and $p_Y i_X = 0_{XY}$; likewise, the pair of maps $(0_{YX}, \mathrm{id}_Y)$ induces a morphism $i_Y : Y \to X \times Y$. Take an object Z with morphisms $f : X \to Z$ and $g : Y \to Z$, and let $\varphi : X \times Y \to Z = f p_X + g p_Y$, such that $\varphi i_X = f p_X i_X + g p_Y i_X = f + g 0_{XY} = f$ and likewise $\varphi i_Y = g$. This construction satisfies the universal property of the coproduct, so $X \times Y$ is both a product and a coproduct. We call it the **biproduct**, and denote it $\oplus$.

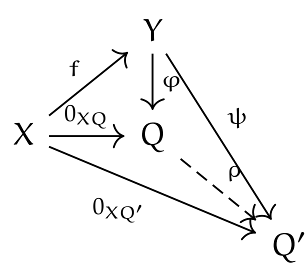
In an arbitrary category C, a **zero object** 0 is, if it exists, an object that is both initial and final. It has the special property that it defines a unique morphism, a **zero morphism**, between any two objects X and Y: this morphism, denoted 0_{XY}, is given by the composition $X \to 0 \to Y$.

We interpret the object 0 as carrying no information, and therefore zero morphisms destroy all information. An arbitrary Ab-category C has zero morphisms in a literal sense: they're the identities of the hom-groups. If C has a zero object 0, then $\mathsf{C}(0, X)$, necessarily being the trivial group, generates these zero morphisms in the manner described above. An Ab-category with a zero object and finite biproducts is known as an **additive category**.

Kernels In the Ab-category $R\text{-}\mathsf{Mod}$, the zero object is simply the zero module. Once we have a zero object, we can take a morphism $f : X \to Y$ and define its **kernel** to be the equalizer of f with 0_{XY}, and its **cokernel** to be the coequalizer of f with 0_{XY}. Specifically, the kernel is an object K along with a morphism $\varphi : K \to X$ such that $f\varphi = 0_{KY}$, and any other K' with a ψ satisfying $f\psi = 0_{K'Y}$ has a unique $\rho : K' \to K$ such that $\psi = \varphi\rho$. In pictures,



In Grp and Ab, K ends up being (isomorphic to) the set of all $x \in X$ that are mapped to 0 by f, with φ the inclusion map from K to X, recovering the normal definition of kernel. (While this case works out very nicely, as do cokernels, it must be emphasized that (co)kernels have not just objects but *morphisms* as well). The cokernel is an object Q along with a morphism $\varphi : Y \to Q$ such that $\varphi f = 0_{XQ}$, and any other Q' with a ψ satisfying $\psi f = 0_{XQ'}$ has a unique $\rho : Q \to Q'$ such that $\psi = \rho\varphi$. Another picture:



In Ab, the cokernel ends up being $Y/\mathrm{Im}(f)$. In summary, if $f\psi = 0$, then ψ factors uniquely through $\ker f$, but if $\psi f = 0$, then ψ factors uniquely through $\operatorname{coker} f$. Zero morphisms restrict the flow of information between two objects X and Y, kernels tell you how difficult it is to silence an X with a morphism $f : X \to Y$, and cokernels tell you how difficult it is to censor Y. The **image** of a morphism φ is defined by $\ker \operatorname{coker} \varphi$, and the **coimage** of φ is $\operatorname{coker} \ker \varphi$.

An additive category A is **abelian** if it has all kernels and cokernels, any monomorphism can be presented as the kernel of some morphism, and any epimorphism can be presented as the cokernel of some morphism.

1.5.2 Chain Complexes

In an abelian category A, a **chain complex** $C_\bullet$ is a collection $\{C_n\}_{n\in\mathbb{Z}}$ along with morphisms $\{d_n : C_n \to C_{n-1}\}_{n\in\mathbb{Z}}$, generally represented as a diagram of the form

$$\cdots \longrightarrow C_{n+1} \xrightarrow{d_{n+1}} C_n \xrightarrow{d_n} C_{n-1} \longrightarrow \cdots$$

We require that $d_n \circ d_{n+1} = 0$ for all n. This implies that $\ker d_n \subseteq \operatorname{im} d_{n+1}$ for all n; if these two submodules of C_n are equal for all n, then the chain complex $C_\bullet$ is said to be **exact**. Dually, a **cochain complex** $C^\bullet$ is a collection of objects $\{C^n\}_{n\in\mathbb{Z}}$ and morphisms $\{d^n : C^{n-1} \to C^n\}$ such that $d^{n+1} \circ d^n = 0$. In specific instantiations of such complexes there may be a specific reason for going in one direction or the other. In the abstract sense, though, flipping the indices is really all we have to do; for this reason, chain and cochain complexes are more or less equivalent, and a chain complex $(C_\bullet, d_\bullet)$ generates a cochain complex $(C^{-\bullet}, d^{-\bullet})$.

Homology An arbitrary chain complex $C_\bullet$ may or may not be exact; the extent to which it fails to be exact at an index n is equivalent to the extent to which $\operatorname{im} d_{n+1}$ fails to be as large as $\ker d_n$. It will always be a submodule, though, so we can record this failure of exactness by taking the quotient module $\ker d_n / \operatorname{im} d_{n+1}$. The **homology** of the chain complex $(C_\bullet, d_\bullet)$ is defined by

$$H_n(C_\bullet) = \ker d_n / \operatorname{im} d_{n+1}$$

and the **cohomology** of a cochain complex $(C^\bullet, d^\bullet)$ is given by

$$H^n(C^\bullet) = \ker d_n / \operatorname{im} d_{n-1}$$

In $\mathsf{R\text{-}Mod}$, elements of $\operatorname{im} d_{n+1}$ are known as the **boundaries** of C_n, and elements of $\ker d_n$ are known as the **cycles** of C_n; $H_n(C_\bullet)$ is then simply the submodule of cycles modulo the relation that identifies two cycles that differ only by a boundary.

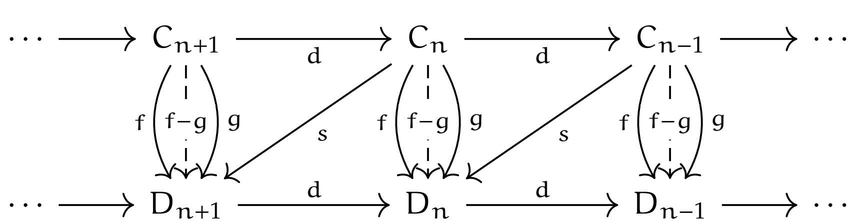
The Category of Chain Complexes A **morphism of chain complexes** $C_\bullet \to D_\bullet$ is a family $u_\bullet$ of morphisms in A such that

$$\begin{array}{ccccccccc}
\cdots & \longrightarrow & C_{n+1} & \xrightarrow{d_{n+1}} & C_n & \xrightarrow{d_n} & C_{n-1} & \longrightarrow & \cdots \\
 & & \downarrow{\scriptstyle u_{n+1}} & & \downarrow{\scriptstyle u_n} & & \downarrow{\scriptstyle u_{n-1}} & & \\
\cdots & \longrightarrow & D_{n+1} & \xrightarrow{d'_{n+1}} & D_n & \xrightarrow{d'_n} & D_{n-1} & \longrightarrow & \cdots
\end{array}$$

is a commutative diagram. The set of all chain complexes on A, along with chain maps between chain complexes, forms a category $\mathsf{Ch}(A)$. This is itself an abelian category, with all kernels, cokernels, sums of morphisms, etc. being computed pointwise. Given a chain map $f : C_\bullet \to D_\bullet$ in $\mathsf{Ch}(R\text{-}\mathsf{Mod})$, we note that if $d_i(g) = 0$ for $g \in C_i$, then $d'_i f_i(g) = f_{i-1} d_i(g) = 0$, and that if $g = d_{i+1}(h)$, then $f_i(g) = f_i d_{i+1}(h) = d'_{i+1} f_{i+1}(h)$; chain maps send boundaries to boundaries and cycles to cycles, and hence induce well-defined maps $H_i(C_\bullet) \to H_i(D_\bullet)$. In this way, the map $H_i : \mathsf{Ch}(R\text{-}\mathsf{Mod}) \to R\text{-}\mathsf{Mod}$, $C_\bullet \mapsto H_i(C_\bullet)$ acts functorially; this holds for an arbitrary abelian category A. Two chain complexes are **quasi-isomorphic** if all of their homology objects are isomorphic; this provides a weaker notion of equivalence than isomorphism.

A chain complex is **bounded** if all but finitely many of the C_n are 0. If C_n is non-zero solely when $n \in [a, b]$, we say that $C_\bullet$ has **amplitude** in $[a, b]$. $C_\bullet$ is **bounded above** if there's a b such that $C_n = 0$ for all $n > b$, and **bounded below** if there's an a such that $C_n = 0$ for all $n < a$. Keeping in line with the identification $C_n = C^{-n}$, a cochain complex is bounded above/below iff its associated chain complex is bounded below/above. These allow us to form full subcategories of Ch(A): the categories of bounded, bounded above, bounded below, and non-negative chain complexes are denoted Ch$(A)_b$, Ch$(A)_-$, Ch$(A)_+$, and Ch$(A)_{\geq 0}$, respectively.

Chain Homotopies A chain complex $C_\bullet$ is **split** if there are maps $s_n : C_n \to C_{n+1}$ such that $d = dsd$. It is **split exact** if it is also exact; equivalently, it is split exact if and only if $ds + sd$ is the identity map. If we have a chain map $f : C_\bullet \to D_\bullet$, f is called **null homotopic** if there are maps $s_n : C_n \to D_{n+1}$ such that $f = ds + sd$. Two chain maps $f, g : C_\bullet \rightrightarrows D_\bullet$ are **chain homotopic** if their difference $f - g$ is null homotopic, i.e. there are maps $s_n : C_n \to D_{n+1}$ such that $f - g = ds + sd$. A diagram:



The maps $\{s_n\}$ are collectively called a **chain homotopy**. We will regard the notion of chain homotopy as an extension of the notion of a homotopy between maps between topological spaces. Correspondingly, we call two chain complexes $C_\bullet$ and $D_\bullet$ **chain homotopy equivalent** if there are maps $f : C_\bullet \to D_\bullet$ and $g : D_\bullet$ to $C_\bullet$ such that gf and fg are equivalent to the identities on $D_\bullet$ and $C_\bullet$, respectively.

1.5.3 Resolutions

Let $F : A \to B$ be an Ab-functor between abelian categories A, B. If, for all exact sequences in A of the form $0 \to X \to Y \to Z \to 0$, F yields an exact sequence $0 \to FX \to FY \to FZ \to 0$, F is known as a **exact functor**. If just $0 \to FX \to FY \to FZ$ is exact, F is known as **left exact**, and if $FX \to FY \to FZ \to 0$ is exact, F is known as **right exact**.

For a fixed $M \in A$, the covariant representable functor $A(M, -)$ is left exact. To see this, let $0 \to X \xrightarrow{f} Y \xrightarrow{g} Z \to 0$ be exact. As in R-Mod, f must be monic and g must be epic. Take the map $f_* := A(X, f)$ sending $\varphi : M \to X$ to $f\varphi : M \to Y$. If $f\varphi = 0_{MY}$, then since f is monic, φ must be 0_{MX}. So f_* is monic, and likewise $g_*f_*(\varphi) = gf\varphi = 0_{XZ}\varphi = 0_{MZ}$, so $g_*f_* = 0_{A(M,X),A(M,Z)}$. Finally, if $\varphi : M \to Y$ satisfies $g_*(\varphi) = 0$, then, since $\operatorname{im}\varphi$ is a subobject of $\operatorname{im} f$, φ factors through f as $\varphi = f\psi = f_*(\psi)$ for some $\psi : M \to X$. So $0 \to A(M, X) \to A(M, Y) \to A(M, Z)$ is exact. Hence, $A(M, -)$ is a left exact functor.

Projective Objects It is not in general true that the final arrow $A(M, Y) \to A(M, Z)$ is an epimorphism, so that we could extend the left exact sequence to an exact sequence. For this to be true, we require the following (equivalent) universal lifting property on M: given any surjection $g : Y \to Z$ in A, and any map $\varphi : M \to Z$, there is a (not necessarily unique) map $\psi : M \to Y$ such that $\varphi = f\psi$. If M had this property, it would follow immediately that $A(M, Y) \to A(M, Z)$ is an epimorphism, and hence that $A(M, -)$ is an exact functor. If M satisfies this universal lifting property, or equivalently if $A(M, -)$ is an exact functor, we call M a **projective object**. For instance, free modules are projective. For some nice rings R, including $\mathbb{Z}$, fields, and division rings, the projective R-modules are the free modules, but this isn't always the case. In general, an R-module is projective if and only if it's a direct summand of a free R-module.

Injectives The dual notion is that of an **injective object**, or an object $M \in A$ such that every monomorphism $f : X \to Y$ and map $\varphi : X \to M$ yields at least one $\psi : Y \to M$ such that $f\psi = \varphi$.

The contravariant functor $\mathsf{A}(-, M)$ is right exact, since it is $\mathsf{A}^{op}(M, -)$ which, A^{op} being abelian, sends exact sequences in A^{op} to left exact sequences in Ab, and hence exact sequences in A to right exact sequences in Ab). $\mathsf{A}(-, M)$ is exact if and only if M is injective. Injective modules are harder to characterize then projective modules, but if $\mathsf{A} = R\text{-}\mathsf{Mod}$ for R a principal ideal domain, then M is injective if and only if for every $r \neq 0 \in r$ and $m \in M$, $m = rm'$ for some $m' \in M$, so that we can "divide" elements of M by nonzero elements of R. For instance, $\mathbb{Q}$ is injective as a $\mathbb{Z}$-module.

It is in general true that left adjoints are right exact and right adjoints are left exact, since left adjoints preserve colimits, and hence cokernels, and right adjoints preserve kernels. In the case $\mathsf{A} = R\text{-}\mathsf{Mod}$, this observation is another way to show that $R(M, -)$ is left exact, and its left adjoint $M \otimes_R -$ is right exact.

Resolutions For some nice rings R, including $\mathbb{Z}$, fields, and division rings, the projective R-modules are the free modules, but this isn't always the case. In general, an R-module is projective if and only if it's a direct summand of a free R-module. $R\text{-}\mathsf{Mod}$ has enough projectives: given an R-module A, take the free R-module on the set of elements of A, $\pi(A) := FJA$. The counit of the $F \dashv J$ adjunction gives us a natural map $\pi(A) \to A$ (that sends a sequence of elements of A to its sum) which is a surjection.

An abelian category A has **enough projectives** if for every $M \in \mathsf{A}$ there is an epimorphism from a projective object P to M, and **enough injectives** if there is a monomorphism from X to an injective object I. A **left resolution** of M is a complex $X_\bullet$ along with a map $\epsilon : X_0 \to M$ such that the following sequence

$$\dots \longrightarrow X_2 \xrightarrow{d_2} X_1 \xrightarrow{d_1} X_0 \xrightarrow{\epsilon} M \longrightarrow 0$$

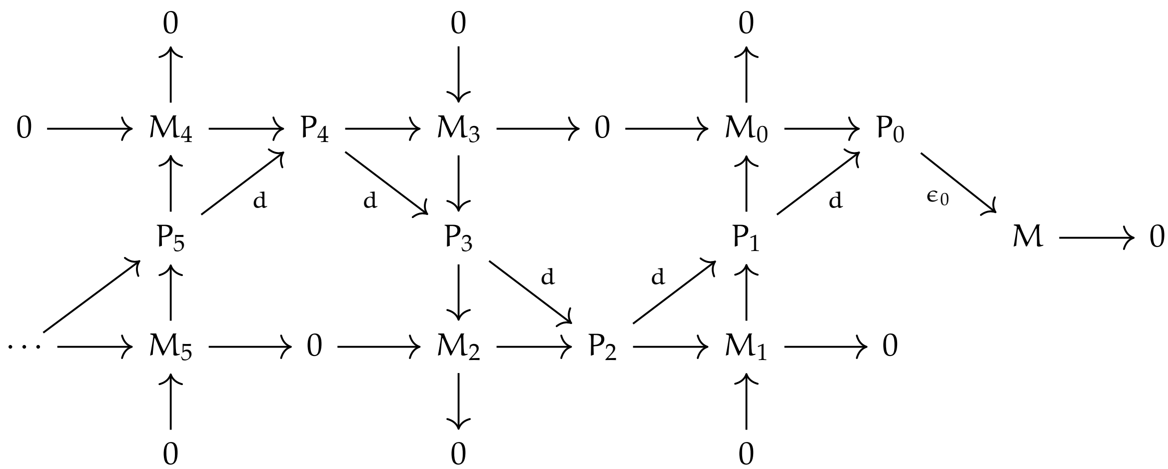
is exact. If furthermore all X_i are projective objects, then $X_\bullet$ is known as a **projective resolution** of M. Dually, a **right resolution** of M is a cochain complex $X^\bullet$ along with a map $\epsilon : M \to X^0$ such that the sequence

$$0 \longrightarrow M \xrightarrow{\epsilon} X^0 \xrightarrow{d^1} X^1 \xrightarrow{d^2} X^2 \longrightarrow \dots$$

is exact. If all X^i are injective, $X^\bullet$ is known as a **injective resolution**.

In an abelian category A with enough projectives (injectives), *every* object $M \in \mathsf{A}$ has a projective (injective) resolution.

Proof. Choosing a projection $\epsilon_0 : P_0 \to M$, we recursively choose a projective P_n and an epimorphism $\epsilon_n : P_n \to M_{n-1}$, set $M_n = \ker \epsilon_n$, and let $d_n : P_n \to P_{n-1}$ be the composition $P_n \to M_{n-1} \to P_{n-1}$. See:



Using our $\pi(A) \to A$ projection as ϵ_0, we see that M_0 consists of all sequences in $\pi(A)$ that sum to 0 (and comes with an injection into P_0), P_1 is $\pi(M_0)$, coming with a canonical ϵ_1, and so on. The kernel of each d is the image of the next, by design, so this is a projective resolution of M.

The proof for injective objects is dual to the above proof. ∎

Maps between objects M, N naturally induce chain maps between projective resolutions. Letting $P_\bullet \xrightarrow{\epsilon} M$, $Q_\bullet \xrightarrow{\eta} N$ be projective resolutions of M and N, and f a morphism $M \to N$, there is a chain map $\alpha : P_\bullet \to Q_\bullet$ that lifts f in the sense that the following diagram commutes:

$$\begin{array}{ccccccccccc} \cdots & \longrightarrow & P_2 & \longrightarrow & P_1 & \longrightarrow & P_0 & \xrightarrow[\epsilon]{} & M & \longrightarrow & 0 \\ & & \downarrow{\scriptstyle \alpha_2} & & \downarrow{\scriptstyle \alpha_1} & & \downarrow{\scriptstyle \alpha_0} & & \downarrow{\scriptstyle f} & & \\ \cdots & \longrightarrow & Q_2 & \longrightarrow & Q_1 & \longrightarrow & Q_0 & \xrightarrow{\eta} & N & \longrightarrow & 0 \end{array}$$

This chain map is unique up to chain homotopy equivalence.

The dual phenomenon is observed with injective objects: an injective resolution $N \xrightarrow{\theta} I^\bullet$ is naturally lifted *by* f to an injective resolution $M \xrightarrow{\zeta} E^\bullet$ in a way that makes the following diagram commute:

$$\begin{array}{ccccccccccc} 0 & \longrightarrow & M & \xrightarrow[\zeta]{} & E^0 & \longrightarrow & E^1 & \longrightarrow & E^2 & \longrightarrow & \cdots \\ & & \downarrow{\scriptstyle f} & & \downarrow{\scriptstyle \alpha^0} & & \downarrow{\scriptstyle \alpha^1} & & \downarrow{\scriptstyle \alpha^2} & & \\ 0 & \longrightarrow & N & \xrightarrow{\theta} & I^0 & \longrightarrow & I^1 & \longrightarrow & I^2 & \longrightarrow & \cdots \end{array}$$

1.5.4 Derived Functors

Left Derived Functors Fix a right exact functor F, and take an R-module M. Given a projective resolution $P_\bullet$ of M, $FP_1 \to FP_0 \to FM \to 0$ is an exact sequence, but the rest of $FP_\bullet$ isn't necessarily exact. The ith homology of $FP_\bullet$ is known as the ith **left derived functor** of F, $L_iF(M) := H_i(FP_\bullet)$. The homology at the zeroth position is given by $L_0F(M) = FM$, so the ith derived functor of F can be seen as the ith "homological extension" of F, with the zeroth extension obviously being F itself. The module $L_iF(M)$ is independent of the projective resolution we choose for M: any two different projective resolutions $P_\bullet, Q_\bullet$ will yield a pair of chain maps $f : P_\bullet \to Q_\bullet$, $g : Q_\bullet \to P_\bullet$ each lifting the identity map id_M, implying that $h = gf$ is a map $P_\bullet \to P_\bullet$ lifting id_M from $P_\bullet$ to itself. Since $\mathrm{id}_{P_\bullet}$ also serves this role, and h is unique up to chain homotopy, h and $\mathrm{id}_{P_\bullet}$ must be chain homotopic, and hence induce equivalent maps on homology, implying that the transformation induced by using $Q_\bullet$ instead of $P_\bullet$ – which is a natural transformation – has an inverse, and hence a natural isomorphism.

Example. Our canonical example of a right exact functor on $R\text{-}\mathsf{Mod}$ is $- \otimes_R N$; its corresponding left derived functors are known as the **Tor** functors, defined by

$$\mathrm{Tor}_i^R(M, N) := L_i(- \otimes_R N)(M)$$

$\mathrm{Hom}_R(-, N)$ is also right exact, and we define the **Ext** functors by

$$\mathrm{Ext}_R^i(M, N) := L_i(\mathrm{Hom}_R(-, N))(M)$$

Right Derived Functors Given a *left* exact functor F and an R-module M with an (again, arbitrary) injective resolution $I^\bullet$, we can define the **right derived functor** $R^iF(M)$ to be the ith cohomology of $FI^\bullet$, $R^iF(M) := H^i(FI^\bullet)$. When $F = \mathrm{Hom}_R(M, -)$, we again arrive at $\mathrm{Ext}_R^i(M, N) := R^i(\mathrm{Hom}_R(M, -))(N)$. Namely, it doesn't matter if we compute the Ext functor via a left or right derived functor, and in the same vein we can show that $L_i(- \otimes_R N)(M) \cong L_i(M \otimes_R -)(N) \cong \mathrm{Tor}_i^R(M, N)$; further exposition can be found in [Weibel, 1995].

A table of correspondences:

Left derived functor L_iF	Right derived functor R^iG
Right exact functor F	Left exact functor G
Projective resolution $P_\bullet \to A$	Injective resolution $A \to I_\bullet$
$L_iF(A) = H_i(F(P))$	$R^iG(A) = H^i(G(P))$
Tor functor	Ext functor

For computational purposes, it's useful to note that Tor_i^R preserves filtered colimits – colimits over what are essentially directed preorders – and in particular directed limits (which are, confusingly, actually colimits) in both variables. In the case of $\mathsf{Ab} = \mathbb{Z}\text{-}\mathsf{Mod}$, since every abelian group G is the direct limit of its finitely generated subgroups, we only need to know a few values of $\mathrm{Tor}_i^{\mathbb{Z}}$, perhaps computed directly via selecting convenient projective resolutions, in order to compute a wide variety of Tor groups.

Example. For an arbitrary abelian group G, we may calculate $\mathrm{Tor}_i^{\mathbb{Z}}(\mathbb{Z}/n\mathbb{Z}, G)$ by selecting the projective resolution $0 \to \mathbb{Z} \xrightarrow{\times n} \mathbb{Z} \to \mathbb{Z}/n\mathbb{Z} \to 0$, which upon tensoring with G becomes $0 \to G \xrightarrow{\times n} G \to 0$. The homology of this complex at the 0th position is G/nG, and the homology at the first position is the n-torsion subgroup ${}_nG = \{g \in G \mid ng = 0\}$. So $\mathrm{Tor}_0^{\mathbb{Z}}(\mathbb{Z}/n\mathbb{Z}, G) = G/nG$, and $\mathrm{Tor}_1^{\mathbb{Z}}(\mathbb{Z}/n\mathbb{Z}, G) = {}_nG$. (The ability of Tor to compute torsion subgroups is where Tor gets its name). In fact, since every abelian group G can be written as the direct limit of its finitely generated subgroups, each of which is either some $\mathbb{Z}^n$ or some $\mathbb{Z}/n\mathbb{Z}$, this approach can be used to show that $\mathrm{Tor}_i^{\mathbb{Z}}(G, H)$ vanishes for $i \geq 2$.

In contrast, Ext is named after its ability to compute extensions of R-modules. An **extension** of M by N is an exact sequence $0 \to N \to X \to M \to 0$, and such an extension **splits** if $X \cong M \oplus N$. If $\mathrm{Ext}_R^1(M, N)$ vanishes, then every extension of M by N splits; Ext^1 therefore tells us what obstruction prevents a given extension of M by N from splitting.

1.5.5 Singular Cohomology

Take a topological space X. Let $\mathrm{Hom}(\Delta^n, X)$ be the set of maps from the space

$$\Delta^n = \{(x_0, \ldots, x_n) \in \mathbb{R}^{n+1} \mid x_0 + \ldots + x_n = 1, x_0, \ldots, x_n \geq 0\}$$

known as the n-**simplex**, to X. The images of maps $\alpha, \beta, \ldots$ in this set are known as **singular** n-**simplices**, and denoted $\alpha|[v_0, \ldots, v_n]$, where each vertex v_i is the image of the vertex e_i of Δ^n. We write $\alpha|[v_0, \ldots, \widehat{v}_i, \ldots, v_n]$ for the singular $(n-1)$-simplex obtained by projecting the regular n-simplex onto the face opposing the ith vertex and sending that to X. Let $C_n(X)$ be the free abelian group on $\mathrm{Hom}(\Delta^n, X)$, whose elements are known as n-**chains**, and $\partial_n : C_n(X) \to C_{n-1}(X)$ the linear map defined on bases as

$$\partial_n(\alpha) = \sum_{i=0}^{n} (-1)^i \alpha|[v_0, \ldots, \widehat{v}_i, \ldots, v_n]$$

known as the **boundary operator**. For instance, ∂_1 sends a singular 1-simplex, or a path in X, to the 0-chain consisting of its end minus its beginning. It's easy to check that $\partial_{n-1}\partial_n = 0$, so $(C_\bullet, \partial_\bullet)$ forms a chain complex of abelian groups. Its homology groups are known as the **singular homology groups** of X.

A map $f : X \to Y$ generates a map $f_\sharp : C_n(X) \to C_n(Y)$ sending $\alpha : \Delta^n \to X$ to $f\alpha : \Delta^n \to Y$. $f_\sharp \partial_n^{(X)} = \partial_n^{(Y)} f_\sharp$, so this map is a chain map, and hence extends to a map $f_* : H_n(X) \to H_n(Y)$ evidencing H_n as a functor $\mathsf{Top} \to \mathsf{Ab}$; homotopic maps induce the same map, so H_n is in fact a map $\mathsf{hTop} \to \mathsf{Ab}$.

Given a group G, let $C^n(X)$ be the set of all homomorphisms $C_n(X) \to G$, known as n-**cochains**, which is itself an abelian group. We may precompose any morphism with ∂_{n+1} to obtain a homomorphism $\delta^{n+1} : C^n(X) \to C^{n+1}(X), \varphi \mapsto \alpha\partial_{n+1}$ known as the **coboundary operator**. Since $\delta^n\delta^{n-1}(\alpha)(\varphi) = \varphi\partial_{n-1}\partial_n = 0$, $(C^\bullet, \delta^\bullet)$ is a cochain complex, whose cohomology groups $H^n(X; G)$ are known as X's **singular cohomology groups with coefficients in** G. The failure of $H^n(X; G)$ to be equivalent to $\mathrm{Hom}_{\mathsf{Ab}}(H_n(X), G)$ is given by the **universal coefficient theorem for homology**, which states that the sequence

$$0 \longrightarrow \mathrm{Ext}(H_{n-1}(X), G) \longrightarrow H^n(X; G) \longrightarrow \mathrm{Hom}_{\mathsf{Ab}}(H_n(X), G)$$

is split exact; this is a purely algebraic fact, but evidences $H^n(-; G)$ as a functor $\mathsf{hTop} \to \mathsf{Ab}$ as well, and is often useful in computing cohomology groups in cases where Ext is easy to calculate. If you'd like to actually do some calculations, see the tools provided by [Hatcher, 2005], such as long exact sequences.

Eilenberg-MacLane Spaces Consider a contravariant functor F from the homotopy category Hotc of pointed, connected CW complexes to Set. If F maps wedge products to products and, for every u, v in some cover U, V of a CW complex X restricting to the same element of $F(U \cap V)$, there is at least one $x \in F(X)$ restricting to U and V, then the **Brown representability theorem** states that F is naturally isomorphic to the functor $[-, X_F]$ for some $X \in \mathsf{Hotc}$. By Yoneda, $F \mapsto X_F$ is a functor. When $F = H^n(-; G)$, both of these properties are indeed satisfied, and the representing space is the Eilenberg-MacLane space $K(G, n)$ whose defining property is that $\pi_i(K(G, n)) = G$ when $i = n$ and 0 otherwise. Note that such a space isn't unique up to isomorphism, but rather only up to weak equivalence. It is expedient to give a few examples: $K(\mathbb{Z}, 1) \simeq S^1$, $\mathbb{RP}^\infty \simeq K(\mathbb{Z}/2\mathbb{Z}, 1)$ and $\mathbb{CP}^\infty \simeq K(\mathbb{Z}, 2)$.

This representability plays a useful role in the theory of vector bundles: consider the functor

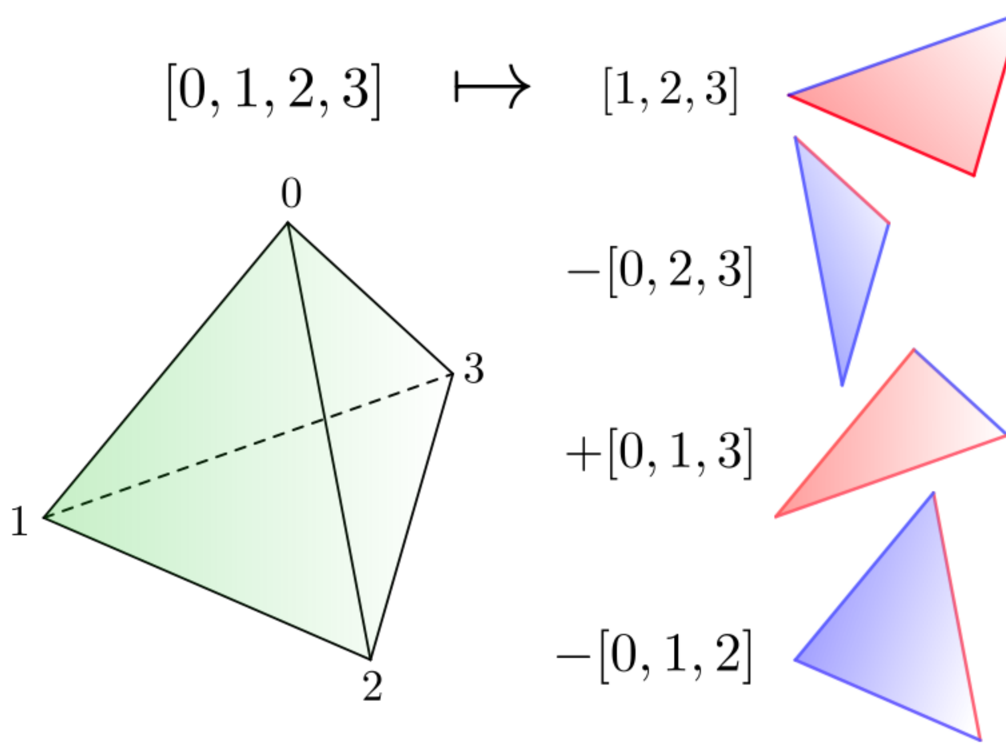
VB_k^n sending a space X to the set $\mathsf{VB}_k^n(X)$ of k-vector bundles of dimension n over some paracompact space X. It is known (see, e.g., [Weibel, 2013, Husemoller, 1975]) that VB_k^n is representable by the infinite Grassmanian Grass_n, which is the space of all n-dimensional subsets of k^∞. Namely, $\mathsf{VB}_k^n(X) \cong [X, \mathsf{Grass}_n]$. In the case $n = 1$, $\mathsf{Grass}_n = k\mathbb{P}^\infty$. Furthermore, we know that if Y is an Eilenberg-MacLane space $K(G, n)$, that $[X, Y] \cong H^n(X; G)$. For $k = \mathbb{R}, \mathbb{C}$, then, we have the following isomorphisms:

$$\mathsf{VB}_{\mathbb{R}}^1(X) \cong [X, \mathbb{RP}^\infty] \cong H^1(X; \mathbb{Z}/2\mathbb{Z})$$

$$\mathsf{VB}_{\mathbb{C}}^1(X) \cong [X, \mathbb{CP}^\infty] \cong H^2(X; \mathbb{Z})$$

So we may send a complex vector bundle $E \xrightarrow{\pi} X$ to an element of the second singular cohomology class of X; this element is known as the **first Chern class** $c_1(X)$.

Figure 1.3: Singular homology's boundary operator.



Chapter 2

Physics

2.1 Classical Mechanics

We'll sketch out the basics, using [Landau and Lifshitz, 2013] as our primary source for classical mechanics in its traditional, analytic sense; [Arnold, 2013] concerns the porting of this theory over to manifolds, which will later allow us to discuss general relativity and more abstract models of mechanics such as those encountered in synthetic differential geometry.

2.1.1 Equations of Motion

Suppose we have a system consisting of N particles in a three-dimensional space. Each particle has an x, y, and z component, and we require $3N$ **degrees of freedom** to express the state of this system at any given moment. Generalizing this, suppose the quantities $q_1, \ldots, q_s$ completely define a system: these q_i are **generalized coordinates**, and their time derivatives $\dot{q}_i$ are their **generalized derivatives**. Heuristically, if all coordinates $q = \{q_i\}$ and velocities $\dot{q}$ are given, the accelerations $\ddot{q}$ are uniquely determined.

The most general formulation of classical mechanics is given by the **principle of least action**, which states that (a) there is a function $L(q, \dot{q}, t)$ (known as the **Lagrangian** of a system's generalized coordinates at a given time (of which q and $\dot{q}$ are themselves functions), and that q and $\dot{q}$ are specified so as to extremize the **action**

$$S = \int_{t_1}^{t_2} L(q, \dot{q}, t)\, dt$$

To play around with this, we'll need some concepts from the calculus of variations. For a functional $F[f]$, the **functional derivative** is given by

$$\frac{\delta F}{\delta f} = \lim_{\epsilon \to 0} \frac{F[f + \epsilon\eta] - F[f]}{\epsilon}$$

For instance, the functional derivative of the action is given by

$$\frac{\delta S}{\delta q(t_0)} = \lim_{\epsilon \to 0} \frac{1}{\epsilon} \int_{t_1}^{t_2} L(q + \epsilon\eta, \dot{q} + \epsilon\dot{\eta}, t) - L(q, \dot{q}, t)\, dt$$

$$= \lim_{\epsilon \to 0} \frac{1}{\epsilon} \int_{t_1}^{t_2} \epsilon\eta \frac{\partial}{\partial q} L(q, \dot{q}, t) + \epsilon\dot{\eta} \frac{\partial}{\partial \dot{q}} L(q, \dot{q}, t) + O(\epsilon^2)\, dt$$

If we set boundary conditions on what $q(t_1), \dot{q}(t_1), q(t_2)$, and $\dot{q}(t_2)$ are, we must also set $\eta(t_1) = \eta(t_2) = 0$, so as not to alter these conditions. Then, applying integration by parts, we get

$$\frac{\delta S}{\delta q} = \int_{t_1}^{t_2} \eta \left[\frac{\partial L}{\partial q} - \frac{d}{dt} \left(\frac{\partial L}{\partial \dot{q}} \right) \right] dt$$

Since the principle of least action implies that q is selected so as to extremize the action, we must be at a peak (or trough) of the action, and $\delta S/\delta q$ must be zero, regardless of what η is; the expression in the brackets must therefore be zero. Therefore, any q obeying the principle of least action must also obey the equation

$$\frac{d}{dt} \left(\frac{\partial L}{\partial \dot{q}} \right) - \frac{\partial L}{\partial q} = 0$$

which is known as the **Euler-Lagrange equation**.

2.1.2 Lagrangians and Hamiltonians

In a vacuum, we can assume by symmetry that we're in a reference frame where space is homogeneous and isotropic (the same regardless of orientation); such a reference frame is called an **inertial frame**. In an inertial frame, the Lagrangian can't refer explicitly to the radius vector, the time, or the direction of the velocity, implying that the Lagrangian for a *free* particle is solely a function of $\vec{v} \cdot \vec{v} = v^2$. Plugging this finding into the Euler-Lagrange equations, we see that $\frac{d}{dt}\frac{\partial L}{\partial \vec{v}} = 0$, so $\partial L/\partial \vec{v}$ is constant; since this is a function of $\vec{v}$ only, it follows that $\vec{v}$ is constant, and therefore that free motion in an inertial frame has a constant velocity: this is known as the **law of inertia**. Heuristically, two inertial frames, perhaps moving at different velocities, are equivalent in all mechanical respects: this is known as **Galileo's relativity principle**.

For a *system* of particles which interact with each other, but which are isolated from exterior forces (a **closed** system), we subtract from the kinetic energy term $T = \sum \frac{1}{2} m_i v_i^2$ a potential energy term U that depends on the locations r_i of the particles, giving us

$$L = \sum \frac{1}{2} m_i v_i^2 - U(r_1, \ldots, r_n)$$

Solving the Euler-Lagrange equations gives us

$$m_i \frac{dv_i}{dt} = -\frac{\partial U}{\partial r_i}$$

Such equations of motion are called **Newton's equations**, and the term on the LHS, $m\dot{v}_i$, is known as the **force**. Note that, since the equations of motion depend entirely on derivatives of the Lagrangian, the potential is effectively only defined up to a constant; we generally choose this constant such that the potential goes to zero as the particles get infinitely far away from one another.

Given a Lagrangian L, we may define the **conjugate momentum** to a coordinate q_i to be $p_i := \frac{\partial L}{\partial \dot{q}_i}$. For instance, when $L = \frac{1}{2} m\dot{q}^2 - U(q)$, $p = m\dot{q}$. If the kinetic energy T is a function of $\dot{q}$ alone and the potential energy a function of q alone, then $\sum_i p_i \dot{q}_i = 2T$, and the quantity $H = \sum_i p_i \dot{q}_i - L$ yields $T + U$, the total energy of the system. This quantity, which is in general conserved, is known as the **Hamiltonian**. While we express the Lagrangian as a function of q, $\dot{q}$, and t, we conventionally express the Hamiltonian as a function of p, q, and t. By matching different expressions for the total differential dH of the Hamiltonian,

$$dH = \frac{\partial H}{\partial p} dp + \frac{\partial H}{\partial q} dq + \frac{\partial H}{\partial t} dt = d(p\dot{q} - L)$$

we can obtain **Hamilton's equations**,

$$\frac{dp}{dt} = -\frac{\partial H}{\partial q} \qquad \frac{dq}{dt} = \frac{\partial H}{\partial p}$$

2.2 Quantum Mechanics

The section on functional analysis is based on [Haase, 2014, Rudin, 1973], and the natural segue into quantum probability theory relies on *many* sources, including [Takhtadzhian, 2008, Meyer, 2006, Holevo, 2003, Rédei and Summers, 2007], each of which tells a small part of a large story. In addition to the sources used in our discussion of functional analysis and quantum probability theory, we use [Sakurai et al., 2014] as a source for quantum mechanics.

2.2.1 Banach Spaces

In the theory of finite dimensional vector spaces, everything goes right. More specifically, every such space V satisfies the following:

- The double dual of V, V^{**}, is canonically isomorphic to V itself.
- All norms on V are equivalent, and induce the same topology.
- With this topology, any linear map from V is continuous.
- An endomorphism on V is injective iff it is surjective.
- The unit ball in V (under any norm) is compact.

The theory of infinite dimensional vector spaces, however, is far more dangerous: *none* of these statements hold, nor can they be easily fixed. In such an infinite dimensional vector space W, the following properties are satisfied:

- As cardinals, $\dim W^{**} > \dim W^{*} > \dim W$, these inequalities being strict.
- W generally has many different topologies of interest.
- Linear maps from W aren't necessarily continuous.
- There are non-surjective injections $W \to W$.
- The unit ball is never compact.

In nature, infinite dimensional vector spaces tend to occur as spaces of functions, hence the name functional analysis. There is a hierarchy of classes of infinite-dimensional vector spaces, with each level of the hierarchy introducing a new structure, or a new condition to be fulfilled by a structure provided at the lower tier. At the bottom rung are simply k-vector spaces, where we assume k is either $\mathbb{R}$ or $\mathbb{C}$.

Normed Spaces The first thing we can do with a vector space V is put a **norm** on it. This is a function $||\cdot|| : V \to [0, \infty)$ which satisfies the following properties:

1. *Homogeneity*: $||cv|| = |c|\,||v||$, for $c \in k$.
2. *Triangle inequality*: $||v + w|| \leq ||v|| + ||w||$
3. *Definiteness*: $||v|| = 0$ iff $v = \vec{0}$.

Equipped with such a norm, V becomes a **normed space**. This norm induces a topology on V whose basis consists of open sets

$$B_r(v) = \{w \in V \mid ||v - w|| < r\}$$

for all $r \in [0, \infty)$ and all $v \in V$. Given two normed vector spaces V, W, we may ask which linear maps $A : V \to W$, also known as **operators**, preserve the norm, in the sense that $||Av||_W \leq c||v||_V$ for all $v \in V$, for some fixed $c \geq 0$. Such an operator is known as a **bounded operator**. It's well known that an operator is bounded if and only if it is continuous: in this sense, the structure on V that a norm provides is equivalent to the structure that the topology induced by the norm itself provides. The smallest such c satisfying $||Av||_W \leq c||v||_V$ is given by $\sup_{||f||_V \leq 1} ||Av||_W$, and is known as the **operator norm** $||A||$. With this norm, the space $\mathcal{B}(V, W)$ of bounded operators $V \to W$, with its natural vector space structure, becomes a normed space itself.

An important family of normed spaces can be constructed as follows: take a measure space $(\Omega, \mathcal{F}, \mu)$ and consider the vector space of measurable functions $\Omega \to k$, $k \in \{\mathbb{R}, \mathbb{C}\}$. Define the p-norm of a function f to be

$$||f||_p := \left(\int_\Omega |f|^p \right)^{1/p}$$

for $1 \leq p < \infty$. The space of functions f for which $||f||_p < \infty$ is a vector space $\mathcal{L}^p(\Omega, \mu)$, but it isn't a normed space, since functions which are 0 almost everywhere have norm zero. The set of all such functions forms a linear subspace of $\mathcal{L}^p(\Omega, \mu)$, though, and quotienting out by it yields a proper normed space $L^p(\Omega, \mu)$, known as an L^p **space**, whose elements aren't strictly measurable functions $\Omega \to k$, but equivalence classes of measurable functions which differ by sets of measure zero [Rudin, 1973]. As $p \to \infty$, $||f||_p$ converges to the essential supremum of $|f|$, since raising $|f|$ to a power $p > 1$ makes a greater change when $|f|$ is large, with the size of p exaggerating this change. This allows us to define $||f||_\infty$ to be the essential supremum of $|f|$ over Ω, and thereby obtain the space $L^\infty(\Omega, \mu)$.

In the special case when μ is the counting measure, which sends a finite $S \subseteq \Omega$ to $|S|$ and an infinite S to ∞, the set $L^p(\mathbb{N}, \mu)$ is known as the ℓ^p **space**; its elements are sequences $\{c_0, c_1, \ldots\}$ and the norm of a sequence $c = \{c_n\}_{n\in\mathbb{N}}$ is just $\left(\sum_{n=0}^{\infty} |c_n|^p\right)^{1/p}$ when $1 \leq p < \infty$, and $\sup c$ when $p = \infty$.

Inner Product Spaces Given a (k-)vector space V, an **inner product** on V is a mapping $\langle\cdot,\cdot\rangle : V \times V \to k$ which is

1. *Conjugate-symmetric*: $\langle v, w\rangle = \overline{\langle w, v\rangle}$
2. *Positive definite*: $\langle v, v\rangle \geq 0$, and $\langle v, v\rangle = 0$ iff $v = \vec{0}$.
3. *Sesquilinear*: Linear in the first argument, and conjugate linear in the second argument.

A vector space equipped with an inner product is known as a **inner product space**. The norm **induced** by the inner product $\langle\cdot,\cdot\rangle$ is given by $||v|| = \sqrt{\langle v, v\rangle}$; it is straightforward to check that this is indeed a norm, and therefore that inner product spaces are a subset of normed spaces. This norm satisfies the **polarization identity**

$$||f + g||^2 - ||f - g||^2 = 4\,\mathrm{Re}\,(\langle f, g\rangle)$$

as well as the **parallelogram law**

$$||f + g||^2 + ||f - g||^2 = 2\left(||f||^2 + ||g||^2\right)$$

In fact, an arbitrary norm on a vector space is induced by an inner product if and only if it satisfies the parallelogram law [Haase, 2014].

Example. As in the finite dimensional case, two vectors v, w on an inner product space are **orthogonal** if $\langle v, w\rangle = 0$. For instance, consider the k-vector space of continuous functions $[0,1] \to k = \mathbb{C}$, with inner product $\langle f, g\rangle = \int_0^1 f\overline{g}\,dx$. For $f_n = e^{2\pi i n x}$, $n \in \mathbb{Z}$, we have

$$\langle f_m, f_n\rangle = \int_0^1 e^{2\pi i(m-n)x}\,dx$$

which when $m = n$ is 1 and when $m \neq n$ is $\frac{1}{2\pi i(m-n)}\left(e^{2\pi i(m-n)} - 1\right) = 0$. So, in fact, $\{f_n\}$ is not only a set of pairwise orthogonal vectors, but an orthonormal set. It is not an orthonormal basis, since an arbitrary $f \in C[0,1]$ cannot be expressed as a finite linear combination of the f_n, but (since this is just a Fourier transform) we know that we can specify coefficients $c_n = \langle f, f_n\rangle$

such that the sum $\sum_{i\in\mathbb{Z}} c_n f_n$ *converges* to f under the norm induced by the inner product. Such a "basis" in which every element of the vector space can be expressed as the limit of a countable sum is known as a **Schauder basis**.

Banach Spaces A **Banach space** is a normed space $(V, ||\cdot||)$ which is complete with respect to its norm, having for each Cauchy sequence $\{v_n\}_{n\in\mathbb{N}}$ a vector v such that $\lim_{n\to\infty} ||v_n - v|| = 0$. This completeness condition ensures that V has "no holes", so that all sequences that should converge (Cauchy sequences) do converge. An incomplete normed space V can be made complete in the following manner: take the set of all Cauchy sequences $\{v_n\}_{n\in\mathbb{N}}$ in V, and, given $v = \{v_n\}, w = \{w_n\}$, define a "metric" on Cauchy sequences by $D(v, w) = \lim_{n\to\infty} ||v_n - w_n||$. If V isn't already complete, this isn't an actual metric: let v and w be the same sequence except at the first element to get $D(v, w) = 0$ with $v \neq w$. To fix this, we declare v and w to be equivalent to be equal if $\lim_{n\to\infty} ||v_n - w_n|| = 0$. This is an equivalence relation by the triangle inequality, and quotienting the set of Cauchy sequences out by it makes D a proper metric on what is now a complete space, which we denote by $\widehat{V}$. Of course, if V is already complete, we can identify Cauchy sequences with the vector they converge to, so $\widehat{V}$ can be identified with V. If not, then V naturally *embeds* into V, this embedding being given by sending a $v \in V$ to the equivalence class of the Cauchy sequence $(v, v, v, \ldots)$. In this way, every normed vector space V naturally embeds into the Banach space $\widehat{V}$ known as the **completion** of V.

$L^p(\Omega, \mu)$ is always a Banach space, a fact often known as the Riesz-Fischer theorem. For V an arbitrary Banach space, the set $\mathcal{B}(V) := \mathcal{B}(V, V)$ of bounded operators on V is, when equipped with the operator norm, a Banach space. This space can be equipped with an associative multiplication given by composition. A Banach space equipped with an associative algebra structure is known as a **Banach algebra**; we also require that $||AB|| \leq ||A||\,||B||$, but this holds trivially for the Banach algebra $\mathcal{B}(V)$. In addition, the normed vector space $V^* := \mathcal{B}(V, k)$ is also a Banach space, known as the **dual** of V.

Example. Banach spaces often appear in the study of differential equations and dynamical systems, since they allow us to use linear algebra in sufficiently nice topological spaces. For instance, let X be a Banach space, and f a continuous map $X \to X$. f is called a **contracting** map if there's a $\lambda < 1$ s.t. $d(f(x), f(y)) \leq \lambda d(x, y)$, where $d(x, y) := ||x - y||$. f and its positive iterates $f^2, f^3, \ldots$ form what is known as a discrete-time topological dynamical system. Of course, $d(f^n(x), f^n(y)) \to 0$ as $n \to \infty$; every $\{f^n(x)\}_{x\in\mathbb{N}}$ is, in fact, a Cauchy sequence, so, given that X is complete by virtue of being a Banach space, there's a unique limit p to which all points converge, known as the **fixed point**.

We verify that it's a Cauchy sequence as follows: for $n \geq m$,

$$d(f^m(x), f^n(x)) \leq d(f^m(x), f^{m+1}(x)) + \ldots + d(f^{n-1}(x), f^n(x))$$

$$\leq (\lambda^m + \lambda^{m+1} + \ldots + \lambda^{n-1}) d(f(x), x) \leq \lambda^m (1 + \lambda + \lambda^2 + \ldots) d(f(x), x) = \frac{\lambda^m}{1-\lambda} d(f(x), x) \overset{m \to \infty}{\longrightarrow} 0$$

In particular, as $m \to \infty$, $d(p, f^n(x)) \to 0$, implying that $f(p) = p$. As $n \to \infty$, we get $d(f^m(x), p) \leq \frac{\lambda^n}{1-\lambda} d(f(x), x)$. We say that two sequences of points $\{x_n\}_{n \in \mathbb{N}}$ and $\{y_n\}_{n \in \mathbb{N}}$ **converge exponentially** to each other if $d(x_n, y_n) < c\lambda^n$ for some $c > 0, \lambda < 1$. In the case that $\{y_n\}_{n \in \mathbb{N}}$ is a constant sequence $y_n = y$, we just say that $\{x_n\}_{n \in \mathbb{N}}$ converges exponentially to y. We therefore have the **Contraction Mapping Principle**: under the action of iterates of a contracting map f on a complete metric space X, all points converge with exponential speed to the unique fixed point of f.

2.2.2 Hilbert Spaces

Hilbert spaces combine the theories of inner product and Banach spaces. In particular, a **Hilbert space** is an inner product space $\mathcal{H}$ which is Banach with respect to the norm induced by its inner product or, equivalently, a Banach space whose norm satisfies the parallelogram law. Among the L^p spaces, this is only satisfied for $p = 2$, in which case the norm on $L^2(\Omega, \mu)$ is induced by the inner product $\langle f, g \rangle = \int_\Omega f\overline{g}\, d\mu$. When a Hilbert space $\mathcal{H}$ has an orthonormal Schauder (countable) basis, it's called **separable**. This is essentially a size restriction on $\mathcal{H}$; every Hilbert space has a possibly uncountable orthonormal basis (assuming the AC), but we will assume that our Hilbert spaces are separable to avoid size issues. We'll also assume that $k = \mathbb{C}$ unless otherwise specified.

C*-Algebras For the purposes of quantum mechanics, we're not interested in Hilbert spaces per se, but in algebras of operators on Hilbert spaces. A Banach algebra of the form $\mathcal{B}(\mathcal{H})$ has a natural involution operation given by taking adjoints: the **adjoint** of an operator $A \in \mathcal{B}(\mathcal{H})$ is an operator $A^\dagger$ satisfying $\langle Av, w \rangle = \langle v, A^\dagger w \rangle$ for all $v, w \in \mathcal{H}$. (In the real or complex finite dimensional case, this simply corresponds to taking the transpose or conjugate transpose, respectively). The fact that adjoints always exist is a consequence of the **Riesz representation theorem**, which states that any $\varphi \in \mathcal{H}^*$ can be represented as $\langle -, v \rangle$ for some $v \in \mathcal{H}$; if we set $\varphi = \langle A-, w \rangle$ for a fixed $w \in \mathcal{H}$, this theorem gives us a v such that $\varphi = \langle -, v \rangle$, and therefore an identification $\langle Ax, w\rangle = \langle x, v\rangle$. This v depends linearly and continuously on w, and hence can be represented as $A^\dagger w$, giving us the adjoint $A^\dagger$. We can check that this really does define an involution on $\mathcal{B}(\mathcal{H})$: $\langle A^{\dagger\dagger}v, w\rangle = \overline{\langle w, A^{\dagger\dagger}v\rangle} = \overline{\langle A^\dagger w, v\rangle} = \langle v, A^\dagger w\rangle = \langle Av, w\rangle$, so $A^{\dagger\dagger} = A$. Furthermore, $\langle ABv, w\rangle = \langle Bv, A^\dagger w\rangle = \langle v, B^\dagger A^\dagger w\rangle$, so $(AB)^\dagger = B^\dagger A^\dagger$. It can also be verified that $||A^\dagger A|| = ||A^\dagger||\,||A||$, and this property, along with the previous two, makes $\mathcal{B}(\mathcal{H})$ a **C*-algebra** when equipped with the involution $(\cdot)^\dagger$. In general, a C*-algebra is a Banach algebra with an involution satisfying $(AB)^\dagger = B^\dagger A^\dagger$ and $||A^\dagger A|| = ||A^\dagger||\,||A||$; the **Gelfand-Naimark theorem** allows us to identify any C*-algebra as a subalgebra of some $\mathcal{B}(\mathcal{H})$.

Observables and Projections Three especially important subsets of $\mathcal{B}(\mathcal{H})$ must be distinguished: first are the **self-adjoint** operators, which satisfy $A^\dagger = A$. (Physicists often call a self-adjoint operator a **Hermitian operator**, or an **observable**). For $\mathcal{H} = \mathbb{R}^n$, these are the symmetric matrices $A = A^T$, and for $\mathcal{H} = \mathbb{C}^n$, these are the conjugate symmetric/Hermitian matrices $A = A^H$. The eigenvalues of a self-adjoint operator, i.e. those $\lambda \in \mathbb{C}$ such that $Av = \lambda v$ for some v known as λ's eigenvector, can easily be seen to be real even if $\mathcal{H}$ is complex:

$$\lambda||v||^2 = \langle \lambda v, v\rangle = \langle Av, v\rangle = \langle v, Av\rangle = \langle v, \lambda v\rangle = \overline{\lambda}||v||^2$$

In addition, the eigenvectors v_1, v_2 of a self-adjoint A are orthogonal given that they have different eigenvalues $\lambda_1 \neq \lambda_2$:

$$\lambda_1\langle v_1, v_2\rangle = \langle Av_1, v_2\rangle = \langle v_1, Av_2\rangle = \lambda_2\langle v_1, v_2\rangle$$

so $(\lambda_1 - \lambda_2)\langle v_1, v_2\rangle = 0$, implying that $\langle v_1, v_2\rangle = 0$. We denote the set of all self-adjoint operators as $\mathcal{O}(\mathcal{H}) \subset \mathcal{B}(\mathcal{H})$; it isn't closed as an algebra, since $(AB)^\dagger = B^\dagger A^\dagger = BA$ isn't necessarily equal to AB, but it is closed under the **commutator** $i[A, B] = i(AB - BA)$, with $(i[A, B])^\dagger = (-i)(B^\dagger A^\dagger - A^\dagger B^\dagger) = i[A, B]$.

The second subset of $\mathcal{B}(\mathcal{H})$ consists of the **positive operators**, for which $\langle v, Av\rangle$ is real and non-negative for all $v \in \mathcal{H}$. Obviously, $\langle v, Av\rangle = \overline{\langle Av, v\rangle}$, suggesting that positive operators are self-adjoint. Given two self-adjoint operators A_1, A_2, we write $A_1 \geq A_2$ if $A_1 - A_2$ is positive; this forms a partial order on $\mathcal{O}(\mathcal{H})$.

Finally, there are the **projection operators**, those operators $P \in \mathcal{B}(\mathcal{H})$ which satisfy $P^2 = P$. These operators are necessarily self-adjoint and positive, satisfying $I \geq P \geq 0$, where I is the identity operator $Iv = v$ and 0 is the zero operator $0v = \vec{0}$. In fact, projections can be characterized as orthogonal projections onto some linear subspace of $\mathcal{H}$. For instance, every $v \in \mathcal{H}$ induces a projection operator $P_v w = v\langle w, v\rangle / \langle v, v\rangle$. Given an operator A, define its **range** to be $R(A) = A\mathcal{H}$ and its **null space** to be $N(A) = \{v \in \mathcal{H} \mid Av = \vec{0}\}$. Given a family $\{P_\alpha\}$ of projections, we can then define the **meet** $\wedge_\alpha P_\alpha$ to be the smallest closed subspace of $\mathcal{H}$ containing $\bigcap_\alpha R(P_\alpha)$, and the **join** $\vee_\alpha P_\alpha$ to be the smallest closed subspace containing $\bigcup_\alpha R(P_\alpha)$. Denoting by $\mathcal{P}(\mathcal{H})$ the subset of $\mathcal{O}(\mathcal{H})$ containing the projections, these are the inf and sup operations with respect to the partial order on $\mathcal{P}(\mathcal{H})$ inherited from $\mathcal{O}(\mathcal{H})$.

Diagonalizability Since we've assumed $\mathcal{H}$ to be separable, we can fix a countable orthonormal basis $(e_1, e_2, \ldots)$, and represent any $v \in \mathcal{H}$ as the converging sum $\sum_{i=1}^{\infty} v_i e_i$, where $v_i = \langle v_i, e_i\rangle$. This allows us to write $\langle v, w\rangle = \langle \sum_i v_i e_i, \sum_j w_j e_j\rangle = \sum_i v_i w_i$, and to express an operator A in terms of a "matrix" $A_{ij} = \langle e_i, Ae_j\rangle$. With this notation, the usual formulas for finite-dimensional vector spaces can be extended: $(Av)_i = \sum_{ij} A_{ij} v_j$, $(AB)_{ij} = \sum_k A_{ik} B_{kj}$, and so on. In particular, A is **diagonal** when $A_{ij} = 0$ for $i \neq j$, and **diagonalizable** when there is an orthonormal countable basis in which A is diagonal. Two self-adjoint operators A, B are **mutually diagonalizable** when there is a single basis in which they're both diagonal; this happens when $[A, B] = 0$.

The **trace** of an operator A is given by $\mathrm{Tr}A = \sum_i A_{ii} = \sum_i \langle e_i, Ae_i\rangle$; this value is independent of the basis chosen, being a property of the operator A itself. This sum may not always converge, but when it does, A is said to be of **trace class**. For instance, we can take the trace of a projection operator of the form P_v:

$$\mathrm{Tr}P_v = \sum_i \langle e_i, P_v e_i\rangle = \sum_i \left\langle e_i, v\frac{\langle e_i, v\rangle}{\langle v, v\rangle}\right\rangle = \frac{1}{\langle v, v\rangle}\sum_i |\langle e_i, v\rangle|^2 = \frac{1}{\sum_i |v_i|^2}\sum_i |v_i|^2 = 1$$

On the set of trace class operators in $\mathcal{B}(\mathcal{H})$, denoted $\mathcal{T}(\mathcal{H})$, the trace generates a norm: take an operator A and define the self-adjoint operator $A^\dagger A$, which has real eigenvalues $\sigma_1, \sigma_2, \ldots$. The **trace norm** of A, denoted variously as $||A||_*$ or $\mathrm{Tr}\sqrt{A^\dagger A}$, is then $\sum_i \sqrt{\sigma_i}$. With this norm, $\mathcal{T}(\mathcal{H})$ is a Banach space; in fact, its dual can be identified with $\mathcal{B}(\mathcal{H})$ itself. We say that $\mathcal{T}(\mathcal{H})$ is the **predual** of $\mathcal{B}(\mathcal{H})$, and write $\mathcal{T}(\mathcal{H}) = \mathcal{B}(\mathcal{H})_*$.

Any operator $\rho \in \mathcal{T}(\mathcal{H})$ with trace 1 is said to be a **state**; the projection operators P_v are special among these, and are called **pure states**. It is a consequence of the Hilbert-Schmidt theorem that an arbitrary state A can be decomposed into a sum of finitely many pure states as $A = \sum_{i=1}^{N} c_i P_{v_i}$, where the $\{v_i\}$ are orthonormal and $\sum_{i=1}^{N} c_i = 1$.

Bra-Ket Notation Let $\mathcal{H}$ be a complex Hilbert space. Dirac's **bra-ket notation** prescribes that we write an element ψ of $\mathcal{H}$ as $|\psi\rangle$, calling them **kets**, and elements ϕ of $\mathcal{H}^*$ as $\langle\phi| := \langle\phi, -\rangle$, calling them **bras**. Note that physicists tend to write the inner product as being conjugate linear in the *first* argument, rather than the second, which is why we've used $\langle\phi, -\rangle$ instead of the $\langle -, \phi\rangle$ above. We'll continue to use this convention for this section. The inner product of two kets $|\phi\rangle$, $|\psi\rangle$ is written as $\langle\phi|\psi\rangle$. The correspondence between $\mathcal{H}$ and $\mathcal{H}^*$ given by the Riesz representation theorem sends a $c|\psi\rangle$ to $\bar{c}\langle\phi|$, and a term of the form $A|\psi\rangle$ to $\langle\phi|A^\dagger$, where we define the action of an operator A on a bra $\langle\phi|$ as $(\langle\phi|A)|\psi\rangle = \langle\phi|(A|\psi\rangle)$. We generally require our bras and kets to be **normalized**, requiring that $\langle\psi|\psi\rangle = 1$; an arbitrary element of $\mathcal{H}$ can be normalized by dividing it by its norm. In contrast to the inner product, bra-ket notation allows us to express the **outer product** of a bra $\langle\phi|$ with a ket $|p\rangle si$, which is the operator $|\psi\rangle\langle\phi|$ that acts on a $|\xi\rangle$ as $(|\psi\rangle\langle\phi|)(|\xi\rangle) = |\psi\rangle\langle\phi|\xi\rangle$. We may also speak of the outer product of two bras $\langle\phi_1|\langle\phi_2|$ or kets $|\psi_1\rangle|\psi_2\rangle$, which is just defined to be the tensor product in $\mathcal{H} \otimes \mathcal{H}$. We'll rewrite a few of our above formulas in this notation: $A_{ij} = \langle e_i|A|e_j\rangle$, $|v\rangle = \sum_i |e_i\rangle\langle e_i|v\rangle$, $\mathrm{Tr} A = \sum_i \langle e_i|A|e_i\rangle$, and $P_v = |v\rangle\langle v|$ (note that $|v\rangle$ is assumed to be normalized). Note that since $v = \sum_i |e_i\rangle\langle e_i|v\rangle = (\sum_i |e_i\rangle\langle e_i|)v$, we can write $\sum_i |e_i\rangle\langle e_i| = I$. This is known as a **resolution of the unity**, and can be inserted anywhere: for instance, $\langle v|w\rangle = \sum_i \langle v|e_i\rangle\langle e_i|w\rangle$. Commonly, $\mathcal{H} = L^2(M, \mathbb{C})$, where M is a Riemannian manifold with metric g and the inner product is $\langle\psi|\phi\rangle = \int_M \overline{\psi}(x)\phi(x)\,\omega$, where ω is the volume form on M.

Resolution of the identity works when we replace the $\{e_i\}$ with an arbitrary orthonormal basis, for instance the eigen*kets* of a self-adjoint operator A, when they form a complete set. In physical systems, we often use the case of $A = H$, where H is an operator representing the Hamiltonian, whose eigenvalues are thought of as the allowed energy levels of the system. The eigenvalue equation $H|\psi\rangle = E|\psi\rangle$ is known as the **time-independent Schrödinger equation**. The Hamiltonian H generates a one-parameter group of operators $U_t := e^{-\frac{i}{\hbar}tH}$, where $e^A := I + A + \frac{1}{2}A^2 + \dots$ satisfies the usual properties of the exponential, and $\hbar$ is a positive constant. We have $U_t U_s = e^{-\frac{i}{\hbar}(t+s)H} = U_{t+s}$ and $U_t^\dagger = e^{\frac{i}{\hbar}tH} = U_{-t}$, so that $U_t U_t^\dagger = U_t^\dagger U_t = I$; operators whose adjoints are their inverses are known as **unitary**, and the group $\{U_t\}_{t\in\mathbb{R}}$ is known as the **unitary group** generated by H. For instance, the unitary group generated by $i\hbar\frac{d}{dx}$ on $L^2(\mathbb{R})$ is given by

$$U_x f(x') = e^{-\frac{i}{\hbar}xi\hbar\frac{d}{dt}}f(x') = e^{x\frac{d}{dx}}f(x') = \left(f + xf' + \frac{x^2}{2}f'' + \dots\right)(x') = f(x' - x)$$

where we've identified the penultimate step as a Taylor expansion. The operator $i\hbar\frac{d}{dx}$ is the quantum analog of momentum, and we correspondingly say that momentum is the generator of translation.

2.2.3 Measurements

As per Dirac, "a measurement always causes the system to jump into an eigenstate of the dynamical variable that is being measured." To illustrate, say an operator A with some corresponding physical variable (e.g., position) has eigenkets $\{|a_i\rangle\}$, where a_i refers to an actual value of the variable. A normalized ket $|\alpha\rangle$ is represented in this basis as $|\alpha\rangle = \sum c_i |a_i\rangle$, where $c_i = \langle a_i|\alpha\rangle$. When we make a measurement of the variable corresponding to A, $|\alpha\rangle$ jumps into *one* of the $|a_i\rangle$, and the probability of a specific ket $|a_i\rangle$ is $|\langle a_i|\alpha\rangle|^2$. Since $|\alpha\rangle$ is normalized, we know that $\sum_i |\langle a_i|\alpha\rangle|^2 = 1$, so the probabilities sum to 1. The **expectation value** of A in the state $|\alpha\rangle$, denoted as $\langle A\rangle_\alpha$ (α is often suppressed, especially when it is some ground state), can be calculated as

$$\langle A\rangle_\alpha = \sum_{a_i} a_i P(a_i) = \sum_{a_i} a_i |\langle a_i|\alpha\rangle|^2 = \sum_{a_i}\sum_{a_j} \langle\alpha|a_j\rangle\langle a_j|A|a_i\rangle\langle a_i|\alpha\rangle = \langle\alpha|A|\alpha\rangle$$

Define the **commutator** of two observables A, B as

$$[A, B] = AB - BA$$

and the **anticommutator** of A and B as

$$\{A, B\} = AB + BA$$

The observables A, B are said to be **compatible** when $[A, B] = 0$, and **incompatible** otherwise.

Suppose A's eigenvalues are nondegenerate and generate a basis, in which the matrix representation of A is diagonal. If B is compatible with A, B is diagonal in A's basis as well. Why? $\langle a_i|[A, B]|a_j\rangle = \langle a_i|0|a_j\rangle = 0 = (a_i - a_j)\langle a_i|B|a_j\rangle$, which by nondegeneracy implies that $\langle a_i|B|a_j\rangle = 0$ unless $i = j$. So, really, the eigenkets of A are the eigenkets of B, though they may have different eigenvalues: they are said to be **simultaneous eigenkets**, and are sometimes denoted by $|a_i, b_i\rangle$. We may also use a **collective index**, $|K_i\rangle = |a_i, b_i\rangle$. Due to the simultaneity of the eigenkets, measurements of A do not interfere with measurements of B, and vice-versa; this can be extended to larger sets of pairwise compatible operators. Of course, if A and B are incompatible, then simultaneous eigenkets generally do *not* exist and successive measurements *do* interfere with each other.

To represent our uncertainty in the result of a measurement, we adopt the statistical notion of variance, calling it **dispersion**: defining $\Delta A = A - \langle A \rangle$, the dispersion, also known as the variance or mean square deviation, is given by the expectation of $(\Delta A)^2$, or

$$\langle (\Delta A)^2 \rangle = \langle (A^2 - 2A\langle A \rangle + \langle A \rangle^2) \rangle = \langle A^2 \rangle - \langle A \rangle^2$$

It is more convenient to denote this by σ_A^2.

For observables A, ΔA is also Hermitian, since the expectation is a real number (implicitly multiplied by the identity) and thus equal to its own adjoint. We can use the fact that an operator can be defined by its action on all possible kets to lift certain identities on vectors in Hilbert spaces to corresponding identities on their operators: for instance, if we assume that operators A and B are Hermitian, we can lift the Cauchy-Schwarz identity $\langle \alpha|\alpha \rangle \langle \beta|\beta \rangle \geq |\langle \alpha|\beta \rangle|^2$ to a corresponding identity $\langle A^2 \rangle \langle B^2 \rangle \geq |\langle AB \rangle|^2$. Since the dispersion operators of observables are Hermitian, this implies that $\sigma_A^2 \sigma_B^2 \geq |\langle \Delta A \Delta B \rangle|^2$ for any observables A, B. In fact, expanding this yields:

$$\sigma_A^2 \sigma_B^2 \geq |\langle \Delta A \Delta B \rangle|^2 = \left| \frac{1}{2}\langle [A, B] \rangle + \frac{1}{2}\langle \{\Delta A, \Delta B\rangle \right|^2 = \frac{1}{4}|\langle [A, B] \rangle|^2 + \frac{1}{4}|\langle \{\Delta A, \Delta B\} \rangle|^2$$

giving us the important inequality

$$\sigma_A \sigma_B \geq \frac{1}{2}|\langle [A, B] \rangle|$$

(Note that "σ_A" is notational trickery, since σ_A^2 itself is not a square, but the expectation value of a square; however, as σ_A^2 corresponds to variance, σ_A corresponds to the standard deviation of A).

2.2.4 Position, Momentum, and Time

We've been dealing with finite-dimensional spaces so far, where spectra are finite and everything converges. Now we'll move to infinite-dimensional spaces, replacing Kronecker deltas by Dirac deltas and sums by integrals: for instance, $\langle a_i | a_j \rangle = \delta_{ij}$ becomes $\langle a_i | a_j \rangle = \delta(i - j)$, and $\sum_i |a_i\rangle\langle a_i| = 1$ becomes $\int |a_i\rangle\langle a_i| \, di = 1$.

Consider a position operator x on one dimension, whose eigenkets $x|x_i\rangle = x_i|x_i\rangle$ form a complete set. An arbitrary physical state $|\alpha\rangle$ can be expanded as $|\alpha\rangle = \int_{-\infty}^{\infty} |x_i\rangle\langle x_i|\alpha\rangle\, dx_i$. Suppose we centered a detector of length ℓ at position x_0: when the detector registers a particle, the state changes:

$$|\alpha\rangle = \int_{-\infty}^{\infty} |x_i\rangle\langle x_i|\alpha\rangle\, dx_i \longrightarrow \int_{x_0-\ell/2}^{x_0+\ell/2} |x_i\rangle\langle x_i|\alpha\rangle\, dx_i$$

The probability of the particle being detected in this range is given by

$$\int_{x_0-\ell/2}^{x_0+\ell/2} |\langle x_i|\alpha\rangle|^2\, dx_i$$

Of course, as $\ell \to \infty$, this probability goes to 1 as long as $|\alpha\rangle$ is normalized.

To consider three dimensions x, y, z, we must be assured that measurement in one dimension does not affect the other two, so $[x, y] = [x, z] = [y, z] = 0$. Defining $\vec{x}$ as a collective index for x, y, z, such that $|\vec{x}\rangle$ is a simultaneous eigenket for the observables x, y, z, consider the **infinitesimal translation** operator $\mathcal{J}(d\vec{x})|\vec{x}\rangle = |\vec{x} + d\vec{x}\rangle$. What properties should we expect such an operator to have? It should preserve normalized eigenkets, implying that $\langle\alpha|\mathcal{J}^\dagger(d\vec{x})\mathcal{J}(d\vec{x})|\alpha\rangle = \langle\alpha|\alpha\rangle = 1$ and therefore that $\mathcal{J}(d\vec{x})$ is unitary. We should also have $\mathcal{J}(d\vec{x}_1)\mathcal{J}(d\vec{x}_2) = \mathcal{J}(d\vec{x}_1 + d\vec{x}_2)$ and $\mathcal{J}(-d\vec{x}) = \mathcal{J}^{-1}(d\vec{x})$. Finally, as $d\vec{x}$ goes to zero, $\mathcal{J}(d\vec{x})$ should go to the identity operator: $\lim_{d\vec{x}\to 0} \mathcal{J}(d\vec{x}) = 1$.

If we take $\mathcal{J}(d\vec{x}) = 1 - i\vec{K} \cdot d\vec{x}$ for some hermitian $\vec{K} = (K_x, K_y, K_z)$, all these properties are satisfied (up to $O((d\vec{x})^2)$, which is good enough, since $d\vec{x}$ is infinitesimal). Accepting this to be the correct form for $\mathcal{J}(d\vec{x})$, we note that $[\vec{x}, \mathcal{J}(d\vec{x})] = d\vec{x}$ and therefore that $[x_i, K_j] = i\delta_{ij}$. This $\vec{K}$ seems to generate translations, so it must be in some way related to momentum. Since $\vec{K} \cdot d\vec{x}$ is dimensionless, $\vec{K}$ has units L^{-1}. We can define it as $\vec{p}$ divided by some constant with the dimension of action, L^2MT^{-1}. Calling this constant $\hbar$, we rewrite $\mathcal{J}(d\vec{x}) = 1 - i\vec{p} \cdot d\vec{x}/\hbar$, assuring that momentum really is the generator of translation. Our commutation relation becomes $[x_i, p_j] = i\hbar\delta_{ij}$, and we can now state the **Heisenberg uncertainty principle** as a special case of the more general relation above:

$$\sigma_x \sigma_{p_x} \geq \frac{\hbar}{2}$$

Note: $[p_i, p_j] = 0$, and we can use $\vec{p} = (p_x, p_y, p_z)$ to create a simultaneous momentum eigenket $|\vec{p}\rangle$. This forms one of the three **canonical commutation relations** of quantum mechanics:

$$[x_i, x_j] = 0 \qquad [p_i, p_j = 0] \qquad [x_i, p_j] = i\hbar\delta_{ij}$$

Time Evolution Suppose a state $|\alpha\rangle$ is pictured at some time t_0. We write this state as $|\alpha, t_0\rangle$, and its evolution to an arbitrary time t we write $|\alpha, t_0; t\rangle$. We want a time evolution operator $\mathcal{U}(t, t_0)|\alpha, t_0\rangle = |\alpha, t_0; t\rangle$ with the same conditions as the above infinitesimal position operator. We again make the choice $\mathcal{U}(t_0 + dt, t_0) = 1 - i\Omega dt$ for some Hermitian Ω. In classical mechanics, the Hamiltonian H is the generator of time evolution, and we correspondingly define $\Omega = H/\hbar$, giving us $\mathcal{U}(t_0 + dt, t_0) = 1 - iH\,dt/\hbar$. We find that

$$\mathcal{U}(t + dt, t_0) - \mathcal{U}(t, t_0) = -i(H/\hbar)\,dt\,\mathcal{U}(t, t_0)$$

and therefore that

$$i\hbar\frac{\partial}{\partial t}\mathcal{U}(t, t_0) = HU(t, t_0)$$

Multiplying both sides by a state ket $|\alpha\rangle$ immediately leads to the **time-dependent Schrodinger equation**,

$$i\hbar\frac{\partial}{\partial t}|\alpha, t_0; t\rangle = H|\alpha, t_0; t\rangle$$

Defining the exponential of an operator A by the Taylor series for the usual exponential, $e^A = 1 + A + A^2/2 + A^3/6 + \ldots$, the solution to this equation is the same as it is for a normal differential equation:

$$\mathcal{U}(t, t_0) = e^{-\frac{i}{\hbar}H(t-t_0)}$$

when H is not a function of time,

$$\mathcal{U}(t, t_0) = e^{-\frac{i}{\hbar}\int_{t_0}^{t} H(t')\,dt'}$$

when H is a function of time but $[H(t_1), H(t_2)] = 0$, and

$$\mathcal{U}(t, t_0) = 1 + \sum_{n=1}^{\infty}\left(-\frac{i}{\hbar}\right)^n \int_{t_0}^{t}\int_{t_0}^{t_1}\cdots\int_{t_0}^{t_{n-1}} H(t_1)H(t_2)\ldots H(t_n)\,dt_n\,dt_{n-1}\ldots dt_1$$

when H is a function of time and $[H(t_1), H(t_2)] \neq 0$. We'll generally deal only with the first case.

Suppose that H is time-independent and generates a complete basis $\{|a_i\rangle\}$, with $H|a_i\rangle = E_{a_i}|a_i\rangle$. Setting $t_0 = 0$ and expanding the time evolution operator in terms of $|a_i\rangle\langle a_i|$, we find that

$$e^{-\frac{i}{\hbar}Ht} = \sum_i \sum_j |a_j\rangle\langle a_j|e^{-\frac{i}{\hbar}Ht}|a_i\rangle\langle a_i| = \sum_i |a_i\rangle e^{-\frac{i}{\hbar}E_{a_i}t}\langle a_i|$$

For an arbitrary ket $|\alpha\rangle = \sum_i |a_i\rangle\langle a_i|\alpha\rangle = \sum_i c_{a_i}|a_i\rangle$, we have

$$|a;t\rangle = e^{-\frac{i}{\hbar}Ht}|\alpha\rangle = \sum_i c_{a_i} e^{-\frac{i}{\hbar}E_{a_i}t}|a_i\rangle$$

So the coefficent $c_{a_i}(t)$ is given by $c_{a_i}(t) = c_{a_i}e^{-\frac{i}{\hbar}E_{a_i}t}$.

How does the expectation value of an observable change over time? Observe:

$$\langle B\rangle_{a_i} = \langle a_i, t|B|a_i, t\rangle = \langle a_i|\mathcal{U}^\dagger(t,0)B\mathcal{U}(t,0)|a_i\rangle = \langle a_i|e^{\frac{i}{\hbar}E_{a_i}t}Be^{-\frac{i}{\hbar}E_{a_i}t}|a_i\rangle = \langle a_i|B|a_i\rangle$$

implying that the expectation values of observables taken with respect to energy eigenstates does *not* change over time. Energy eigenstates are correspondingly known as **stationary states**. In general, this does not hold true for expectation values taken with respect to superpositions of energy eigenstates, which are correspondingly known as **nonstationary states**.

The above exposition is an example of the **Schrodinger picture** of quantum dynamics, in which state kets are postulated to change over time while observables stay constant. We can view this in another way, though: state kets are constant, while observables change. This is known as the **Heisenberg picture**, and relies on the following mathematical equality: consider two state kets $|\beta\rangle$ and $|\alpha\rangle$ and an observable U. Since observables are unitary, $\langle\beta|\alpha\rangle = \langle\beta|U^\dagger U|\alpha\rangle$. For an operator X, consider the action of a unitary transformation $X \mapsto U^\dagger XU$ on $\langle\beta|U|\alpha\rangle$. We have

$$\langle\beta|X|\alpha\rangle \mapsto \langle\beta|U^\dagger XU|\alpha\rangle$$

But we can view this in two equivalent ways:

$$(\langle\beta|U^\dagger)\,X\,(U|\alpha\rangle) \;=\; \langle\beta|\,(U^\dagger XU)\,|\alpha\rangle$$

So either the bras and kets change as $|\alpha\rangle \mapsto U|\alpha\rangle$, or the operator itself changes as $X \mapsto U^\dagger XU$. These two pictures have different physical interpretations, but are entirely equivalent; in the case that $U = \mathcal{U}$, the time evolution operator, we recover the Schrodinger-Heisenberg distinction.

2.3 Special Relativity

As in classical mechanics, to talk about nature we need a reference frame, or coordinate system. We would like moving bodies not acted upon by external forces to move at constant velocities; a reference frame in which this holds is known as an inertial reference frame. We can have multiple reference frames, each attached to a distinguished point serving as the origin; if one is inertial, and the other moves uniformly relative to the first, the other is inertial. Galileo's principle of relativity states that laws of nature are identical in all inertial reference frames. This principle, however, was formulated with the idea of instantaneous transmission of physical signals in mind; in experiment, we find that this doesn't happen, and that the maximum velocity of propagation is a finite constant known as the speed of light, $c \approx 3 \times 10^8$ m/s. Einstein's principle of relativity states that physical laws are invariant under choice of inertial reference frame; in particular, they all measure the same c. Theories of mechanics built upon this principle are called relativistic.

2.3.1 Intervals

In special relativity, the primitive objects of study are events, or points in spacetime ($\mathbb{R}^4$). Suppose two events happen with spacetime coordinates in a reference frame K given by (x_1, y_1, z_1, t_1) and (x_2, y_2, z_2, t_2), respectively, corresponding to the emission and receiving of a light-speed signal, respectively. The signal covers a distance $c(t_2 - t_1)$ which is equal to $\sqrt{(x_2 - x_1)^2 + (y_2 - y_1)^2 + (z_2 - z_1)^2}$, so we can write

$$(x_2 - x_1)^2 + (y_2 - y_1)^2 + (z_2 - z_1)^2 - c^2(t_2 - t_1)^2 = 0$$

In a system K' where the coordinates of the two events are (x'_1, y'_1, z'_1, t'_1) and (x'_2, y'_2, z'_2, t'_2), respectively, the velocity c^2 is still the same due to the principle of invariance, so we have

$$(x'_2 - x'_1)^2 + (y'_2 - y'_1)^2 + (z'_2 - z'_1)^2 - c^2(t'_2 - t'_1)^2 = 0$$

In general, in a reference frame K where two events have coordinates (x_1, y_1, z_1, t_1) and (x_2, y_2, z_2, t_2), the interval between those two coordinates is given by

$$s_{12}^2 = c^2(t_2 - t_1)^2 - (x_2 - x_1)^2 - (y_2 - y_1)^2 - (z_2 - z_1)^2$$

We've deduced that if the interval is zero in any one reference frame, it's zero in all reference frames. If two events are infinitely close to each other, the interval ds between them is given by

$$ds^2 = c^2\,dt^2 - dx^2 - dy^2 - dz^2$$

If we measure the same interval in two different reference frames K and K' to get ds and ds', it follows from the facts that (1) if $ds = 0$ then $ds' = 0$ and (2) ds and ds' are infinitesimals of the same order, that ds and ds' are proportional to each other: $ds = a\,ds'$. Since space and time are homogeneous and isotropic, the constant of proportionality cannot depend on the coordinates or the time, nor can it depend on the direction of the relative velocity. Therefore, $ds' = a\,ds$, with the *same* constant of proportionality. It follows that $ds = a^2\,ds$, so $a^2 = 1$ and $a = \pm 1$. a obviously can't be -1, since moving between *three* reference frames would give us $ds = -ds$, so we must have $a = 1$. Therefore, $ds = ds'$ and $s = s'$. The interval between two events is independent of the frame of reference.

The Light Cone Suppose we have two events in spacetime, viewed from a reference frame K, and you, a massive object (no offense) want to get from one to the other by traveling along a straight line. Were we to attach a reference frame K' to you, putting you at the origin, we'd find that both events have the same space coordinates in K'. Introducing the notation $t_{12} = t_2 - t_1$ and $l_{12}^2 = (x_2 - x_1)^2 + (y_2 - y_1)^2 + (z_2 - z_1)^2$, the intervals in K and K' are $s_{12}^2 = c^2 t_{12}^2 - l_{12}^2$ and $s_{12}'^2 = c^2 t_{12}'^2 - l_{12}'^2$. Since $l_{12}'^2 = 0$ and $s_{12}^2 = s_{12}'^2$, we have $s_{12}^2 = c^2 t_{12}^2 - l_{12}^2 = c^2 t_{12}'^2 > 0$. So you can get from one to the other if $s_{12}^2 > 0$. We call such an interval **timelike**, since all that's keeping you from traveling along it is time. If we want the two events to happen at the same time, we require $s_{12}^2 < 0$, and call the interval **spacelike**, since you'd have to *teleport* through space to get from one to the other. Because of the invariance of intervals, the spacelike/timelike divide is an absolute division, independent of reference frames; at any point p in a coordinate system there is a cone defined by $x^2 + y^2 + z^2 - c^2 t^2 = 0$ known as the **light cone**, any point outside of which is absolutely remote relative to p, and any point inside which is either in the absolute past or absolute future relative to p, where $t < 0$ and $t > 0$, respectively.

Proper Time Suppose that we're at the center of an inertial reference frame K, we have two clocks C and C', and we chuck C' away at an arbitrary velocity. During an infinitesimal period of time dt as measured by our clock C, C' will travel a distance $\sqrt{dx^2 + dy^2 + dz^2}$. Because of the invariance of intervals, $ds^2 = c^2dt^2 - dx^2 - dy^2 - dz^2 = c^2dt'^2$, so

$$dt' = \frac{ds}{c} = dt\sqrt{1 - \frac{dx^2 + dy^2 + dz^2}{c^2\,dt^2}} = dt\sqrt{1 - \frac{v^2}{c^2}}$$

Integrating this expression, we see that over a time interval $t_2 - t_1$ measured by C, C′ experiences a time interval

$$t_2' - t_1' = \int_{t_1}^{t_2} \sqrt{1 - \frac{v^2}{c^2}}\,dt$$

Since this interval is less than $t_2 - t_1$, C′ is seen as lagging. Paradoxically, however, from C′′s reference frame, C is lagging!

The proper time for an object is the time read by a clock moving along with that object, which is the integral $\int_a^b \frac{ds}{c}$ taken along the world line of the clock. For two points separated by a timelike interval, this integral has the maximum value when taken along the straight world line joining these two points.

2.3.2 Lorentz Transformations

We want to translate the set of coordinates (x, y, z, t) in the reference frame K to another set of coordinates (x', y', z', t') in a reference frame K′. Supposing K′ moves along K's x axis at a velocity V, in classical mechanics we'd set $x' = x + Vt, y' = y, z' = z, t' = t$, which is known as the Galilean transformation, but this fails to leave intervals invariant, making it unacceptable for relativistic mechanics.

Setting $\tau = ict$, such that $s^2 = x^2 + y^2 + z^2 + \tau^2$, and changing coordinates to (x, y, z, τ), what we're looking for is precisely an isometry of this space. It's then either a parallel displacement or a rotation. Displacement doesn't matter, since it only changes the origin, so we want a rotation: every rotation can be broken up into six rotations in the $xy, zy, xz, \tau x, \tau y, \tau z$ planes. We don't care about $xy, zy, xz, \tau y$, or τz rotations, so this must be a τx rotation, changing coordinates as $x = -\tau' \sin\psi$, $\tau = \tau' \cos\psi$. From this it follows that $\tan\psi = iV/c$, so simple algebra leads us to the change of coordinates

$$x = \frac{x' + Vt'}{\sqrt{1 - \frac{V^2}{c^2}}} \quad y = y' \quad z = z' \quad t = \frac{t' + V\frac{x'}{c^2}}{\sqrt{1 - \frac{V^2}{c^2}}}$$

This transformation is known as the Lorentz transformation. Clearly, it yields the Galilean transformation as $c \to \infty$. As a consequence, suppose a rod moving along the x axis at velocity V relative to us measures its own length as $\Delta x'$: we will then measure its length as

$$\Delta x = \frac{\Delta x'}{\sqrt{1 - \frac{V^2}{c^2}}}$$

In other words, the faster it goes, the shorter it appears to us. This is known as Lorentz contraction.

By considering such a transformation for infinitesimal dx, dt, we can find formulas for the transformation of velocities: under the same conditions as above, we have

$$v_x = \frac{v'_x + V}{1 + v'_x \frac{V}{c^2}} \quad v_y = \frac{v'_y \sqrt{1 - \frac{V^2}{c^2}}}{1 + v'_y \frac{V}{c^2}} \quad v_z = \frac{v'_z \sqrt{1 - \frac{V^2}{c^2}}}{1 + v'_z \frac{V}{c^2}}$$

Again, as $c \to \infty$, we get the classical transformation, in which $v_x = v'_x + V$.

We generally denote the factor $\frac{1}{\sqrt{1 - \frac{V^2}{c^2}}}$ as γ, the Lorentz factor. So, for instance, we can restate Lorentz contraction and time dilation as $\Delta x = \gamma \Delta x'$ and $\Delta t = \gamma \Delta t'$, respectively.

2.3.3 Four-vectors

We'll set $c = 1$ from now on; if you want, you can figure out where it's been hidden via dimensional analysis. In the four dimensional spacetime manifold in which relativistic mechanics take place, Minkowski space, vectors have three space components and one time component, and are known as four-vectors. The inner product on this space is given by

$$a \cdot b = a^0 b^0 - a^1 b^1 - a^2 b^2 - a^3 b^3$$

We can write this neatly by introducing a metric tensor η_{ij} on this manifold, given by

$$\eta_{ij} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{bmatrix}$$

So $a \cdot b = \eta_{ij} a^i b^j$. We can restate several of the above developments in sleeker ways: the infinitesimal interval (or line element) is given by $ds^2 = -\eta_{ij} dx^i dx^j$, the path length and proper time are given by

$$\Delta s = \int \sqrt{-\eta_{ij}\frac{dx^i}{d\lambda}\frac{dx^j}{d\lambda}}\, d\lambda \qquad \Delta\tau = \int \sqrt{\eta_{ij}\frac{dx^i}{d\lambda}\frac{dx^j}{d\lambda}}\, d\lambda$$

Recall the Einstein summation notation: (i) when the same index appears in both a raised and a lowered position, we implicitly sum over it, e.g. $v_i w^i = \sum_{i=1}^4 v_i w^i$ (ii) we use the metric to raise and lower indices at will, e.g. $v^i = \eta^{ij} v_j$, and (iii) putting indices in square (curly) brackets indicates that we wish to take their commutator (anticommutator), e.g. $v_{[i,} w_{j]} = v_i w_j - v_j w_i$. By rewriting everything in terms of tensors, we can express relationships without invoking any sort of reference frame; doing this makes an equation, relationship, or theory covariant (which has nothing to do with covariance/contravariance of tensors).

The velocity of a particle x^i, parametrized by its proper time, is given by $v^i = \partial_\tau x^i$; since $d\tau^2 = \eta_{ij} dx^i dx^j$, we have $\eta_{ij} v^i v^j = 1$, the interpretation being that we're *always* traveling at the same speed through spacetime (light-speed, really; examining units, the 1 yields a hidden c), and that moving faster through space just means moving slower through time. The momentum of a particle is given by $p^i = \gamma m v^i$, and the energy is γm. The force on a particle is given by $f^i = \partial_\tau u^i$.

2.4 General Relativity

General relativity is far more subtle, though a significant portion of the legwork was performed in the previous discussion of Riemannian geometry. We postulate that gravitational force on an observer is equivalent to the "pseudo"-force experienced by an observer in an accelerating reference, a postulate known as the *equivalence principle*. Our sources include [Wald, 2007, Carroll, 2019, Misner et al., 1973]. The differential geometry book [Kühnel, 2015] discusses general relativity as well, focusing in particular on "Einstein manifolds", or Riemannian manifolds whose metrics are solutions to the vacuum Einstein field equations.

2.4.1 Pseudo-Riemannian Manifolds

We begin by recapping some constructions on a pseudo-Riemannian manifold (M, g). The *Levi-Civita connection* ∇_i is the unique connection on M that preserves g and has vanishing torsion tensor, and its difference from the ordinary derivative ∂_i is given by the *Christoffel symbols*,

$$\Gamma^i{}_{jk} := \frac{1}{2} g^{i\ell} \left(\partial_k g_{\ell j} + \partial_j g_{\ell k} - \partial_\ell g_{jk} \right)$$

Having written these down, we can express the action of ∇_i on a vector v^j as

$$\nabla_i v^j = \partial_i v^j + \Gamma^j{}_{ik} v^k$$

In local coordinates, the Christoffel equations give us second-order differential equations for the position x^i of a "particle" traveling on a geodesic, known as the *geodesic equations*:

$$\frac{d^2 x^i}{dt^2} + \Gamma^i{}_{jk} \frac{dx^j}{dt} \frac{dx^k}{dt} = 0$$

(Compare this with the result that the geodesics in a flat space are straight lines, i.e. $\ddot{x} = 0$). For any given initial position x^i and velocity $\frac{dx^i}{dt}$, the theory of ordinary differential equations tells us that a unique solution exists to the geodesic equations.

Given an infinitesimal square with sides v^i and w^i, parallel transport of a vector x^i around the square generally fails to leave x^i unaltered. The difference, as a vector, is linear in v^i, w^i, and x^i, and hence is given by $y^\ell = R^\ell{}_{ijk} v^j w^k x^i$ for some tensor $R^\ell{}_{ijk}$ known as the *Riemann curvature tensor*. In terms of the Christoffel symbols, this tensor can be given as

$$R^\ell{}_{ijk} = \partial_j \Gamma^\ell{}_{ki} - \partial_k \Gamma^\ell{}_{ji} + \Gamma^\ell{}_{jm} \Gamma^m{}_{ki} - \Gamma^\ell{}_{km} \Gamma^m{}_{ji}$$

Contracting it yields the *Ricci curvature* R_{ij} and *scalar curvature* R:

$$R_{ij} = R^\ell{}_{i\ell j} \qquad\qquad R = R^i{}_i$$

We define the *Einstein tensor* G_{ij} by

$$G_{ij} = R_{ij} - \frac{1}{2} R g_{ij}$$

A metric g_{ij} which solves the equations $G_{ij} = 0$ is one which distributes the curvature of M "most evenly" [Kühnel, 2015]. A key property of the Einstein tensor is its vanishing divergence: $\nabla^i G_{ij} = 0$.

The Stress-Energy Tensor General relativity historically has its roots in an attempt to generalize the Poisson equation, a field-theoretic version of Newtonian gravity. Given a mass density

ρ and a gravitational field $\vec{g}$ expressed as the gradient of a scalar potential ϕ, Gauss's law reads $\nabla \cdot \vec{g} = -4\pi G\rho$, where G is a gravitational constant. Plugging in $\vec{g} = -\nabla\phi$, we obtain *Poisson's equation*,

$$\nabla^2\phi = 4\pi G\rho$$

To generalize this to the framework of special relativity, we first need to figure out how to replace ρ with something that respects mass-energy equivalence and transforms like a tensor. The solution is a symmetric tensor T_{ij} known as the **stress-energy tensor**. An observer with velocity v^i will measure a mass-energy per unit volume of $T_{ij}v^iv^j$. Given an x^j orthogonal to v^μ, the component $-T_{ij}x^jv^i$ is interpreted as the momentum density of the matter in the x^j direction. A y^k also orthogonal to v^i can be plugged in along with x^j, and $T_{ij}x^iy^j$ is interpreted as the x^i-y^j component of the stress tensor for a point in an arbitrary material body. To summarize, the stress-energy-momentum tensor T_{ij} gives us *stress* when we plug in two position vectors, *momentum* when we plug in a position vector and an orthogonal velocity vector, and *energy* when we plug in one velocity vector twice. Conservation of energy implies that the stress-energy tensor has vanishing divergence: $\nabla^iT_{ij} = 0$.

2.4.2 The Einstein Field Equations

We've identified the mass density ρ with the mass-energy density $T_{ij}v^iv^j$. Now we have to replace $\nabla^2\phi$ with a tensorial quantity as well; it should have at most second-order derivatives of the metric, and it should be divergence-free.

A first guess is given by the observation that the differential acceleration of two nearby particles with separation vector x is given by $-(x \cdot \nabla)\nabla\phi$. However, since their world lines will be geodesics, and a fortiori curves on our spacetime manifold, we know that this same acceleration is given by $-R^\ell{}_{jik}v^jv^kx^i$. So let's make the correspondence $R^\ell{}_{jik}v^jv^k = \partial_i\partial^\ell\phi$, and therefore $\partial^2\phi = R^\ell{}_{j\ell k} = R_{jk}$, and conclude that the correct covariant generalization of the Poisson equation is given by $R_{ij}v^iv^j = 4\pi GT_{ij}v^iv^j$, or more concisely $R_{ij} = 4\pi GT_{ij}$.

This was, in fact, one of Einstein's guesses. It is wrong. It is in general true that $\nabla^iG_{ij} = \nabla^i(R_{ij} - \frac{1}{2}Rg_{ij}) = 0$, and hence the divergence of R_{ij} is given by $\nabla^i\frac{1}{2}Rg_{ij} = \frac{1}{2}\nabla_jR$. Hence, divergence-freeness of R_{ij} implies that $\nabla_iR = 0$, i.e. that R and hence $T = T^i{}_i$ are constant throughout the universe! The correct solution to the problem is contained within the problem itself: we replace R_{ij} with $\frac{1}{2}G_{ij}$, which we already know to be divergence-free. This yields the

Einstein field equations:

$$G_{ij} = 8\pi G T_{ij}$$

Comparing units, we see that there's a hidden c^{-4} on the right-hand side; it is convenient to define **Einstein's constant** by $\kappa = 8\pi G/c^4$ and simply write $G_{ij} = \kappa T_{ij}$.

The Lagrangian Formulation In Lagrangian mechanics, we associate to a physical system a function of time $L(t)$ known as the Lagrangian, which governs the dynamics of the system; the Lagrangian is allowed to operate on the positions and velocities of the particles, e.g. as $L(t) = L(q(t), \dot{q}(t)) = \frac{1}{2}m\dot{q}(t)^2 - mgq(t)$. In a field-theoretic context, such as general relativity, we may also consider the Lagrangian as a function of fields ϕ and their first derivatives $\partial_\mu\phi$, e.g. as $L(t) = \int \frac{1}{2}\partial_\mu\phi\partial^\mu\phi - \frac{1}{2}m^2\phi^2 \, d^3x$. In this case, we refer to the term which is integrated over space to get the Lagrangian as the Lagrangian density $\mathcal{L}$. Integrating the Lagrangian over time yields the action, $S = \int L \, dt$; the principle of least action states that the positions/fields involved in the Lagrangian are chosen so as to minimize the variation of the action under an arbitrary variation in said positions/fields $\delta S = 0$.

A covariant formulation of Lagrangian mechanics requires us to replace ∂_μ with the covariant derivative ∇_μ, so as to make all terms appearing in the Lagrangian tensorial; further, if we wish to work on an n-dimensional Riemannian manifold (M, g), we must integrate the scalar Lagrangian density $\mathcal{L}$ with respect to the volume form $\sqrt{|g|}\, d^n x$, where $d^n x := dx_1 \wedge \ldots \wedge dx_n$ and $|g|$ is the determinant of the metric tensor.

In a vacuum, the Einstein-Hilbert action of general relativity is given by the Lagrangian density $\mathcal{L}_V = R/2\kappa$:

$$S_V = \int \frac{R}{2\kappa}\sqrt{|g|}\, d^4x$$

Upon variation of the metric, this yields

$$\delta S_V = \int \left(R_{\mu\nu} - \frac{1}{2}Rg_{\mu\nu}\right)\delta g^{\mu\nu}\sqrt{|g|}\, d^4x$$

(A detailed derivation is given in [Carroll, 2019]). Since this must be zero for all variations of the metric, we obtain $R_{\mu\nu} - \frac{1}{2}Rg_{\mu\nu} = G_{\mu\nu} = 0$, Einstein's equations for a vacuum.

To add mass-energy fields, we add an arbitrary density $\mathcal{L}_M$ to the Lagrangian density, which by the linearity of integration splits the action S into $S_V + S_M$, the sum of the vacuum and mass-energy actions. Working in reverse, we define the stress-energy tensor as

$$T_{\mu\nu} = -\frac{2}{\sqrt{|g|}}\frac{\delta S_M}{\delta g^{\mu\nu}}$$

guaranteeing that the principle of least action reduces to Einstein's equation, $G_{\mu\nu} = \kappa T_{\mu\nu}$.

2.5 Quantum Field Theory

This section discusses the Lorentz covariant generalization of quantum mechanics to fields known as quantum field theory. Our sources for vanilla quantum field theory are [Peskin, 2018, Lancaster and Blundell, 2014, Ticciati et al., 1999]; the two-volume series [Deligne et al.,] delivers mathematical rigor to the field. Being especially confusing, we have tried to root our discussion of spinors in representation theory, for which the books [Weinberg, 1995, Bleecker, 2005] are useful.

2.5.1 Classical Field Theory

The setup for studying classical field theories in Minkowski space with metric $\eta^{\mu\nu} = \text{diag}(+1, -1, -1, -1)$ is as follows:

1. Take the *Lagrangian density* of the theory. As an example, we will work with the Lagrangian of a free scalar theory,

$$\mathcal{L} = \frac{1}{2}\partial_\mu\phi\partial^\mu\phi - \frac{1}{2}m^2\phi^2$$

2. Plug $\mathcal{L}$ into the *Euler-Lagrange equations*,

$$\partial_\mu\left(\frac{\partial\mathcal{L}}{\partial\left(\partial_\mu\phi\right)}\right) - \frac{\partial\mathcal{L}}{\partial\phi} = 0$$

This yields four equations ($\mu = 0, 1, 2, 3$) for each free variable; our theory only has one free variable (ϕ). Plugging the above $\mathcal{L}$ into the Euler-Lagrange equations, we calculate

$$\frac{\partial(\partial_\mu\phi\partial^\mu\phi)}{\partial(\partial_\mu\phi)} = \frac{\partial(\eta^{\mu\nu}\partial_\mu\phi\partial_\nu\phi)}{\partial(\partial_\mu\phi)} = \eta^{\mu\nu}\frac{\partial(\partial_\mu\phi)}{\partial(\partial_\mu\phi)}\partial_\nu\phi+\eta^{\mu\nu}\partial_\mu\phi\frac{\partial(\partial_\nu\phi)}{\partial(\partial_\mu\phi)} = \partial^\mu\phi+(\partial^\nu\phi)\delta^\mu_\nu = 2\partial^\mu\phi$$

and hence obtain the equation

$$\partial_\mu\partial^\mu\phi + m^2\phi = 0$$

Writing $\partial_\mu \partial^\mu$ as ∂^2, this becomes the **Klein-Gordon equation**

$$(\partial^2 + m^2)\phi = 0$$

3. If we want more information, we may calculate the *Hamiltonian density* of the theory. In a theory with n free variables $\phi_1, \dots, \phi_n$, this is first done by associating to each ϕ_i a *conjugate momentum*

$$\Pi_i^\mu = \frac{\partial \mathcal{L}}{\partial(\partial_\mu \phi_i)}$$

and then deriving the Hamiltonian as

$$\mathcal{H} = \left(\sum_{i=1}^{n} \Pi_i^0 \partial_0 \phi_i \right) - \mathcal{L}$$

For our free scalar theory, we have

$$\Pi^\mu = \frac{\partial \mathcal{L}}{\partial(\partial_\mu \phi)} = \partial^\mu \phi$$

and hence

$$\mathcal{H} = \partial^0 \phi \partial_0 \phi - \mathcal{L} = \frac{1}{2}\partial^0 \phi \partial_0 \phi - \frac{1}{2}\sum_{i=1}^{3} \partial_i \phi \partial^i \phi + \frac{1}{2} m^2 \phi^2 = \frac{1}{2}\left[\left(\frac{\partial \phi}{\partial t} \right)^2 + (\nabla \phi)^2 + m^2 \phi^2 \right]$$

4. Alternatively, we can define the *stress-energy tensor* ${T^\mu}_\nu$ of the theory, given by

$${T^\mu}_\nu = \frac{\partial \mathcal{L}}{\partial\left(\partial_\mu \phi\right)} \partial_\nu \phi - \mathcal{L} \delta_\nu^\mu$$

This gives rise to four conserved quantities,

$$P^i = \int T^{0i} \, d^3 x$$

For the Klein-Gordon Lagrangian, we obtain a stress energy tensor of

$${T^\mu}_\nu = \partial^\mu \phi \partial_\nu \phi - \frac{1}{2} \delta_\nu^\mu \left(\partial_\rho \phi \partial^\rho \phi - m^2 \phi^2 \right)$$

For $\mu = \nu = 0$, we reclaim the Hamiltonian, and for $\mu = 0, \nu \neq 0$, we obtain

$$T^{0i} = \sum_{j=1}^{3} \eta^{ij} T^{0}_{\ j} = -T^{0}_{\ i} = \partial^0 \phi \partial_i \phi$$

This gives us a set of tools for the analysis of classical fields.

Another example is given by classical electromagnetism. Setting $c = 1$, define the electromagnetic four-potential A_μ to have as its timelike component the electric potential ϕ and as its spacelike components the magnetic vector potential $\vec{A}$. The exterior derivative of this one-form is known as the **electromagnetic tensor** $F_{\mu\nu}$, and as a matrix looks like

$$\begin{bmatrix} 0 & E_x & E_y & E_z \\ -E_x & 0 & -B_z & B_y \\ -E_y & B_z & 0 & -B_x \\ -E_z & -B_y & B_x & 0 \end{bmatrix}$$

The Lagrangian of classical electromagnetism is given by

$$\mathcal{L} = \overbrace{-\frac{1}{4\mu_0} F^{\mu\nu} F_{\mu\nu}}^{\text{field}} - \overbrace{A_\mu J^\mu}^{\text{source}}$$

where $J^\mu = (\rho, \vec{j})$ is a four-current. With some effort, we may show that the Euler-Lagrange equations read

$$\partial_\mu F^{\mu\nu} = \mu_0 J^\nu$$

For $\nu = 0$ this reduces to $\nabla \cdot \vec{E} = \mu_0 \rho = \rho / \varepsilon_0$, Gauss's law. For $\nu = 1, 2, 3$, we obtain $\nabla \times \vec{B} = \mu_0 \vec{j} + \frac{\partial \vec{E}}{\partial t}$, or Ampere's law.

2.5.2 Canonical Quantization

To quantize a classical field theory with position variables $\phi_1, \ldots, \phi_n$ and conjugate momenta $\Pi^\mu_1, \ldots, \Pi^\mu_n$, we turn the position and momentum variables into operators $\widehat{\phi}_1, \ldots, \widehat{\phi}_n, \widehat{\Pi}^\mu_1, \ldots, \widehat{\Pi}^\mu_n$, and impose the **equal-time commutation relations**

$$[\widehat{\phi}_i(t, \vec{x}), \widehat{\Pi}^0_j(t, \vec{y})] = i\delta^{(3)}(\vec{x} - \vec{y})\delta_{ij}$$

with all commutators among $\widehat{\phi}$s and among $\widehat{\Pi}$s being zero. The Hamiltonian $\mathcal{H}$, being a function of ϕ and Π, becomes an operator $\widehat{\mathcal{H}}$ as well, as does $H = \int \mathcal{H}\, d^3x$.

Fundamentally, quantizing $\widehat{H}$ gives it a quantized spectrum. In the case where we have one variable ϕ with no self-interactions (i.e., the Euler-Lagrange equations are linear in ϕ), we have a lowest-energy **vacuum state** $|0\rangle$ to which we can add a "particle" with momentum $\vec{p}$ via the **creation operator** $\widehat{a}^\dagger_{\vec{p}}$, and remove a particle with momentum $\vec{q}$ via the **annihilation operator** $\widehat{a}_{\vec{q}}$.

Additional variables will define additional pairs of annihilation and creation operators, generally denoted $(\widehat{b}^\dagger_{\vec{p}}, \widehat{b}_{\vec{q}}), (\widehat{c}^\dagger_{\vec{p}}, \widehat{c}_{\vec{q}})$, and so on. We may reconstruct $\widehat{\phi}$ from the annihilation and creation operators by means of a **mode expansion** which, in the case of the Klein-Gordon field, is given by

$$\widehat{\phi}(t, \vec{x}) = \int \frac{d\vec{p}}{(2\pi)^{3/2}} \frac{1}{\sqrt{2E_{\vec{p}}}} \left(\widehat{a}_{\vec{p}} e^{-ip\cdot x} + \widehat{a}^\dagger_{\vec{p}} e^{ip\cdot x} \right)$$

where $p \cdot x = (t, \vec{p}) \cdot (t, \vec{x}) = t^2 - \vec{p} \cdot \vec{x}$, and $E_{\vec{p}} = \sqrt{\vec{p}^2 + m^2}$. We interpret $\widehat{\phi}(x)$ as creating a particle at position x. We define the state $|\vec{p}\rangle$ consisting of one particle with momentum $\vec{p}$ by $|\vec{p}\rangle = \widehat{a}^\dagger_{\vec{p}}|0\rangle$, so that $\langle \vec{p}|\vec{q}\rangle = \delta^{(3)}(\vec{p} - \vec{q})$.

In general, though, our theory will not be free from self-interactions, so we have to replace the vacuum state $|0\rangle$ with a more mysterious ground state $|\Omega\rangle$. While acting on $|0\rangle$ with $\widehat{a}^\dagger_{\vec{p}}$ yields a state with a single particle of momentum $\vec{p}$, acting on $|\Omega\rangle$ with $\widehat{a}^\dagger_{\vec{p}}$ guarantees nothing but a superposition of particles whose momenta sum to $\vec{p}$.

The dynamics of a quantum field theory can be analyzed via its **correlation functions**, numbers of the form

$$\langle \Omega | \widehat{\phi}(x_1) \dots \widehat{\phi}(x_n) \phi(y_1)^\dagger \dots \widehat{\phi}(y_n)^\dagger | \Omega \rangle$$

which express the probability for particles created at positions $y_1, \dots, y_n$ to travel to positions $x_1, \dots, x_n$. To evaluate these, we need some additional machinery.

Green's Functions Given a linear differential operator L, e.g. $Lx(t) = m\frac{d^2}{dt^2}x(t) + cx(t)$, we define the **Green's function** of L to be a function $G(t, u)$ such that $LG(t, u) = \delta(t - u)$. Given a differential equation $Lx(t) = f(t)$, we may use G to solve for x as

$$x(t) = \int G(t, u) f(u)\, du$$

noting that

$$Lx(t) = L\left(\int G(t,u)f(u)\,du\right) = \int LG(t,u)f(u)\,du = \int \delta(t-u)f(u)\,du = f(t)$$

Normal and Time Ordering When we have a series of scalar fields $\widehat{\phi}(x_1), \widehat{\phi}(x_n)$ being multiplied, we define the **time-ordering symbol** T by $T\widehat{\phi}(x_1)\cdot\ldots\cdot\widehat{\phi}(x_n) = \widehat{\phi}(x_{i_1})\cdot\ldots\cdot\widehat{\phi}(x_{i_n})$, where the x_{i_j} are such that $x^0_{i_j} \leq x^0_{i_k}$ iff $j \geq k$; T simply orders the scalar fields from latest to earliest in time. Similarly, the **normal ordering symbol** N puts all creation operators on the left, e.g. as $N\widehat{a}_{\vec{p}}\widehat{a}^\dagger_{\vec{q}}\widehat{a}_{\vec{r}} = \widehat{a}^\dagger_{\vec{q}}\widehat{a}_{\vec{p}}\widehat{a}_{\vec{r}}$ (note that $\widehat{a}_{\vec{p}}$ and $\widehat{a}_{\vec{r}}$ commute, so it doesn't matter what order they're placed in). We define the **contraction** of two operators as

$$\contraction{}{\widehat{A}}{}{\widehat{B}}\widehat{A}\widehat{B} = \langle 0|T\widehat{A}\widehat{B}|0\rangle$$

So, for instance,

$$\contraction{}{\widehat{A}}{\widehat{B}\widehat{C}\widehat{D}}{\widehat{E}}\contraction[2ex]{\widehat{A}}{\widehat{B}}{\widehat{C}}{\widehat{D}}\contraction[1ex]{\widehat{A}\widehat{B}}{\widehat{C}}{\widehat{D}\widehat{E}}{\widehat{F}}\widehat{A}\widehat{B}\widehat{C}\widehat{D}\widehat{E}\widehat{F} = \widehat{A}\widehat{E}\langle 0|T\widehat{B}\widehat{D}|0\rangle\langle 0|T\widehat{C}\widehat{F}|0\rangle$$

Wick's theorem states that applying T to a given string of operators is equivalent to applying N to that string plus all of its possible contractions. For instance,

$$T\widehat{A}\widehat{B}\widehat{C}\widehat{D} = N\widehat{A}\widehat{B}\widehat{C}\widehat{D} + \langle 0|T\widehat{A}\widehat{B}|0\rangle N\widehat{C}\widehat{D} + \langle 0|T\widehat{A}\widehat{C}|0\rangle N\widehat{B}\widehat{D} + \ldots + \langle 0|T\widehat{A}\widehat{B}|0\rangle\langle 0|T\widehat{C}\widehat{D}|0\rangle + \ldots$$

where we first list the term with zero contractions, then those with one contraction, then with two. As a particular case, this allows us to evaluate terms of the form $\langle 0|T\widehat{A}\widehat{B}\widehat{C}\ldots|0\rangle$: since

Since taking $\langle 0|N\widehat{A}\widehat{B}\ldots|0\rangle$ always yields zero, we see that this simplifies to the sum of all terms which contract *all* elements.

Propagators We define the **Feynman propagator** by

$$G(x,y) = \langle\Omega|T\widehat{\phi}(x)\widehat{\phi}^\dagger(y)|\Omega\rangle$$

The interpretation of this is as follows: starting from the ground state $|\Omega\rangle$, create a particle at spacetime point y, wait a while, and then attempt to annihilate it at spacetime point x; the extent to which the state no longer resembles $|\Omega\rangle$ is given by taking its product against $\langle\Omega|$. When we're in a free theory with $|\Omega\rangle = |0\rangle$, $G(x, y)$ is known as the **free propagator**

$$\Delta(x, y) = \langle 0|T\widehat{\phi}(x)\widehat{\phi}^\dagger(y)|0\rangle$$

Perturbation Expansions To see this machinery in action, we need a non-free, interacting field theory. One such theory is given by the "ϕ^4" theory, with Lagrangian

$$\mathcal{L} = \frac{1}{2}\partial_\mu\phi\partial^\mu\phi - \frac{1}{2}m^2\phi^2 - \frac{\lambda}{4!}\phi^4$$

This is similar to the Klein-Gordon Lagrangian, except for the ϕ^4 term which induces a non-linear Euler-Lagrange equation

$$(\partial^2 + m^2)\phi = -\frac{\lambda}{3!}\phi^3$$

The quantized Hamiltonian $\widehat{\mathcal{H}}$ is similar to that of the Klein-Gordon Hamiltonian, but with an extra "interaction" term $\frac{\lambda}{4!}\widehat{\phi}^4$. We correspondingly decompose $\widehat{\mathcal{H}}$ as $\widehat{\mathcal{H}}_0 + \widehat{\mathcal{H}}'$, where $\widehat{\mathcal{H}}_0$ is the Klein-Gordon Hamiltonian and $\widehat{\mathcal{H}}'$ is this interaction term. When λ is small, we can approximate the evolution of an arbitrary operator $\widehat{O}$ as $\widehat{O}_I(t) = e^{i\widehat{H}_0 t}\widehat{O}e^{-i\widehat{H}_0 t}$, where the subscript I denotes that we're working in the "interaction picture". We define the S-**matrix** by

$$\widehat{S} = T\left[e^{-i\int_{-\infty}^{\infty}\widehat{\mathcal{H}}_I \, d^4x}\right]$$

Since this is generally insoluble, we expand in powers of $-i\int_{-\infty}^{\infty}\widehat{\mathcal{H}}_I \, d^4x$:

$$\widehat{S} = T\left[1 - i\int\widehat{\mathcal{H}}_I(x)\, d^4x + \frac{(-i)^2}{2}\int\widehat{\mathcal{H}}_I(x)\widehat{\mathcal{H}}_I(y)\, d^4x\, d^4y + \dots\right] = T\left[1 + \sum_{n=1}^{\infty}\frac{(-i)^n}{n!}\int\prod_{m=1}^{n}\widehat{\mathcal{H}}_I(x_m)\, d^4x_m\right]$$

We can analyze the probability that a particle with momentum $\vec{p}$ turns into a particle with momentum $\vec{q}$ by plugging the two probabilities into the S-matrix: for instance, in the ϕ^4 theory, we obtain

$$\langle\vec{q}|\widehat{S}|\vec{p}\rangle \propto \langle 0|\widehat{a}_{\vec{q}}\widehat{S}\widehat{a}^\dagger_{\vec{p}}|0\rangle =$$

$$T\left[\langle 0|\widehat{a}_{\vec{q}}\widehat{a}^\dagger_{\vec{p}}|0\rangle + (-i)\left(\frac{\lambda}{4!}\right)\int\langle 0|\widehat{a}_{\vec{q}}\widehat{\phi}(x)^4\widehat{a}^\dagger_{\vec{p}}|0\rangle\, d^4x + \frac{(-i)^2}{2}\left(\frac{\lambda}{4!}\right)^2\int\langle 0|\widehat{a}_{\vec{q}}\widehat{\phi}(x)^4\phi(y)^4\widehat{a}^\dagger_{\vec{p}}|0\rangle\, d^4x\, d^4y + \dots\right]$$

$$= \langle 0|\widehat{a}_{\vec{q}}\widehat{a}^{\dagger}_{\vec{p}}|0\rangle + \sum_{n=1}^{\infty} \frac{(-i)^n}{n!}\left(\frac{\lambda}{4!}\right)^n \int \langle 0|T\left[\widehat{a}_{\vec{q}}\left(\prod_{m=1}^{n}\widehat{\phi}(x_m)^4\right)\widehat{a}^{\dagger}_{\vec{p}}\right]|0\rangle \prod_{m=1}^{n} d^4x_m$$

Thus, the higher-order corrections to $\langle 0|\widehat{a}_{\vec{q}}\widehat{S}\widehat{a}^{\dagger}_{\vec{p}}|0\rangle$ arise in powers proportional to λ.

Let's analyze the first-order correction, given by

$$\frac{-i\lambda}{4!}\int \langle 0|T\,\widehat{a}_{\vec{q}}\widehat{\phi}(x)\widehat{\phi}(x)\widehat{\phi}(x)\widehat{\phi}(x)\widehat{a}^{\dagger}_{\vec{p}}|0\rangle\, d^4x$$

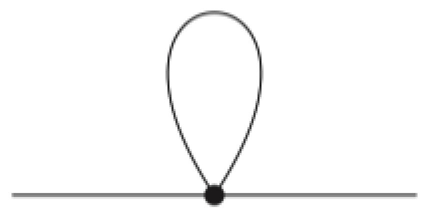
As stated above, the integrand can be reduced to the sum of all total contractions over its six members. Given $2n$ operators, there are $\frac{(2n)!}{2^n n!}$ distinguishable ways to contract all operators (i.e., form n pairs); $2n = 6$ here, there are 15 terms to consider. In each of these, either the annihilation and creation operators have been contracted with one another, or they have not. The cases in which they have number $\frac{4!}{2^2 \cdot 2!} = 3$, and the cases in which they have not, so that each one is contracted with a $\widehat{\phi}(x)$, number $15 - 3 = 12$. The three terms are of the form $\langle 0|\widehat{a}_{\vec{q}}\widehat{a}^{\dagger}_{\vec{p}}|0\rangle = \delta^{(3)}(\vec{q}-\vec{p})$, and we may also calculate $\langle 0|\widehat{\phi}(x)\widehat{a}^{\dagger}_{\vec{p}}|0\rangle = \frac{1}{(2\pi)^{3/2}}\frac{1}{\sqrt{2E_{\vec{p}}}}e^{-ip\cdot x}$, $\langle 0|\widehat{a}_{\vec{q}}\widehat{\phi}(x)|0\rangle = \frac{1}{(2\pi)^{3/2}}\frac{1}{\sqrt{2E_{\vec{q}}}}e^{iq\cdot x}$.

We can represent each nth order term in the S-matrix expansion via a **Feynman diagram**, where vertices represent particles and lines between vertices represent contractions. Each theory has its own rules for drawing Feynman diagrams. For the ϕ^4 theory, the rules are as follows: an nth order term has n vertices, one for each field $\hat{\phi}(x_i)$, $i = 1, \ldots, n$, with four outgoing lines for each vertex, each representing a possible contraction of one of the four $\hat{\phi}(x_i)$s. $\widehat{a}_{\vec{q}}$ is drawn as an outgoing line, and $\widehat{a}^{\dagger}_{\vec{p}}$ as an incoming line. Contractions between operators are represented by connecting lines. For instance, the contraction

$$\langle 0|\widehat{a}_{\vec{q}}\widehat{\phi}(x)\widehat{\phi}(x)\widehat{\phi}(x)\widehat{\phi}(x)\widehat{a}^{\dagger}_{\vec{p}}|0\rangle = \left(\frac{1}{(2\pi)^{3/2}}\frac{1}{\sqrt{2E_{\vec{q}}}}e^{iq\cdot x}\right)\Delta(x-x)\left(\frac{1}{(2\pi)^{3/2}}\frac{1}{\sqrt{2E_{\vec{p}}}}e^{-ip\cdot x}\right)$$

$$= \frac{e^{i(q-p)\cdot x}\Delta(0)}{16\pi^3\sqrt{E_{\vec{q}}E_{\vec{p}}}}$$

(which must still be integrated and multiplied by the appropriate factor to yield a term of the S-matrix) has the following Feynman diagram:



Read from left to right, we see that two of the center vertex's outgoing lines have been attached to one another, one has been attached to the incoming $\widehat{a}^\dagger_{\vec{p}}$ line, and one has been attached to the outgoing $\widehat{a}_{\vec{q}}$ line.

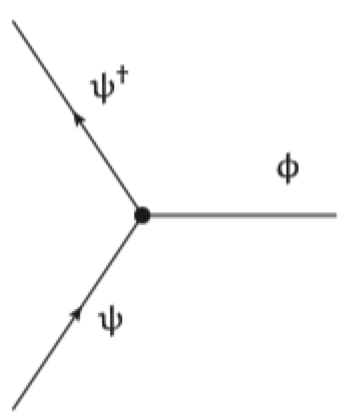
Here's a more interesting theory with one complex scalar field ψ and one real scalar field ϕ interacting with each other:

$$\mathcal{L} = \partial^\mu \psi^\dagger \partial_\mu \psi - m^2 \psi^\dagger \psi + \frac{1}{2}\partial^\mu \phi \partial_\mu \phi - \frac{1}{2}\mu^2 \phi^2 - g\psi^\dagger \psi \phi$$

The interaction part is given by $-g\psi^\dagger\psi\phi$, known as a **Yukawa interaction**. This theory displays psions with annihilation and creation operators $\widehat{a}_{\vec{p}}, \widehat{a}^\dagger_{\vec{p}}$, antipsions with operators $\widehat{b}_{\vec{p}}, \widehat{b}^\dagger_{\vec{p}}$, and phions with operators $\widehat{c}_{\vec{p}}, \widehat{c}^\dagger_{\vec{p}}$. What's the likelihood that one psion goes in with momentum $\vec{p}$ and one psion comes out with momentum $\vec{q}$? It is proportional to $\langle 0|\widehat{a}_{\vec{q}}\widehat{S}\widehat{a}^\dagger_{\vec{p}}|0\rangle$, so our expansion looks like

$$\langle 0|\widehat{a}_{\vec{q}}\widehat{S}\widehat{a}^\dagger_{\vec{p}}|0\rangle = \langle 0|\widehat{a}_{\vec{q}}\widehat{a}^\dagger_{\vec{p}}|0\rangle + \sum_{n=1}^{\infty} \frac{(-i)^n}{n!} g^n \int \langle 0|T\left[\widehat{a}_{\vec{q}}\left(\prod_{m=1}^{n} \widehat{\psi}^\dagger(x_m)\widehat{\psi}(x_m)\widehat{\phi}(x_m)\right)\widehat{a}^\dagger_{\vec{p}}\right]|0\rangle \prod_{m=1}^{n} d^4x_m$$

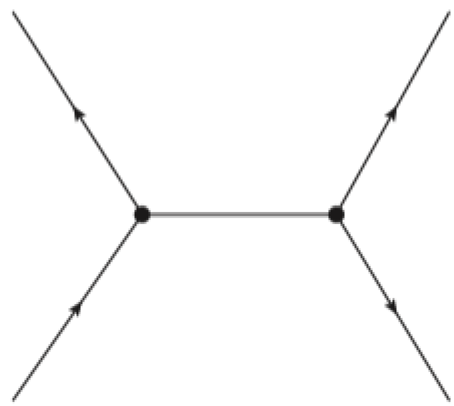
Each vertex looks like



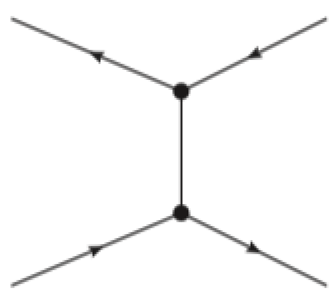
where we have drawn psions with arrows going forwards in time and antipsions with arrows going backwards in time. A more complicated scenario: what is the probability that a psion and antipsion with momenta $\vec{p}_1, \vec{p}_2$ will become psions and antipsions with momenta $\vec{q}_1, \vec{q}_2$? This probability, which is calculated via decompositions of

$$\langle 0|\widehat{b}_{\vec{q}_1}\widehat{a}_{\vec{q}_2}\widehat{\psi}^\dagger(x)\widehat{\psi}(x)\widehat{\phi}(x)\widehat{\psi}^\dagger(y)\widehat{\psi}(y)\widehat{\phi}(y)\widehat{a}^\dagger_{\vec{p}_2}\widehat{b}^\dagger_{\vec{p}_1}|0\rangle$$

has several interesting factors, one of which is represented by the following diagram:



We think of this as a psion and antipsion meeting at the left vertex and annihilating one another to produce a phion, which travels for a bit before becoming another psion and antipsion. Here is another term:



Here, a psion and antipsion travel independently of another, until one fires a phion at the other, changing the momenta of both particles.

Functional Integration Consider a one-dimensional quantum particle moving from point $x_a = (t_a, \vec{x}_a)$ to $x_b = (t_b, \vec{x}_b)$. The propagator for this particle is given by $G = \langle \vec{x}_b|\widehat{U}(t_b, t_a)|\vec{x}_a\rangle =$

$\langle\vec{x}_b|e^{-\frac{i}{\hbar}\widehat{H}(t_b-t_a)}|\vec{x}_a\rangle$. We may write $\widehat{H} = \widehat{p}^2/2m + \widehat{V}$; while the action of $\widehat{p}$ on a momentum state is less clear, $e^{-\frac{i}{\hbar}\widehat{V}}|\vec{x}\rangle$ is just $e^{-\frac{i}{\hbar}V(\vec{x})}$. Since $\widehat{U}(t_c - t_b)\widehat{U}(t_b - t_a) = \widehat{U}(t_c - t_a)$, we may take a partition $(x_a = x_0, x_1, \ldots, x_{N-1}, x_N = x_b)$, where $t_k - t_{k-1} = \Delta t$, and write $G = \langle\vec{x}_b|e^{-\frac{i}{\hbar}\widehat{H}\Delta t} \cdot \ldots \cdot e^{-\frac{i}{\hbar}\widehat{H}\Delta t}|\vec{x}_a\rangle$. Suppressing $\hbar$ and inserting resolutions of the unity between each mini-operator $\widehat{1} = \int|\vec{x}_k\rangle\langle\vec{x}_k|d\vec{x}_k$, we obtain

$$G = \langle\vec{x}_b|e^{-i\widehat{H}\Delta t}\left(\int|\vec{x}_{N-1}\rangle\langle\vec{x}_{N-1}|\,d\vec{x}_{N-1}\right)e^{-i\widehat{H}\Delta t}\ldots e^{-i\widehat{H}\Delta t}|\vec{x}_a\rangle = \left(\int d\vec{x}_{N-1}\ldots d\vec{x}_1\right)\prod_{k=1}^{N-1}\langle\vec{x}_k|e^{-i\widehat{H}\Delta t}|\vec{x}_{k-1}\rangle$$

We evaluate each term of the product as

$$\langle\vec{x}_k|e^{-i\widehat{H}\Delta t}|\vec{x}_{k-1}\rangle = \langle\vec{x}_k|e^{-i(\widehat{p}^2/2m)\Delta t}|\vec{x}_{k-1}\rangle e^{-iV(\vec{x}_{k-1})\Delta t} = \langle\vec{x}_k|e^{-i(\widehat{p}^2/2m)\Delta t}\left(\int|\vec{p}\rangle\langle\vec{p}|\vec{x}_{k-1}\rangle\,d\vec{p}\right)$$

$$= e^{-iV(\vec{x}_{k-1})\Delta t}\int\frac{d\vec{p}}{\sqrt{2\pi}}\langle\vec{x}_k|e^{-i(\widehat{p}^2/2m)\Delta t}|\vec{p}\rangle\langle\vec{p}|\vec{x}_{k-1}\rangle = e^{-iV(\vec{x}_{k-1})\Delta t}\int e^{-i(p^2/2m)\Delta t}\langle\vec{x}_k|\vec{p}\rangle\langle\vec{p}|\vec{x}_{k-1}\rangle\,d\vec{p}$$

$$= e^{-iV(\vec{x}_{k-1})\Delta t}\int e^{-i(p^2/2m)\Delta t}\frac{e^{i\vec{p}\cdot\vec{x}_k}}{\sqrt{2\pi}}\frac{e^{-i\vec{p}\cdot\vec{x}_{k-1}}}{\sqrt{2\pi}}\,d\vec{p} = e^{-iV(\vec{x}_{k-1})\Delta t}\int e^{-i(p^2/2m)2\Delta t+i\vec{p}\cdot(\vec{x}_k-\vec{x}_{k-1})}\frac{d\vec{p}}{2\pi}$$

Evaluated exactly by completing the square and comparing to the 1-dimensional Gaussian integral, we get

$$G = \int\exp\left[\sum_{k=1}^{N-1}\frac{i}{\hbar}\Delta t\left(\frac{m}{2}\frac{(\vec{x}_k-\vec{x}_{k-1})^2}{(\Delta t)^2} - V(\vec{x}_{k-1})\right)\right]d\vec{x}_1\ldots d\vec{x}_{N-1}$$

Taking $N \to \infty$, the partition $(\vec{x}_0 = \vec{x}_a, \ldots, \vec{x}_N = \vec{x}_b)$ becomes a *trajectory* $x(t)$ and the sum becomes an integral:

$$G = \int e^{\frac{i}{\hbar}\int\frac{1}{2}m\dot{x}(t)-V(x(t))\,dt}\,\mathcal{D}x = \int e^{\frac{i}{\hbar}S[x]}\,\mathcal{D}x$$

where the **integration measure** $\mathcal{D}x$ is defined as the limit as $N \to \infty$ of the product $\frac{d\vec{x}_k}{\xi}$ for some constant ξ, whose purpose is to keep things from blowing up. This integral iterates over all paths from x_a to x_b, and is hence known as the **path integral**. Each trajectory makes an infinitesimal contribution $e^{\frac{i}{\hbar}S}$, and the interference between contributions leads to a single propagator value. Since the probability is a function of the absolute value of the propagator, the portion of trajectory space that makes the largest contribution to the probability of a specific observation is simply the portion where S changes the least. For very small $\hbar$, small changes in S over a certain portion of trajectory space will lead to massive destructive interference, zeroing out the contribution from that portion; hence, the propagator will *nearly* behave as though it were following the principle of least action $\delta S = 0$, but will display small quantum contributions.

To calculate this, compare its form to that of the n-dimensional Gaussian integral

$$\int e^{-\frac{1}{2}\vec{x}^T A\vec{x}+\vec{b}^T\vec{x}}\, d\vec{x} = \sqrt{\frac{(2\pi)^N}{\det A}}e^{-\frac{1}{2}e^{\vec{b}^T A^{-1}\vec{b}}}$$

As $N \to \infty$, the vector $m \mapsto \vec{x}_m$ becomes a function $t \mapsto x(t)$, and matrices $m, n \mapsto A_{mn}$ become functions $s, t \mapsto \widehat{A}(s,t)$, $(\widehat{A}f)(s) = \int A(s,t)f(t)\, dt$. The dot product of vectors becomes an integral of functions, and the inverse matrix A^{-1} becomes an inverse kernel $\int \widehat{A}(s,t)\widehat{A}^{-1}(t,u)\, dt = \delta(s-u)$, i.e. a Green's function of $\widehat{A}$. The determinant of $\widehat{A}$ remains the product of its eigenvalues; although this may diverge, dividing it by a certain anti-blowing-up constant ξ as $N \to \infty$ keeps things from blowing up. In general, writing f^T for its corresponding functional $f^T g = \int f(x)g(x)\, dx$, we write

$$\int e^{-\frac{1}{2}\int \phi(x)\widehat{A}(x,y)\phi(y)\, dx\, dy+\int b(x)\phi(x)\, dx}\mathcal{D}\phi = \int e^{-\frac{1}{2}\phi^T\widehat{A}\phi+b^T\phi}\mathcal{D}\phi = \sqrt{\frac{(2\pi)^N}{\det A}}e^{-\frac{1}{2}b^T\widehat{A}^{-1}b}$$

Given a theory, we define a generating functional known as the **partition function** Z, acting on a function J which we interpret as an operator $(J\phi)(x) = J(x)\phi(x)$, as

$$Z[J] = \int e^{i\int \mathcal{L}[\phi(x)]+(J\phi)(x)\, d^4x}\mathcal{D}\phi$$

The **normalized partition function** $\mathcal{Z}[J]$ is given by $Z[J]/Z[J=0]$, or

$$\frac{\int e^{i\int \mathcal{L}[\phi(x)]+(J\phi)(x)\, d^4x}\mathcal{D}\phi}{\int e^{i\int \mathcal{L}[\phi(x)]\, d^4x}\mathcal{D}\phi}$$

We will generally find that the denominator (which is often written $Z_0[J]$) cancels out the quantities that fail to converge. For instance, consider the free scalar field theory $\mathcal{L} = \frac{1}{2}\partial_\mu\phi\partial^\mu\phi -$

$\frac{1}{2}m^2\phi^2$. First, we calculate the denominator: we can use integration by parts to write

$$\int \frac{1}{2}\partial_\mu\phi\partial^\mu\phi \, d^4x = -\int \frac{1}{2}\phi\partial^2\phi$$

and hence

$$\int e^{i\int \mathcal{L}[\phi(x)]\, d^4x} \, \mathcal{D}x = \int e^{\frac{i}{2}\int \phi\left[-(\partial^2+m^2)\right]\phi \, d^4x} \, \mathcal{D}x$$

$$= \sqrt{\frac{(2\pi)^N}{\det\left[-(\partial^2 + m^2)\right]}}$$

This quantity, which makes up the denominator, will also appear in the numerator, and is therefore cancelled out, leaving only the component $e^{-\frac{1}{2}J\left[-(\partial^2+m^2)\right]^{-1}J}$. We can identify $\left[-(\partial^2 + m^2)\right]^{-1}$ with $-i\Delta(x, y)$, where $\Delta(x, y)$ is the free propagator. Hence, the normalized partition function for the free scalar theory is given by

$$\mathcal{Z}[\widehat{J}] = e^{-\frac{1}{2}\int J(x)\Delta(x-y)J(y)\, d^4x \, d^4y}$$

In general, we may calculate correlation functions as

$$\langle\Omega|T\,\phi(x_1)\cdot\ldots\cdot\phi(x_n)|\Omega\rangle = \frac{1}{Z_0}\int \phi(x_1)\cdot\ldots\cdot\phi(x_n)e^{\frac{i}{\hbar}S[\phi]}\,\mathcal{D}\phi$$

Writing $\langle\phi(x_1)\cdot\ldots\cdot\phi(x_n)\rangle \coloneqq \langle\Omega|T\,\phi(x_1)\cdot\ldots\cdot\phi(x_n)|\Omega\rangle$, we may equivalently express this as

$$\langle\phi_1(x)\cdot\ldots\cdot\phi_n(x)\rangle = \frac{(-i)^n}{Z_0}\frac{\delta^n Z}{\delta J(x_1)\cdot\ldots\cdot\delta J(x_n)}$$

We see that knowledge of the partition function gives us knowledge of all correlation functions, which is more or less all we want to know about a given theory.

Path integration offers another approach to perturbative expansions: let $\mathcal{L} = \mathcal{L}_0$ minus some interaction term $\mathcal{L}_I$. We have

$$e^{i\int \mathcal{L}\, dx^4} = e^{i\int \mathcal{L}_0\, d^4x}\left(1 - i\int \mathcal{L}_I[\phi(x)]\, d^4x + (-i)^2\int \mathcal{L}_I[\phi(x)]\mathcal{L}_I[\phi(y)]\, d^4x\, d^4y + \ldots\right)$$

which, when combined with the above equation, yields another way to expand S-matrix terms.

2.5.3 Representations of the Lorentz Group

Recall that the distance between two points x^μ, y^μ of Minkowski space X is given by

$$\left(\eta_{\mu\nu}x^\mu y^\nu\right)^{1/2} = \sqrt{(x^0-y^0)^2-(x^1-y^1)^2-(x^2-y^2)^2-(x^3-y^3)^2}$$

An **isometry** of Minkowski space is a continuous map $X \to X$ preserving the distance between points; the set of all such isometries is a Lie group known as the **Poincaré group**. It is ten-dimensional, with 4 dimensions dedicated to translations, three to rotations (x-y, x-z, y-z), and three to boosts, or rotations involving the time dimension (t-x, t-y, t-z).

Discarding the translations gives us a six-dimensional Lie group known as the **Lorentz group** $L = O(1,3)$; its objects are all linear maps, and hence can be written as matrices $\Lambda^\mu{}_\nu$ satisfying

$$\eta_{\mu\nu}\Lambda^\mu{}_\sigma\Lambda^\nu{}_\rho x^\sigma y^\rho = \eta_{\mu\nu}x^\mu y^\nu$$

In matrix notation, such a Λ satisfies $x^T\eta y = (\Lambda x)^T\eta(\Lambda y)$ for all x, y, and hence $\Lambda^T\eta\Lambda = \eta$. It follows that $\det(\Lambda^T\eta\Lambda) = -(\det\Lambda)^2 = \det\eta = -1$, so that $\det\Lambda \in \pm 1$. Also, letting $e^0 = (1,0,0,0)$, we have

$$1 = (e^0)^T\eta(e^0) = (\Lambda e^0)^T\eta(\Lambda e^0) = (\Lambda^0{}_0)^2 - (\Lambda^1{}_0)^2 - (\Lambda^2{}_0)^2 - (\Lambda^3{}_0)^2$$

so that $(\Lambda^0{}_0)^2 \geq 1$, implying that either $\Lambda^0{}_0 \geq 1$ or $\Lambda^0{}_0 \leq 1$. It follows that L is composed of four connected components, each consisting of all transformations Λ with a specified determinant and sign of $\Lambda^0{}_0$. We write these components as

$$L_+^\uparrow = \{\Lambda \in L \mid \det\Lambda = 1, \Lambda^0{}_0 \geq 1\} \qquad L_-^\downarrow = \{\Lambda \in L \mid \det\Lambda = -1, \Lambda^0{}_0 \leq 1\}$$

and likewise for $L_-^\uparrow, L_+^\downarrow$. $L_+^\uparrow$, which contains the identity, is often known as the **restricted** or **proper orthochronous** Lorentz group, $SO^+(1,3)$. Defining the space inversion and time reversal operators $P = \mathrm{diag}(+1,-1,-1,-1)$ and $T = \mathrm{diag}(-1,+1,+1,+1)$ gives the structure of the Klein four-group $\{I_4, P, T, PT\}$ to these four connected components.

Since the exponentiation operator e^- from a Lie algebra $\mathfrak{g}$ to its Lie group G is continuous, and hence has an image contained in one connected component, $\mathfrak{g}$ depends solely on the special component of G containing the identity. Thus, the Lie algebras of $L = O(1,3), SO(1,3)$, and $L_+^\uparrow = SO^+(1,3)$ are all the same. This algebra is generally written as $\mathfrak{so}(1,3)$.

Fix a Lie group G and Lie algebra $\mathfrak{g}$. A **Lie group representation** of G is a smooth homomorphism $\Pi : G \to GL(n;\mathbb{C})$ for some n. A **Lie algebra representation** of $\mathfrak{g}$ is a Lie algebra homomorphism $\pi : \mathfrak{g} \to \mathfrak{gl}(n; \mathbb{C}) \cong \mathrm{End}(\mathbb{C}^n)$. Since the Lie algebra of a Lie group is the tangent space to its identity, the pushforward of any Lie group representation defines a homomorphism between Lie algebras; this homomorphism preserves brackets, so that Lie group representations induce Lie algebra representations. If $\mathfrak{g}$ is the Lie algebra of G, it isn't true in general that (Lie algebra) representations of $\mathfrak{g}$ come from (Lie group) representations of G, but, if G is connected, we may find a group G_1 fitting into a short exact sequence of groups

$$1 \longrightarrow \pi_1(G) \longrightarrow G_1 \xrightarrow{\varphi} G \longrightarrow 1$$

known as the **universal covering group** of G. Representations of $\mathfrak{g}$ are in bijection with representations of G_1 rather than G.

Define the 2×2 Hermitian **Pauli matrices** as

$$\sigma^0 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \quad \sigma^1 = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \quad \sigma^2 = \begin{bmatrix} 0 & -i \\ i & 0 \end{bmatrix} \quad \sigma^3 = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$$

(Generally, σ^0 is omitted, giving us three Pauli matrices). These obviously span the space $H(2, \mathbb{C})$ of 2×2 Hermitian matrices, and in fact we have a pair of isomorphisms $\underset{\sim}{-}, \tilde{-} : \mathbb{R}^4 \to H(2, \mathbb{C})$ defined by $\underset{\sim}{x} = \delta_{\mu\nu} x^\mu \sigma^\nu$, $\tilde{x} = \eta_{\mu\nu} x^\mu \sigma^\nu$. We can computationally verify that $\det \underset{\sim}{x} = \det \tilde{x} = x \cdot x$, and $\tilde{x}\underset{\sim}{x} = \underset{\sim}{x}\tilde{x} = (x \cdot x) I_2$. It follows that, for an arbitrary determinant 1 complex matrix A, the linear map $\varphi(A)(x) = (\underset{\sim}{-})^{-1}(A \underset{\sim}{x} A^\dagger)$ defines a homomorphism $\varphi : SL(2; \mathbb{C}) \to L$; in fact, we can show that it is a surjection $SL(2; \mathbb{C}) \to L_+^\uparrow$ with kernel $\varphi^{-1}(I_4) = \{\pm I_2\} \cong \mathbb{Z}/2\mathbb{Z}$.

Topologically, $L_+^\uparrow$ is equivalent to $\mathbb{R}^3 \times SO(3)$, and therefore $\pi_1(L_+^\uparrow) = \pi_1(SO(3)) = \mathbb{Z}/2\mathbb{Z}$. It follows that the homomorphism $\varphi : SL(2; \mathbb{C}) \to L_+^\uparrow$ fits into a short exact sequence

$$1 \to \mathbb{Z}/2\mathbb{Z} \to SL(2; \mathbb{C}) \to L_+^\uparrow \to 1$$

evidencing $SL(2; \mathbb{C})$ as the universal covering group of $L_+^\uparrow$.

Given a Lie group or algebra representation M, a subspace V of $\mathbb{C}^n$ mapped into itself by all $\Pi(g)$ is known as **invariant**; $\{\vec{0}\}$ and $\mathbb{C}^n$ are trivially invariant, but any representation with no nontrivial invariant subspaces is known as **irreducible**. Every representation of $SL(2; \mathbb{C})$ decomposes as the direct sum of irreducible representations, i.e. $\Pi(g) = \Pi_1(g) \oplus \Pi_2(g) \oplus \ldots \oplus \Pi_k(g)$ with each $\Pi_j(g)$ an $n_j \times n_j$ matrix, where $\sum_{j=1}^k n_j = n$. We define a pair of representations $\Pi^{(1/2,0)}, \Pi^{(0,1/2)} : SL(2; \mathbb{C}) \to GL(2; \mathbb{C})$ given by

$$\Pi^{(1/2,0)}(A) = A \qquad \Pi^{(0,1/2)}(A) = (A^\dagger)^{-1}$$

For $\mu, \nu \in \{0, 1/2, 1, 3/2, \ldots\}$, we define $\Pi^{(\mu,\nu)} : \mathrm{SL}(2;\mathbb{C}) \to \mathrm{GL}(4^{\mu+\nu};\mathbb{C})$ by

$$\Pi^{(\mu,\nu)}(A) = \left(\bigotimes_{i=1}^{2\mu} \Pi^{(1/2,0)}(A)\right) \otimes \left(\bigotimes_{i=1}^{2\nu} \Pi^{(0,1/2)}(A)\right)$$

The $\Pi^{(\mu,\nu)}$ are the irreducible representations of $\mathrm{SL}(2;\mathbb{C})$. Every irreducible representation of the Lorentz algebra can be recovered as the pushforward of some $\Pi^{(\mu,\nu)}$, which we denote $\pi^{(\mu,\nu)}$. Under an infinitesimal Lorentz transformation, or an element $g \in \mathfrak{so}(1,3)$, an n-component complex field $\Phi = (\phi_1, \ldots, \phi_n)$ described by a Lorentz covariant theory must experience an infinitesimal change described by a matrix $M(g) \in \mathfrak{gl}(n;\mathbb{C})$, where M is a representation of $\mathfrak{so}(1,3)$ and thus decomposes as $M = \bigoplus_{i=1}^{k} \pi^{(\mu_i,\nu_i)}$. The largest $\mu_i + \nu_i$ is known as the **spin** of Φ.

Spinors The Lorentz algebra is a 6-dimensional vector space, with three rotation dimensions and three boost dimensions. It is spanned by the set $J^{\mu\nu} = i(x^\mu \partial^\nu - x^\nu \partial^\mu)$ of tangent vectors (since $J^{\mu\nu} = -J^{\nu\mu}$, there really are only six), and satisfy the commutation relations

$$[J^{\mu\nu}, J^{\rho\sigma}] = i(\eta^{\nu\rho} J^{\mu\sigma} - \eta^{\mu\rho} J^{\nu\sigma} - \eta^{\nu\rho} J^{\mu\rho} + \eta^{\mu\sigma} J^{\nu\rho})$$

Any set of six $n \times n$ matrices $S^{\mu\nu}$ satisfying the same commutation relations (in particular, $[S^{\mu\nu}, S^{\nu\mu}] = 0$, so that $S^{\nu\mu} = -S^{\mu\nu}$) defines a Lie algebra homomorphism $\mathfrak{so}(1,3) \to \mathfrak{gl}(n;\mathbb{C})$, and hence a representation of the Lorentz algebra.

Any set of four $n \times n$ matrices γ^μ γ^μ such that $\gamma^\mu\gamma^\nu + \gamma^\nu\gamma^\mu = 2\eta^{\mu\nu} I_n$ yields a set of matrices $S^{\mu\nu} = \frac{i}{4}[\gamma^\mu, \gamma^\nu]$ satisfying these relations. One such set of **gamma matrices** is given in block diagonal form by

$$\gamma_0 = \begin{bmatrix} 0 & I_2 \\ I_2 & 0 \end{bmatrix} \quad \gamma_i = \begin{bmatrix} 0 & \sigma^i \\ -\sigma^i & 0 \end{bmatrix}$$

This yields matrices

$$S^{0i} = -\frac{i}{2} \begin{bmatrix} \sigma^i & 0 \\ 0 & -\sigma^i \end{bmatrix} \quad S^{ij} = \frac{1}{2} \varepsilon^{ijk} \begin{bmatrix} \sigma^k & 0 \\ 0 & \sigma^k \end{bmatrix}$$

and, for a family of scalars $c_{\mu\nu}$, gives the representation $c_{\mu\nu} J^{\mu\nu} \mapsto c_{\mu\nu} S^{\mu\nu}$, known as the **chiral representation**. This representation decomposes as $\pi^{(1/2,0)} \oplus \pi^{(0,1/2)}$; complex 2-dimensional vector fields transforming according to $\pi^{(1/2,0)}$ and $\pi^{(0,1/2)}$ are known as the **left-handed** and **right-handed Weyl spinors**, whereas a 4-dimensional complex vector field transforming according to $\pi^{(1/2,0)} \oplus \pi^{(0,1/2)}$ is known as a **Dirac spinor**.

The most notable property of Dirac spinors is their behavior under rotations: consider for instance the action of an infinitesimal θ degree rotation in the xy-plane on a Dirac spinor, which we obtain by exponentiating its representation:

$$\begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & \theta i & 0 \\ 0 & -\theta i & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \mapsto \begin{bmatrix} e^{i\theta/2} & 0 & 0 & 0 \\ 0 & e^{-i\theta/2} & 0 & 0 \\ 0 & 0 & e^{i\theta/2} & 0 \\ 0 & 0 & 0 & e^{-i\theta/2} \end{bmatrix}$$

Under a full $360° = 2\pi$ revolution, a Dirac spinor doesn't return to its original state, but picks up a minus sign; it takes a $720° = 4\pi$ rotation to return the spinor to its original state. Dirac spinor fields are spin 1/2 fields, as opposed to scalar fields, which transform under the trivial representation of the Lorentz algebra and are hence spin 0. In general, a spin $n > 0$ field requires a $2\pi/n$ degree rotation to return to its original state; spin 0 fields are invariant under any rotation.

2.5.4 The Dirac Field

While Dirac spinors are four-component vectors, they will be treated analogously to the scalars seen in previous field theories: we will generally not give them indices. Consequently, four-component vectors of four-component vectors, or 4×4 matrices, will have one index. To refer to the space-like components, or the in the case of matrices the latter three components, though, we may still use vector notation (or, in the case of ∂, $\nabla = (\partial^1, \partial^2, \partial^3)$). For a four-component object x_μ, we write the contraction $\gamma^\mu x_\mu$ as $\not{x}$; note that $\not{x}^2 = \gamma^\mu\gamma^\nu x_\mu x_\nu = \frac{1}{2}(\gamma^\mu\gamma^\nu + \gamma^\nu\gamma^\mu)x_\mu x_\nu$ (because we are summing over all μ, ν) $= \eta^{\mu\nu}x_\mu x_\nu = x^2$.

A **Dirac field** is a Dirac spinor field ψ with Lagrangian

$$\overline{\psi}(i\not{\partial} - m)\psi = 0$$

where $\overline{\psi} = \psi^\dagger\gamma^0$, $m = mI_4$, and ∂_μ acts on ψ coordinate-wise. The Euler-Lagrange equation for ψ yields the **Dirac equation**

$$(i\not{\partial} - m)\psi = 0$$

where $m = mI_4$. It follows that $(-i\not{\partial} - m)(i\not{\partial} - m)\psi = (\not{\partial}^2 + m^2)\psi = (\partial^2 + m^2)\psi = 0$, so that the Dirac equation implies the Klein-Gordon equation in each coordinate. The Hamiltonian is given by $\mathcal{H} = -\overline{\psi}(i\vec{\gamma}\cdot\nabla - m)\psi$, so the conjugate momentum of ψ is given by $\Pi^{\mu}_{\psi} = i\overline{\psi}\gamma^{\mu}$ and the conjugate momentum of $\overline{\psi}$ is given by $\Pi^{\mu}_{\overline{\psi}} = 0$.

Splitting ψ into left-handed and right-handed Weyl spinor fields as $\psi = (\psi_L, \psi_R)$, or equivalently by separating it into eigenvalues of the **chirality operator** $\gamma^5 = i\gamma^0\gamma^1\gamma^2\gamma^3 = \begin{bmatrix} -I_2 & 0 \\ 0 & I_2 \end{bmatrix}$, we see that the mass operator leaves ψ_L and ψ_R in their place, whereas gamma operators switch them. In general, this causes the two fields to interact with one another, but when $m = 0$, they do not, and the Dirac equation splits into two separate equations known as the **Weyl equations**:

$$i(\partial_0 - \vec{\sigma}\cdot\nabla)\psi_L = 0 \qquad i(\partial_0 + \vec{\sigma}\cdot\nabla)\psi_R = 0$$

The solutions to the Dirac equation are given by waves of the form $\psi(x) = \begin{bmatrix} \xi\sqrt{p\cdot\sigma} \\ \xi\sqrt{p\cdot\overline{\sigma}} \end{bmatrix} e^{-ip\cdot x}$ for positive energy, and $\psi(x) = \begin{bmatrix} \eta\sqrt{p\cdot\sigma} \\ -\eta\sqrt{p\cdot\overline{\sigma}} \end{bmatrix} e^{ip\cdot x}$ for negative energy. The ξ and η forming the spinors $u(p) = \begin{bmatrix} \xi\sqrt{p\cdot\sigma} \\ \xi\sqrt{p\cdot\overline{\sigma}} \end{bmatrix}$ and $v(p) = \begin{bmatrix} \eta\sqrt{p\cdot\sigma} \\ -\eta\sqrt{p\cdot\overline{\sigma}} \end{bmatrix}$ are arbitrary, so we choose to normalize, setting $\xi^\dagger\xi = \eta^\dagger\eta = 1$. We write u^i for ξ^i, $i = 1, 2$, and likewise for v^i. We can write down some useful properties of the u^i and v^i: $u^\dagger(p)u(p) = v^\dagger(p)v(p) = 2E_{\vec{p}}$, $\sum_j u^j(p)\overline{u}^j(p) = \gamma\cdot p + m$, $\sum_j v^j(p)\overline{v}^j(p) = \gamma\cdot p - m$.

Quantization To quantize the Dirac field, we can *not* impose the equal-time commutation relation $[\psi(x), i\psi^\dagger(y)] = i\delta^{(4)}(x-y)$. The particles described by any field with half-integer spin are *fermions*, meaning that interchanging the position of any two fermions adds a negative sign to the state of the field. In particular, any state with two fermions occupying the same position in spacetime must be zero. This is in contrast to particles described by integer spin fields, such as the spin 0 Klein-Gordon equation, which can be stacked on top of one another indefinitely; these particles are known as *bosons*. Hence, we impose equal-time *anti*commutation relations,

$$\{\widehat{\psi}_j(\vec{x}), i\widehat{\psi}^\dagger_j(\vec{y})\} = i\delta^{(3)}(\vec{x}-\vec{y})$$

where j indexes the components of each spinor.

The mode expansions for $\widehat{\psi}$ and $\widehat{\overline{\psi}}$ can be given as

$$\widehat{\psi}(x) = \int \frac{d^3p}{(2\pi)^{3/2}} \frac{1}{\sqrt{2E_{\vec{p}}}} \sum_{j=1}^{2} u^j(p)\widehat{a}_{j\vec{p}}e^{-ip\cdot x} + v^j(p)\widehat{b}^\dagger_{j\vec{p}}e^{ip\cdot x}$$

$$\widehat{\overline{\psi}}(x) = \int \frac{d^3p}{(2\pi)^{3/2}} \frac{1}{\sqrt{2E_{\vec{p}}}} \sum_{j=1}^{2} \overline{u}^j(p)\widehat{a}^\dagger_{j\vec{p}} e^{ip\cdot x} + \overline{v}^j(p)\widehat{b}_{j\vec{p}} e^{-ip\cdot x}$$

The interpretation is that $\widehat{a}^\dagger_{s\vec{p}}$ creates a fermion with momentum $\vec{p}$ and handedness given by j, whereas $\widehat{b}^\dagger_{s\vec{p}}$ creates an *anti*fermion.

Quantum Electrodynamics The Dirac equation obviously has a global U(1) symmetry, since the Dirac Lagrangian $\mathcal{L}$ remains invariant under phase shifts $\psi \mapsto \psi e^{i\alpha}$, $\alpha \in \mathbb{R}$. We're going to outline a procedure by which we can turn global symmetries of Lagrangians into local symmetries, and then analyze the Dirac Lagrangian with local U(1) invariance.

In general, given a principal G-bundle $E \xrightarrow{\pi} X$ with specified connection one-form ω, we write v^V and v^H for the restrictions of an arbitrary vector field v to its vertical and horizontal components, which satisfy $\partial i_*(v^V) = \omega(v^H) = 0$ and $v^V + v^H = v$. Given a (possibly $\mathfrak{g}$-valued) k-form η on E, we define $\eta^H(v_1, \ldots, v_k) := \eta(v_1^H, \ldots, v_k^H)$ and likewise for η^V. The **exterior covariant derivative** on the bundle with connection $(E \xrightarrow{\pi} X, \omega)$ is given by $D\eta := (d\eta)^H$. The **curvature** of the connection form ω is given by $\Omega := D\omega$. Cartan's structure equation states that $\Omega = d\omega + \frac{1}{2}[\omega, \omega]$, where $[\omega, \omega](v, w) = [\omega(v), \omega(w)] - [\omega(w), \omega(v)] = 2[\omega(v), \omega(w)]$. It follows that $d\Omega = d(d\omega + \frac{1}{2}[\omega, \omega]) = \frac{1}{2}d[\omega, \omega] = \frac{1}{2}([d\omega, \omega] - [\omega, d\omega]) = [d\omega, \omega]$. Since $[[\omega, \omega], \omega] = 0$, we can write $d\omega = [\Omega, \omega]$. A locally U(1) invariant version of the Dirac equation, in which E is spinors and X is spacetime, has $d\psi = \partial_\mu\psi = (d\psi)^H + (d\psi)^V = D\psi + (d\psi)^V$. Hence, the gauge covariant derivative $D_\mu\psi$ differs from ∂_μ by a one-form: we will write $D_\mu\psi = \partial_\mu\psi + iqA_\mu\psi$, where q is a constant and A_μ is known as the **gauge field**, transforming under a shift $\psi \mapsto \psi e^{i\alpha}$ as $A_\mu \mapsto A_\mu - \frac{1}{q}\partial_\mu\alpha$.

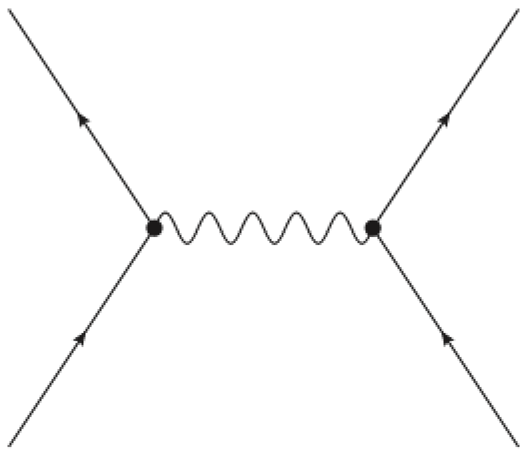
To make the Dirac equation as we know it *locally* U(1) invariant, we will simply make the derivative covariant, replacing ∂_μ with $D_\mu = \partial_\mu + iqA_\mu$. This gives us a U(1) gauge theory $\mathcal{L} = \overline{\psi}(i\not{D} - m)\psi = \overline{\psi}(i\not{\partial} - m)\psi - q\overline{\psi}\not{A}\psi$. In order to use this to model electromagnetism, we simply *add* the Lagrangian of classical electromagnetism, obtaining a Lagrangian

$$\mathcal{L} = -\frac{1}{4}F_{\mu\nu}F^{\mu\nu} + \overline{\psi}(i\not{D} - m)\psi$$

Note that $F_{\mu\nu} = dA_\mu$, so that this is a restriction on the gauge field itself. Hence, A_μ serves two purposes: it both enforces local U(1) invariance and serves as an electromagnetic current.

$\mathcal{L} = -\frac{1}{4}F_{\mu\nu}F^{\mu\nu} + \overline{\psi}(i\not{D} - m)\psi$ is the Lagrangian of **quantum electrodynamics**. The current density is recovered from the interacting part as $J^\mu = \overline{\psi}\gamma^\mu\psi$. ψ creates fermions (electrons),

$\overline{\psi}$ creates antifermions (positrons), and A_μ is a massless boson (photon) field interacting with electrons via the interaction term $\mathcal{L}_I = -q\overline{\psi}\not{A}\psi$. S-matrix terms see photons interacting with pairs of electrons and fermions, creating many of the same interactions seen in the previously encountered Yukawa interaction theory: the diagram



encountered in evaluating the amplitude of an electron and positron yielding an electron and positron, represents a process whereby the electron and positron annihilate, yielding a photon, which photon then transforms into an electron-positron pair. In evaluating this term we integrate over all possible photon momenta, so this photon, which clearly cannot be observed by experiment, can have arbitrary mass; it is said to be a **virtual photon**, as it cannot and should not exist as an actual photon.

Chapter 3

Synthetic Differential Geometry

In synthetic differential geometry, we develop geometry from an axiomatic point of view. This is done using categories with enough structure to discuss the notions of smoothness and infinitesimality fundamental to geometry, namely *elementary topoi*. Such categories have their own internal logic, and we can add axioms in order to enforce certain properties on our smooth objects. We will first introduce elementary topoi, the universes in which synthetic differential geometry takes place; our exposition follows the sources [MacLane and Moerdijk, 2012, Johnstone, 2014]. We have also relied on the sources [Moerdijk and Reyes, 2013, Kock, 2006, Kostecki, 2009] in discussing synthetic differential geometry itself.

3.1 Grothendieck Topologies

A Grothendieck topology is a way of generalizing the structure a topology provides to a set X to arbitrary categories. This section is based largely off of [MacLane and Moerdijk, 2012], with additional topos theoretic details filled in from the works by Johnstone [Johnstone, 2014, Johnstone, 2002].

3.1.1 Subobjects

In many concrete categories, monomorphisms can be interpreted as inclusions.

- In $\mathsf{R}\text{-}\mathsf{Mod}$, for instance, a monomorphism $M \to N$ evidences the image of M, which is isomorphic to M itself, as a submodule of M.

- In Set, monomorphisms $X \to Y$ are simply injective functions, and can be interpreted as subset inclusions.
- In Top, monomorphisms are continuous inclusions, and their images are subspaces.

Clearly, interpreting monomorphisms as inclusions gives us a notion of "subobject" in each of the above categories, and we would like to generalize this to arbitrary categories. We must take care to identify monomorphisms yielding the same subobject with one another, though, by setting up the appropriate equivalence relation.

We define a **subobject** of an object X in a category C to be an equivalence class of monomorphisms $Y \to X$, where two monomorphisms $f : Y' \to X$ and $g : Y'' \to X$ are identified if there is a pair of maps $h : Y' \to Y''$ and $k : Y'' \to Y'$ such that g factors as fk and h factors as gh. When the two monomorphisms factor through one another, we consider them to be the "same" as inclusions, and hence the same subobject.

Let's take this definition for a test drive: given a finite set X of cardinality n, say $X = \{0, 1, \ldots, n-1\}$, any injection $f : Y \to X$, where Y is by necessity isomorphic to the m-element set $\{0, \ldots, m-1\}$, $0 \leq m \leq n$, is determined as a subobject uniquely by its image in X: if this image consists of p different elements, we set $Y' = \{0, \ldots, p-1\}$ and let $g : Y' \to X$ send 0 to the smallest element in the image, ..., $p-1$ to the largest. We can factor g through f by sending $k \in Y'$ to any element of $f^{-1}(g(k))$ and factor f through g likewise. It follows that there's one subobject for every possible *image* of a monomorphism into X; these are in bijection with its subsets, so subobjects in Set are subsets.

If we have two monomorphisms $f : Y \to X$, $g : Y' \to X$ such that f factors through g but not necessarily vice versa, then we say that $f \subseteq g$, or that f contains g. $\subseteq$ is compatible with the previous equivalence relation on monomorphisms, and hence makes the collection of subobjects of an object X in a category C, denoted by $\mathrm{Sub}_C(X)$, into a poset. Posets are categories in natural ways, and the properties of these categories tell us much about subobjects. Abusing notation to write the subobject associated to a monic $U \to X$ simply as U, and writing $U \subseteq X$ for the subobject inclusion, the product of two subobjects $U, V \subseteq X$ is the smallest subobject that is not only contained in both U and V, but contains any other subobject that is also contained in U and V; this can be interpreted variously as the *greatest common denominator*, *meet*, or *intersection*. The coproduct is the *least common multiple*, *join*, or *union*.

Subobject Classifiers In Set, a subset $S \subset X$ is equivalent to a choice over all $x \in X$ as to whether $x \in S$ or not, i.e. a map $X \to 2$. This interpretation defines a **characteristic function** $\chi_S : X \to 2 = \{0, 1\}$ sending $x \in X$ to 0 if $x \notin S$ and 1 if $x \in S$. We've made the choice of using the subobject $1 = \{0\} \subset 2$ to capture truth, which choice we encode by a monomorphism $\mathsf{true} : 1 \to 2, 0 \mapsto 1$. This allows us to express S as the pullback of true along χ_S. In Set, we can give this pullback explicitly: it is

$$\{(x, b) \in X \times 1 \mid \chi_S(x) = \mathsf{true}(b)\} = \{x \in X \mid \chi_S(x) = 1\}$$

It follows that χ_S is the *unique* function $X \to \Omega$ which yields S upon taking the pullback.

Generalizing, in a category C with finite limits, a **subobject classifier** is an object Ω along with a monomorphism $\mathsf{true} : 1 \to \Omega$ such that every monic $f : S \to X$ admits a unique $\chi_f : X \to \Omega$ such that S is the pullback of true along χ_f. That finite limits exist means that an arbitrary morphism $f : X \to \Omega$ yields by pullback a morphism $f' : X \times_\Omega 1 \to X$: it is in general true that monomorphisms pull back to monomorphisms, so that f' is a monomorphism and hence defines a subobject of X which is in turn classified uniquely by f. When C is locally small, this yields an isomorphism $\mathrm{Sub}_C(X) \cong C(X, \Omega)$ between subobjects of X and morphisms $X \to \Omega$.

Sieves For an arbitrary small category C, we may define a **subfunctor** of a functor $F : C \to D$ as a subobject of F in the category D^C; in the case that F is a presheaf $C^{op} \to \mathsf{Set}$, or an object in the presheaf category $\widehat{C} = \mathsf{Set}^{C^{op}}$, a subfunctor G of F is another presheaf $C^{op} \to \mathsf{Set}$ such that $GX \subseteq FX$ for all X, and each $Gf : GY \to GX$ induced by an $f : X \to Y$ is the corresponding restriction of Ff from elements of FY to elements of GY. In $\widehat{C}$, monics are determined pointwise: a natural transformation $\alpha : F \to G$ is a monomorphism in $\widehat{C}$ iff each $\alpha_X : FX \to GX$ is a monomorphism in C. For $\widehat{C}$ to have a subobject classifier Ω, Ω must in particular classify each representable presheaf $よX = C(-, X)$. By Yoneda's lemma $\widehat{C}(C(-, X), F) \cong FX$, we have

$$\mathrm{Sub}_{\widehat{C}}(C(-, X)) \cong \widehat{C}(C(-, X), \Omega) \cong \Omega X$$

This means that, when $\widehat{C}$ has a subobject classifier, its action as a functor is to send an object X to the set of subfunctors of $C(-, X)$. Its action on a morphism $f : X \to Y$ is to send the subfunctor S of $C(-, Y)$, which we can regard as a collection of morphisms $\{g_\lambda : Z_\lambda \to Y\}_{\lambda \in \Lambda}$, to the subfunctor $(\Omega f)(S)$ of $C(-, X)$ which, as a collection of morphisms, is the set of all morphisms h into X such that $fh \in S$. We define a **sieve** on an object $X \in C$ to be a subfunctor of $C(-, X)$, explicitly considered as a collection of morphisms into X. Given a sieve S on Y and a map $f : X \to Y$, the map $f^* := \Omega f$ sends S to, as above, the set of all morphisms h into X such that fh is in S. The functor $C(-, X)$, trivially considered as a subobject of itself, corresponds to the set of *all* arrows into X, which we write as t_X.

3.1.2 Sites

As stated previously, we can generalize the notion of a topological structure on a space X to a topological structure on arbitrary categories by studying $Op(X)$; associating to each object U of this category a "covering" of sets of functions with codomain U such that the images of the elements of any set cover U as an open set, we find that the notion of sheaf becomes a primarily categorical one.

The right generalization is given by that of a **Grothendieck topology**, or an assignment to each object X of a category C a collection $J(X)$ of sieves $\{X_\lambda \to X\}_{\lambda\in\Lambda}$ such that

1. $t_X \in J(X)$.
2. For $S \in J(X)$ and $f : Y \to X$, $f^*(S) \in J(Y)$.
3. For $S = \{f_\lambda : X_\lambda \to X\} \in J(X)$, if an arbitrary sieve S' on X satisfies $f^*_\lambda(S') \in J(X_\lambda)$ for all f_λ in S, then $S' \in J(X)$ as well.

If $S \in J(X)$, we say that S **covers** X; if $f^*(S)$ covers Y for a morphism $f : Y \to X$, then we say that S **covers** f as well. A sieve S which contains all morphisms that it covers is known as a **closed sieve**. If C has pullbacks, we can define a **basis** of a Grothendieck category to be an assignment to each $X \in C$ a set $B(X)$ of families of morphisms $\{X_\lambda \to X\}_{\lambda\in\Lambda}$ such that

1. Every singleton $\{f : Y \cong X\}$ is in $B(X)$.
2. If $\{f_\lambda : X_\lambda \to X\}_{\lambda\in\Lambda} \in B(X)$ and $g : Y \to X$, then $\{(\pi_D)_\lambda : X_\lambda \times_X Y \to Y\}_{\lambda\in\Lambda} \in B(X)$, where $(\pi_D)_\lambda$ is the pullback of f_λ along g.
3. If $\{f_\lambda : X_\lambda \to X\}_{\lambda\in\Lambda} \in B(X)$, then for any $\{g_{\lambda\xi} : X_{\lambda\xi} \to X_\lambda\}_{\xi\in\Xi_\lambda} \in B(X_\lambda)$, $\{f_\lambda \circ g_{\lambda\xi} : X_{\lambda\xi} \to X\}_{\lambda\in\Lambda,\xi\in\Xi_\lambda} \in B(X)$.

A basis $B(X)$ generates a topology $J(X)$ by the rule $S \in J(X)$ if there's an $S' \in B(X)$ contained in S.

Sites and Sheaves A category C equipped with a Grothendieck topology J is known as a **site**. The natural example is $C = Op(X)$ for some topological space X; defining $J(U)$ to be the set of open covers of U yields a Grothendieck topology known as the **classical topology**. Expanding, we can define a Grothendieck topology on Top itself known as the **big classical topology** by defining $J(X)$ to be the collections of families $\{X_i \to X \mid i \in I\}$ of open embeddings whose unioned images form open covers of X.

Given a category C with pullbacks and a Grothendieck topology J on C, a presheaf $P \in \widehat{C}$ is a **sheaf** on the site (C, J) when $\forall X \in C$, $\forall S \in J(X)$, the following is an equalizer diagram:

$$P(X) \longrightarrow \prod_{f:X_\lambda \to X \in S} P(X_\lambda) \longrightarrow \prod_{\substack{f:X_\lambda \to X \in S \\ g:X_\mu \to X \in S}} P(X_\lambda \times_X X_\mu)$$

More comprehensibly, P is a sheaf if, for all objects $X \in C$ and all covering sieves $S \in J(X)$, every natural transformation $S \Rightarrow P$ uniquely extends to a natural transformation $S \Rightarrow C(-, X) \Rightarrow P$, so that $\widehat{C}(-, P)$ turns the inclusion $S \to C(, -)$ into an isomorphism.

For instance, consider the classical topology on Op(X). Writing $X(U, V) := \mathrm{Hom}_{\mathrm{Op}(X)}(V, U)$ for convenience, a sieve S on U, being a subfunctor of $X(-, U)$, associates to each $V \in \mathrm{Op}(X)$ either the singleton set $\{V \subseteq U\}$ or the empty set. A natural transformation from $X(-, U)$ to P is an assignment for each $V \subseteq U$ of an element $s|_V \in P(V)$ which is compatible with inclusions: the morphism $X(V, U)$ to $X(W, U)$ induced by the inclusion $W \subseteq V$ sends $s|_V$ to $s|_W$. In particular, every $s|_V$ must be induced from $s|_U$ by the inclusion $V \subseteq U$. So, we can think of a natural transformation $X(-, U) \Rightarrow P$ as an element of $P(U)$. A natural transformation $f : S \Rightarrow P$, in contrast, only yields an element $t|_V = f_V(\{V \subseteq U\})$ if $S(V)$ is nonempty, and yields nothing otherwise. We're only guaranteed that the V for which we get an element $t|_V$ form an open cover of U, and agree on intersections. It follows that the *existence* of a natural transformation $S \Rightarrow X(-, U)$ factoring any natural transformation $f : S \Rightarrow P$ leads to the traditional gluing axiom for sheaves, and the *uniqueness* of such a natural transformation leads to the locality axiom.

As in the topological case, the category of sheaves on a site, denoted by $\mathrm{Sh}(C, J)$, forms a full, reflective subcategory of $\widehat{C}$, whose inclusion functor $i : \mathrm{Sh}(C, J) \to \widehat{C}$ admits a left adjoint $(-)^{sh} : \hat{C} \to \mathrm{Sh}(C, J)$ known as **sheafification**: every morphism from a presheaf P to a sheaf Q extends to a unique sheaf morphism $P^{sh} \to Q$. In addition, $\mathrm{Sh}(C, J)$ has exponentials: if P is a presheaf and F is a sheaf on C, then the exponential presheaf F^P, which associates to an object $X \in C$ the set of natural transformations $C(-, X) \times P \Rightarrow F$, is a sheaf. The subobject classifier in $\mathrm{Sh}(C, J)$ is the sheaf that sends $X \in C$ to the set of closed sieves on X. For any function $f : Y \to X$ and closed sieve S on X, the sieve f^*S is closed on Y, and the restriction map $\Omega X \to \Omega Y$ induced by f simply takes S to f^*S. The maximal sieve t_X is closed, and the morphism $X \mapsto t_X$ yields the natural transformation true : id $\Rightarrow \Omega$.

The amount of sheaves on the site (C, J) is roughly analogous to the coarseness of the topology; we generally want all representable presheaves, of the form $C(-, X)$, to be sheaves, and any Grothendieck topology J satisfying this is known as a **subcanonical** topology. The **canonical** topology on (C, J) is the largest subcanonical topology.

Grothendieck Topoi A **Grothendieck topos** is a category $\mathcal{C}$ which is equivalent to some $\mathrm{Sh}(C, J)$. A **geometric morphism** $f : \mathcal{C} \to \mathcal{D}$ is defined to be a pair consisting of a **direct image** functor $f_* : \mathcal{C} \to \mathcal{D}$ and a **inverse image** functor $f^* : \mathcal{D} \to \mathcal{C}$ such that $f^* \dashv f_*$ and f^* preserves limits. Grothendieck topoi and geometric morphisms form a category Topos; when equipped with pairs of natural transformations between the functors comprising geometric morphisms, this becomes a 2-category.

The archetypal example of a Grothendieck topos is Set: this is the category of sheaves on the trivial category $\bullet$, since every set S corresponds to a presheaf $P(\bullet) = S$, every presheaf is trivially a sheaf, and natural transformations between sheaves correspond to morphisms between sets. Further, Set is the *terminal* object in Topos. Just as morphisms from the terminal object in Set, $\{\bullet\}$, correspond to points, or elements, of sets, a geometric morphism from Set to a topos $\mathcal{C}$ is known as a **point** of $\mathcal{C}$.

It is in general hard to find a site (C, J) evidencing $\mathcal{C}$ as a Grothendieck topology, but **Giraud's theorem** gives us a concrete way of identifying topoi: a category C is a Grothendieck topos if (1) it is locally small, has all finite limits, is cocomplete, (2) there is a set $\{S_\lambda\}$ of objects of C such that, for every $f, g : X \rightrightarrows Y$, if $fh = gh$ for every $h : S_\lambda \to X$, then $f = g$ (a small set of **generators**), (3) coproduct inclusions are monic and their pullback is an initial object (coproducts are **disjoint**), (4) small colimits are preserved under pullback (colimits are **universal**), and (5) every internal equivalence relation on an object X yields an internal quotient object (equivalence relations are **effective**).

3.2 Topos Theory

3.2.1 Grothendieck Topoi

Direct Image Functors Consider a topological space X, and its corresponding category $\mathrm{Sh}(X)$ of sheaves of sets. A continuous morphism $f : X \to Y$ generates a pair of adjoint functors:

- On the right, the direct image functor $f_* : \mathrm{Sh}(X) \to \mathrm{Sh}(Y)$, which sends a sheaf F on X to the sheaf $(f^*F)(V) = F(f^{-1}(V))$.

- On the left, the inverse image functor $f^* : \mathrm{Sh}(Y) \to \mathrm{Sh}(X)$, which sends a sheaf G on Y to the sheaf $(f_*G)(U) = \varinjlim_{V \supseteq f(U)} G(V)$.

By their adjunction, f^* preserves all colimits while f_* preserves all limits. f^* preserves finite limits, in fact, as it is a general fact that filtered colimits such as $\varinjlim$ preserve finite limits. If X and Y are sober[1], such that every point $x \in X$ can be deduced from the lattice of open subsets containing x (and likewise for Y), then in fact any such adjunction $f^* \dashv f_* : \mathrm{Sh}(X) \to \mathrm{Sh}(Y)$ whose left adjoint preserves finite limits comes from a continuous map $f : X \to Y$.

An instructive case is given by setting $X = \{*\}$, the vacuously Hausdorff and hence sober one-point space, since the category $\mathrm{Sh}(X)$ is equivalent to Set. Points of Y are equivalent to morphisms $X \to Y$, and hence equivalent to limit preserving left adjoints $f^* : \mathrm{Sh}(Y) \to \mathrm{Set}$. On the other hand, the fact that X is terminal in Top gives us a unique functor $f_* : \mathrm{Sh}(Y) \to \mathrm{Set}$ for any morphism $f : Y \to X$; this is the global sections functor, and its inverse image is the constant sheaf functor.

Geometric Morphisms Let $\mathcal{E} = \mathrm{Sh}(C, J)$ and $\mathcal{F} = \mathrm{Sh}(D, K)$ be Grothendieck topoi. An adjunction $f^* \dashv f_* : \mathcal{E} \to \mathcal{F}$ with f^* preserving finite limits is known as a **geometric morphism** $\mathcal{E} \to \mathcal{F}$, with f^* and f_* being called the direct and inverse images, respectively. This will be the topos-theoretic generalization of the above observation that morphisms $f : X \to Y$ generate adjoints $f^* \dashv f_* : \mathrm{Sh}(X) \to \mathrm{Sh}(Y)$. Similarly, we define a **point** of $\mathcal{E}$ to be a geometric morphism $p : \mathrm{Set} \to \mathcal{E}$. We form Grothendieck topoi and their geometric morphisms into a category Topos, whose terminal object is Set; the unique morphism $\Gamma : \mathcal{E} \to \mathrm{Set}$ has as its direct image the global sections functor.

If f^*, which preserves finite limits, preserves all small limits, then by the special adjoint functor theorem it has a further left adjoint $f_! : \mathcal{E} \to \mathcal{F}$, which we can compute as $f_! Y = \int^{X \in \mathcal{E}} \coprod_{f^* X \to Y} X$; an adjunction $f_! \dashv f^* \dashv f_* : \mathcal{E} \to \mathcal{F}$ characterizes an **essential geometric morphism**.

Many useful properties of Grothendieck topos are defined by analogy to topological spaces[2].

[1] Sobriety is a relatively weak condition, as it is implied by Hausdorffness (and hence present for manifolds, CW complexes, and so on); all affine schemes (and hence all schemes) are sober as well. So it holds in *most* practical cases.

[2] Or, more technically, *locales*, though we will note that sober topological spaces embed fully and faithfully into locales.

For instance, take X sober, and let $p : \mathrm{Sh}(X) \to \mathrm{Set}$. Connectedness of X is equivalent to fullness and faithfulness of $p^* : \mathrm{Set} \to \mathrm{Sh}(X)$. Hence, we call an arbitrary geometric morphism $f : \mathcal{E} \to \mathcal{F}$ **connected** if f^* is full and faithful, and call $\mathcal{E}$ itself connected if $\Gamma : \mathcal{E} \to \mathrm{Set}$ is connected (so that $\Gamma^* : \mathrm{Set} \to \mathcal{E}$ is full and faithful). Connected morphisms are necessarily essential, their identifying property being that $f_!$ preserves the terminal object.

3.2.2 Elementary Topoi

An **elementary topos** is a category $\mathcal{E}$ which is cartesian closed, has finite limits, including a terminal object 1, and a subobject classifier Ω. We define the contravariant **power object** functor as $\mathcal{P} := \Omega^-$, which due to the hom-exponential adjunction satisfies

$$\mathrm{Sub}_{\mathcal{E}}(X \times Y) = \mathcal{E}(X \times Y, \Omega) \cong \mathcal{E}(X, \mathcal{P}Y)$$

As with Grothendieck topoi, the canonical elementary topos is Set; as we will see, constructions in Set directly inspire many definitions of structures in elementary topoi.

3.2.3 Set-like Properties of Topoi

Set as a Topos Set is a topos with the following data:

- The subobject classifier is given by $\Omega = 2 = \{0, 1\}$.
- The true morphism is given by the inclusion $1 \hookrightarrow 2$.
- The exponential $[X, Y]$ is simply the set of all maps from X to Y. Hence, $[-,-] = \mathrm{Set}(-,-)$.
- The evaluation morphism $\mathrm{ev}_{X,Y} : [X, Y] \times X \to Y$ takes a map $\varphi : X \to Y$ and element $x \in X$ and sends it to $\varphi(x) \in Y$ (hence the name evaluation).
- The coevaluation morphism $\mathrm{coev}_{X,Y} : X \to [Y, X \times Y]$ sends x to the map sending y to $x \times y$.
- The classifying arrow of an inclusion $f : X \hookrightarrow Y$ is given by $\chi_f(y) = [y \in \mathrm{im} f]$.

These examples will serve as our intuition for how these gadgets work in arbitrary elementary topoi; they will also serve as a foundation for us to characterize more "Set-like" gadgets.

Membership In Set, subsets of a set X are in bijection with morphisms $X \to 2$: an $S \subseteq X$ is mapped to the morphism $\underline{S}(x) = [x \in S]$. Hence, in any topos $\mathcal{E}$ we define the **power object functor** $\mathcal{P} = [-, \Omega] : \mathcal{E}^{\mathrm{op}} \to \mathcal{E}$. In Set, the contravariant action sends a morphism $f : X \to Y$ to the morphism $\mathcal{P}f$ sending a $\underline{V} : Y \to 2$ to the composition $\underline{V} \circ f : X \to Y \to 2$, which is equivalent to the inverse image $f^{-1}(V)$; it therefore gives us an inverse image in $\mathcal{E}$.

Now, $\mathrm{ev}_{X,\Omega}$ gives a map $\mathcal{P}X \times X \to \Omega$ which in Set sends $U \subseteq X$ and $x \in X$ to $[x \in U]$; in $\mathcal{E}$ we denote $\mathrm{ev}_{X,\Omega}$ by $\in_X$, calling it the **membership map** (or predicate). Note that this map is obtained by adjunction from $\mathrm{id}_{\mathcal{P}X}$, and we therefore call it the $\mathcal{P}$-**transpose** of $\mathrm{id}_{\mathcal{P}X}$; the $\mathcal{P}$-transpose of a general map $f : X \times Y \to \Omega$ is the adjunct map $\omega_{X,Y,\Omega}(f) : X \to \mathcal{P}Y$, and the $\mathcal{P}$-transpose of a map $g : X \to \mathcal{P}Y$ is similarly $\omega^{-1}_{X,Y,\Omega}(g) : X \times Y \to \Omega$. For convenience we simply denote transposition by $\widehat{\cdot}$.

Equality Given an $X \in \mathcal{E}$, the universal property of the product $X \times X$ ensures for any pair of arrows $f, g : Y \to X$ an arrow $h : Y \to X \times X$ yielding f and g upon projection. If $f = g = \mathrm{id}_X$, we get an arrow $\Delta_X : X \to X \times X$ with $\pi_X \Delta_X = \mathrm{id}_X$. This is known as the **diagonal morphism**; if for $f, g : Y \to X$ we have $\Delta_X f = \Delta_X g$, then $\pi_X \Delta_X f = \pi_X \Delta_X g$ and therefore $f = g$, forcing Δ_X monic. A similar construction gives us the epic **codiagonal** $\nabla_X : X \amalg X \to X$.

The classifying map of Δ_X is written as $\delta_X : X \times X \to \Omega$. In Set, $\delta_X(x, x') = [x = x']$, so δ_X is in general referred to as the **equality map** (or predicate). Its $\mathcal{P}$-transpose $\widehat{\delta}_X : X \to \mathcal{P}X$ will in Set send $x \in X$ to $\{x\}$, and is in general referred to as the **singleton map**.

Images Given a monic $f : X \to Y$, we will construct a direct image morphism $\exists_f : \mathcal{P}X \to \mathcal{P}Y$. Pull $t : 1 \to \Omega$ back along $\in_X$ to obtain a monic $g : Z \to \mathcal{P}X \times X$. Compose g with the monic $\mathrm{id}_{\mathcal{P}X} \times f$ to get a monic $Z \to \mathcal{P}X \times Y$, take the characteristic map $\mathcal{P}X \times Y \to \Omega$, and transpose to get a map $\exists_f : \mathcal{P}X \to \mathcal{P}Y$. In Set, $Z = \{(U, x) \in \mathcal{P}X \times X \mid x \in U\}$, so the monic $Z \to \mathcal{P}X \times Y$ sends (U, x) to $(U, f(x))$, and its characteristic map sends (U, y) to $[y \in f(U)]$; the transpose of this map sends U to $\{y \in Y \mid y \in f(U)\}$, justifying our interpretation of $\exists_f$ as a **direct image** map.

Now we will construct the image of an arbitrary morphism $f : X \to Y$ as a subobject of Y. First, push f out along itself to get a pair of morphisms $g, g' : Y \to Y +_X Y$ with $gf = g'f$. Take the equalizer of g with g' to get a monic $h : Z \to Y$ with $gh = g'h$; its universal property yields for any $h' : Z' \to Y$ with $gh' = g'h'$ a morphism $k : Z' \to Z$ with $h' = hk$. For f, this universal property gives an epic $k : X \to Z$ with $f = hk$. By the fact that this construction involves only universal properties, this gives a factorization of any morphism $f : X \to Y$ into an epic $X \to Z$

followed by a monic $Z \to Y$, the latter of which is known as the **image** of f.

Logic We can construct many logical operators using the categorical properties of Ω. While $\mathtt{true} : 1 \to \Omega$ is given by definition, we may define $\mathtt{false} : 1 \to \Omega$ to be the classifying arrow of the monic initial arrow $0 \to 1$. Negation $\neg : \Omega \to \Omega$ is given by $\chi_{\mathtt{false}}$, $\wedge : \Omega \times \Omega \to \Omega$ by δ_Ω, $\implies : \Omega \times \Omega \to \Omega$ by $\chi_{\leq}$, and $\vee : \Omega \times \Omega \to \Omega$ by $(\mathtt{true} \times \mathrm{id}_\Omega) \amalg (\mathrm{id}_\Omega \times \mathtt{true})$.

Furthermore, we may define the existential and universal quantifiers $\exists$ and $\forall$ as "internal" adjoints to the power object functor $\mathcal{P}$. Given $f : X \to Y$, we can construct for each $Z \in \mathcal{E}$ a map $\mathrm{Hom}_{\mathcal{E}}(Z, \mathcal{P}Y) \to \mathrm{Hom}_{\mathcal{E}}(Z, \mathcal{P}X)$ in the functorial manner; an internal left (right) adjoint is a left (right) natural inverse. By Yoneda, existence of such inverses implies existence of natural maps $\exists_f, \forall_f : \mathcal{P}X \to \mathcal{P}Y$ (internally) adjoint to $\mathcal{P}f : \mathcal{P}Y \to \mathcal{P}X$.

In Set, this works as follows: $\exists_f(S)$ is the set $\{y \in Y \mid \exists x \in X \text{ with } f(x) = y \text{ and } x \in S\}$, i.e. the direct image of S. $\forall_f(S)$ is the set $\{y \in Y \mid \forall x \in X, \text{ if } f(x) = y \text{ then } x \in S\}$; there can be no element of $\forall_f(S)$ that is mapped to by an element outside of S. Consider for instance the mapping $f : \mathbb{Z} \to \mathbb{Z}, n \mapsto n^2$. $\exists_f(\mathbb{N})$ will return the non-negatives, while $\forall_f(\mathbb{N})$ will return $\{0\}$, as 0 is the only integer for which $x^2 = 0 \implies x \in \mathbb{N}$.

To summarize, we have defined:

- The *power object functor* $\mathcal{P} = [-, \Omega] : \mathcal{E}^{\mathrm{op}} \to \mathcal{E}$
- The *membership map* $\in_X = \mathrm{ev}_{X,\Omega}$
- *Transposition* $\widehat{\cdot} : \mathcal{E}(X \times Y, \Omega) \cong \mathcal{E}(X, \mathcal{P}Y)$.
- The *diagonal morphism* $\Delta_X : X \to X \times X$
- The *equality map* $\delta_X = \chi_{\Delta_X} : X \times X \to \Omega$
- The *singleton map* $\{\cdot\}_X = \widehat{\delta}_X : X \to \mathcal{P}X$
- The *direct image map* $\exists_f : \mathcal{P}X \to \mathcal{P}Y$
- The *image factorization* $X \twoheadrightarrow \mathrm{im} f \rightarrowtail Y$
- The *logical operators* $\wedge, \vee, \implies : \Omega \times \Omega \to \Omega$ and $\neg : \Omega \to \Omega$.
- The *existential quantifiers* $\forall_f, \exists_f : \mathcal{P}X \to \mathcal{P}Y$ induced by an $f : X \to Y$.

3.2.4 Mitchell-Bénabou Language

The language of an elementary topos $\mathcal{E}$ consists of the following data:

- For every $1 \to X$, a constant c of *type* X. This is often written $c : X$.
- For every X, variables $\{x_n : X\}_{n \in \mathbb{N}}$.

In the interpretation of this language, a term of type X with free variables of type $X_1, \ldots, X_n$ will be given by a morphism $X_1 \times \ldots \times X_n \to X$. The terms of the language are defined inductively: first, we proclaim every constant and variable of type X to be a term of type X, variables being terms with one free variable. We shall write terms as $\alpha, \beta, \ldots$.

- `true` and `false` are terms of type Ω, also known as **formulas**; they have no free variables, and are interpreted as their corresponding constants.
- (Membership predicate) If $\alpha : X$ and $\beta : \mathcal{P}X$ have the same free variables $x_1, \ldots, x_n$, $\alpha \in \beta$ is a formula with the same free variables $x_1, \ldots, x_n$, interpreted as the arrow $\in_X \circ (\beta \times \alpha)$.
- (Equality predicate) If $\alpha, \beta : X$ have the same free variables $x_1, \ldots, x_n$, then $\alpha = \beta$ is a formula with the same free variables $x_1, \ldots, x_n$, interpreted as the arrow $\delta_X \circ (\alpha \times \beta)$.
- (Application) If α is a term of type X and $f : X \to Y$ a morphism, then $f(\alpha)$ is a term of type Y, interpreted as $f \circ \alpha$.
- (Composition) If α is a term of type X with free variables $x_1, \ldots, x_n$ of types $X_1, \ldots, X_n$, and $y_1, \ldots, y_n$ are terms of types $X_1, \ldots, X_n$ sharing no bound variables with α, and each with free variables $y_1^1, \ldots, y_1^{m_1}, \ldots, y_n^1, \ldots, y_n^{m_n}$, then $\alpha(y_1, \ldots, y_n)$ is a term of type X with free variables $y_1^1, \ldots, y_n^{m_n}$, interpreted as $\alpha \circ (\Pi_i y_i)$.
- (Evaluation) Given $\alpha : X$ and $\beta : Y^X$, $\beta(\alpha)$ is a term of type Y, interpreted as $ev_{X,Y} \circ (\beta \times \alpha)$. ($\in_X$ is a special case of this).
- (Currying) Given a term α of type X with a free variable y of type Y, $\lambda y.\alpha$ is a term of type X^Y, interpreted as the transpose of α.
- (Logic) If ϕ, ψ are formulas, then so are $\phi \implies \psi$, $\phi \wedge \psi$, $\phi \vee \psi$, $\neg\phi$, and so on. These are interpreted in the obvious way.

- (Quantification) If ϕ is a formula with free variables $y, x_1, \ldots, x_n$ of types $Y, X_1, \ldots, X_n$, then $(\exists y \in Y)\,\phi$ and $(\forall y \in Y)\,\phi$ are formulas with free variables $x_1, \ldots, x_n$. These are interpreted by binding y via $\lambda y.\phi : X_1 \times \ldots \times X_n \to \mathcal{P}Y$, and composing with the $\forall_p$ and $\exists_p : \mathcal{P}Y \to \Omega = \mathcal{P}1$ generated by the terminal morphism $p : Y \to 1$.

We can define further shortcuts using these symbols, such as the uniqueness quantifier $\exists!$:

$$(\exists! x \in X)(\phi(x)) \iff (\exists x \in X)\,(\phi(x) \wedge (\forall x' \in X)(\phi(x') \implies x = x'))$$

the $\notin$ and $\neq$ predicates ($x \notin X \iff \neg(x \in X), x \neq x' \iff \neg(x = x')$) (though $\neg(x \notin x)$ isn't necessarily equivalent to $x \in X$ and likewise for $\neq$), and so on. We may also rewrite quantifiers when they are obvious from convention or usage, e.g. rewriting $(\forall x \in X)(\exists y \in Y)$ as $\forall x \exists y$ and $(\forall x_1 \in X)(\forall x_2 \in X)$ as $\forall x_1, x_2$.

A formula ϕ with free variable $x : X$, which we may also write as $\phi(x)$, is equivalent via interpretation to a morphism $X \to \Omega$, and therefore (by $\mathrm{Sub}_{\mathcal{E}}(X) \cong \mathrm{Hom}_{\mathcal{E}}(X, \Omega)$) a *subobject* of X. We write this subobject as $\{x \in X \mid \phi(x)\}$, or just $\{x \mid \phi\}$. Consider for instance the subobject of X^Y given by

$$\mathrm{Inj}(Y, X) = \{f \in X^Y \mid (\forall y, y')(f(y) = f(y') \implies y = y')\}$$

which nominally classifies "injective" maps $Y \to X$. We will translate this: the term $f(y) = f(y') \implies y = y'$ is the arrow

$$(\Rightarrow) \times ((\delta_X \circ (\mathrm{ev}_{X,Y} \times \mathrm{ev}_{X,Y})) \times \delta_Y) \circ \Gamma : X^Y \times Y \times Y \to \Omega$$

where Γ is the purely logistical morphism morally sending (f, y, y') to (f, y, f, y', y, y'). Call this arrow ϕ. We transpose ϕ to get a morphism $X^Y \times Y \to \mathcal{P}Y$, apply $\forall_p$ to get a morphism $X^Y \times Y \to \Omega$, transpose to get $X^Y \to \mathcal{P}Y$, apply $\forall_p$ to get $X^Y \to \Omega$, and then take the fibered product with $\mathtt{true} : 1 \to \Omega$ to get the desired subobject $\mathrm{Inj}(Y, X) \rightarrowtail X^Y$.

We will consider two other examples: for $A, B : \mathcal{P}X$, let $A \cup B$ be the subobject $\{S \in \mathcal{P}X \mid (\forall s \in S)(s \in A \vee s \in B)\}$.

In Set, for instance, ϕ takes a map $f : Y \to X$ and two elements y, y' of Y. It turns this triplet into the sextuplet (f, y, f, y', y, y') via Γ, applies $\mathrm{ev}_{X,Y}$ to the first two pairs to obtain the quadruplet $(f(y), f(y'), y, y')$, then applies δ_X and δ_Y to each pair to obtain the pair $([f(y) = f(y')], [y = y'])$ of truth values, which it applies $\Rightarrow$ to. Transposition and application of $\forall_p$ returns the morphism sending an $f : X \to Y$ to the truth of whether it satisfies $\phi(f, y, y')$ for all $y, y' \in Y$, and pullback returns the subset of all $f : X \to Y$ that do satisfy this. The internal language allows us to reason about things such as injective functions as though they "really" existed.

A **first-order formula** in $\mathcal{E}$ is any formula that can be formed via these rules. We may include rules allowing for infinitary conjunction and disjunction, leading to the **infinitary first-order formulas**. A **geometric formula** is an infinitary first-order formula that does not involve negation, implication, or infinitary conjunction; these are called geometric because their truth is preserved by pullback along geometric morphisms $f^* \dashv f_* : \mathcal{E} \to \mathcal{F}$. Logical morphisms preserve the truth of all first-order formulas.

3.2.5 Kripke-Joyal Semantics

Semantics Every formula $\phi(x)$ with free variable $x : X$ has a corresponding subobject $\{x \mid \phi\}$. Every morphism $f : U \to X$ also has a corresponding subobject $\mathrm{im} f$; if $\mathrm{im} f \leq \{x \mid \phi\}$, such that f factors through the subobject $\{x \mid \phi\}$, we say that U **forces** ϕ on the "generalized element" f, written as $U \Vdash \phi(f)$, where $\phi(f) := \phi \circ f$. Given this, the following relations on $\Vdash$, which state the **Kripke-Joyal semantics** of $\mathcal{E}$, hold:

1. $U \Vdash \phi(f) \wedge \psi(f)$ iff $U \Vdash \phi(f)$ and $U \Vdash \psi(f)$.
2. $U \Vdash \phi(f) \vee \psi(f)$ iff there are arrows $g : V \to U$, $h : W \to U$ such that $g \amalg h : V \amalg W \to U$ is epi, with $V \Vdash \phi(fg)$ and $W \Vdash \phi(fh)$.
3. $U \Vdash \phi(f) \implies \psi(f)$ iff for any $g : V \to U$ such that $V \Vdash \phi(fg)$, V also forces $\psi(fg)$.
4. $U \Vdash \neg\phi(f)$ if for any $g : V \to U$ such that $V \Vdash \phi(fg)$, V is the initial object.
5. $U \Vdash \exists y\, \phi(f, y)$ (for some formula $\phi : X \times Y \to \Omega$ and generalized element $f : U \to X$) iff there's an epic $e : V \to U$ and generalized element $g : V \to Y$ such that $V \Vdash \phi(fe, g)$.
6. $U \Vdash \forall y\, \phi(f, y)$ iff for *every* arrow $h : V \to U$ and generalized element $g : V \to Y$ we have $V \Vdash \phi(fh, g)$.

We say that a formula $\phi(x_1, \ldots, x_n)$ is *true* in $\mathcal{E}$, writing $\mathcal{E} \models \phi$, if the morphism $1 \to \Omega$ given by $\forall x_1, \ldots, \forall x_n\, \phi(x_1, \ldots, x_n)$ is equal to the arrow $\texttt{true} : 1 \to \Omega$, or equivalently if we have $1 \Vdash \forall x_1, \ldots, \forall x_n\, \phi(x_1, \ldots, x_n)$.

The language and semantics of a topos admit several rules for inference that we can use in order to think about this language independent from its arrow-theoretic nature: for instance, we have a *modus ponens rule*: if $U \Vdash \phi(f)$ and $U \Vdash \phi(f) \implies \psi(f)$, then, since $\mathrm{id}_U : U \to U$ has $U \Vdash \phi(f \circ \mathrm{id}_U) = \phi(f)$, it follows that $U \Vdash \psi(f)$. In general, we can carry out *intuitionistic logic*, which is more or less the same as classical logic save for a lack of the PEM. So it is not generally true in a non-Boolean topos $\mathcal{E}$ that $\mathcal{E} \models \phi \vee \neg\phi$, nor is it true that $\mathcal{E} \models \neg\neg\phi \implies \phi$.

Axioms in Topoi There are many useful axioms we can assume our topos $\mathcal{E}$ to have, which using $\mathcal{E}$'s internal logic we can state precisely. We may have, for instance, the *(internal) principle of excluded middle* (PEM):

$$\mathcal{E} \models (\forall p \in \Omega)(p \vee \neg p)$$

If this holds, we call $\mathcal{E}$ a *Boolean topos*; in such a topos we can obtain for every subobject $S \rightarrowtail X$ a complement $S^c \rightarrowtail X$.

The *internal axiom of choice* (IAC) is the internal statement that "every surjection has a section", which in Set really is equivalent to the axiom of choice:

$$\mathcal{E} \models (\forall f \in Y^X)\left[(\forall y \in Y)(\exists x \in X)(f(x) = y) \implies (\exists s \in X^Y)(\forall y' \in Y)(f(s(y')) = y')\right]$$

This is strictly stronger than the PEM, but *weaker* than the external AC: the IAC can be true in $\mathcal{E}$ without the actual statement "every surjection has a section" being true in $\mathcal{E}$.

The *axiom of infinity* is not phrased in the internal language, but is far-reaching nevertheless: it postulates the existence of a **natural numbers object** (n.n.o.), or an object $\mathbb{N} \in \mathcal{E}$ equipped with two morphisms $s : \mathbb{N} \to \mathbb{N}$, $z : 1 \to \mathbb{N}$ which is universal in the sense that for any $1 \xrightarrow{x} X \xrightarrow{f} X$, there's a unique $h : \mathbb{N} \to X$ with $hz = x$ and $hs = fh$.

Given an n.n.o. $\mathbb{N}$, we can define an addition map $+ : \mathbb{N} \times \mathbb{N} \to \mathbb{N}$: this is the unique map such that the following diagram is commutative:

$$\begin{array}{ccccc} 1 \times \mathbb{N} & \xrightarrow{z \times \mathrm{id}_{\mathbb{N}}} & \mathbb{N} \times \mathbb{N} & \xrightarrow{s \times \mathrm{id}_{\mathbb{N}}} & \mathbb{N} \times \mathbb{N} \\ \| & & \downarrow + & & \downarrow + \\ \mathbb{N} & \xrightarrow{\mathrm{id}_{\mathbb{N}}} & \mathbb{N} & \xrightarrow{s} & \mathbb{N} \end{array}$$

To get this map, apply the universal property of $\mathbb{N}$ to the diagram $1 \to \mathbb{N}^{\mathbb{N}} \to \mathbb{N}^{\mathbb{N}}$, where the first map is the transpose of the identity and the second is $s^{\mathbb{N}}$; this gives us a map $\widehat{+} : \mathbb{N} \to \mathbb{N}^{\mathbb{N}}$ with $\widehat{+} \circ z = \mathrm{id}_{\mathbb{N}}$ and $s^{\mathbb{N}} \circ \widehat{+} = \widehat{+} \circ s$, which by transpose corresponds to a map $+ : \mathbb{N} \times \mathbb{N} \to \mathbb{N}$ making the above diagram commutative.

Given an n.n.o. $\mathbb{N}$, it is straightforward to mimic the construction of $\mathbb{Z}$ and $\mathbb{Q}$. Recall that in Set, $\mathbb{Z}$ is defined to be $\mathbb{N} \times \mathbb{N}$ modulo the relation that $(a, b) \sim (c, d)$ if $a + d = b + c$. In $\mathcal{E}$, we can take the pullback of $+$ along itself to get an object X morally representing all pairs of pairs of integers with equal sums, along with projections $\pi_1, \pi_2 : X \to \mathbb{N} \times \mathbb{N}$. Taking the two projections $\pi_1', \pi_2' : \mathbb{N} \times \mathbb{N} \to \mathbb{N}$, we quotient by the equivalence relation by taking the coequalizer of $\pi_1'\pi_1 \times \pi_2'\pi_2$ with $\pi_2'\pi_1 \times \pi_1'\pi_2$, giving us an integers object $\mathbb{Z}$. We can similarly define a multiplication $* : \mathbb{Z} \to \mathbb{Z}$ and use it to create a rational numbers object $\mathbb{Q} \in \mathcal{E}$.

It is not as easy to get a real numbers object $\mathbb{R}$, though; there are many different possible constructions, and while these are equivalent in Set, they are not generally equivalent in elementary topoi. We shall use the Dedekind real numbers, which is the "largest" among many popular constructions. A Dedekind cut in a topos $\mathcal{E}$ with rational numbers object $\mathbb{Q}$ is a pair of subobjects $L, U \rightarrowtail \mathbb{Q}$ such that the following hold in $\mathcal{E}$:

- (Non-emptiness) $(\exists x \in \mathbb{Q})(x \in L)$ and $(\exists x \in \mathbb{Q})(x \in R)$
- (Disjointness) $(\forall x)(\neg(x \in L \wedge x \in U))$
- (Order) $(\forall x, y)(x < y \wedge y \in L \implies y \in L)$ and $(\forall x, y)(x < y \wedge x \in U \implies y \in U)$
- (Dichotomy) $(\forall x, y)(x < y \implies (x \in L \vee y \in U))$
- (Openness) $(\forall x)(x \in L \implies (\exists y)(y \in L \wedge x < y))$ and $(\forall x)(x \in U \implies (\exists y)(y \in U \wedge y < x))$.

Taking the conjunction of all of these gives a formula φ on $\mathcal{P}\mathbb{Q}\times\mathcal{P}\mathbb{Q}$, the corresponding subobject $\{(L, U) \mid \varphi\}$ of which is known as the (Dedekind) real numbers object $\mathbb{R}$.

Objects in Topoi Given an object $G \in \mathcal{E}$, we may stipulate internal axioms amounting to the existence of an algebraic structure on G: for instance, suppose we equip G with a morphism $0 : 1 \to G$ and a morphism $+ : G \times G \to G$ written infix, and assume that $\mathcal{E}$ models the following sentences:

- $(\forall g \in G)(0 + g = g + 0 = g)$
- $(\forall g, h, k)((g + h) + k = g + (h + k))$.
- $(\forall g \exists h)(g + h = 0)$.
- $(\forall g, h)(g + h = h + g)$.

This will be an abelian group from $\mathcal{E}$'s point of view, and since the theory of abelian groups can be expressed intuitionistically, objects which are abelian groups according to the internal logic are also internal abelian groups; this holds for most similar theories, including rings and modules.

We shall make particular use of a certain kind of object known as a *Weil algebra*. Given a ring object R in a topos $\mathcal{E}$ (or a **ringed topos** $(\mathcal{E}, R)$), a **Weil algebra** is a local ring $(W, \mathfrak{m})$ with an R-algebra structure, such that W is finite-dimensional as an R-module and can be written as the direct sum $R \oplus \mathfrak{m}$. In the ringed topos $(\mathsf{Set}, \mathbb{R})$, Weil algebras are equivalent to $\mathbb{R}$-algebras, finite-dimensional as vector spaces, of the form $C_0^\infty(\mathbb{R}^n)/I$, where C_0^∞ denotes smooth functions vanishing at 0. For instance, $C_0^\infty(\mathbb{R})/(x^2)$ is the ring of dual numbers $\mathbb{R}[\varepsilon] := \mathbb{R}[x]/(x^2)$. With R-algebra homomorphisms mapping maximal ideals into maximal ideals, Weil algebras form a category $\mathsf{W}(\mathcal{E})$.

3.3 Infinitesimals

3.3.1 The Kock-Lawvere Axiom

Given a commutative ring object R in a topos $\mathcal{E}$, we define the subobject of **infinitesimals** of R by $D := \{x \in R \mid x^2 = 0\}$. The **Kock-Lawvere axiom** for R reads

$$(\forall f \in R^D)(\exists! c \in R)\,((\forall \epsilon \in D)(f(\epsilon) = f(0) + c\epsilon))$$

Clearly $0 \in D$, so $0 : 1 \to R$ factors through D. As a consequence, we have that if $c_1\epsilon = c_2\epsilon$ for all $\epsilon \in D$, then $c_1 = c_2$ (let $f(\epsilon) = c_1\epsilon$). The KL axiom allows us to work with infinitesimals as though they actually exist, using them to define derivatives around points. However, this comes at a cost: we cannot in general exhibit non-zero infinitesimals.

In order to work with the KL axiom, we must explicitly reject the principle of excluded middle: to see this, define a map $f : D \to R$ which sends ϵ to 0 if $\epsilon = 0$ and to 1 otherwise; the KL axiom implies that there's a unique $c \in R$ such that $f(\epsilon) = c \cdot \epsilon$ for all $\epsilon \in D$. Assuming the LEM, either D contains only 0 or D contains other elements. If D contains only 0, then c cannot be unique; hence, it contains an $\epsilon \neq 0$, and a unique c such that $c\epsilon = 1$. It follows that $0 = (c\epsilon)^2 = 1^2 = 1$, a contradiction. Hence, we must throw out the LEM, and work constructively. Another consequence of this is the *undecidability* of R: the sentence $(\forall x, y)(x = y \wedge x \neq y)$ is not true. In particular, $\mathcal{E}$ cannot show that infinitesimals are non-zero.

This is in part because the KL axiom is very strong: fixing an $x \in R$, $f : R \to R$, and $k : D \to R$ sending 0 to $f(x)$ and ϵ to $k(\epsilon) = f(x + \epsilon)$, the KL axiom gives a unique c_x in R such that $f(x + \epsilon) = f(x) + c\epsilon$. We write $f'(x) := c_x$ to get a function $f' : R \to R$ known as the derivative of

f, and state **Taylor's formula**:

$$\forall \epsilon \in D(f(x+\epsilon) = f(x) + \epsilon f'(x))$$

So KL implies that every function $f : R \to R$ is differentiable.

An Alternative Statement Here's another statement equivalent to the KL axiom: take the R-algebra $R[\varepsilon] = R \times R$ with multiplication $(a, b) \cdot (c, d) = (ac, ad + bc)$. Then (KL2), the map $\alpha : R[\varepsilon] \to R^D$, $\alpha(a, b)(\epsilon) = a + \epsilon b$ is an R-algebra isomorphism.

It's clear that $(\alpha(a, b)\alpha(c, d))(\epsilon) = \alpha(ac, ad + bc)(\epsilon)$, as well as that this statement, KL2, implies the original statement (KL1). To see the converse, assume KL1. Then, not only is every function f of the form $\alpha(f(0), c)$, but for every $\alpha(a, b)$ there is a unique $c \in R$ such that $a + b\epsilon = \alpha(a, b)(\epsilon) = \alpha(a, b)(0) + c\epsilon = a + c\epsilon$ for all ϵ; b obviously satisfies this, and hence is the only element of R that satisfies this, making it, and hence the pair (a, b) recoverable from the function $\alpha(a, b)$. So KL1 is equivalent to KL2.

Spectra Given an arbitrary R-algebra $A \in \mathcal{E}$ and a finitely generated R-algebra $B = R[x_1, \ldots, x_n]/I$, for instance a Weil algebra, the **spectrum** $\mathrm{Spec}_A(B)$ is a subobject of A^n consisting of those $a = (a_1, \ldots, a_n)$ such that $P(a) = 0$ for all $P \in I$. For instance, $\mathrm{Spec}_R(R[x]/(x^2)) = \{x \in R \mid x^2 = 0\} = D$. For W a Weil algebra, the object $\mathrm{Spec}_R(W)$ is known as the **formal infinitesimals** object of R (with respect to W). The process of taking spectra with respect to R is functorial: a morphism $\psi : W \to W'$ of Weil algebras generates a morphism $\Psi : \mathrm{Spec}_R(W') \to \mathrm{Spec}_R(W)$

A third formulation of the KL axiom states that (KL3) the R-algebra homomorphism $\alpha : W \to R^{\mathrm{Spec}_R(W)}$, $\alpha(P)(x_1, \ldots, x_n) = P(x_1, \ldots, x_n)$, is an isomorphism. In the topos $\mathcal{E}$, every Weil algebra W yields a functor $(-)^{\mathrm{Spec}_R W}$ which is right adjoint to the functor $- \times \mathrm{Spec}_R W$. If each W satisfies the KL axiom and $(-)^{\mathrm{Spec}_R W}$ is always a left adjoint as well, $\mathcal{E}$ is known as a **smooth topos**. The right adjoint, known as the **amazing right adjoint**, is denoted $(-)^{1/\mathrm{Spec}_R W}$.

Differentiation The differentiation given by the KL axiom satisfies the usual properties: for instance, consider two functions $g, f : R \to R$. $(gf)(x + \epsilon)$ is equal to $(gf)(x) + \epsilon(gf)'(x)$, but also equal to $g(f(x) + \epsilon f'(x))$, which since $\epsilon f'(x)$ is an infinitesimal is itself equal to $(gf)(x) + \epsilon f'(x)g'(f(x))$, implying that $(gf)'(x) = f'(x)(g'f)(x)$, i.e. the chain rule. Similarly, differentiation satisfies the product rule, is R-linear, sends constants to 0, and sends id_R to 1.

We define D_n to be the set of all nth order infinitesimals, or elements $x \in R$ such that $x^{n+1} = 0$.

(In particular, $D = D_1$). D_∞ is defined to be the set of all nilpotent elements, or $x \in R$ such that $x^n = 0$ for some $n \geq 1$. Supposing that $2, 3, \ldots$ are invertible in R, the higher order extensions of the KL axiom are as follows:

$$\forall f \in R^{D_n} \exists! c_1, \ldots, c_n \in R \left(\forall \epsilon \in D_n (f(\epsilon) = f(0) + c_1\epsilon^1 + c_2\epsilon^2 + \ldots + c_n\epsilon^n)\right)$$

and the corresponding Taylor formulas are

$$\forall \epsilon \in D_n \left(f(x+\epsilon) = f(x) + \epsilon f'(x) + \frac{\epsilon^2}{2} f''(x) + \ldots + \frac{\epsilon^n}{n!} f^{(n)}(x) \right)$$

An R-module V satisfying the following vector version of the KL axiom is known as a **Euclidean R-module**:

$$\forall f \in V^D \exists! v \in V (\forall \epsilon \in D (f(\epsilon) = f(0) + \epsilon \cdot v))$$

When $V \cong R^n$, we can write $\vec{x} = (x_1, \ldots, x_n)$, and we have for a function $g : R^n \to R^n$ such that $g(\vec{x} + \epsilon \cdot \vec{y}) = f(\epsilon)$ a $\vec{z} \in R^n$ such that $g(\vec{x} + \epsilon \cdot \vec{y}) = g(\vec{x}) + \epsilon \cdot \vec{z}$. We define the **directional derivative** $\partial_{\vec{y}} g$ of g in the direction $\vec{y}$ to be this $\vec{z}$, and the ith **partial derivative** $\partial_i f$ to be the directional derivative in the direction of the ith unit vector. The map $\vec{y} \to \partial_{\vec{y}} g$ is known as the **differential** g' of g.

3.3.2 Differential Geometry

Microlinear Spaces Given a topos $\mathcal{E}$ and a commutative ring object R satisfying the KL axiom, take the nested categories $\mathsf{Weil} \subseteq R\text{-}\mathsf{Alg}_{\mathsf{FP}} \subseteq R\text{-}\mathsf{Alg}$ of Weil algebras, finitely presented R-algebra objects, and R-algebra objects, respectively. We have a pair of functors $R^- : \mathcal{E}^{op} \to \mathcal{E}$ and $\mathsf{Spec}_R : R\text{-}\mathsf{Alg}_{\mathsf{FP}}^{op} \supseteq \mathsf{Weil}^{op} \to \mathcal{E}$. Given a finite limit diagram $\mathcal{J}$ of Weil algebras, $\mathcal{D} = \mathsf{Spec}_R(\mathcal{J})$ is, while not necessarily a colimit, at least a cocone. An object $M \in \mathcal{E}$ is a **microlinear space** if $M^{\mathcal{D}}$ is a limit diagram for every $\mathcal{J}$. Microlinear spaces will serve as our generalized manifolds. These spaces contain R, are closed under limits (e.g., arbitrary products), and contain exponentials: if M is microlinear and X an arbitrary object, M^X is again microlinear. Thus, we already have a rich abundance of microlinear spaces. A **Lie group** is a group internal to $\mathcal{E}$ which is also a microlinear space; again, the trivial example is R.

Tangent Vectors Given a microlinear space M, a **vector bundle** over M is an epic $E = \pi : E \to M$ such that $\pi^{-1}(x)$ is a Euclidean R-module, and a **section**, also known as an E-**vector field**, of the vector bundle E is a morphism $s : M \to E$ such that $\pi s = \mathrm{id}_M$. The **tangent bundle** of a microlinear space M is the object M^D equipped with a map $\pi : M^D \to M, t \mapsto t(0)$; its elements are **tangent vectors**, and the **tangent space** of M at a point x is the collection M^D_x of $t \in M^D$ with $\pi(t) = t(0) = x$. We write $TM = M^D, T_xM = M^D_x$, and think of elements of M^D as probings of M in infinitesimal directions, hence tangent vectors. A TM-vector field, just known as a vector field, is a map $M \to M^D$ satisfying the above properties; by cartesian closure, we can look at a vector field X not just as a map $M \to M^D$, but as a map $M \times D \to M$, or even as a map $D \to M^M$ taking an infinitesimal d and giving us an infinitesimal deformation X_d of M. Using this definition, the object $\mathfrak{X}(M)$ of all vector fields on M becomes an R-module under the action $(rX)_d = X_{rd}$. This definition also allows isomorphisms φ to act on vector fields X: we define $(\varphi_* X)_d = \varphi X_d \varphi^{-1}$. If φ is an *endo*morphism, we may define $(\varphi^*\omega)(v) = \omega(\varphi \circ v)$.

Given a $v \in M^{D^n}$, which we think of as a function taking in n infinitesimals and outputting an element of the microlinear space M, as well as an $r \in R$, we define $r_k v(d_1, \dots, d_n) = v(d_1, \dots, rd_k, \dots, d_n)$. Given a $\sigma \in S_n$, we define $v^\sigma(d_1, \dots, d_n) = v(d_{\sigma 1}, \dots, d_{\sigma n})$. An n**-form** on M is a map $\omega : M^{D^n} \to R$ such that $\omega(r_k v) = r\omega(v)$ and $\omega(v^\sigma) = (-1)^\sigma \omega(v)$. The object $\Lambda^n(M)$ of all n-forms on M is a microlinear space as well as a Euclidean R-module. We denote by $X * v$ the element of $M^{D^{n+1}}$ given by $(X * v)(d_1, \dots, d_{n+1}) = X_{d_1}(v(d_2, \dots, d_{n+1}))$, and by $i_X\omega$ the $(n-1)$-form acting on a $w \in M^{D^{n-1}}$ by $(i_X\omega)(w) = \omega(X * w)$.

For $X, Y \in \mathfrak{X}(M)$, we define $[X, Y]_{d_1 d_2} = Y_{-d_2} X_{-d_1} Y_{d_2} X_{d_1}$; the vector field $[X, Y]$ is also written $L_X Y$, and is equivalently the unique vector field such that $(X_{-d})_* Y - Y = dL_X Y$. The **exterior derivative** of an n-form ω is given by

$$(d\omega)(v) = \sum_{i=1}^{n+1} (-1)^{i+1} (F^i_v)'(0)$$

where $F^i_v(e) = \omega(v(d_1, \dots, d_{i-1}, e, d_{i+1}, \dots, d_n))$; as expected, it satisfies $d^2 = 0$. With this in mind, we state Cartan's three "magical formulae" without proof: $L_{[X,Y]} = L_{[X,}L_{Y]}$, $i_{[X,Y]} = L_{[X,}i_{Y]}$, and $L_X = di_X + i_X d$.

Formal Manifolds More specific than the microlinear spaces are the *formal manifolds*, which take some effort to set up. A morphism $f : X \to Y$ is **étale** if for every element $x : 1 \to X$ and morphism g from an infinitesimal object $\mathrm{Spec}_R W$ to Y, there is a unique arrow $h : \mathrm{Spec}_R W \to X$ which maps $0 \in \mathrm{Spec}_R W$ to x and satisfies $fh = g$, i.e. makes the diagram below commutative.

$$\begin{array}{ccc} 1 & \xrightarrow{0} & \mathrm{Spec}_R W \\ {\scriptstyle x}\downarrow & \overset{h}{\swarrow} & \downarrow{\scriptstyle g} \\ X & \xrightarrow[f]{} & Y \end{array}$$

If $Y = R^n$ and f is monic, X is said to be an n-dimensional **model object**. An object M is an n-dimensional **formal manifold** if there are étale monics $X_i \to M$, where each X_i is an n-dimensional monic object, whose coproduct is a regular epic morphism $\amalg_i X_i \to M$.

3.3.3 Smooth Algebras

Let CartSp be the subcategory of Diff consisting of the cartesian spaces $\{\mathbb{R}^n\}_{n\in\mathbb{N}}$. A C^∞**-ring**, or a **smooth algebra**, is a product-preserving functor CartSp $\to$ Set, and a C^∞-ring **homomorphism** is a natural transformation of functors. These form a category which we will denote C^∞-Alg. Intuitively, C^∞-rings are modeled on (but not restricted to) rings of the form $C^\infty(M)$, for some smooth manifold M; for such a ring, we may define $\Phi_f(\varphi_1, \ldots, \varphi_n)(p) = f(\varphi_1(p), \ldots, \varphi_n(p))$ to get a C^∞-ring.

Given a C^∞-ring A : CartSp $\to$ Set, we may endow $A(\mathbb{R})$, and hence all $A(\mathbb{R}^n)$, with the structure of an $\mathbb{R}$-algebra by using the images of the morphisms $+ : \mathbb{R}^2 \to \mathbb{R}$ and $c \cdot - : \mathbb{R} \to \mathbb{R}$: for $x, y \in A(\mathbb{R})$ and $c \in \mathbb{R}$, we denote by $x + y$ the image of $(x, y) \in \mathbb{R}^2$ under the morphism $A(+) : A(\mathbb{R}^2) = A(\mathbb{R})^2 \to A(\mathbb{R})$, and we denote by cx the image of x under the morphism $A(c \cdot -) : A(\mathbb{R}) \to A(\mathbb{R})$. That the necessary $\mathbb{R}$-algebra identities hold in CartSp imply that they hold in Set as well. Hence, we may associate to every C^∞-ring an underlying $\mathbb{R}$-algebra $A(\mathbb{R})$. We will often *identify* A with $A(\mathbb{R})$, though we can't identify any given $\mathbb{R}$-algebra X with a C^∞-ring: it's necessary that X lifts morphisms $\mathbb{R}^n \to \mathbb{R}^m$ to morphisms $X^n \to X^m$ in a nice way. Specifically, we require an operation $\Phi_f : X^n \to X$ for every smooth map $f : \mathbb{R}^n \to \mathbb{R}$ such that, for $h(x_1, \ldots, x_n) = g(f_1(x_1, \ldots, x_n), \ldots, f_m(x_1, \ldots, x_n))$, we have $\Phi_h(x_1, \ldots, x_n) = \Phi_g(\Phi_{f_1}(x_1, \ldots, x_n), \ldots, \Phi_{f_m}(x_1, \ldots, x_n))$ as well as $\Phi_{\pi_i}(x_1, \ldots, x_n) = x_i$.

Finitely Generated Ideals Of particular consequence is when A is equivalent to $C^\infty(\mathbb{R}^n)/I$ for some ideal I of $C^\infty(\mathbb{R}^n)$: when this happens, A is said to be **finitely generated**, and when $I = (i_1, \ldots, i_m)$ is finitely generated as an ideal, A is said to be **finitely presented**. Every C^∞-ring of the form $C^\infty(M)$ for a smooth manifold M is finitely presented, for instance. If A is local as a normal ring, it's known as a **local** C^∞-ring. The primary example is, as encountered in algebraic geometry, the stalk of the sheaf of smooth functions on $\mathbb{R}^n$, written $C^\infty_p(\mathbb{R}^n)$.

We define the category L^{op} to be the subcategory of C^∞-Alg consisting of the finitely generated algebras; the objects of L are known as **loci**, and written as $\ell A, \ell B, \ldots$ (where A, B are finitely generated smooth algebras). A morphism $\ell B \to \ell A$ of L is a morphism $A \to B$, or, if $B = C^\infty(\mathbb{R}^m)/J$ and $A = C^\infty(\mathbb{R}^n)/I$, an equivalence class $[\varphi]$ of functions $\mathbb{R}^m \to \mathbb{R}^n$ acting as $\varphi(f) = f \circ \varphi$; we require each φ to satisfy $f \in I \implies \varphi(f) \in J$, so that φ extends to a function $C^\infty(\mathbb{R}^n)/I \to C^\infty(\mathbb{R}^m)/J$, $f + (I) \mapsto \varphi(f) + (J)$, and write $\varphi \sim \psi$ if each $\pi_i \circ (\varphi - \psi) : \mathbb{R}^n \to \mathbb{R}$ is in I.

$\mathsf{Set}^{\mathsf{L}^{op}}$ is a Grothendieck topos (by equipping L with the indiscrete topology in which all presheaves are sheaves). The functor $s : \mathsf{Diff} \to \mathsf{L}$ sending a smooth manifold M to $\ell C^\infty(M)$ is full and faithful, and when combined with the full and faithful Yoneda embedding $よ : \mathsf{L} \to \mathsf{Set}^{\mathsf{L}^{op}}$ evidences Diff as a subcategory of $\mathsf{Set}^{\mathsf{L}^{op}}$. So, $\mathsf{Set}^{\mathsf{L}^{op}}$ can be thought of as a category of "generalized" smooth spaces, and at the same time as a category of "variable" sets. For a functor $P \in \mathsf{Set}^{\mathsf{L}^{op}}$, we say that a **element of** P **at stage** ℓA is an element x of the set $P(\ell A)$. By Yoneda, these can be identified with natural transformations from ℓA to P (where we have silenced the Yoneda embedding). A map $\varphi : A \to B$ in L yields a map $\varphi : \ell B \to \ell A$ in $\mathsf{Set}^{\mathsf{L}^{op}}$, and hence maps elements of P at stage ℓA to elements of P at stage ℓB by composition; this is known as **restriction**, and written as $x|_\varphi$.

Smooth Reals In the topos $\mathsf{Set}^{\mathsf{L}^{op}}$, the smooth real line R can be identified as the functor $R = \ell C^\infty(\mathbb{R})$; elements of R at stage ℓA, or natural transformations $\ell A \to R$, are just called reals at stage ℓA. For $A = C^\infty(\mathbb{R}^n)/I$, this is an equivalence class $f(x) \bmod I$, where $f : \mathbb{R}^n \to \mathbb{R}$. The internal ring structure on R derives from a ring structure on each set of reals at a given stage ℓA given by simply taking pointwise addition and multiplication of functions mod I. The terminal object ("point") is given by $1 = \ell(C^\infty(\mathbb{R})/(x))$, and the object of nth order infinitesimals is $\ell(C^\infty(\mathbb{R})/(x^{n+1}))$. The **smooth interval object** $[a, b]$ is given by $\ell(C^\infty(\mathbb{R})/m^\infty_{[a,b]})$, where $m^\infty_{[a,b]}$ is the ideal consisting of functions that vanish on $[a, b]$. Again, we may analyze these objects by their elements at stage ℓA for $A = C^\infty(\mathbb{R}^n)/I$: for instance, the nth order infinitesimals are those smooth functions f such that $f^{n+1} \in I$. To prove all of this, we state the Kripke-Joyal semantics for $\mathsf{Set}^{\mathsf{L}^{op}}$: letting x be an element of X at stage ℓA, we have

- $\ell A \Vdash \psi(x) \wedge \phi(x)$ (resp. $\psi(x) \vee \phi(x)$) iff $\ell A \Vdash \psi(x)$ and (resp. or) $\ell A \Vdash \phi(x)$.
- $\ell A \Vdash \phi(x) \implies \psi(x)$ iff for every $f : \ell B \to \ell A$ in L, $\ell B \Vdash \phi(x|_f)$ implies $\ell B \Vdash \psi(x|_f)$.
- $\ell A \Vdash \exists y \in F\, \phi(x, y)$ iff there's an element y_0 of F at stage ℓA such that $\ell A \Vdash \phi(x, y_0)$.
- $\ell A \Vdash \forall y \in F\, \phi(x, y)$ iff for every $f : \ell A \to \ell B$ in L and element y_0 of F at stage ℓB we have

$\ell B \Vdash \phi(x|_f, y_0)$.

This allows us to prove that the KL axiom $\forall f \in R^D \exists! c \in R\, (\forall \epsilon \in D(f(\epsilon) = f(0) + c\epsilon))$ is valid for R, as well as the following **integration axiom**:

$$\forall f \in R^{[0,1]} \exists! F \in R^{[0,1]}\, (F' = f \wedge F(0) = 0)$$

The function F whose derivative is f is known as the integral of f.

While L^{op} consists of the finitely generated smooth algebras, we define G^{op} to consist of finitely generated smooth algebras whose ideals are determined by germs. The category G, then, consists of loci of the form $\ell(C^\infty(\mathbb{R}^n)/I)$, where I is such that $f \in I$ iff the germ of f at an arbitrary point $x \in Z(I)$ (i.e., $g(x) = 0$ for all $g \in I$) is in the germ of I. (The $\Rightarrow$ part is trivial, whereas the $\Leftarrow$ part is the real restriction, and where the name "ideal determined by germs" comes from). A second subcategory $F^{op} \subset L^{op}$ is given by smooth algebras of the form $C^\infty(\mathbb{R}^n)/I$, where I is **closed**, or such that if for every $x \in Z(I)$, the Taylor series of a function f at x resembles the Taylor series of some element of I at x, then $f \in I$. Finally, an ideal I of $C^\infty(\mathbb{R}^n)$ is **point determined** if $Z(f) \supseteq Z(I) \implies f \in I$. These generate the subcategory E^{op}.

Since the germ of a function contains its Taylor series, closed ideals are germ determined, so that $F^{op} \subset G^{op}$ and hence $F \subset G \subset L$; furthermore, since the Taylor series of f in particular tells us about its vanishing points, point determined ideals are closed, and hence $E \subset F \subset G \subset L$. Every ideal I of $C^\infty(\mathbb{R}^n)$ admits a smallest germ determined ideal $\tilde{I}$ given by the set of all f whose germ is an element of the germ of I at all points $x \in Z(I)$; this assignment is functorial, and is in fact left adjoint to the inclusion $G^{op} \to L^{op}$. The same formula gives us left adjoints to the inclusions $E^{op} \to F^{op}$, $F^{op} \to G^{op}$, and hence a sequence of *co*reflective subcategory inclusions $E \to F \to G \to L$. The right adjoints $L \to E$, $L \to F$, $L \to G$ are customarily denoted by γ, κ, and λ, respectively; we'll also denote the right adjoints $G \to E$, $G \to F$, and $F \to E$ by γ, κ, and γ, so that γ makes a finitely generated ideal in any of these categories point determined, κ makes an ideal closed, and λ makes an ideal germ determined.

Given a function $f \in C^\infty(\mathbb{R}^n)$, the most general solution to providing $C^\infty(\mathbb{R}^n)$ with an inverse of f is given by the smooth algebra $C^\infty(f^{-1}(\mathbb{R} - \{0\}))$. We write this algebra as $C^\infty(\mathbb{R}^n)\{f^{-1}\}$, and associate to it a canonical morphism $\eta_f : C^\infty(\mathbb{R}^n) \to C^\infty(\mathbb{R}^n)\{f^{-1}\}$ restricting a smooth g on $\mathbb{R}^n$ to the subset of $\mathbb{R}^n$ on which f doesn't vanish. We define $(C^\infty(\mathbb{R}^n)/I)\{f^{-1}\} = C^\infty(\mathbb{R}^n)\{f^{-1}\}/\eta_f(I)$; while this construction doesn't necessarily map elements of G^{op} to elements of G^{op}, $C^\infty(\mathbb{R}^n)/\{f^{-1}\}/\eta_f(I)$ will be finitely generated so long as

$C^\infty(\mathbb{R}^n)/I$ is, and hence we can obtain a germ determined locus $\lambda\ell((C^\infty(\mathbb{R}^n)/I)\{f^{-1}\})$ equipped with a canonical morphism into $\ell(C^\infty(\mathbb{R}^n)/I)$.

The Topos $\mathcal{G}$ We define a Grothendieck topology J on G as follows: a family $\{f_\alpha : \ell A_\alpha \to \ell A\}_{\alpha\in\Omega}$ is a covering family if for every $\alpha \in \Omega$ there's a function $b_\alpha \in A$ such that f_α factors through the canonical map $\lambda\ell(A\{b_\alpha^{-1}\}) \to \ell A$, and the family $\{\gamma f_\alpha\}_{\alpha\in\Omega}$ covers $\gamma\ell A$. J sends ℓA to its collection of covering families. The Grothendieck topos Sh(G, J) is denoted $\mathcal{G}$. As usual, we have a sheafification functor $-^{sh} : \mathrm{Set}^{G^{op}} \to \mathcal{G}$ left adjoint to the inclusion functor $\mathcal{G} \to \mathrm{Set}^{G^{op}}$, as well as a global sections functor $\Gamma : \mathcal{G} \to \mathrm{Set}$, $\Gamma(F) = F(1)$, right adjoint to the sheafification of the constant presheaf functor $\Delta(S)(\ell A) = S$. Writing $A = C^\infty(\mathbb{R}^n)/I$, this sheafification sends ℓA to the set of locally constant functions $Z(I) \to S$. Γ is also left adjoint to the functor B sending a set S to the sheaf sending ℓA to the set of all functions $Z(I) \to S$.

The Kripke-Joyal semantics for $\mathcal{G}$ are equivalent to those of $\mathrm{Set}^{L^{op}}$ for the operators $\wedge$, $\implies$, and $\forall$, but differ for the other connectives.

- $\ell A \Vdash \varphi(x) \vee \psi(x)$ iff there's a covering family $\{f_\alpha : \ell A_\alpha \to \ell A\}$ such that, for each α, $\ell A_\alpha \Vdash \varphi(x|_{f_\alpha})$ or $\ell A_\alpha \Vdash \psi(x|_{f_\alpha})$.
- $\ell A \Vdash \exists y \in F\,\phi(x, y)$ iff there's a covering family $\{f_\alpha : \ell A_\alpha \to \ell A\}$ such that, for each α, there's an element y_α of F at stage ℓA_α (i.e., $y_\alpha \in F(\ell A_\alpha)$) with $\ell A_\alpha \Vdash \phi(x, y_\alpha)$.
- $\ell A \Vdash \neg\phi(x)$ iff for every $f : \ell B \to \ell A$ such that $\ell B \Vdash \phi(x|_f)$, $B = 0$.

Just as in $\mathrm{Set}^{L^{op}}$, $R = \mathcal{G}(-, \ell C^\infty(\mathbb{R}))$ is a commutative ring object with orders $<, \leq$. The difference is that, in $\mathcal{G}$, R satisfies the following additional properties: $\mathcal{G} \models \neg(0 = 1)$, $\mathcal{G} = \forall x, y \in R(x + y \in U(R) \implies x \in U(R) \vee y \in U(R))$, and $\mathcal{G} \models \forall x \in R\,\exists n \in \mathbb{N}(x < n)$. Here, $\mathbb{N}$ is the natural numbers object/sheaf sending ℓA to the set of locally constant functions $\ell A \to \mathbb{N}$. The first two statements state that R is a local ring, and the third states that R is Archimedean. Furthermore, R satisfies the **field axiom**

$$\forall x_1, \ldots, x_n \in R\,(\neg(x_1 = 0 \wedge \ldots \wedge x_n = 0) \implies (x_1 \in U(R) \vee \ldots \vee x_n \in U(R)))$$

as well as the Kock-Lawvere and integration axioms from $\mathrm{Set}^{L^{op}}$. Locality is often studied in the form of an **apartness relation** # whereby $x\#y$ if $x - y \in U(R)$, or equivalently if $x < y \vee x > y$.

If we replace G with F and λ in the definition of a covering family with κ, we obtain a Grothendieck topology J on F whose corresponding Grothendieck topos Sh(F, J) is denoted $\mathcal{F}$; the entirety of the above discussion of $\mathcal{G}$ holds for $\mathcal{F}$.

3.3.4 Cohesive Topoi

Cohesion A topos **over** a base topos $\mathcal{B}$ is a topos $\mathcal{E}$ equipped with a geometric morphism $f = (f^* : \mathcal{B} \to \mathcal{E} \dashv f_* : \mathcal{E} \to \mathcal{B})$. For instance, if $\mathcal{E} = \mathrm{Sh}(C, J)$ is a Grothendieck topos and $\mathcal{B} = \mathsf{Set}$, there is a natural geometric moprhism which has as its left adjoint the sheafification of the constant presheaf functor and as its right adjoint the global sections functor $\Gamma(F) = \mathcal{E}(1, F)$.

The topos $\mathcal{E}$ over $\mathcal{B}$ is **cohesive** if f^* has a further left adjoint $f_!$ which preserves all finite products, including the terminal object, and f_* has a further right adjoint $f^!$, so that we have an adjoint quadruple $f_! \dashv f^* \dashv f_* \dashv f^!$. The canonical example is when $\mathcal{B} = \mathsf{Set}$ and f_* is the global sections functor Γ. As such, we often denote f_* by Γ, f^* by Disc, $f^!$ by coDisc, and $f_!$ by Π_0, giving us an adjunction

$$\Pi_0 \dashv \mathrm{Disc} \dashv \Gamma \dashv \mathrm{coDisc}$$

The idea is that the global sections functor sends an object $X \in \mathcal{E}$ to its set of (global) elements, and its left and right adjoints send a set to its corresponding *discrete* and *codiscrete*, or indiscrete, spaces in $\mathcal{E}$. Π_0 sends X to its set of connected components, a la $\pi_0 : \mathsf{Top} \to \mathsf{Set}$. This adjoint quadruple induces an adjoint triple $\mathrm{Disc} \circ \Pi_0 \dashv \mathrm{Disc} \circ \Gamma \dashv \mathrm{coDisc} \circ \Gamma$, all of which are endofunctors on $\mathcal{E}$. The functor $\mathrm{Disc} \circ \Pi_0$ drops information internal to connected components while identifying each connected component, keeping the shape of an object X: it is known as the **shape modality** $\int$. $\mathrm{Disc} \circ \Gamma$ sends X to the discrete topology on its points, detaching its points: it is known as the **flat modality** $\flat$. $\mathrm{coDisc} \circ \Gamma$ does the opposite, dissolving the structure of X into a cohesive "blob": it is known as the **sharp modality** $\sharp$. The $\mathrm{Disc} \dashv \Gamma$ adjunction has a counit $\eta : \mathrm{Disc} \circ \Gamma \to 1$ yielding for every X a canonical morphism $\epsilon^\flat_X : \flat X \to X$: in this way, $\flat$ is not just an endofunctor but an idempotent comonad. In the same way, $\int$ and $\sharp$ are idempotent monads on $\mathcal{E}$, with units $\eta^\sharp : 1 \to \sharp$ and $\eta^\int : 1 \to \int$. Objects for which $\epsilon^\flat_X : \flat X \cong X$ are known as **discrete**, and objects for which $\eta^\sharp_X : X \cong \sharp X$ are known as **codiscrete**. If $\eta^\sharp_X$ is at least a monomorphism, X is known as **concrete**. Since $\flat$ and $\sharp$ are both idempotent, the subcollection of (co)discrete objects of $\mathcal{E}$ assembles into a subcategory given by the image $\flat\mathcal{E}$ ($\sharp\mathcal{E}$).

For an example of a cohesive topos, put the following Grothendieck topology on CartSp: a **differentiably good open cover** of $\mathbb{R}^n$ is a covering $\{f_i : U_i \to \mathbb{R}^n\}$, where each $U_i \subseteq \mathbb{R}^{n_i}$, such that each non-empty finite intersection of the $f_i(U_i)$ is diffeomorphic to $\mathbb{R}^n$. With J sending $\mathbb{R}^n$ to its set of differentiably good open covers, we define $\mathrm{Sh}(\mathsf{CartSp}, J)$ to be the topos of **smooth sets**, $\mathcal{S}mooth\mathcal{S}et$. A smooth set X can be thought of as a collection of sets $\{X^n\}_{n \in \mathbb{N}}$, where

$X^n = X(\mathbb{R}^n)$ is thought of as the set of n-dimensional *plots* of X, along with maps $Xf : X^n \to X^m$ for every smooth "coordinate transformation" $f : \mathbb{R}^m \to R^n$ satisfying the sheaf conditions. For M a smooth manifold, the functorial assignment $M, \mathbb{R}^n \mapsto \mathsf{Diff}(\mathbb{R}^n, M)$ yields a full and faithful embedding of Diff into $\mathcal{S}mooth\mathcal{S}et$, of which we have as a special case the Yoneda embedding $\text{よ}(\mathbb{R}^m) = \mathsf{CartSp}(-, M) = \mathsf{Diff}(-, M)$. An important subcategory of $\mathcal{S}mooth\mathcal{S}et$ is given by its concrete objects, which are known as **diffeological spaces**. These are characterized by the property that they can be identified with an actual set X, with $X(\mathbb{R}^n)$ being a subset of $\mathsf{Set}(\mathbb{R}^n, X)$.

The adjoints Disc and Γ between $\mathcal{S}mooth\mathcal{S}et$ and Set are given by the constant sheaf and global sections functor, as usual. For a set S, coDisc(S) is the sheaf that sends $\mathbb{R}^n$ to $\mathsf{Set}(\mathbb{R}^n, S)$; for a smooth set X, $\Pi_0(X)$ is given by the colimit over the $X(\mathbb{R}^n)$.

Elasticity Given a cohesive topos $(\mathcal{E}, \Pi_{\mathcal{E}} \dashv \mathrm{Disc}_{\mathcal{E}} \dashv \Gamma_{\mathcal{E}} \dashv \mathrm{coDisc}_{\mathcal{E}})$ over Set, take a cohesive topos $(\mathcal{F}, \Pi_{\mathcal{F}} \dashv \mathrm{Disc}_{\mathcal{F}} \dashv \Gamma_{\mathcal{F}} \dashv \mathrm{coDisc}_{\mathcal{F}})$ over Set, and equip $\mathcal{F}$ with a functor $\iota_{inf} : \mathcal{E} \to \mathcal{F}$ with a series of left adjoints

$$\iota_{inf} \dashv \Pi_{inf} \dashv \mathrm{Disc}_{inf} \dashv \Gamma_{inf}$$

such that $\Pi_{\mathcal{F}} = \Pi_{\mathcal{E}} \circ \Pi_{inf}$ and likewise for $\mathrm{Disc}_{\mathcal{E}}$ and $\Gamma_{\mathcal{E}}$. We say that $\mathcal{F}$ is an **elastic topos** over $\mathcal{E}$, or **differentially cohesive**. So, the situation is as follows:

$$\begin{array}{ccccc}
 & & & \xrightarrow{\ \iota_{inf}\ } & \\
 & \xleftarrow{\ \Pi_{\mathcal{E}}\ } & & \xleftarrow{\ \Pi_{inf}\ } & \\
\mathsf{Set} & \xrightarrow{\ \mathrm{Disc}_{\mathcal{E}}\ } & \mathcal{E} & \xrightarrow{\ \mathrm{Disc}_{inf}\ } & \mathcal{F} \\
 & \xleftarrow{\ \Gamma_{\mathcal{E}}\ } & & \xleftarrow{\ \Gamma_{inf}\ } & \\
 & \xrightarrow{\ \mathrm{coDisc}_{\mathcal{E}}\ } & & & \\
 & \multicolumn{3}{c}{\xrightarrow{\qquad \mathrm{coDisc}_{\mathcal{F}} \qquad}} &
\end{array}$$

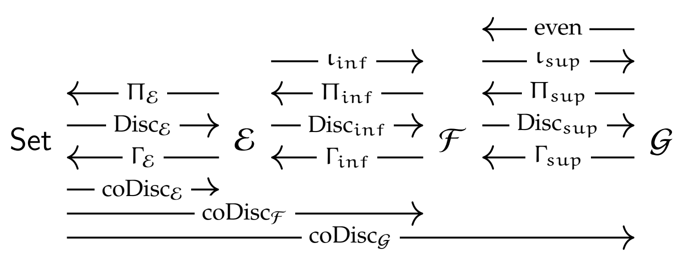
Again, each of $\mathcal{E}$ and $\mathcal{F}$ have their own co/monads $(\int_{\mathcal{E}}, \flat_{\mathcal{E}}, \sharp_{\mathcal{E}})$, $(\int_{\mathcal{F}}, \flat_{\mathcal{F}}, \sharp_{\mathcal{F}})$, but we now have an addiitional adjoint triple $\iota_{inf} \circ \Pi_{inf} \dashv \mathrm{Disc}_{inf} \circ \Pi_{inf} \dashv \mathrm{Disc}_{inf} \circ \Gamma_{inf}$ of endofunctors on $\mathcal{F}$. $\iota_{inf} \circ \Pi_{inf}$ is an idempotent comonad known as the **reduction modality** $\Re$, $\mathrm{Disc}_{inf} \circ \Pi_{inf}$ an idempotent monad known as the **infinitesimal shape modality** $\Im$, and $\mathrm{Disc}_{inf} \circ \Gamma_{inf}$ an idempotent comonad known as the **infinitesimal flat modality** $\&$. We have a category $\Re\mathcal{F}$ of **reduced** objects and a category $\Im\mathcal{F}$ of **coreduced** objects. It is relatively straightforward to show that $\&\flat X \cong X$, and therefore $\&\mathcal{F} \supseteq \flat\mathcal{F}$, and likewise $\Im\mathcal{F} \supseteq \int\mathcal{F}$. We write these relations as $\& > \flat$ and $\Im > \int$.

For an example, consider the category $\mathsf{FormalCartSp}$ whose objects are smooth loci of the form $\mathbb{R}^n \times D$, where ℓW is the formal dual of a Weil algebra, and whose morphisms are smooth maps; these are known as **infinitesimally thickened Cartesian spaces**. The coverings are of the form $\{f_i \times \mathrm{id} : U_i \times D \to \mathbb{R}^n \times D\}$ for $\{f_i : U_i \to \mathbb{R}^n\}$ a covering of $\mathbb{R}^n$. The sheaf topos over FormalCartSp is known as the **Cahiers topos** $\mathcal{CT}$; its objects are known as **formal smooth sets**. The inclusion of CartSp into FormalCartSp induces via left Kan extension an inclusion functor $\iota_{inf} : \mathcal{SmoothSet} \to \mathcal{CT}$: namely, a smooth set X sends an infinitesimally thickened Cartesian space $\mathbb{R}^n \times D$ to the set $\coprod_{m \in \mathbb{N}} \mathrm{FormalCartSp}(\mathbb{R}^n \times D, \mathbb{R}^m) \times X(\mathbb{R}^m)$ quotiented by the relation identifying $(\alpha : \mathbb{R}^n \times D \to \mathbb{R}^m, \beta : \mathbb{R}^m \to X)$ with $(\alpha' : \mathbb{R}^n \times D \to \mathbb{R}^{m'}, \beta' : \mathbb{R}^{m'} \to X)$ if there's an $f : \mathbb{R}^m \to \mathbb{R}^{m'}$ such that $f\alpha = \alpha', \beta' f = \beta$. The right adjoint Π_{inf} to this inclusion functor just restricts a formal smooth set $\mathcal{X}$ to the smooth set $X(\mathbb{R}^n) = \mathcal{X}(\mathbb{R}^n)$. Disc_{inf} sends a smooth set X to the formal smooth set $\mathcal{X}(\mathbb{R}^n \times D) = X(\mathbb{R}^n)$, and Γ_{inf} is given by the right Kan extension of a formal smooth set X along Disc_{inf}. The elastic topos $\mathcal{CT}$ is, unlike $\mathcal{SmoothSet}$, suited for synthetic differential geometry, due to the addition of infinitesimals.

Solidity Take a cohesive topos $\mathcal{E}$ over Set and an elastic topos $\mathcal{F}$ over $\mathcal{E}$, with the same notation as before. We now add the third layer of cohesion: take a cohesive topos $\mathcal{G}$ over Set which is elastic over $\mathcal{E}$, bearing a functor $\Gamma : \mathcal{G} \to \mathcal{E}$ fitting in an adjoint quadruple $\iota \dashv \Pi_0 \dashv \mathrm{Disc} \dashv \Gamma$. Equip $\mathcal{G}$ with a functor even : $\mathcal{G} \to \mathcal{F}$ fitting in an adjoint quintuple

$$\mathrm{even} \dashv \iota_{sup} \dashv \Pi_{sup} \dashv \mathrm{Disc}_{sup} \dashv \Gamma_{sup}$$

such that $\iota = \iota_{sup} \circ \iota_{inf}$, likewise for Π, Disc, and Γ, and $\Pi_{\mathcal{G}} = \Pi_{\mathcal{E}} \circ \Pi_{inf} \circ \Pi_{sup}$, likewise for $\mathrm{Disc}_{\mathcal{G}}$ and $\Gamma_{\mathcal{G}}$. The situation is as follows:



We again have a triplet of endofunctors: the idempotent monad $\iota_{sup} \circ$ even known as the **fermionic modality** $\rightrightarrows$, the idempotent comonad $\iota_{sup} \circ \Pi_{sup}$ known as the **bosonic modality** $\rightsquigarrow$, and the idempotent monad $\mathrm{Disc}_{sup} \circ \Pi_{sup}$ known as the **rheonomy modality** Rh. The topos $\mathcal{G}$ is known as **solid**, or super-differentially cohesive, over $\mathcal{F}$. By being elastic over $\mathcal{E}$ and cohesive over Set, it also has the two previous triplets of modalities, and admits the relations $\rightsquigarrow > \Re$ and $\mathrm{Rh} > \Im$. We therefore have three generations of modalities, which [nLab authors, 2020] arranges into the progression

$$\begin{array}{ccccccccc}
\mathrm{id} & \dashv & \mathrm{id} & & & & & & \\
\vee & & \vee & & & & & & \\
\rightrightarrows & \dashv & \rightsquigarrow & \dashv & \mathrm{Rh} & & & & \text{solidity} \\
 & & \vee & & \vee & & & & \\
 & & \Re & \dashv & \Im & \dashv & \& & & \text{elasticity} \\
 & & & & \vee & & \vee & & \\
 & & & & \int & \dashv & \flat & \dashv & \sharp \quad \text{cohesion} \\
 & & & & & & \vee & & \vee \\
 & & & & & & \varnothing & \dashv & *
\end{array}$$

including the trivial adjunctions $\mathrm{id} \dashv \mathrm{id}$ and $\varnothing \dashv *$, where $\varnothing$ and $*$ are the constant endofunctors on the initial and terminal objects, respectively.

Our example, building on the previous two examples, is inspired by supersymmetry: in physics, fermions are represented by (for now, real) numbers $\psi_i, \psi_j, \ldots$ which *anticommute*: $\psi_i\psi_j = -\psi_j\psi_i$. Bosons, on the other hand, are reals $\theta_i, \theta_j, \ldots$ which commute, $\theta_i\theta_j = \theta_j\theta_i$. Define the real **Grassmann algebra** $\Lambda^\bullet\mathbb{R}^q$ to be the $\mathbb{R}$-algebra freely generated by $\{\psi_1, \ldots, \psi_q\}$ under the relations $\psi_i\psi_j = -\psi_j\psi_i$. We define the **super-Cartesian space** $\mathbb{R}^{p|q}$ by the relation $C^\infty(\mathbb{R}^{p|q}) = C^\infty(\mathbb{R}^p) \otimes_{\mathbb{R}} \Lambda^\bullet\mathbb{R}^q$. This is a commutative algebra over $\mathbb{R}$, and hence an object of the category $\mathsf{CAlg}_{\mathbb{R}}$, and is $\mathbb{Z}/2\mathbb{Z}$ graded, as we can split it into "fermions" with degree 1 and "bosons" with degree 0. The degree of an object x is denoted $|x|$; we have $|xy| = |x||y| \bmod 2$ (note: $(\psi_i\psi_j)\psi_k = -\psi_i\psi_k\psi_j = \psi_k(\psi_i\psi_j)$) and $xy = (-1)^{|x||y|}yx$. These relations define a **commutative superalgebra**, an object of a category $\mathsf{sCAlg}_{\mathbb{R}}$. Just as Diff embeds fully and faithfully into $\mathsf{CAlg}_{\mathbb{R}}^{op}$ via $C^\infty(-)$, $\mathbb{R}^{p|q}$ can be identified within $\mathsf{sCAlg}_{\mathbb{R}}^{op}$ as the formal dual of $C^\infty(\mathbb{R}^{p|q})$. The set of all $\mathbb{R}^{p|q}, p, q \in \mathbb{N}$, forms the subcategory $\mathsf{SuperCartSp}$. The category $\mathsf{SuperFormalCartSp}$ is defined in a manner completely analogous to $\mathsf{FormalCartSp}$, as well as the topoi $\mathcal{SSS} := \mathcal{SuperSmoothSet}$ and $\mathcal{SFSS} := \mathcal{SuperFormalSmoothSet}$. $\mathcal{SFSS}$ is solid over the Cahiers topos $\mathcal{CT}$, with the functor even stripping the degree 1 part from a super formal smooth set.

3.4 Physical Models

3.4.1 General Relativity

Synthetic differential geometry allows us to construct an intuitionistic theory of spacetime in which general relativity can be constructed; we will use the model of SDG provided by the topos $\mathcal{G}$ of sheaves over the site of finitely generated smooth algebras with germ determined ideals. Our plan will be to set up the elements of classical Riemannian geometry (connections, curvature, and so on) in a synthetic manner, and study the interpretation of Einstein's equations in $\mathcal{G}$.

Connections and Curvature An **infinitesimal** n**-cube** on an object M is an element of $M^{D^n} \times D^n$, and an **infinitesimal** n**-chain** is an element of the free R-module $C_n(M)$ generated by all infinitesimal n-cubes on M. Writing $I = [0,1]$, a **finite** (or "big") n**-cube** on M is a morphism $I^n \to M$, and a **finite** n**-chain** an element of the free R-module $\Gamma_n(M)$ generated by finite n-cubes.

A **affine connection** on a microlinear space M is a bilinear morphism $\nabla : TM \times_M TM \to M^{D \times D}$ (where the pullback is taken over the morphisms $v \mapsto v(0)$, so these are two tangent vectors at the same point) such that $\nabla(v,w)(d_1,0) = v(d_1)$ and $\nabla(v,w)(0,d_2) = w(d_2)$. If $\nabla(v,w)(d_1,d_2) = \nabla(w,v)(d_2,d_1)$, ∇ is said to be **torsion-free**. From a connection ∇ on M, we may define another function τ which associates to each $(v,d) \in TM \times D$ a **parallel transport** $\tau_d(v,-) : \pi^{-1}(v(0)) \cong \pi^{-1}(v(d))$; this map is linear in both v and its argument, is the identity for $d = 0$, and $\tau_d(\lambda v, -) = \tau_{\lambda d}(v,-)$. We identify $\tau_d(v,w)$ with the parallel transport of w along v for an infinitesimal period of time d. Specifically, $\tau_{d_1}(v,w)(d_2)$ is defined to be $\nabla(v,w)(d_1,d_2)$.

Given a connection ∇ on a microlinear space M, we would like to define the Riemann curvature tensor in terms of the parallel transport of a vector along the boundary of an infinitesimal 2-chain. Given such a 2-chain $(\gamma, d_1, d_2) \in M^{D^2} \times D^2$ based at a point $x = \gamma(0,0)$, we do this as follows: take a vector v and transport it along $\gamma(-,0)$ for a period of d_1 "seconds". Transport the new vector along $\gamma(d_1,-)$ for a period of d_1 seconds, transport backwards along $\gamma(0,-)$ for d_2 seconds and finally transport backwards along $\gamma(-,d_2)$ for d_2 seconds, before subtracting v from the result. This gives a preliminary map

$$R'(\gamma, d_1, d_2, v) = \tau_{d_2}^{-1}(\gamma(-,d_2), \tau_{d_2}^{-1}(\gamma(0,-), \tau_{d_1}(\gamma(d_1,-), \tau_{d_1}(\gamma(-,0), v)))) - v$$

Being bilinear in both d_1 and d_2, we may define a map $\varphi(d_1, d_2) = R'(\gamma, d_1, d_2, v)$ which induces by microlinearity of T_xM a function $\psi : D \to T_xM$ such that $\psi(d_1d_2) = \varphi(d_1, d_2)$. By KL, this can be written as $\psi(d) = d\hat{v}$ for a unique $v \in T_xM$. We define $R'' : M^{D\times D} \times_M TM \to TM$ to send a pair (γ, v) to this $\hat{v}$, and define the **Riemann curvature tensor** $R : TM\times_M TM\times_M TM \to TM$ by $R(v_1, v_2, v_3) = R''(\nabla(v_1, v_2), v_3)$. If M is a formal manifold, we may work in local coordinates: the connection ∇ becomes a function that takes in a point $x \in M$ along with two vectors $v, w \in R^n$, and returns an element of $M \times R^n \times R^n \times R^n$. The fourth component of this tuple is denoted ∇_4, and used to define the **Christoffel symbols**: in a basis $\{e_1, \ldots, e_n\}$ of R^n, these are given by ${\Gamma^i}_{jk}(x) = \pi_i(\nabla_4(x, e_k, e_j))$. The Riemann curvature tensor decomposes into components in the usual manner: ${R^\ell}_{ijk} = \partial_j {\Gamma^\ell}_{ki} - \partial_k {\Gamma^\ell}_{ji} + {\Gamma^\ell}_{jm}{\Gamma^m}_{ki} - {\Gamma^\ell}_{km}{\Gamma^m}_{ji}$ (again, at every point).

Hence, to a formal manifold $M \in \mathcal{G}$ we may associate a Riemann curvature tensor ${R^\ell}_{ijk}$ to a connection ∇. This gives us a Ricci curvature tensor $R_{ik} = {R^\ell}_{i\ell k}$ and, with a Riemannian metric g_{ij}, a scalar curvature $R = g^{ij}R_{ij}$ and Einstein tensor $G_{ij} = R_{ij} - \frac{1}{2}Rg_{ij}$.

Einstein's Equations Consider R^4 filled with dust with 4-velocity u^i and density ρ. The classical Einstein equations read $G_{ij} = T_{ij} = \kappa c^2 \rho u_i u_k$, where κ is Einstein's constant. In $\mathcal{G}$, real numbers become elements of R at stage ℓA for $A = C^\infty(\mathbb{R}^n)/I$; these are natural transformations from $よ(\ell A)$ to $R = よ(\ell C^\infty(\mathbb{R}))$, which by Yoneda are in bijection with smooth functions $\varphi : \mathbb{R}^n \to \mathbb{R}$ modulo I. So, using $\mathcal{G}$ as a model for SDG, an arbitrary real number $r \in R$ at stage ℓA is really a "parametrized" element of $\mathbb{R}$, changing smoothly as we vary the point $v \in \ell A$. Similarly, an event, or element of R^4, at stage ℓA is really a smooth function $\mathbb{R}^n \to \mathbb{R}^4, v \mapsto (x^0(v), x^1(v), x^2(v), x^3(v)) \bmod I$. Taking the reals at stage $1 = よ(C^\infty(\{*\}))$ recovers the usual set $\mathbb{R}$. So, in SDG, the Einstein equations $G_{ij}(x) = T_{ij}(x), x \in R^4$ carry over without modification at stage 1, stating that two pairs of 16 reals coincide at every point in $\mathbb{R}^4$ ($G_{00}(x)(*) = T_{00}(x)(*)$ and so on). At stage $\ell C^\infty(\mathbb{R})$, the equations state that two pairs of 16 *smooth curves* through $\mathbb{R}^4$, assigned to each point in $\mathbb{R}^4$, coincide; at stage $\ell C^\infty(\mathbb{R}^2)/I$, they become surfaces $\varphi : \mathbb{R}^2 \to \mathbb{R}^4$ modulo the ideal I, and so on. [Guts and Zvyagintsev, 2000] interprets the Einstein equations for a dusty universe at various stages.

This interpretation of general relativity can be carried out in any other smooth topos, thereby inheriting its internal logic instead of $\mathcal{G}$'s logic; to quote [Guts and Grinkevich, 1996],

> "The resulting space-time theory will be non-classical, different from that of the Minkowski space-time. This is a *new* theory of space-time, created in a purely logical manner. It will reflect the real space-time properties to the same extent as the development of mathematical

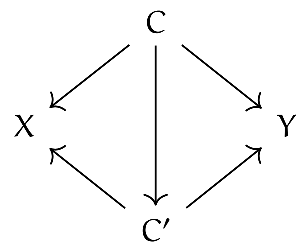
abstractions accompanies the development of the real world."

3.4.2 Classical Mechanics

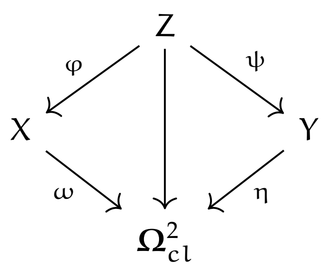
Here's where we bring in the language of cohesive topoi. Let $\mathcal{S} = \mathcal{S}mooth\mathcal{S}et$ be the cohesive topos of smooth sets, constructed above as the sheaf topos on CartSp with the differentiably good open cover topology. Letting $\Omega^p_{cl}(M)$ be the set of closed p-forms on a manifold M, we define a smooth set $\mathbf{\Omega}^p$ by $\mathbf{\Omega}^p(\mathbb{R}^n) = \Omega^p(\mathbb{R}^n)$, as well as a morphism $\mathbf{d} : \mathbf{\Omega}^p \to \mathbf{\Omega}^{p+1}$, $d_{\mathbb{R}^n} = d : \Omega^p(\mathbb{R}^n) \to \Omega^{p+1}(\mathbb{R}^n)$. This smooth set is a "universal moduli space" for p-forms, in the sense that for any smooth manifold M, considered as a smooth set, there's a natural bijection between morphisms $M \to \mathbf{\Omega}^p$ and p-forms on M. Note that the machinery of smooth sets is necessary to solve this moduli problem: $\mathbf{\Omega}^p$ is *not* the image of a smooth manifold, nor is it even a diffeology. However, this anomaly allows us to lift the definition of p-forms from manifolds to smooth sets: given an arbitrary smooth set X, a p-form ω on X is a morphism $X \to \mathbf{\Omega}^p$, and if $d\omega := \mathbf{d} \circ \omega = 0$, ω is **closed**. There is an object $\mathbf{\Omega}^p_{cl}$ of closed p-forms given by $\mathbf{\Omega}^p_{cl}(\mathbb{R}^n) = \{\text{closed p-forms on } \mathbb{R}^n\}$.

Presymplectic Sets A **presymplectic smooth set** is a pair (X, ω), where X is a smooth set and ω a closed 2-form on X. (While ω is closed, we haven't said anything about nondegeneracy, hence *pre*symplectic), or equivalently a morphism $X \to \mathbf{\Omega}^2_{cl}$. A p-form on X is really just an assignment to every plot $\phi \in X(\mathbb{R}^n)$ of a p-form $\omega_{\mathbb{R}^n}(\phi)$ on $\mathbb{R}^n$, so we can add and multiply them, and in particular we can take the **tensor product** of presymplectic sets $(X, \omega) \otimes (Y, \eta)$, which assigns to every product plot $\phi \times \psi X(\mathbb{R}^n) \times Y(\mathbb{R}^n)$ the sum $\omega_{\mathbb{R}^n}(\phi) + \eta_{\mathbb{R}^n}(\psi)$. A **symplectomorphism** between presymplectic sets (X, ω) and (X', ω') is just a morphism $\phi : X \to X'$ such that $\omega'\phi = \omega$. Hence, presymplectic sets assemble into the *slice topos* $\mathcal{S}/\mathbf{\Omega}^2_{cl}$. A presymplectic subset of a presymplectic set (X, ω) is simply a subobject $\phi : X' \rightarrowtail X$, which induces by composition a presymplectic set $(X', \omega|_{X'} := \omega\phi)$. If $(\omega\phi)_{\mathbb{R}^n} : X'(\mathbb{R}^n) \to \mathbf{\Omega}^2_{cl}(\mathbb{R}^n) = \Omega^2_{cl}(\mathbb{R}^n)$ is the constant morphism $x \mapsto 0$, and the dimension of X′ is half that of X, we call X′ a **Lagrangian subset** of X.

Given two objects X, Y, we define a **correspondence** to be a diagram of the form $X \leftarrow C \to Y$, and a **equivalence** of correspondences to be an isomorphism $C \cong C'$ forming a commutative diagram



Given two correspondences $X \leftarrow C \rightarrow Y \leftarrow C' \rightarrow Z$, their composition along Y is defined to be the correspondence $X \leftarrow C \times_Y C' \rightarrow Z$. Hence, we can for an arbitrary topos $\mathcal{E}$ define a 2-category $\mathsf{Corr}(\mathcal{E})$ of correspondences whose 1-morphisms $X \rightarrow Y$ are correspondences $X \leftarrow C \rightarrow Y$ and whose 2-morphisms are morphisms between correspondences. The category $\mathsf{Corr}(\mathcal{S}/\mathbf{\Omega}^2_{cl})$, for instance, has as its objects commutative squares



This is a symmetric monoidal category under the tensor product $(X, \omega) \otimes (Y, \eta) = (X \times Y, \omega + \eta)$ and unit $(*, 0)$.

Smooth Groupoids Suppose that instead we would like $X(\mathbb{R}^n)$ to capture not just plots of $\mathbb{R}^n$ in X, but gauge transformations – nontrivial isomorphisms – between plots. To do this, we need a groupoid structure on each $X(\mathbb{R}^n)$. A **smooth groupoid** is a functor $X : \mathsf{CartSp}^{op} \rightarrow \mathsf{Grpd}$ such that both the set of objects of $X(\mathbb{R}^n)$, denoted $X_0(\mathbb{R}^n)$, and the set of morphisms, denoted $X_1(\mathbb{R}^n)$, assemble into smooth sets. The category of smooth groupoids is denoted $\mathcal{S}mooth\mathcal{G}rpd$; this is just a "refinement" of $\mathcal{S}mooth\mathcal{S}et$, and we'll also denote it $\mathcal{S}$. We may obtain smooth groupoids by taking a smooth set X with an action of a smooth group G, and taking the **smooth homotopy quotient** $X//G$, whose objects $(X//G)_0(\mathbb{R}^n)$ are the objects of $X(\mathbb{R}^n)$, and whose morphisms are of the form $x \rightarrow gx$. For X an arbitrary one-point space, $X//G$ is a groupoid with a single object and an automorphism for each $g \in G$, with composition of morphisms given by composition of group elements. This groupoid is known as $\mathbf{B}G$. We define $\mathbf{B}U(1)_{conn}$ to be the smooth groupoid to send $\mathbb{R}^n$ to the groupoid $\Omega^1(\mathbb{R})//\mathsf{Diff}(\mathbb{R}^n, U(1))$ (where the composition of two smooth functions $f, g : \mathbb{R}^n \rightarrow U(1)$ is $(f \cdot g)(v) = f(v) \cdot g(v)$).

3.4.3 Quantum Mechanics

Take a smooth topos $\mathcal{E}$ with smooth real line R, and denote by U(R) the subobject of invertible (non-infinitesmal) elements of R. Assume that R satisfies the **field axiom**,

$$\forall x_1, \ldots, x_n \in R\,(\neg(x_1 = 0 \wedge \ldots \wedge x_n = 0) \implies (x_1 \in U(R) \vee \ldots \vee x_n \in U(R)))$$

(For instance, we can again let $\mathcal{E} = \mathcal{G}$). In particular, for $n = 1$ we have $\forall x \in R\,(x \neq 0 \implies x \in U(R))$. Denoting by C the complex numbers object (a 2-dimensional R-algebra, which also satisfies the field axiom), we define a **inner product** on an R-module V to be a symmetric, bilinear map $\langle -, - \rangle : V \times V \to C$ satisfying $v \neq 0 \implies \langle v, v \rangle > 0$. Note that, for $V = R$, we have for $x \neq 0$ that $\langle x, x \rangle = x^2 \langle 1, 1 \rangle > 0$, implying that $x^2 = 0$ and hence $x \in U(R)$; it follows that the existence of an inner product on R relies on the field axiom for $n = 1$.

We'll analyze the case of a spin 1/2 interaction, first in the classical case studied in [Sakurai et al., 2014], and then in the case of SDG, exposited in [Fearns, 2002].

The Stern-Gerlach Experiment In the Stern-Gerlach experiment, silver atoms are shot at a target, passing through an inhomogeneous magnetic field $\vec{B}$ which splits the silver atoms along the z axis. The electron shell structure of silver is 2, 8, 18, 18, and 1: four full shells, followed by a fifth shell with a single electron. The first four shells cancel each other out magnetically, so the magnetic moment $\vec{\mu}$ of the atom is proportional to the spin $\vec{S}$ of the one electron. If the electron behaved classically, the magnetic moment of the atom along the z axis, μ_z, would be distributed anywhere between $-|\vec{\mu}|$ and $|\vec{\mu}|$, resulting in the silver atoms forming a continuous interval on the target. What we observe in practice is two distinct spots on the target, indicating that the electron spin along the z axis is either fully up, $S_z = \hbar/2$, or fully down, $S_z = -\hbar/2$. The same holds when we reorient the machine to split the atoms along the x or y axes, suggesting that the electron's spin, when measured along a given axis, will take either an up or down spin along that axis. We model this as follows: we have three axes x, y, z and three operators S_x, S_y, S_z, each of which has two eigenvectors with eigenvalues $\pm\hbar/2$. We can model these operators as elements of $\mathbb{C}^{2\times 2}$: recalling the definition of the Pauli matrices

$$\sigma^1 = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \quad \sigma^2 = \begin{bmatrix} 0 & -i \\ i & 0 \end{bmatrix} \quad \sigma^3 = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$$

we write $S_{x_i} = \frac{\hbar}{2}\sigma^i$. So the spin of an electron with spin up along the z axis is modeled by the ket $|S_z;+\rangle = [1,0]^T$, and likewise $|S_y;+\rangle = [1,i]^T/\sqrt{2}$, $|S_x;+\rangle = [1,1]^T/\sqrt{2}$.

Microlinear Lie Groups Moving to a smooth topos $\mathcal{E}$, define the microlinear group $G = SO(3)$ to be the subobject of $R^{3\times3}$ consisting of the orthogonal matrices with determinant 1. With matrix multiplication, this is a Lie group internal to $\mathcal{E}$ with identity $e = I_3$. The fiber T_eG, consisting of all $f : D \to G$ such that $f(0) = e$, then has a bilinear operation $[-,-] : T_eG \times T_eG \to T_eG$ given as $[v,w](d_1d_2) = w(-d_2)v(-d_1)w(d_2)v(d_1)$. This is antisymmetric and satisfies the Jacobi identity, so we call it the Lie algebra $\mathfrak{g}$ associated to the Lie group G. $\mathfrak{so}(3)$ is, in fact, isomorphic to the Lie algebra $\mathfrak{su}(2)$ generated by the Pauli matrices, implying that we can consider these matrices, and hence the spin operators themselves, as elements of T_eG.

Now, suppose we have a system consisting of two interacting electrons, the total energy being encapsulated in a unitary Hamiltonian operator H. The classical time-dependent Schrödinger equation expressing the evolution of a time-dependent state $|\psi;t\rangle$ is $i\hbar\frac{d}{dt}|\psi;t\rangle = H|\psi;t\rangle$. In SDG, we take $t \in R$, $d \in D$, and instead write $|\psi;t+d\rangle = |\psi;t\rangle - \frac{id}{\hbar}H|\psi;t\rangle$. As proven in the paper [Kock, 1986], if $\mathcal{E}$ is **well-adapted**, possessing a full and faithful functor $\mathsf{Diff} \to \mathcal{E}$, then we have the following integration axiom for a Lie group G with Lie algebra $\mathfrak{g}$:

$$\forall f \in \mathfrak{g}^R \exists! F \in G^R \left(F(0) = e \wedge \forall t \in R \forall d \in D \left(F(t+d)F(t)^{-1} = f(t)(d)\right)\right)$$

The Hamiltonian is a member of the Lie group $U(4)$, and an infinitesimal perturbation to it, as expressed by the SDG Schrödinger equation, is a member of $\mathfrak{u}(4)$; by the integration axiom, this can be integrated to obtain a unique time evolution of $|\psi\rangle$.

While computing actual results in a well-adapted topos such as $\mathcal{G}$ would be tedious, this result is a proof of concept that well-adapted topoi have the necessary structure required to formulate quantum mechanics.

Chapter 4

Quantum Logic

Topos quantum theory is a separate attempt to recast physics in the lanuage of topoi, and is not nearly as geometrically inspired. Its focus is on the logical aspects of quantum mechanics, in particular quantum contextuality, a strange feature of quantum systems that separate them from classical ones.

Our sources include the four-part series of articles by Döring and Isham [Döring and Isham, 2008a, Döring and Isham, 2008b, Döring and Isham, 2008c, Döring and Isham, 2008d] as well as the two-part textbook series by Flori [Flori, 2013a, Flori, 2018]. The review [Flori, 2013b] is also useful at conveying a broad overview of the topic. The talk [Isham, 2002] analyzes the role of topos quantum theory in developing models of quantum gravity, one of the original inspirations for the subject.

4.1 Quantum Contextuality

4.1.1 Realism

In classical physics, a system is endowed with a state space S and a set of physical quantities $\mathcal{O} = \{A_\lambda : S \to \mathbb{R}\}_{\lambda \in \Lambda}$ in a deterministic, context-free way; that is, for a subset $U \subseteq \mathbb{R}$, there is a $\Lambda_U \subseteq \Lambda$ indexing over all the A_λ mapped to U by S (e.g., systems whose energy lies in a given interval). The underlying logic of such a system is Boolean, in that it is either *true* or *false* that $A_\lambda(s) \in U$, i.e. the law of excluded middle holds. Hence, we can say that a system in a given state has definite values of its physical quantities – a particle has a definite position, energy, and so on. Because of this, classical physics is said to be *realist*. The topos of sets naturally accommodates realist theories, such as models of classical physics, but it has trouble accommodating non-realist theories. Quantum physics is such a non-realist theory, as exhibited by the Kochen-Specker theorem; this theorem is fundamental to understanding topos quantum theory, so we will study it in detail.

4.1.2 The Kochen-Specker Theorem

Classical Logic In classical physics, we may apply a physical quantity $A \in \mathcal{O}$ to a state $s \in S$ to obtain a real number. Hence, there is a map $f : \mathcal{O} \times S \to \mathbb{R}$, which by the cartesian closure of sets corresponds both to a map $\mathcal{O} \to (S \to \mathbb{R}), A \mapsto f_A$ and a map $S \to (\mathcal{O} \to \mathbb{R}), s \mapsto V_s$. The map V_s associating to an observable A its value in the state s is known as a **valuation function**. We require such valuation functions to satisfy the reasonable property that $V_s(h \circ f_A) = h(V_s(A))$ for any sufficiently nice (generally Borel) function $h : \mathbb{R} \to \mathbb{R}$. For instance, if our system has a single particle with position x, measuring $\sin(x)$ should yield the same result as taking the sine of a measurement of x. This property is known as the **functional composition condition** (FUNC).

Valuation Functions In a quantum system with Hilbert space $\mathcal{H}$, the physical quantities are self-adjoint operators $A^\dagger = A$, but applying such an operator to an arbitrary state $|\psi\rangle$ doesn't have to result in a real number unless $|\psi\rangle$ is an eigenvector of A. Hence, we postulate a more flexible definition: a valuation function is a function $V : \mathcal{O}(\mathcal{H}) \to \mathbb{R}$ such that $V(A)$, which is identified somehow as the value of A, is an eigenvalue of A, and the FUNC $V(h(A)) = h(V(A))$ holds. Here, we calculate $h(A)$ by taking the eigenvector (spectral) decomposition

$$A = \sum_{n=1}^{N} A|c_n\rangle\langle c_n| = \sum_{n=1}^{N} c_n P_{c_n}$$

and writing $h(A) = \sum_{n=1}^{N} h(c_n) P_{c_n}$. A consequence of the FUNC is that $V(A+B) = V(A)+V(B)$, and when $[A,B] = 0$, $V(AB) = V(A)V(B)$. In particular, since $[P_\psi, P_\psi] = 0$ for projection operators P_ψ, we must have $V(P_\psi)^2 = V(P_\psi^2) = V(P_\psi)$, and hence $V(P_\psi) \in \{0,1\}$. Since propositions about the state of a quantum system can be formulated as projection operators, the FUNC implies that a valuation function necessarily imposes a Boolean logic on propositions.

The Kochen-Specker Theorem In order to have a realist model of quantum physics, it is necessary that we be able to assign values to all self-adjoint operators simultaneously in a way that respects the FUNC. The **Kochen-Specker theorem** states that this is impossible when $\dim \mathcal{H} > 2$: namely, if we can construct a valuation function $V : \mathcal{O}(\mathcal{H}) \to \mathbb{R}$, then V cannot satisfy the FUNC. We will not prove this in general, but instead give a special case. Given a Hilbert space $\mathcal{H}$ of dimension $n > 2$, we may take an orthogonal basis $|e_1\rangle, \ldots, |e_n\rangle$ and construct projection operators $P_{e_i} = |e_i\rangle\langle e_i|$. Since $\sum_i P_{e_i} = I$, it follows that a valuation function V must satisfy $V(\sum P_{e_i}) = \sum_i V(P_{e_i}) = 1$, which since each $V(P_i)$ is either 0 or 1 implies that exactly one of the P_{e_i} is 1. We will exploit this property by constructing several different orthonormal bases and showing that it is impossible to consistently assign values of 0 or 1 to each vector in such a way that the vectors in each base sum to 1. In $\mathcal{H} = \mathbb{R}^4$, we choose the following 11 bases:

	1	2	3	4	5	6
e_1	1,0,0,0	1,0,0,0	1,0,0,0	1,0,0,0	-1,1,1,1	-1,1,1,1
e_2	0,1,0,0	0,1,0,0	0,0,1,0	0,0,0,1	1,-1,1,1	1,1,-1,1
e_3	0,0,1,0	0,0,1,1	0,1,0,1	0,1,1,0	1,1,-1,1	1,0,1,0
e_4	0,0,0,1	0,0,1,-1	0,1,0,-1	0,1,-1,0	1,1,1,-1	0,1,0,-1

	7	8	9	10	11
e_1	1,-1,1,1	1,1,-1,1	0,1,-1,0	0,0,1,-1	1,0,1,0
e_2	1,1,-1,1	1,1,1,-1	1,0,0,-1	1,-1,0,0	0,1,0,1
e_3	0,1,1,0	0,0,1,1	1,1,1,1	1,1,1,1	1,1,-1,-1
e_4	1,0,0,-1	1,-1,0,0	1,-1,-1,1	1,1,-1,-1	1,-1,-1,1

The goal is to assign a 1 to exactly one member of each column. To see that this is impossible, note that each vector appears an even number of times, so we'll end up assigning 1 to an even number of vectors, rather than the required 11.

It follows that we have to throw out either the Boolean logic which assigns a truth value $x \in \{0, 1\}$ to each projection, or the FUNC. We will dispose of the former by moving to an intuitionistic topos.

4.2 Topoi of Contexts

4.2.1 Von Neumann Algebras

A C^*-algebra $\mathcal{M}$ is a **von Neumann algebra** if it has a predual. By Gelfand-Naimark, we can always assume that $\mathcal{M} \subseteq \mathcal{B}(\mathcal{H})$ for some $\mathcal{H}$. Defining the **commutant** of an arbitrary unital C^*-subalgebra $\mathcal{M} \subseteq \mathcal{B}(\mathcal{H})$ to be

$$\mathcal{M}' \coloneqq \{A \in \mathcal{B}(\mathcal{H}) \mid AB = BA \text{ for all } B \in \mathcal{M}\}$$

von Neumann's **double commutant theorem** states that $\mathcal{M}$ is a von Neumann algebra if and only if $\mathcal{M} = \mathcal{M}''$. Note that if a von Neumann algebra $\mathcal{M}$ is contained in its commutator $\mathcal{M}'$, it must be abelian; if it is in fact equal to its commutator, we call it **maximally abelian**. On the other hand, a commutator might be called "maximally noncommutative" if $\mathcal{M}$ and $\mathcal{M}'$ are as disjoint as possible, having only in common scalar multiples of the identity. Such a von Neumann algebra for which $\mathcal{M} \cap \mathcal{M}' = \{zI \mid z \in \mathbb{C}\}$ is known as a **factor**. The most obvious example is $\mathcal{B}(\mathcal{H})$ itself.

Any von Neumann algebra can be reconstructed from its set of projections $\mathcal{P}(\mathcal{M})$, as $\mathcal{M} = \mathcal{P}(\mathcal{M})''$. In this way, we can study $\mathcal{M}$ simply by studying its projections which, as noted previously, form a lattice with meets and joins. We may put an equivalence relation on $\mathcal{P}(\mathcal{M})$, whereby $A \sim B$ if there's an $X \in \mathcal{M}$ satisfying $X^\dagger X = A$ and $XX^\dagger = B$. This generates a partial ordering on $\mathcal{P}(\mathcal{M})$, whereby $A \leq B$ if there is some A' with $R(A') \subseteq R(B')$ and $A \sim A'$. We can "approximate" arbitrary operators $P \in \mathcal{P}(\mathcal{H})$ from the perspective of an arbitrary von Neumann algebra $\mathcal{M}$ by taking its **outer $\mathcal{M}$-support**, or the smallest operator in $\mathcal{M}$ greater than or equal to P:

$$\delta^o(P)_{\mathcal{M}} = \bigwedge \{Q \in \mathcal{P}(\mathcal{M}) \mid Q \geq P\}$$

We may also take its **inner $\mathcal{M}$-support**, or the largest operator in $\mathcal{M}$ less than or equal to P:

$$\delta^i(P)_{\mathcal{M}} = \bigvee \{Q \in \mathcal{P}(\mathcal{M}) \mid Q \leq P\}$$

In a von Neumann algebra $\mathcal{M}$, there's a natural embedding $\mathcal{M}_* \to \mathcal{M}^*$ given by taking a $\phi \in \mathcal{M}_*$ and defining its action on $\mathcal{M}$ as $\phi(A) = A(\Phi)$. If a $\phi \in \mathcal{M}^*$ can be obtained in this way, and it is a state, it is known as a **normal state**. Normal states can additionally be characterized by the following continuity property: for any countable family $\{P_n\}$ of mutually orthogonal projections in $\mathcal{M}$, $\phi\left(\bigvee P_n\right) = \sum \phi(P_n)$. On the von Neumann algebra $\mathcal{B}(\mathcal{H})$, every normal state

ϕ acts on operators A as $\phi(A) = \mathrm{Tr}(\Phi A)$ for some unique state $\Phi \in \mathcal{T}(\mathcal{H}) \subset \mathcal{B}(\mathcal{H})$; in this context, Φ is known as the **density operator** corresponding to ϕ.

Gelfand Representations Given an abelian von Neumann algebra $\mathcal{M}$, denote by $\Sigma_{\mathcal{M}}$ the set of $\mathbb{C}$-algebra homomorphisms $\lambda : \mathcal{M} \to \mathbb{C}$ such that $\lambda(I) = 1$, known as its **Gelfand spectrum**. With the weak-* topology, $\Sigma_{\mathcal{M}}$ is a compact Hausdorff space. The **Gelfand representation theorem** states that $\mathcal{M}$ is isomorphic as a C*-algebra to the C*-algebra of continuous complex functions on $\Sigma_{\mathcal{M}}$; this construction, which is functorial, is in fact half of a contravariant equivalence between the categories of unital C*-algebras and compact Hausdorff spaces. The isomorphism sends an operator $A \in \mathcal{M}$ to a continuous function $\overline{A} : \Sigma_{\mathcal{M}} \to \mathbb{C}$, $\overline{A}(\lambda) = \lambda(A)$, known as its **Gelfand transform**; if $A = A^{\dagger}$, then $\overline{A} = \overline{A}^{\dagger}$, implying that self-adjoint operators are transformed into real functions.

Of particular interest is the image of projections $P \in \mathcal{M}$ under the Gelfand transform, $\overline{P}(\lambda) = \lambda(P)$. Since $\lambda(P)^2 = \lambda(P^2) = \lambda(P)$ for any $\lambda \in \Sigma_{\mathcal{M}}$, the range of $\overline{P}$ must be $\{0, 1\}$. The function λ judges a projection P either *true* or *false*, and the transformed projection $\overline{P}$ judges a function λ as λ judges $\overline{P}$. We denote by S_P the set of $\lambda \in \Sigma_{\mathcal{M}}$ on which $\overline{P}$ is 1; since $\overline{P}$ is continuous, S_P is closed, being $\overline{P}^{-1}(\{1\})$, and open, being the complement of $\overline{P}^{-1}(\{0\})$, making it a clopen subset of $\Sigma_{\mathcal{M}}$.

4.2.2 Daseination

Given a Hilbert space, consider the poset category $\mathcal{V}(\mathcal{H})$ of abelian von Neumann subalgebras of $\mathcal{B}(\mathcal{H})$, where $\mathcal{M} \to \mathcal{N}$ if $\mathcal{M} \subseteq \mathcal{N}$. A quantum system is analyzed by means of self-adjoint operators, which represent observable quantities, and a morphism $\mathcal{M} \to \mathcal{N}$ in general increases the number of self-adjoint operators, giving us *more physical information* about the system; we correspondingly identify the objects of $\mathcal{V}(\mathcal{H})$ as **contexts** from which one can view the system. Elements of the presheaf category $\mathsf{Set}^{\mathcal{V}(\mathcal{H})^{\mathrm{op}}}$, then, are assignments of set-valued data to each context in a manner consistent under restriction.

The **spectral presheaf** Σ on $\mathcal{V}(\mathcal{H})$ sends $\mathcal{M}$ to its Gelfand spectrum $\Sigma_{\mathcal{M}}$, and an inclusion $\mathcal{M} \subseteq \mathcal{N}$ to the restriction morphism $\Sigma_{\mathcal{N}} \to \Sigma_{\mathcal{M}}$; we think of an element of $\Sigma_{\mathcal{M}}$, or a function $\lambda : \mathcal{M} \to \mathbb{C}$, as a measurement taken in the *context* of $\mathcal{M}$. From this point of view, the question of contextuality comes down to the following question: can we assign to each operator $A \in \mathcal{M} \subset \mathcal{B}(\mathcal{H})$ a measurement $\lambda(A) \in \mathbb{C}$ in a way that doesn't depend on the context $\mathcal{M}$? Such an assignment is a natural transformation $\Lambda : 1 \Rightarrow \Sigma$, i.e. a global element $\Lambda \in \Gamma\Sigma$; as such, the Kochen-Specker theorem should morally be equivalent to the statement that $\Gamma\Sigma$ is empty. The FUNC can be used to show this more concretely: let $\Lambda : 1 \Rightarrow \Sigma$ be such a global element, and let $\lambda_{\mathcal{M}} = \Lambda_{\mathcal{M}}(*) : \mathcal{M} \to \mathbb{C}$ be the measurement function it picks out for each $\mathcal{M}$. Pick a pair $\mathcal{M} \subset \mathcal{N} \in \mathcal{V}(\mathcal{H})$, and a self-adjoint operator A in $\mathcal{N}$ that's not in $\mathcal{M}$. In general, we may find a function $f : \mathbb{R} \to \mathbb{R}$ and self-adjoint operator B on $\mathcal{M}$ such that $f(A)$ restricts to B, from which it follows that $f(\lambda_{\mathcal{N}}(A)) = \lambda_{\mathcal{M}}(f(B))$, and thereby assign values to all self-adjoint operators simultaneously in a way that respects the FUNC. The Kochen-Specker theorem disallows this, and is hence equivalent to the statement that $\Gamma\Sigma = \varnothing$.

The **outer presheaf** O sends $\mathcal{M}$ to its set of projections $\mathcal{P}(\mathcal{M})$, and sends an inclusion $\mathcal{M} \subseteq \mathcal{N}$ to the $\mathcal{M}$-support function $\delta^o(-)_{\mathcal{M}} : \mathcal{P}(\mathcal{N}) \to \mathcal{P}(\mathcal{M})$. For a fixed $P \in \mathcal{P}(\mathcal{H})$, we can consider $\delta^o(P)_{\mathcal{M}}$ to be the $\mathcal{M}$ component of a natural transformation $\delta^o(P) : 1 \Rightarrow O$, giving us a map $\mathcal{P}(\mathcal{H}) \to \Gamma O, P \mapsto \delta(P)$. From the natural transformation $\delta(P) \in \Gamma O$ we may obtain a subfunctor $S^o \subseteq \Sigma$ given by $S^o(P)(\mathcal{M}) = S_{\delta^o(P)_{\mathcal{M}}}$, the clopen subset of $\Sigma_{\mathcal{M}}$ consisting of those λ sending $\delta^o(P)_{\mathcal{M}}$ to 1. Being a subfunctor which is at every object of $\mathcal{V}(\mathcal{H})$ a clopen subset of $\Sigma_{\mathcal{M}}$, we call S a **clopen subfunctor**. The set of all clopen subfunctors of Σ is denoted $\text{Sub}_{cl}(\Sigma)$, and the map $\delta^o : \mathcal{P}(\mathcal{H}) \to \text{Sub}_{cl}(\Sigma), P \mapsto S^o(P)$ is known as (outer) **daseination**. Daseination sends a projection P on the state space $\mathcal{H}$ to the set of all measurements on each context that judge the restriction of P to that context to be true, "bringing it into existence"; the concept of *dasein*, central to Heidegger's existential philosophy, roughly translates into "existence". Daseination maps the empty projection $\varnothing$ to the empty subobject $\subset \Sigma$, and the identity projection I to the trivial subobject $\Sigma \subseteq \Sigma$. It is injective, losing no information about P.

We may repeat the same process with δ^i, defining the **inner presheaf** I as sending $\mathcal{M}$ to $\mathcal{P}(\mathcal{M})$ and $\mathcal{M} \subseteq \mathcal{N}$ to $\delta^i(-)_{\mathcal{M}} : \mathcal{P}(\mathcal{N}) \to \mathcal{P}(\mathcal{M})$. Taking $\delta^i : \mathcal{P}(\mathcal{H}) \to \text{Sub}_{cl}(\Sigma), \delta^i(P)(\mathcal{M}) = S_{\delta^i(P)_{\mathcal{M}}}$ gives us **inner daseination**.

To daseinize an arbitrary self-adjoint operator $A \in \mathcal{O}(\mathcal{H})$ (recall that $\mathcal{O}(\mathcal{H})$ consists of the self-adjoint operators on $\mathcal{H}$), it is first necessary to construct a **spectral family** of A. This is an $\mathbb{R}$-indexed right-continuous family $\{A_\alpha\}$ of projection operators such that $\alpha \leq \beta \implies A_\alpha \preceq A_\beta$, $\lim_{\alpha\to\infty} A_\alpha = I, \lim_{\alpha\to-\infty} A_\alpha = 0$, and $\int_{\mathbb{R}} \alpha \, dA_\alpha = A$. The spectral theorem asserts the existence of such a family for all $A \in \mathcal{O}(\mathcal{H})$, so we may construct an ordering $\preceq_s$ on $\mathcal{O}(\mathcal{H})$: $A \preceq_s B$ if $A_\alpha \preceq B_\alpha$ for all $\alpha \in \mathbb{R}$. We then define outer and inner daseination in the usual manner:

$$\delta^o(A)_{\mathcal{M}} = \bigwedge \{B \in \mathcal{O}(\mathcal{M}) \mid B \succeq_s A\} \qquad \delta^i(A)_{\mathcal{M}} = \bigvee \{B \in \mathcal{O}(\mathcal{M}) \mid B \preceq_s A\}$$

The corresponding outer and inner presheaves, which send von Neumann algebras to their self-adjoint operators and inclusions to daseinations, are known as the **outer and inner de Groote presheaves** $\mathbb{O}$ and $\mathbb{I}$.

4.2.3 Measurement

While $\mathrm{Set}^{\mathcal{V}(\mathcal{H})}$ has a real numbers object given by the constant presheaf on $\mathbb{R}$, the Kochen-Specker theorem advises us against using this object for the purposes of measurement. For any poset $(\mathcal{P}, \leq)$, however, we can set up a better system. For posets $(\mathcal{P}, \leq), (\mathcal{Q}, \leq)$, let $\mathcal{OP}(\mathcal{P}, \mathcal{Q})$ be the set of order-preserving functions $\mathcal{P} \to \mathcal{Q}$, $\mathcal{OR}(\mathcal{P}, \mathcal{Q})$ the set of order-reversing functions, and, for $X \in \mathcal{P}$, define the sets

$$\downarrow X = \{X' \in \mathcal{P} \mid X' \leq X\} \qquad \uparrow X = \{X' \in \mathcal{P} \mid X' \geq X\}$$

$\mathcal{P}$ generates the presheaf $\mathcal{P}^{\succeq}$ of order-reversing functions which sends $\mathcal{M} \in \mathcal{V}(\mathcal{H})$ to $\mathcal{OR}(\downarrow \mathcal{M}, \mathcal{P})$, and sends an inclusion $\mathcal{M} \subseteq \mathcal{N}$ to the map $\mathcal{OR}(\downarrow \mathcal{N}, \mathcal{P}) \to \mathcal{OR}(\downarrow \mathcal{M}, \mathcal{P})$, $\mu \mapsto \mu|_{\downarrow \mathcal{M}}$. Likewise, the presheaf $\mathcal{P}^{\preceq}$ of order-preserving functions is given by replacing $\mathcal{OR}$ with $\mathcal{OP}$. We define a map $\breve{\delta}^{o}$ sending $A \in \mathcal{O}(\mathcal{H})$ to a natural transformation $\Sigma \Rightarrow \mathbb{R}^{\succeq}$ as follows: $\breve{\delta}^{o}(A)_{\mathcal{M}}$ sends a $\lambda : \mathcal{M} \to \mathbb{C} \in \Sigma_{\mathcal{M}}$ to the order-reversing function $\mu : \downarrow \mathcal{M} \to \mathbb{R}$, $\mathcal{M}' \mapsto \lambda(\delta^{0}(A)_{\mathcal{M}'})$. (Since A is self-adjoint, the range of λ is $\mathbb{R}$). $\breve{\delta}^{i}$ sends A to a natural transformation $\Sigma \Rightarrow \mathbb{R}^{\preceq}$, $\breve{\delta}^{i}(A)_{\mathcal{M}}(\lambda)(\mathcal{M}') = \lambda(\delta^{i}(A)_{\mathcal{M}'})$.

Finally, we define our quantity-value object by combining the two forms of daseination: $\mathbb{R}^{\leftrightarrow}$ is the presheaf sending $\mathcal{M}$ to $\mathcal{OP}(\downarrow \mathcal{M}, \mathbb{R}) \times \mathcal{OR}(\downarrow \mathcal{M}, \mathbb{R})$ and sending inclusions to restrictions, and $\breve{\delta}$ sends A to the natural transformation $\Sigma \Rightarrow \mathbb{R}^{\leftrightarrow}$. $\breve{\delta}$ sends A to the natural transformation $\delta^{i}(A) \times \delta^{o}(A)$. For a self-adjoint operator A in a context $\mathcal{M}$, $\breve{\delta}(A)_{\mathcal{M}}$ sends $\lambda : \mathcal{M} \to \mathbb{C}$ to the set of possible measurements of A.

4.2.4 Quantum Systems

Given a quantum system S with state space $\mathcal{H}_S$, we have seen how to define a topos $\mathcal{E}(S) = \mathrm{Set}^{\mathcal{V}(\mathcal{H}_S)^{\mathrm{op}}}$ and endow it with a state object Σ and quantity-value object $\mathcal{R} = \mathbb{R}^{\leftrightarrow}$, as well as how to daseinize observables, realizing them as natural transformations $\Sigma \Rightarrow \mathcal{R}$ via the map $\breve{\delta}$. Now, we will introduce a language $\mathcal{L}(S)$ for reasoning about physical quantities in S, which admits a model in $\mathcal{E}(S)$, separate from its Mitchell-Benabou language.

The basic type symbols of the language $\mathcal{L}(S)$ are $1, \Omega$, the state object Σ, and the quantity-value object $\mathcal{R}$, all of which are represented by their corresponding objects in $\mathcal{E}(S)$; we close these under finite products and the power object operation $T \mapsto \mathcal{P}T = \Omega^T$. To each type symbol T is associated a countable set of variables t of type T, as well as a special symbol $*$ of type 1. To each pair of type symbols T, T' there's a set $F_{\mathcal{L}(S)}(T, T')$ of function symbols written as $f : T \to T'$. These are represented by natural transformations. As in the Mitchell-Benabou language, we can take $t_1, t_2 : T, t : \mathcal{P}T$, and $\omega : \Omega$, and form the terms $t_1 = t_2 : \Omega$, $t_1 \in t : \Omega$, $\{t_1 | \omega\} : \mathcal{P}T$. We can also "evaluate" an $A : T \to T'$ at a $t : T$ to get an $A(t) : T'$.

The Kochen-Specker theorem in a system S is given by the statement that $F_{\mathcal{L}(S)}(1, \Sigma)$ is empty. It is conjectured that there are many other ways in which the representation of the local logics $\mathcal{L}(S)$ in the topoi $\mathsf{Set}^{\mathcal{V}(\mathcal{H}_S)^{op}}$ resemble quantum mechanics.

Chapter 5

Topological Quantum Field Theory

We have seen that the partition function Z associated to a quantum field theory is, in a sense, all we need to know about it. We calculate this function for a given Riemannian manifold (M, g) by integrating over the fields $\phi : \Sigma \to M$; in general, $Z(M) = \int e^{-iS[\phi]/\hbar} \mathcal{D}\phi$ changes as we change g, as the action warps along with the fabric of spacetime. Certain quantum field theories, however, have actions which are independent of the metric, and thus compute *topological invariants*. Such QFTs, which are called background independent, are known as **topological quantum field theories**. The usefulness of background independence appears in many scenarios: general relativity, for example, is diffeomorphism invariant, making it a topological (non-quantum) field theory which we'd like to couple a quantum field theory to[1].

We will first study topological quantum field theories from a 1-categorical point of view, largely following [Aspinwall, 2009], before moving on to the ∞-categorical point of view studied in [Kapustin, 2010, Lurie, 2009b].

5.1 Categorical Organization

5.1.1 Functorial Quantum Field Theory

From a sufficiently abstract point of view, we may view a topological quantum field theory as a functor from a "geometric" category to a "linear" category. There are many variations on this theme: we will first explore the case in which our geometric category $\mathcal{Cob}^R(n)$ has as its objects

[1]In fact, it is very difficult to consistently couple GR to quantum field theories; this is what makes the study of quantum gravity difficult, and TQFTs especially important to it.

Riemannian $(n-1)$-manifolds (including $\varnothing$) and as its morphisms Riemannian n-bordisms[2], and our linear category is $\mathbb{C}$-Vect. Both of these categories are symmetric monoidal, $\mathcal{Cob}^R(n)$ under the coproduct (disjoint union), so we may define a quantum field theory to be a symmetric monoidal functor $Z : \mathcal{Cob}^R(n) \to \mathbb{C}$-Vect.

In the simplest case, $n = 1$, all elements of $\mathcal{Cob}^R(n)$ are clusters of points, of the form $*^{\amalg n}$, so Z sends an arbitrary bordism $*^{\amalg n} \to *^{\amalg m}$ to a map $\mathcal{H}^{\otimes n} \to \mathcal{H}^{\otimes m}$, where $\mathcal{H} := Z(*)$. This case is easily seen to have four defining features:

1. The bordism $[0, t] : * \to *$ is sent to a linear operator $\mathcal{H} \to \mathcal{H}$; functoriality ensures that splitting this interval up into bordisms $[t, t_{n-1}] \circ \ldots \circ [t_1, t_0] : * \to * \to \ldots \to *$ does not change this operator, and therefore that $Z([0, t])$ is of the form e^{-tH} for some self-adjoint operator H.

2. This bordism can also be interpreted as going from $* \amalg *$ to $\varnothing$, which yields a linear operator $\mathcal{H} \otimes \mathcal{H} \to \mathbb{C}$, or equivalently a bilinear operator $\mathcal{H} \times \mathcal{H} \to \mathcal{H}$ which we may interpret as multiplication.

3. The bordism that connects $* \amalg *$ to $*$ via a Y-shaped graph yields a linear operator $\mathrm{Tr} : \mathcal{H} \otimes \mathcal{H} \to \mathcal{H}$, which we may interpret as taking a trace of operators.

4. An interval can also be interpreted as a bordism from $*$ to $\varnothing$, giving us a "trace" of elements $\mathrm{Tr} : \mathcal{H} \to \mathbb{C}$.

5.1.2 Topological Quantum Field Theory

The setup used in [Lurie, 2009b] to describe topological quantum field theories from a functorial point of view is similar, but we make some changes to emphasize the topological nature of the theory. Let $\mathsf{Cob}(n)$ denote the category whose objects are oriented compact smooth $(n-1)$-manifolds without boundary[3], and whose morphisms are oriented bordisms[4]. With the same symmetric monoidal structure as previously, we define a TQFT to be a functor $Z : \mathsf{Cob}(n) \to k$-Vect. Again, many interesting phenomena can be immediately observed by the consideration

[2] Tragically, the word 'bordism' is synonymous with 'cobordism'. However, the notation $\mathcal{Cob}(n)$ is in common use.

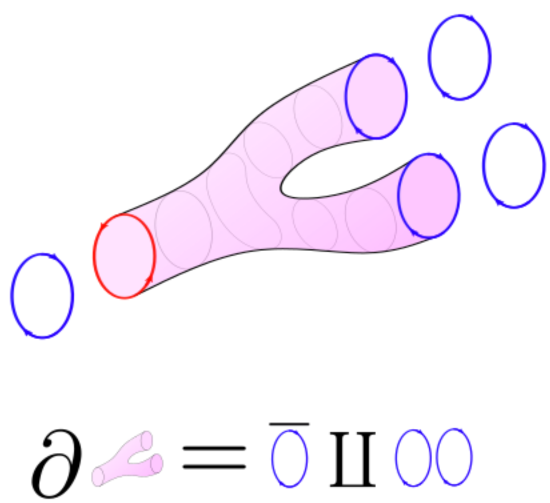
[3] We will assume that all manifolds are compact and smooth.

[4] Given oriented $(n-1)$-manifolds M and N, this is a manifold X whose boundary ∂X admits an orientation-preserving diffeomorphism to $\overline{M} \amalg N$, where $\overline{(\cdot)}$ inverts orientation

of elementary bordisms: for instance, every oriented manifold M has a cylinder $M \times [0,1]$ with boundary $\overline{M} \amalg M$, which can be considered as (a) the identity bordism on $\overline{M}$, (b) a bordism $\overline{M} \amalg M \to \varnothing$ which generates a canonical map $Z(\overline{M}) \otimes Z(M) \to k$, and (c) a bordism $\varnothing \to \overline{M} \amalg M$ which generates a canonical map $k \to Z(\overline{M}) \otimes Z(M)$. In fact, $Z(\overline{M})$ is isomorphic to the dual space of $Z(M)$, with the map $Z(\overline{M}) \otimes Z(M) \to k$ being interpreted as function evaluation; for this reason, we call this map the evaluation map ev_M, and its dual the coevaluation map $coev_M$.

As previously, the case of $n = 1$ is easily evaluated; the case of $n = 2$ is slightly more interesting. All 1-dimensional compact smooth oriented manifolds without boundaries are disjoint unions of S^1, and the fact that there is an orientation reversing diffeomorphism on S^1 provides an isomorphism $Z(S^1) \cong Z(S^1)^*$, so that in particular the vector space $A = Z(S^1)$ is finite-dimensional. The "pair of pants" bordism $S^1 \coprod S^1 \to S^1$ yields a morphism $A \otimes A \to A$ which endows A with a commutative, associative multiplication the unit of which can be found as the image of 1 under the morphism associated to the bordism $D^2 : \varnothing \to S^1$, and the bordism $D^2 : S^1 \to \varnothing$ yields a map $A \to k$ again known as the trace.

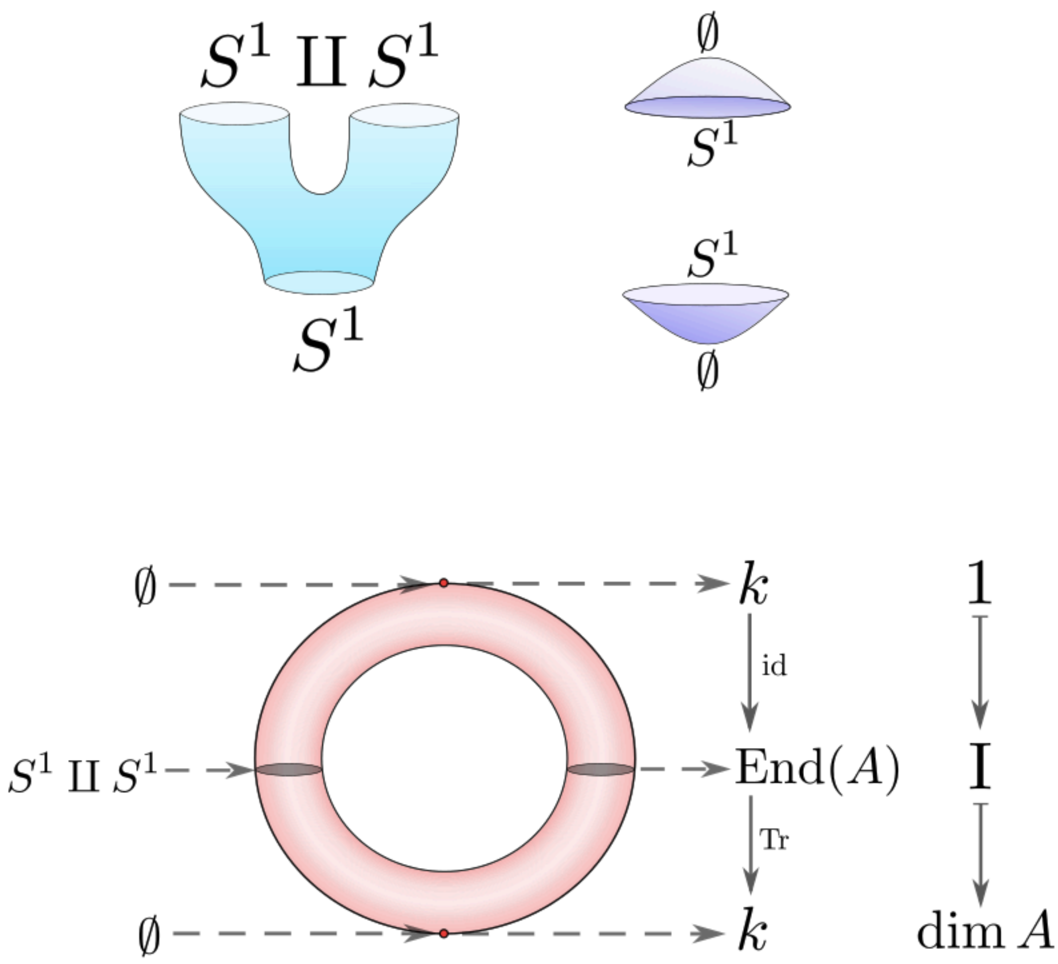
Figure 5.1: A bordism between S_1 and $S_1 \amalg S_1$.



Every 2-manifold M can be interpreted as a bordism $\varnothing \to \varnothing$, and in particular gives us an endomorphism $Z(M) : k \to k$, which is uniquely determined by $Z(M)(1)$; in this way, we can think of an n-dimensional TQFT as an association of a diffeomorphism invariant element of k to each n-manifold. In the case $n = 2$, the classification theorem of closed surfaces guarantees that we simply need to know the genus g of a manifold M to find this element: when $g = 1$, for instance, we have $M \cong T^2$. As a bordism $\varnothing \to \varnothing$, this is equivalent to the composition

$\varnothing \to S^1 \coprod S^1 \to \varnothing$, which after Z yields the composition $k \to A \otimes A \to k$. The first map sends 1 to id_A, and the second sends id_A to $\mathrm{Tr}(\mathrm{id}_A) = \dim A$; we see that a 2-dimensional TQFT Z associates the dimension of its underlying vector space to T^2. Note the method used here: we break up the n-manifold T^2 into a collection of simpler $(n-1)$-manifolds whose behavior we understand. This method allows us to completely understand the behavior of 2-dimensional TQFTs, but fails for higher dimensions: the n-manifolds grow incredibly complicated, as do the $(n-1)$-manifolds.

Figure 5.2: The evaluation of the bordism $T^2 : \varnothing \to \varnothing$ as the trace operator.



5.1.3 Higher Categorical Organization

This prompts the question: what additional structure on an n-dimensional TQFT is required to be able to "triangulate" arbitrary n-manifolds? The answer: move to a higher category where our n-bordisms are between $(n-1)$-bordisms between $(n-2)$-bordisms between.... More precisely, turn $\mathsf{Cob}(n)$ into an $(\infty, 1)$-category $\mathcal{Cob}(n)$ by declaring 2-morphisms to be orientation-preserving diffeomorphisms between bordisms, 3-morphisms isotopies between diffeomorphisms[5], and so on. We can combine these into a single (∞, n)-category $\mathcal{Bord}_n$ as follows: objects are *unoriented* 0-manifolds, 1-morphisms are bordisms, ..., n-morphisms are bordisms between $(n-1)$-bordisms, $(n+1)$-morphisms are diffeomorphisms, $(n+2)$-morphisms are isotopies, and so on. We will define what these terms mean.

5.2 Higher Categories

There are many different ways to view higher categories, each suggesting their own terminology and notation, and as such this chapter is a chimera blended from many sources. These include [Leinster, 2004, Lurie, 2009a, Riehl and Verity, 2018], as well as the more topologically focused [Cisinski, 2019, Lurie, 2009b]. [Riehl, 2014] discusses a lot of the necessary background, including enrichment and lifting problems.

5.2.1 Simplices

The **simplex category** Δ consists of all finite non-empty (von Neumann) ordinals, considered as ordered sets. We use the notation $[n]$ to denote the ordered set $(0, 1, \ldots, n)$. The morphisms are order-preserving set-maps; in particular, there are the **elementary face operators** $\delta^i : [n-1] \to [n]$ and the **elementary degeneracy operators** $\sigma^i : [n+1] \to [n]$. These act on an ordered set as follows: $\delta^i((0, 1, \ldots, n{-}1)) = (0, 1, \ldots, i{-}1, i{+}1, \ldots, n{-}1, n)$ increments i and everything above it, while $\sigma^i((0, 1, \ldots, n{+}1)) = (0, 1, \ldots, i{-}1, i, i, i{+}1, \ldots, n)$, decrementing $i{+}1$ and everything above it. We define a face operator to be a composite of elementary face operators and a degeneracy operator to be a composite of elementary degeneracy operators. In Δ, the epimorphisms are the surjective maps are the degeneracy operators, whereas the monomorphisms are the injective maps are the face operators. *Every* morphism in Δ can be factored as a degeneracy operator followed by a face operator (the factorization can be determined algorithmically in the obvious way).

A **simplicial object** in a category C is a contravariant functor $\Delta^{op} \to C$; these can be organized

[5]These are maps between homotopies $X \times I \to Y \times I$ such that each fiber $X \times \{t\}$ is mapped homeomorphically onto $Y \times \{t\}$.

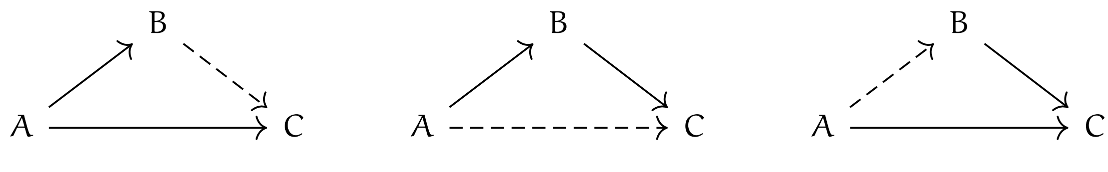
into the category of simplicial objects over C, $sC = C^{\Delta^{op}}$. In particular, the category of **simplicial sets** is given by the functor category $\mathcal{sSet} := \mathsf{Set}^{\Delta^{op}}$. The **standard** n**-simplex** is the object $\Delta^n := h_{[n]} \in \mathcal{sSet}$; by the Yoneda lemma, we have for an arbitrary simplicial set X that $X_n := X([n]) \cong \mathcal{sSet}(\Delta^n, X)$. We can characterize the simplicial set X more directly, as an $\mathbb{N}$-graded set $S = \amalg_n X_n$ with maps $d_i := X\delta^i : X_n \to X_{n-1}$ and $s_i := X\sigma^i : X_n \to X_{n+1}$, required to satisfy commutativity conditions which arise in Δ itself [6].

Every topological space X has a corresponding simplicial set $\mathrm{Sing}X$, whose n-simplices are the usual singular n-simplicies, i.e. continuous maps $\Delta^n_{top} \to X$. $\mathrm{Sing}X$ characterizes X up to weak homotopy equivalence, and the functor $\mathrm{Sing} : \mathsf{Top} \to \mathcal{sSet}$ has a left adjoint $|\cdot| : \mathcal{sSet} \to \mathsf{Top}$ known as **geometric realization**. Two simplicial sets are said to be **weakly equivalent** if their geometric realizations are weakly equivalent.

Kan Complexes The **horn** Λ^n_k is the subfunctor (simplicial subset) of Δ^n obtained by removing both the interior and the face opposite the kth vertex. A simplicial set K is a **Kan complex** if, for any $0 \leq k \leq n$, any morphism $\Lambda^n_k \to K$ extends to a morphism $\Delta^n \to K$. Since $|\Lambda^n_k|$ is weakly homotopy equivalent to Δ^n (retract it), any $\mathrm{Sing}X$ is a Kan complex. Geometrically, K is not "too complicated", in the sense that the image of any Λ^n_k is enough to determine an entire Δ^n.

The **nerve** of a category C is the simplicial set $N(C)$, where $N(C)_n$ is defined to be the set of all functors $[n] \to C$ ($[n]$ denoting the category $*_0 \to *_1 \to \ldots \to *_n$). So $N(C)_n$ is the set of all sequences of morphisms $X_0 \xrightarrow{f_1} X_1 \xrightarrow{f_2} \ldots \xrightarrow{f_n} X_n$. The face map d_i cuts out X_i by composing f_{i+1} with f_i, whereas s_i doubles X_i by inserting an id_{X_i}. C can be recovered up to isomorphism from $N(C)$ by regarding $N(C)_0$ as the vertices (applying s_0 to get their identity morphisms), $N(C)_1$ as the morphisms, and $N(C)_2$ as the associative composition data. In fact, we can characterize the simplicial sets that are isomorphic to nerves of categories: they are the simplicial sets K such that any map $\Lambda^n_k \to K$ has a *unique* extension $\Delta^n \to K$. These neither contain or are contained by the Kan complexes. However, we may define a **weak Kan complex** by requiring that any map $\Lambda^n_k \to K$ can be extended to a map $\Delta^n \to K$ only for $0 < k < n$. The nerve of any category is a weak Kan complex, but not vice-versa; philosophically, this originates from the fact that weak Kan complexes should come from categories where composition doesn't hold up to *equality* but up to some form of equivalence. We will define an $(\infty, 1)$**-category** to be a weak Kan complex, though there are different characterizations. Some examples of extensions of horns:

[6] In particular, we have (1) $d_i d_j = d_{j-1} d_i$ when $i < j$, (2) $s_i s_j = s_{j+1} s_i$ if $i \leq j$, (3) $d_i s_j = s_{j-1} d_i$ if $i < j$, (4) $d_j s_j = d_{j+1} s_j = \mathrm{id}_{X_n}$, and (5) $d_i s_j = s_j d_{i-1}$ when $i > j + 1$.



Λ_0^2 Λ_1^2 Λ_2^2

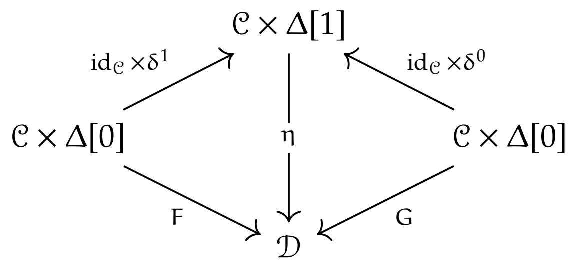
Of course, in the categorical setting, only the extension of Λ_1^2 should be possible (via composition); extension of Λ_0^2 and Λ_2^2 are only possible if we can invert morphisms, and hence should be interpreted as particular to groupoids.

Kan complexes are, by definition, complexes where all such extensions $\Lambda_i^n \to \Delta^n$ are possible; weak Kan complexes are those where i must be in $1, \ldots, n-1$ for an extension to be possible. Categorically, extension of Λ_1^2 just represents composition of morphisms. Extension of Λ_1^3 and Λ_2^3 represent associativity: $(A \to B \to C) \to D = A \to (B \to C \to D)$ can be witnessed by either an $(A \to B) \Rightarrow (A \to C)$ or a $(B \to C) \Rightarrow (B \to D)$. These two 2-morphisms yield a 2-isomorphism between $(A \to B \to C) \to D$ and $A \to (B \to C \to D)$. (In particular, associativity is witnessed by *isomorphism*, not strict equality). Such 2-morphisms can themselves be composed, and are associative up to 3-isomorphism. The idea is that the vertices of the simplicial set are the objects of a category, the 1-simplices form the usual 1-morphisms, and the recursive nature of simplicial sets provides us with higher morphisms; the weak Kan complex (extension) condition we require of the simplicial set amounts to enforcing composition and higher associativity conditions.

Any Kan complex is a weak Kan complex, and (by definition) an $(\infty, 1)$-category. This includes the singular complex $\operatorname{Sing} X$. Nerves of categories are $(\infty, 1)$-categories as well. Given a simplicial set S, we define its opposite to act on elements of Δ as $S^{op}(a_{i_1} \to \ldots \to a_{i_k}) = S(a_{i_k} \to \ldots \to a_{i_1})$. S extends Λ_i^n to Δ^n if and only if S^{op} extends Λ_{n-i}^n to Δ^n, so S is an $(\infty, 1)$-category if and only if S^{op}, the opposite $(\infty, 1)$-category, is.

Functors Reverting to the interpretation of ∞-categories as weak Kan complexes, we define the category ∞-$\mathcal{C}at$ as the corresponding full subcategory of $\mathrm{Set}^{\Delta^{op}}$; in particular, an ∞-functor between ∞-categories is simply a morphism of simplicial sets, or equivalently a natural transformation of the underlying functors $\Delta^{op} \to \mathrm{Set}$. An ∞-functor $F : \mathcal{C} \to \mathcal{D}$ is essentially surjective when the induced $hF : h\mathcal{C} \to h\mathcal{D}$ is essentially surjective, and fully faithful when hF is fully faithful as an $h\mathrm{CW}$-enriched functor, i.e. an isomorphism $h\mathcal{C}(X, Y) \to h\mathcal{D}(FX, FY)$. A fully faithful naturally surjective ∞-functor is an equivalence of ∞-categories.

A natural transformation $\eta : F \Rightarrow G$ between ∞-functors $\mathcal{C} \to \mathcal{D}$ is given by a simplicial homotopy between the simplicial set maps F and G, which is a simplicial set map $\mathcal{C} \times \Delta[1] \to \mathcal{D}$. Noting that $\mathcal{C} \times \Delta[0] \equiv \mathcal{C}$, we require that $\eta \circ (\mathrm{id}_{\mathcal{C}} \times \delta^1) = F$ and $\eta \circ (\mathrm{id}_{\mathcal{C}} \times \delta^0) = G$.



This structure generalizes in the obvious way, giving us for every pair $(\mathcal{C}, \mathcal{D})$ of $(\infty, 1)$-categories an $(\infty, 1)$-category of functors $\mathcal{C} \to \mathcal{D}$, denoted $\mathrm{Fun}(\mathcal{C}, \mathcal{D})$. A pair of $(\infty, 1)$-functors $F : \mathcal{C} \to \mathcal{D}$ and $G : \mathcal{D} \to \mathcal{C}$ form an ∞-categorical adjunction if $\mathrm{Map}_{\mathcal{D}}(FX, Y)$ is equivalent as an ∞-groupoid to $\mathrm{Map}_{\mathcal{C}}(X, GY)$ for all $X \in \mathcal{C}$, $Y \in \mathcal{D}$.

Having described the basic idea of ∞-categories and their functors via one model, we will introduce two more notions of an $(\infty, 1)$-category: topological categories and simplicial categories, and then show that all three of them are equivalent, giving us multiple ways to think about $(\infty, 1)$-categories.

5.2.2 Topological and Simplicial Categories

Topological Categories A topological category is a CG-enriched category, where CG is the convenient category of compactly generated Hausdorff spaces. Hence, in a topological category $\mathcal{C}$, the set $\mathcal{C}(X, Y)$ has the structure of a compactly generated space, which we denote the **mapping space** $\mathrm{Map}_{\mathcal{C}}(X, Y)$. We equip these mapping spaces with associative composition laws α_{XYZ} : $\mathrm{Map}(X, Y) \times \mathrm{Map}(Y, Z) \to \mathrm{Map}(X, Z)$ (where the product is taken in CG). A functor $F : \mathcal{C} \to \mathcal{D}$ between topological categories is a **strong equivalence** if it is essentially surjective and induces homeomorphisms $\mathrm{Map}_{\mathcal{C}}(X, Y) \cong \mathrm{Map}_{\mathcal{D}}(FX, FY)$ (i.e., it's an equivalence of categories that respects the enriched structure). The **homotopy category** $h\mathcal{C}$ of the topological category $\mathcal{C}$ has the same objects, but $h\mathcal{C}(X, Y) := \pi_0 \mathrm{Map}_{\mathcal{C}}(X, Y)$.

A functor $F : \mathcal{C} \to \mathcal{D}$ between topological categories in particular contains a family of CG-morphisms $\mathrm{Map}_{\mathcal{C}}(X, Y) \to \mathrm{Map}_{\mathcal{D}}(FX, FY)$, all of which are continuous and hence send connected components to connected components; in this way, functors between topological categories descend to functors between their homotopy categories. F is a **weak equivalence** if the induced $hF : h\mathcal{C} \to h\mathcal{D}$ is an equivalence of categories. Strong equivalences are weak equivalences, and we can characterize weak equivalences as weaker in the sense that they only induce weak homotopy equivalences $\mathrm{Map}_{\mathcal{C}}(X, Y) \cong \mathrm{Map}_{\mathcal{D}}(FX, FY)$ and are only essentially surjective in the corresponding homotopy categories. Two topological categories are **equivalent** if there is a weak equivalence between them.

Simplicial Categories The category *sSet* is cartesian monoidal, so we may consider *sSet*-enriched categories, known as **simplicial categories**. With *sSet*-enriched functors serving as morphisms, we have a category sCat of *sSet*-enriched categories. The term simplicial category here may be misleading: while a simplicial set is a functor $\Delta^{op} \to \mathsf{Set}$, simplicial categories are not equivalent to functors $\Delta^{op} \to \mathsf{Cat}$. In particular, a simplicial object X in Cat is a simplicial category if and only if $\mathrm{Obj}(X_0) = \mathrm{Obj}(X_1) = \ldots$.

Equivalences Restricting Sing to a functor $\mathsf{CG} \to sSet$, we have an adjoint pair $|\cdot| \dashv \mathrm{Sing}$, both of which commute with finite products; the unit and counit of this adjunction are both weak homotopy equivalences.

Given a simplicial category $\mathcal{C}$, we may define a topological category $|\mathcal{C}|$ by applying $|\cdot|$ to all hom-simplicial sets; if we have a topological category $\mathcal{D}$, we can apply Sing to each hom-space to get a simplicial category $\mathrm{Sing}\mathcal{D}$. In fact, the category obtained by inverting weak homotopy equivalences in CG is equivalent to the category obtained by inverting weak homotopy equivalences in *sSet*, so $h\mathcal{C} \cong h|\mathcal{C}|$ and $h\mathcal{D} \cong h\mathrm{Sing}\mathcal{D}$. It follows that the unit and counit are not just weak homotopy equivalences but isomorphisms on homotopy categories; if we wish to work with categories up to equivalence, this gives us a way to swap simplicial and topological categories freely.

5.2.3 Segal Spaces

A **Segal space** is a simplicial topological space $X_\bullet = \{X_n\}$ such that X_{m+n} is weakly equivalent to the space $X_m \times_{X_0} X_0^I \times_{X_0} X_n$ (where the map $X_0^I \to X_0$ is given by evaluation at 0). This space is known as the **homotopy pullback**, and often denoted $X_m \times^R_{X_0} X_n$; it is the homotopy-theoretic analogue of the ordinary pullback, in that it's given by weakening commutativity and isomorphism requirements into weak equivalences.

The idea behind this definition is that X_0 yields the objects of the $(\infty, 1)$-category, and the mapping spaces $\mathrm{Map}_{X_\bullet}(x, y)$ are given as $\{x\} \times^R_{X_0} X_1 \times^R_{X_0} \{y\}$; X_1 is the "generalized" space of morphisms, from which ordinary morphisms can be extracted as connected components.

Letting $\mathrm{Hom}(x,y) := \pi_0(\mathrm{Map}(x,y))$ gives us the homotopy category $hX_\bullet$; given an $f \in X_1$ which is mapped by the zeroth and first boundary operators to x and y, we can obtain an $[f] \in \mathrm{Hom}_{hX_\bullet}(x,y)$ by sending the image of the composite map $\{f\} \to \{x\} \times_{X_0} X_1 \times_{X_0} \{y\} \to \mathrm{Map}_{X_\bullet}(x,y)$, where the first map is induced by the universal property of the pullback and the second by the universal property of the homotopy pullback, to its connected component. f is invertible if this $[f]$ is an isomorphism.

The Segal space $X_\bullet$ is called a **complete Segal space** when X_0 is weakly equivalent to the subset of invertible elements of X_1, with weak equivalence given by the degeneracy operator $\delta_0 : X_0 \to \mathrm{im}(X_0) \subseteq X_1$; since $[\delta(x)] = \mathrm{id}_x$ in $hX_\bullet$, this allows us to identify isomorphisms with paths, and extract an $(\infty,0)$-groupoid from $X_\bullet$ by discarding non-invertible 1-morphisms. We can therefore use complete Segal spaces as models for $(\infty,1)$-categories.

Because Cat is cartesian closed, a simplicial object in a category $C^{\Delta^{op}}$ of simplicial objects is equivalent to a functor $\Delta^{op} \times \Delta^{op} \to C$. Inspired by this, we say that a functor $\prod_{i=1}^{n} \Delta^{op} \to C$ is an n**-fold simplicial object** $X = X_{\bullet,\dots,\bullet}$ of C; equivalently, an n-fold simplicial object of C is a simplicial object in the category of $(n-1)$-simplicial objects of C. Such an object comes equipped with n different boundary and degeneracy operators, one for each of the n coordinate indices. In general, properties of n-fold simplicial objects may be considered at a coordinate-wise level (i.e., at the level of C): for instance, two n-fold simplicial spaces are weakly equivalent if they are coordinate-wise weakly equivalent, and a homotopy-commutative square of n-fold simplicial spaces is a homotopy pullback square if it is so coordinate-wise.

An n-fold simplicial space X is essentially constant if all $X_{k_1,\dots,k_n}$ are weakly equivalent to $X_{0,\dots,0}$, and constant if this weak equivalence is witnessed by the face operator(s) $X_{0,\dots,0} \to X_{k_1,\dots,k_n}$. Given an n-fold simplicial space X regarded as a simplicial object $X_\bullet$ of $(n-1)$-fold simplicial spaces, we call X an n-fold Segal space if each X_k is an $(n-1)$-fold Segal space, $X_{k+\ell} \cong X_k \times^R_{X_0} X_\ell$ (that is, the associated square is a homotopy pullback square as defined above), and X_0 is essentially constant; X is complete if, recursively, each X_n is complete, and if the simplicial space $Y_\bullet$ given as $Y_k = X_{k,0,\dots,0}$ is complete. We use complete n-fold Segal spaces as models for (∞,n)-categories.

5.3 Higher Bordism Categories

Earlier, we declared $\mathcal{Cob}(n)$ to be an $(\infty,1)$-category and $\mathcal{Bord}_n$ an (∞,n)-category. We will also equip these categories with a symmetric monoidal structure with (in the case of $\mathcal{Cob}(n)$) duals.

First, we must make precise what this means in the (∞, n)-categorical setting.

Earlier, we declared $\mathcal{Cob}(n)$ to be an $(\infty, 1)$-category and $\mathcal{Bord}_n$ an (∞, n)-category. We will also equip these categories with a symmetric monoidal structure with (in the case of $\mathcal{Cob}(n)$) duals. First, we must make precise what this means in the (∞, n)-categorical setting.

Letting a symmetric monoidal (∞, n)-category be a commutative monoid object in the category of (∞, n)-categories, we define a dualizable object in a symmetric monoidal (∞, n)-category $\mathcal{C}$ to be a dualizable object in the homotopy category $h\mathcal{C}$, or an object V that admits an object V^* and evaluation/coevaluation maps $ev_V : V \otimes V^* \to 1$, $coev_V : 1 \to V^* \otimes V$ such that $(ev_V \otimes id_V) \circ (id_V \otimes coev_V) : V \to V \otimes V^* \otimes V \to V$ and $(id_V \otimes coev_V) \circ (ev_V \otimes id_V) : V^* \to V^* \otimes V \otimes V^* \to V^*$ reduce to the identities on V and V^*. In the symmetric monoidal category Vect, the dualizable objects are precisely the finite-dimensional vector spaces; in general, dualizability is a categorical generalization of the notion of "finiteness".

Given a symmetric monoidal (∞, n)-category $\mathcal{C}$, not necessarily with duals, we may consider the slice category of symmetric monoidal (∞, n)-categories *with* duals over $\mathcal{C}$. This category has a terminal object, or a symmetric monoidal (∞, n)-category with duals $\mathcal{C}^{fd}$ equipped with a symmetric monoidal functor $i : \mathcal{C}^{fd} \to \mathcal{C}$, such that all other symmetric monoidal functors from categories with duals into $\mathcal{C}$ factor through i. Passing from $\mathcal{C}$ to $\mathcal{C}^{fd}$ is essentially equivalent to making every element of $\mathcal{C}$ dualizable in the most efficient possible way. For instance, in k-Vect, viewed as an $(\infty, 1)$-category, $k\text{-Vect}^{fd}$ consists of the finite-dimensional vector spaces. An element $X \in \mathcal{C}$ is **fully dualizable** if it is isomorphic to $i(X_0)$ for some $X_0 \in \mathcal{C}^{fd}$.

Let $\mathcal{Bord}_n^{fr}$ be the (∞, n)-subcategory of $\mathcal{Bord}_n$ whose objects/k-morphisms are n-framed manifolds, or manifolds M with stably trivial tangent bundles $TM \oplus \mathbb{R}^{n-\dim M} \cong \mathbb{R}^n$. The **cobordism hypothesis** states that, for a symmetric monoidal (∞, n)-category with duals $\mathcal{C}$, the (∞, n)-category $\mathrm{Fun}^{\otimes}(\mathcal{Bord}_n^{fr}, \mathcal{C})$ of symmetric monoidal (∞, n)-functors $\mathcal{Bord}_n^{fr} \to \mathcal{C}$ is equivalent to the underlying ∞-groupoid of $\mathcal{C}$ given by discarding non-invertible morphisms, with equivalence given by the evaluation functor $Z \mapsto Z(*)$ (evaluation on the one-point manifold, which is trivially n-framed). In particular, $\mathrm{Fun}^{\otimes}(\mathcal{Bord}_n^{fr}, \mathcal{C})$ is an ∞-groupoid.

Appendix A

Sets and Categories

A.1 Large Categories and Sets

Previously, we brushed over set-theoretic issues in our discussion of category theory, using proper classes with reckless abandon; this has severe implications for the consistency of category theory. We will first demonstrate such implications, and then introduce modern set-theoretic constructs designed to deal with them. For set theory, our primary source is [Jech, 2013], and for its implications for category theory we use [Shulman, 2008] and [Muller, 2001]; [McLarty, 2010] provides an interesting discussion on the use of Grothendieck universes in algebraic geometry, and in particular on the provability of Fermat's Last Theorem in ZFC.

So far, we have assumed that our categories are locally small, if not just small; as we will show, these assumptions are essential, the primary reason being that if we want to work within a reasonable system of axioms such as ZFC, we need to limit ourselves to set-theoretic reasoning. We will survey specific instances of this, before discussing the limitations of ZFC and alternative foundations of category theory.

Adjoint Functor Theorems Adjoint functor theorems make clear the necessity of size issues in category theory.

For instance, Freyd's Special Adjoint Functor Theorem states that if a locally small, complete category C has (a) for each object a set's worth of subobjects (is *well-powered*), (b) a set $Q = \{Q_\lambda\}$ of objects such that whenever $f, g : X \rightrightarrows Y$ are distinct morphisms, there's an $h : Y \to Q_\lambda$ with $hf \neq hg$ (a *cogenerating set*), and (c) for each set $\{X_\lambda\}$ of subobjects of X a pullback, then a functor $R : \mathsf{C} \to \mathsf{D}$ is a right adjoint if and only if it preserves small limits and pullbacks of families of monomorphisms.

A stronger version: Freyd's Adjoint Functor theorem states that a functor R from a locally small, complete category C to a category D is a right adjoint iff it preserves all small limits and satisfies the *solution set condition*: every $Y \in D$ admits a set of arrows $\{f_\lambda : Y \to RX_\lambda\}$ such that every $g : Y \to RX$ factors as $Gt \circ f_i : Y \to RX_\lambda \to RX$. If we only have a proper class $\{f_\lambda\}$, then the deal is off, and R fails to be a right adjoint!

Categories of Sets Set is, strictly, the category whose objects are sets. What a set is, however, depends on your axiomatic system. There is no **true** Set; rather, its nature depends on the system of choice. ZFC is generally the default, but it presents difficulties: for one, we cannot reason about proper classes from within it, since its axioms only apply to sets. Indeed, we can't even state that Set exists from within ZFC, since the collection of its objects is not a set.

A.2 Set Theory

A.2.1 Axiomatic Set Theory

Definition: A set is an object whose existence can be deduced from an axiomatic set theory.

Clearly, this definition is useless without an axiomatic set theory to plug in. The most commonly used theory is ZFC, or Zermelo-Fraenkel set theory with the axiom of choice. The alphabet of the first-order language $\mathcal{L}_\in$ of ZFC consists of

- The logical symbols for universal and existential quantification, $\forall$ and $\exists$, as well as those for conjunction ($\wedge$), disjunction ($\vee$), negation ($\neg$), and one/two-sided implication ($\implies$ and $\iff$).
- The *non*-logical symbols $=$ and $\in$ denoting equality and set membership. These binary relations are the primitives of ZFC.

The axioms of ZFC are as follows:

1. (Extensionality) If two sets X and Y have the same elements, then $X = Y$.
2. (Pairing) For any two sets a and b, there is a a pair set $\{a, b\}$.
3. (Separation Schema) For any formula $\phi(x)$ in $\mathcal{L}_\in$ with one free variable x, and any set X, there is a set $\{x \in X \mid \phi(x)\}$.

4. (Power Set) For any set X, there is a power set $\mathcal{P}(X)$ whose elements are subsets of X.
5. (Union) For any set X, there is a set $\cup_{x \in X} x$ given by taking the union of all elements of X.
6. (Infinity) There exists an infinite set.
7. (Replacement Schema) The image of a set under a set function is also a set.
8. (Regularity) Every non-empty set X contains an element disjoint from X.
9. (Choice) We can pick a single representative for each set in a family of arbitrarily large sets through a choice function.

(The schemata each represent infinitely many axioms, one for each formula ϕ; this works around the fact that we cannot directly iterate over the formulae of $\mathcal{L}_{\in}$). For instance, the existence of the empty set $\varnothing$ can be deduced from the infinite set X postulated by the axiom of infinity and the axiom of separation for the fallacious formula $\phi(x) := (x \in x) \wedge \neg(x \in x)$ applied to X. Any class (collection of sets) whose existence cannot be proved by ZFC is known as a proper class. The prototypical example is the "set of all sets" S, whose existence is contradicted by ZFC: the pair "set" $\{S, S\}$ obviously has no elements disjoint from itself, violating the axiom of regularity.

A.2.2 The Von Neumann Universe

An especially important family of sets is given by the *ordinals*: an ordinal is a set α such that every $x \in \alpha$ is a subset of α, and α is well-ordered by $\in$. The successor of an ordinal is given by $\alpha + 1 := \alpha \cup \{\alpha\}$; an ordinal which is the successor of another ordinal is known as a successor ordinal, and an ordinal which is neither empty nor a successor ordinal is known as a limit ordinal.

The class Ord of all ordinals is well ordered by the relation $\alpha < \beta := \alpha \in \beta$, so limit ordinals can be thought of as "jumps" in this ordinal hierarchy. In fact, an arbitrary ordinal α is *equivalent* to the set of all ordinals β that are less than α. The first ordinal is trivially $\varnothing$, and we can proceed to define the von Neumann ordinals as $0 = \varnothing$, $1 = \{0\} = \{\varnothing\}$, $2 = \{0, 1\} = \{\varnothing, \{\varnothing\}\}$, and so on. The first limit ordinal is the limit of the von Neumann ordinals, $\omega = \{0, 1, 2, \ldots\}$.

Using ordinals, we can construct a *cumulative hierarchy* $\{V_\alpha\}$ of sets, which is built up in stages, one stage for each ordinal number. We start by defining V_0 as $\varnothing$ and, for each successor ordinal $\alpha + 1$, define $V_{\alpha+1} := \mathcal{P}(V_\alpha)$. For each limit ordinal β, we define $V_\beta := \bigcup_{\alpha<\beta} V_\alpha$. Finally, we define the (proper) class V to be the union of all stages: $V := \bigcup_\alpha V_\alpha$. The rank of a set is defined to be the ordinal at which it is introduced in this hierarchy. This is the standard set-theoretic approach to building a universe of sets, and is useful in discussing the category Set of sets – which, by definition, is dependent on one's idea of what a "set" is supposed to be. In other set theories, e.g. ZFC with additional axioms, we will have a different Set.

A.2.3 Large Cardinals

Bijection is an equivalence relation on the proper class of all sets; naively, we may quotient the proper class of sets by this relation to obtain a notion of the cardinality, or size, of a set. Unfortunately, the equivalence classes are not in general sets. A slightly subtler definition which relies on the axiom of choice fixes this: a cardinal is an ordinal that is not in bijection with any of its proper subsets. The cardinality $|S|$ of a set S is the least ordinal α admitting a bijection with S.

The natural numbers are all cardinals, and ω is the first infinite cardinal; since $|\omega| = |\omega+1| = \ldots$, we write this cardinal as $\aleph_0$ rather than ω, though cardinals still admit well-orderings as ordinals.

An important property of a cardinal κ is its cofinality $\mathrm{cf}(\kappa)$, defined to be the smallest cardinality among the subsets of κ all of whose sets have maximal cardinality in κ; the definition generalizes to any well-ordered set, ordinals in particular. Example: the cofinality of any nonzero finite ordinal is 1. An ordinal α such that $\mathrm{cf}(\alpha) = \alpha$ is known as a regular ordinal; all successor ordinals are regular.

Cantor's theorem states that $|S| < |\mathcal{P}(S)|$ for every set S [1], giving us an infinite hierarchy of cardinals $\beth_0 := \aleph_0, \beth_n := 2^{\beth_{n-1}} := |\mathcal{P}(\beth_{n-1})|$. Another infinite hierarchy is given by the *successor cardinal* operation, which associates to a cardinal κ the next largest cardinal κ^+; we have $\aleph_{n+1} := \aleph_n^+$. $\aleph_0$ and the natural numbers are the only countable cardinals; all other cardinals are called uncountable. A successor cardinal is a cardinal which is some cardinal's successor. As with ordinals, we can define limit cardinals, but we must define two flavors: a weak limit cardinal κ is a cardinal which is neither a successor cardinal nor zero. A strong limit cardinal λ is a cardinal such that $\rho < \lambda \implies 2^\rho < \lambda$.

Strong limit cardinals are weak limit cardinals, since obviously $\rho^+ \leq 2^\rho$, and $\aleph_0$ is the first

[1] Proof: suppose there were a bijection f, use replacement to construct the set $T = \{s \in S \mid s \notin f(s)\} \in \mathcal{P}(S)$, and attempt to find an $s \in S$ with $f(s) = T$; we have $s \in T \iff s \notin T$, a contradiction.

strong limit cardinal. For limit ordinals λ, we define $\aleph_\lambda := \bigcup_{\rho<\lambda} \aleph_\rho$, which is in general a weak limit cardinal.

So far, we have stayed within what is provable from ZFC alone. However, weak limit cardinals are as far as ZFC can go; in this sense, such cardinals measure the "strength" of ZFC. We may postulate stronger conditions on the size of a cardinal κ, but there is no guarantee that ZFC can prove the existence of κ. Such cardinals are known as large cardinals. The first condition, or large cardinal property, is given by inaccessibility: a cardinal κ is weakly inaccessible if it is an uncountable regular weak limit cardinal, and strongly inaccessible if it is an uncountable regular strong limit cardinal.

ZFC can neither prove nor disprove the existence of weakly or strongly inaccessible cardinals; in fact, the existence of a weakly inaccessible cardinal would prove the consistency of ZFC.

A.3 Alternatives to ZFC

A.3.1 Von Neumann-Bernays-Gödel Set Theory

In ZFC, we cannot directly talk about classes; they are an informal notion. The von Neumann-Bernays-Gödel (NBG) set theory fixes this by formalizing the notion of a class. NBG is built from ZFC by making the primitive notion that of a class, rather than a set, and introducing a predicate $M(S)$ stating that S is a set. We modify the axiom of extensionality to act on classes, generalize images, unions, power sets, and functions to classes, and we also add a few axioms and axiom schemata:

- (Class comprehension schema) For every formula $\phi(x) \in \mathcal{L}_\epsilon$ quantifying over sets, there is a class $C = \{S \mid \phi(S)\}$.
- (Separation) Every subclass of a set is a set.
- (Global choice) There is a choice function that chooses an element from *every* non-empty set. Equivalently, V is well-ordered.

Though it seems different, the axiom of global choice is a proper class-based generalization of the axiom of choice: it is equivalent to the statement that every class is well-ordered (and hence strictly stronger than choice). Global choice implies another axiom, the axiom of the limitation of size, which characterizes the proper classes: a class is a proper class if and only if it is in bijection with the von Neumann universe V.

NBG allows us to reason about classes, and thus coherently talk about large categories. For instance, we can use global choice to collapse isomorphism classes of proper classes, and thus construct skeletons of large categories, choose for every Y in a large category "a" product $X \times Y$ rather than an isomorphism class (evidencing $X \times -$ as a functor), and so on. We can speak of the category of sets Set, as well as the category of small categories Cat; unfortunately, however, there is no category of *all* categories CAT, since proper classes in NBG cannot contain other proper classes. For the same reason, we cannot speak of a functor category D^{C} when both C and D are large.

A.3.2 Grothendieck Universes

Grothendieck universes were invented by Grothendieck as a convenient way to side-step set theoretic issues in category theory. A Grothendieck universe is a set U which is closed under set-indexed union, power set, pair formation, and is transitive, in the sense that any element of an element of U is itself an element of U.

ZFC cannot prove the existence of Grothendieck universes, but the extent to which it can't can be measured precisely by a large cardinal property: a set U is a Grothendieck universe if and only if there is a strongly inaccessible cardinal κ such that $U = V_\kappa$, where V_κ is the set at stage κ in the von Neumann hierarchy. In order to work properly in a Grothendieck universe, we must introduce to ZFC an axiom stating that a strongly inaccessible cardinal exists, and hence assume that ZFC is consistent.

Grothendieck universes play the starring role in Tarski-Grothendieck (TG) set theory. This is, like NBG, an extension of ZFC, though unlike NBG it goes further, in being able to prove things from within $\mathcal{L}_\in$ that ZFC can't prove; the main extension is given by Tarski's axiom, which states that every set belongs to some Grothendieck universe.

Appendix B

A Categorical Bestiary

We will catalogue many of the categories we commonly encounter, defining their objects, morphisms, and many of their more salient properties.

Set-Like Categories

Set: Objects are the sets which can be proven to exist within some axiomatic set theory (which we assume to be at least as strong as ZFC), and morphisms are functions. This category is complete, cocomplete, and cartesian closed. The product is the cartesian product, the coproduct is the disjoint union, the initial object is the empty set, and the terminal object is any singleton. If necessary, we may choose to work in a Grothendieck universe U, in which case the objects of Set are the sets in U.

FinSet: Objects are finite sets, morphisms are functions. Neither complete nor cocomplete, but has the same finite limits and colimits as Set, and is cartesian closed.

Rel: Objects are sets, morphisms are relations $R \subseteq X \times Y$. Composition of relations $R : X \to Y$ and $S : Y \to Z$ is given by $S \circ R = \{(x, z) \in X \times Z \mid \exists y \in Y\, ((x, y) \in R \wedge (y, z) \in S)\}$. Equivalent to its opposite, and hence all limits are colimits and vice versa. Neither complete nor cocomplete. The biproduct is the disjoint union.

Δ: Objects are finite non-empty von Neumann ordinals $[n] = \{0, \ldots, n\}$ and morphisms are order-preserving set maps $[n] \to [m]$. Also known as the simplex category; contravariant functors $\Delta^{op} \to \mathsf{C}$ are known as simplicial objects of C, and assemble into a category $s\mathsf{C} = \mathsf{C}^{\Delta^{op}}$ which usually has all the limits and colimits that C does, calculated pointwise[1].

Categorical Categories

Cat: Objects are small categories, morphisms are functors. Complete, cocomplete, and cartesian closed. The product is the product of categories, the coproduct is the "disjoint union" (place the categories side by side), the initial object is the empty category, and the terminal object is the trivial category with one object and one identity morphism.

CAT: Objects are small and large categories, morphisms are functors. To avoid Russell's paradox, this must be not a large category but a "very large" category.

Grpd: Objects are groupoids, morphisms are functors. Complete, cocomplete, and cartesian closed. Both a reflective and coreflective subcategory of Cat, and hence limits and colimits are the same as in Cat when they exist in Grpd.

Algebraic Categories

Grp: Objects are groups, morphisms are group homomorphisms. A concrete category. Complete and cocomplete. The product is the direct product $G \times H$, and the coproduct is the free product $G * H$. The zero object is the trivial group 0. Not an abelian category, as it is not additive.

R-Mod: Objects are modules over a commutative ring R, and morphisms are R-module homomorphisms. The product is the direct product $M \times N$, which coincides with the direct sum in the finite case, and the coproduct is the direct sum $M \oplus N$. The zero object is the trivial module 0. Has a closed monoidal structure with tensor product $\otimes_R$ and unit R. An abelian category.

[1] In general, functor categories C^{D} bear all the (co)limits of C, computed pointwise, and if C is (co)complete we're done; if not, C^{D} may bear (co)limits not found in C.

Ab: equivalent to $\mathbb{Z}$-Mod. Objects are abelian groups, morphisms are group homomorphisms. A concrete category and reflective subcategory of Grp. Complete and cocomplete.

CRing: Objects are commutative rings, morphisms are ring homomorphisms. Not balanced. The product is the product of rings, and the coproduct is the tensor product of rings. The opposite category of the geometric category Aff.

R-Alg: Objects are algebras over a commutative ring R, morphisms are R-algebra homomorphisms.

FdVect: Objects are finite-dimensional vector spaces over a field k, morphisms are k-linear maps.

Hilb: Objects are Hilbert spaces over a field $k = \mathbb{R}$ or $\mathbb{C}$, morphisms are operators.

FdHilb: Objects are finite-dimensional Hilbert spaces, morphisms are operators. Symmetric monoidal under the tensor product $\otimes_k$ and dagger compact with $A^\dagger$ being the adjoint of A. The direct sum $\oplus$ is the finite biproduct.

$\mathsf{VB}_k(X)$: Objects are k-vector bundles on a topological space X, morphisms are vector bundle homomorphisms. A symmetric monoidal category under the tensor product $\otimes$.

Geometric Categories

Categories of Topological Objects

Top: Objects are topological spaces, morphisms are continuous maps. A concrete category. Complete and cocomplete. The initial object is the empty set, the terminal object is any singleton, the product is given by the product topology, the coproduct by the disjoint union topology, the equalizer by the subspace topology, and the coequalizer by the quotient topology.

CG: Objects are compactly generated (weak hausdorff k-)spaces, morphisms are continuous maps. Complete and cocomplete, contains the CW complexes, and cartesian closed, and hence a convenient category of topological spaces. Contains the compact and locally compact spaces, as well as the (topological) manifolds.

hTop: Objects are topological spaces, morphisms are homotopy classes of continuous maps. Neither concrete, complete, or cocomplete.

Man^p: Objects are C^p manifolds, morphisms are C^p maps. Concrete, but neither complete nor cocomplete. The product is the product of manifolds, and the coproduct of a countable family of manifolds exists when all manifolds share the same dimension, and is the disjoint union.

Diff: Objects are smooth manifolds, morphisms are smooth maps. See Man^p for $p = \infty$.

CartSp: Objects are cartesian spaces, or smooth manifolds of the form $\mathbb{R}^n$, morphisms are smooth maps.

Categories of Algebro-Geometric Objects

Sch: Objects are schemes, morphisms are morphisms of schemes.

Aff: Objects are affine schemes, morphisms are morphisms of schemes. This is the opposite category of CRing, and hence bears all duals of its properties.

$\mathsf{Sh}(X)$: Objects are sheaves on a topological space X, morphisms are morphisms of sheaves. A reflective subcategory of the presheaf category $\mathsf{Set}^{\mathsf{Op}(X)^{op}}$.

$\mathsf{QCoh}(X)$: Objects are quasi-coherent sheaves on a topological space or variety X, morphisms are morphisms of sheaves.

$\mathsf{Coh}(X)$: Objects are coherent sheaves on a topological space or variety X, morphisms are morphisms of sheaves.

Topoi

For reference, we list the Kripke-Joyal semantics for an elemetary topos $\mathcal{E}$: for $f : U \to X$ and a formula ϕ, we have the following:

1. $U \Vdash \phi(f) \wedge \psi(f)$ iff $U \Vdash \phi(f)$ and $U \Vdash \psi(f)$.
2. $U \Vdash \phi(f) \vee \psi(f)$ iff there are arrows $g : V \to U$, $h : W \to U$ such that $g \amalg h : V \amalg W \to U$ is epi, with $V \Vdash \phi(fg)$ and $W \Vdash \phi(fh)$.
3. $U \Vdash \phi(f) \implies \psi(f)$ iff for any $g : V \to U$ such that $V \Vdash \phi(fg)$, V also forces $\psi(fg)$.
4. $U \Vdash \neg\phi(f)$ if for any $g : V \to U$ such that $V \Vdash \phi(fg)$, V is the initial object.
5. $U \Vdash \exists y\, \phi(f, y)$ (for some formula $\phi : X \times Y \to \Omega$ and generalized element $f : U \to X$) iff there's an epic $e : V \to U$ and generalized element $g : V \to Y$ such that $V \Vdash \phi(fe, g)$.
6. $U \Vdash \forall y\, \phi(f, y)$ iff for *every* arrow $h : V \to U$ and generalized element $g : V \to Y$ we have $V \Vdash \phi(fh, g)$.

Set: The aforementioned category of sets and set maps. The subobject classifier is given by $\Omega = 2 = \{0, 1\}$, and the exponential is given by $Y^X = \mathrm{Hom}_{\mathsf{Set}}(X, Y)$. Its logic is classical logic.

Sh(C, J): Objects are functors $P : C^{op} \to \mathsf{Set}$ satisfying the sheaf condition, morphisms are natural transformations between functors. The subobject classifier Ω sends an object $U \in C$ to the set $\Omega(U)$ of closed sieves on U. The exponential is given by $Q^P(X) = \mathrm{Hom}(h_X \times P, Q)$. The natural transformation $\mathtt{true} : 1 \to \Omega$ sends $1(X) = 1$ to the maximal sieve t_X. When J is indiscrete, such that $\mathrm{Sh}(C, J) = \mathsf{Set}^{C^{op}}$, all sieves are closed, so $\Omega(X)$ is simply the set of all sieves on X.

Semantics:

1. (Unchanged)
2. $U \Vdash \phi(f) \vee \psi(f)$ iff there's a covering $\{f_\lambda : U_\lambda \to U\}_{\lambda \in \Lambda}$ with either $U_\lambda \Vdash \phi(f)$ or $U_\lambda \Vdash \psi(f)$ for all $\lambda \in \Lambda$.
3. (Unchanged)
4. $U \Vdash \neg\phi(f)$ if for any $g : V \to U$ such that $V \Vdash \phi(fg)$,

$\mathsf{Set}^{L^{op}}$:

$\mathcal{G}$:

$\mathcal{SmoothSet}$:

$\mathsf{Set}^{\mathcal{V}(\mathcal{H})}$:

Higher Categories

$\mathcal{Cob}(n)$: an $(\infty, 1)$-category whose objects are oriented compact smooth $(n-1)$-dimensional manifolds without boundary, 1-morphisms are oriented bordisms, 2-morphisms are diffeomorphisms between bordisms, 3-morphisms are isotopies between diffeomorphisms, and so on.

$\mathcal{Bord}_n$: an (∞, n)-category whose objects are unoriented 0-manifolds, 1-morphisms are 1-bordisms, 2-morphisms are 2-bordisms between 1-bordisms, . . ., n-morphisms are n-bordisms between $(n-1)$-bordisms, $(n+1)$-morphisms are diffeomorphisms, $(n+2)$-morphisms are isotopies between diffeomorphisms, and so on.

Bibliography

[Abramsky and Coecke, 2009] Abramsky, S. and Coecke, B. (2009). Categorical quantum mechanics. *Handbook of quantum logic and quantum structures*, 2:261–325.

[Aluffi, 2009] Aluffi, P. (2009). *Algebra: chapter 0*, volume 104. American Mathematical Soc.

[Arnold, 2013] Arnold, V. I. (2013). *Mathematical methods of classical mechanics*, volume 60. Springer Science & Business Media.

[Aspinwall, 2009] Aspinwall, P. (2009). *Dirichlet branes and mirror symmetry*, volume 4. American Mathematical Soc.

[Atiyah, 1988] Atiyah, M. F. (1988). Topological quantum field theory. *Publications Mathématiques de l'IHÉS*, 68:175–186.

[Authors, 2018] Authors, S. P. (2018). Stacks project.

[Baez and Crans, 2003] Baez, J. C. and Crans, A. S. (2003). Higher-dimensional algebra vi: Lie 2-algebras. *arXiv preprint math/0307263*.

[Binosi and Theussl, 2004] Binosi, D. and Theussl, L. (2004). Jaxodraw: A graphical user interface for drawing feynman diagrams. *Computer Physics Communications*, 161(1-2):76–86.

[Bleecker, 2005] Bleecker, D. (2005). *Gauge theory and variational principles*. Courier Corporation.

[Carroll, 2019] Carroll, S. M. (2019). *Spacetime and geometry*. Cambridge University Press.

[Cisinski, 2019] Cisinski, D.-C. (2019). *Higher categories and homotopical algebra*, volume 180. Cambridge University Press.

[Coecke, 2010] Coecke, B. (2010). Quantum picturalism. *Contemporary physics*, 51(1):59–83.

[Deligne et al.,] Deligne, P., Etingof, P. I., and Freed, D. S. *Quantum fields and strings: a course for mathematicians*, volume 1.

[Döring and Isham, 2008a] Döring, A. and Isham, C. J. (2008a). A topos foundation for theories of physics: I. formal languages for physics. *Journal of Mathematical Physics*, 49(5):053515.

[Döring and Isham, 2008b] Döring, A. and Isham, C. J. (2008b). A topos foundation for theories of physics: Ii. daseinisation and the liberation of quantum theory. *Journal of Mathematical Physics*, 49(5):053516.

[Döring and Isham, 2008c] Döring, A. and Isham, C. J. (2008c). A topos foundation for theories of physics: Iii. the representation of physical quantities with arrows δ o (a): σ r . *Journal of Mathematical Physics*, 49(5):053517.

[Döring and Isham, 2008d] Döring, A. and Isham, C. J. (2008d). A topos foundation for theories of physics: Iv. categories of systems. *Journal of Mathematical Physics*, 49(5):053518.

[Fearns, 2002] Fearns, J. D. (2002). A physical quantum model in a smooth topos. *arXiv preprint quant-ph/0202079*.

[Flori, 2013a] Flori, C. (2013a). *A first course in topos quantum theory*, volume 868. Springer.

[Flori, 2013b] Flori, C. (2013b). Review of the topos approach to quantum theory. *Canadian Journal of Physics*, 91(6):471–473.

[Flori, 2018] Flori, C. (2018). *A second course in topos quantum theory*. Springer.

[Fong and Spivak, 2018] Fong, B. and Spivak, D. I. (2018). Seven sketches in compositionality: An invitation to applied category theory. *arXiv preprint arXiv:1803.05316*.

[Guts and Grinkevich, 1996] Guts, A. K. and Grinkevich, E. B. (1996). Toposes in general theory of relativity. *arXiv preprint gr-qc/9610073*.

[Guts and Zvyagintsev, 2000] Guts, A. K. and Zvyagintsev, A. A. (2000). Interpretation of intuitionistic solution of the vacuum einstein equations in smooth topos. *arXiv preprint gr-qc/0001076*.

[Haase, 2014] Haase, M. (2014). *Functional analysis: an elementary introduction*, volume 156. American Mathematical Society.

[Hartshorne, 2013] Hartshorne, R. (2013). *Algebraic geometry*, volume 52. Springer Science & Business Media.

[Hatcher, 2003] Hatcher, A. (2003). Vector bundles and k-theory. *Im Internet unter http://www. math. cornell. edu/~ hatcher*.

[Hatcher, 2005] Hatcher, A. (2005). *Algebraic topology*.

[Holevo, 2003] Holevo, A. S. (2003). *Statistical structure of quantum theory*, volume 67. Springer Science & Business Media.

[Husemoller, 1975] Husemoller, D. (1975). *Fiber bundles*. Springer.

[Isham, 2002] Isham, C. (2002). Some reflections on the status of conventional quantum theory when applied to quantum gravity. *arXiv preprint quant-ph/0206090*.

[Jech, 2013] Jech, T. (2013). *Set theory*. Springer Science & Business Media.

[Johnstone, 2002] Johnstone, P. T. (2002). Sketches of an elephant (2 vols.), volume 43 of oxford logic guides.

[Johnstone, 2014] Johnstone, P. T. (2014). *Topos theory*. Courier Corporation.

[Kapustin, 2010] Kapustin, A. (2010). Topological field theory, higher categories, and their applications. In *Proceedings of the International Congress of Mathematicians 2010 (ICM 2010) (In 4 Volumes) Vol. I: Plenary Lectures and Ceremonies Vols. II–IV: Invited Lectures*, pages 2021–2043. World Scientific.

[Kock, 1986] Kock, A. (1986). Lie group valued integration in well-adapted toposes. *Bulletin of the Australian Mathematical Society*, 34(3):395–410.

[Kock, 2006] Kock, A. (2006). *Synthetic differential geometry*, volume 333. Cambridge University Press.

[Kostecki, 2009] Kostecki, R. P. (2009). Differential geometry in toposes.

[Kühnel, 2015] Kühnel, W. (2015). *Differential geometry*, volume 77. American Mathematical Soc.

[Lancaster and Blundell, 2014] Lancaster, T. and Blundell, S. J. (2014). *Quantum field theory for the gifted amateur*. OUP Oxford.

[Landau and Lifshitz, 2013] Landau, L. D. and Lifshitz, E. M. (2013). *Mechanics and electrodynamics*. Elsevier.

[Lawvere, 1997] Lawvere, F. W. (1997). Toposes of laws of motion. *American Mathematical Society, Transcript from video, Montreal-September*, 27:1997.

[Lawvere, 2002] Lawvere, F. W. (2002). Categorical algebra for continuum micro physics. *Journal of Pure and Applied Algebra*, 175(1-3):267–287.

[Leinster, 2004] Leinster, T. (2004). *Higher operads, higher categories*, volume 298. Cambridge University Press.

[Lurie, 2009a] Lurie, J. (2009a). *Higher topos theory*. Princeton University Press.

[Lurie, 2009b] Lurie, J. (2009b). On the classification of topological field theories. *Current developments in mathematics*, 2008:129–280.

[Mac Lane, 2013] Mac Lane, S. (2013). *Categories for the working mathematician*, volume 5. Springer Science & Business Media.

[MacLane and Moerdijk, 2012] MacLane, S. and Moerdijk, I. (2012). *Sheaves in geometry and logic: A first introduction to topos theory*. Springer Science & Business Media.

[May, 1999] May, J. P. (1999). *A concise course in algebraic topology*. University of Chicago press.

[May and Sigurdsson, 2006] May, J. P. and Sigurdsson, J. (2006). *Parametrized homotopy theory*. Number 132. American Mathematical Soc.

[McLarty, 2010] McLarty, C. (2010). What does it take to prove fermat's last theorem? grothendieck and the logic of number theory. *Bulletin of Symbolic Logic*, 16(3):359–377.

[Meyer, 2006] Meyer, P. A. (2006). *Quantum probability for probabilists*. Springer.

[Misner et al., 1973] Misner, C. W., Thorne, K. S., Wheeler, J. A., et al. (1973). *Gravitation*. Macmillan.

[Moerdijk and Reyes, 2013] Moerdijk, I. and Reyes, G. E. (2013). *Models for smooth infinitesimal analysis*. Springer Science & Business Media.

[Muller, 2001] Muller, F. A. (2001). Sets, classes, and categories. *The British journal for the philosophy of science*, 52(3):539–573.

[Munkres, 2018] Munkres, J. R. (2018). *Elements of algebraic topology*. CRC Press.

[nLab authors, 2020] nLab authors (2020). geometry of physics – categories and toposes. http://ncatlab.org/nlab/show/geometry%20of%20physics%20--%20categories%20and%20toposes. Revision 122.

[Peskin, 2018] Peskin, M. E. (2018). *An introduction to quantum field theory*. CRC Press.

[Rédei and Summers, 2007] Rédei, M. and Summers, S. J. (2007). Quantum probability theory. *Studies in History and Philosophy of Science Part B: Studies in History and Philosophy of Modern Physics*, 38(2):390–417.

[Riehl, 2014] Riehl, E. (2014). *Categorical homotopy theory*, volume 24. Cambridge University Press.

[Riehl, 2017] Riehl, E. (2017). *Category theory in context*. Courier Dover Publications.

[Riehl and Verity, 2018] Riehl, E. and Verity, D. (2018). Elements of infinity-category theory. *Preprint available at www. math. jhu. edu/ẽriehl/elements. pdf*.

[Rudin, 1973] Rudin, W. (1973). Functional analysis.

[Sakurai et al., 2014] Sakurai, J. J., Napolitano, J., et al. (2014). *Modern quantum mechanics*, volume 185. Pearson Harlow.

[Schreiber, 2013] Schreiber, U. (2013). Differential cohomology in a cohesive infinity-topos. *arXiv preprint arXiv:1310.7930*.

[Schreiber, 2016] Schreiber, U. (2016). Modern physics formalized in modal homotopy type theory. *Ncatlab presentation notes*.

[Shulman, 2008] Shulman, M. A. (2008). Set theory for category theory. *arXiv preprint arXiv:0810.1279*.

[Takhtadzhian, 2008] Takhtadzhian, L. A. (2008). *Quantum mechanics for mathematicians*, volume 95. American Mathematical Soc.

[Ticciati et al., 1999] Ticciati, R. et al. (1999). *Quantum field theory for mathematicians*, volume 72. Cambridge University Press.

[Vakil, 2017] Vakil, R. (2017). The rising sea: Foundations of algebraic geometry. *preprint*.

[Wald, 2007] Wald, R. M. (2007). *General relativity*. University of Chicago Press (Chicago, 1984).

[Weibel, 1995] Weibel, C. A. (1995). *An introduction to homological algebra*. Number 38. Cambridge university press.

[Weibel, 2013] Weibel, C. A. (2013). *The K-book: An introduction to algebraic K-theory*, volume 145. American Mathematical Society Providence, RI.

[Weinberg, 1995] Weinberg, S. (1995). *The quantum theory of fields*, volume 1. Cambridge university press.