Compendium of Mathematics & Physics

Nicolas Dewolf

October 16, 2022

Contents

Chapter 1

Category theory

For the general theory of categories, the classical reference is [15]. The main reference for (co)end calculus is [111], while a thorough introduction to the theory of enrichment is given in [12]. For the theory of higher categories and its applications to topology and algebra, the reader is referred to the book by *Baez et al.* [30]. A good starting point for bicategories (and more) is the paper by *Leinster* [86].

1.1 Categories

Definition 1.1.1 (Category). A category $\mathbf{C}$ consists of two collections, the objects $\mathrm{ob}(\mathbf{C})$ and the morphisms $\hom(\mathbf{C})$ or $\mathrm{mor}(\mathbf{C})$, that satisfy the following conditions:

1. **Source and target**: For every morphism $f \in \hom(\mathbf{C})$ there exist two objects $s(f), t(f) \in \mathrm{ob}(\mathbf{C})$, the source and the target. The collection of all morphisms with source x and target y is denoted by $\mathrm{Hom}_{\mathbf{C}}(x, y)$ or $\mathbf{C}(x, y)$.
2. **Composition**: For every two morphisms $f \in \mathbf{C}(y, z)$ and $g \in \mathbf{C}(x, y)$, the composite $f \circ g$ is an element of $\mathbf{C}(x, z)$. Moreover, composition is required to be associative.
3. **Identity**: For every $x \in \mathrm{ob}(\mathbf{C})$, there exists an identity morphism $\mathbb{1}_x \in \mathbf{C}(x, x)$. Identity morphisms are required to satisfy $f \circ \mathbb{1}_x = f = \mathbb{1}_y \circ f$ for all morphisms $f \in \mathbf{C}(x, y)$.

Remark 1.1.2. One technically does not need to consider objects as a separate notion since every object can be identified with its identity morphism (which exists by definition) and, hence, one can work solely with morphisms. It should be noted that for higher categories this remark can be omitted since the objects are always regarded as 0-morphisms in that context.

Definition 1.1.3 (Subcategory). Consider two categories $\mathbf{C}$ and $\mathbf{S}$. $\mathbf{S}$ is called a subcategory of $\mathbf{C}$ if $\mathrm{ob}(\mathbf{S})$ and $\mathrm{Hom}(\mathbf{S})$ are subcollections of $\mathrm{ob}(\mathbf{C})$ and $\mathrm{Hom}(\mathbf{C})$, respectively.

A subcategory is said to be **full** if for every two objects $x, y \in \mathrm{ob}(\mathbf{S})$:

$$\mathbf{S}(x, y) = \mathbf{C}(x, y). \tag{1.1}$$

A subcategory is said to be **wide** or **lluf** if it contains all objects:

$$\mathrm{ob}(\mathbf{S}) = \mathrm{ob}(\mathbf{C}). \tag{1.2}$$

Definition 1.1.4 (Replete subcategory). A subcategory $\mathbf{S} \subseteq \mathbf{C}$ such that if $x \in \mathrm{ob}(\mathbf{S})$ and $f : x \cong y \in \hom(\mathbf{C})$, then also $y \in \mathrm{ob}(\mathbf{S})$ and $f \in \hom(\mathbf{S})$. This can be stated even more concisely in terms of *arrow categories* (Definition 1.3.4).

Definition 1.1.5 (Small category). A category $\mathbf{C}$ for which both $\mathrm{ob}(\mathbf{C})$ and $\hom(\mathbf{C})$ are sets. A category $\mathbf{C}$ is said to be **locally small** if for every two objects $x, y \in \mathrm{ob}(\mathbf{C})$ the collection of morphisms $\mathbf{C}(x, y)$ is a set. A category *equivalent* (see further down below) to a small category is said to be **essentially small**.

Definition 1.1.6 (Opposite category). Let $\mathbf{C}$ be a category. The opposite category $\mathbf{C}^{op}$ is constructed by reversing all arrows in $\mathbf{C}$, i.e. a morphism in $\mathbf{C}^{op}(x, y)$ is a morphism in $\mathbf{C}(y, x)$.

Property 1.1.7 (Involution). From the definition of the opposite category it readily follows that *op* is an involution:

$$(\mathbf{C}^{op})^{op} = \mathbf{C}. \tag{1.3}$$

1.2 Functors

Definition 1.2.1 (Covariant functor). Let $\mathbf{A}, \mathbf{B}$ be categories. A (covariant) functor is an assignment $F : \mathbf{A} \to \mathbf{B}$ satisfying the following conditions:

1. F maps every object $x \in \mathrm{ob}(\mathbf{A})$ to an object $Fx \in \mathrm{ob}(\mathbf{B})$.
2. F maps every morphism $\phi \in \mathbf{A}(x, y)$ to a morphism $F\phi \in \mathbf{B}(Fx, Fy)$.
3. F preserves identities, i.e. $F\mathbb{1}_x = \mathbb{1}_{Fx}$.
4. F preserves compositions, i.e. $F(\phi \circ \psi) = F\phi \circ F\psi$.

Remark 1.2.2 (Category of categories). Small categories, together with (covariant) functors between them, form a category **Cat**. The restriction to small categories is important since otherwise one would obtain an inconsistency similar to *Russell's paradox*. In certain foundations one can also consider the "category" **CAT** of all categories, but this would not be a large category anymore. It would be something like a "very large" category.

Definition 1.2.3 (Contravariant functor). Let $\mathbf{A}, \mathbf{B}$ be categories. A contravariant functor is an assignment $F : \mathbf{A} \to \mathbf{B}$ satisfying the following conditions:

1. F maps every object $x \in \mathrm{ob}(\mathbf{A})$ to an object $Fx \in \mathrm{ob}(\mathbf{B})$.
2. F maps every morphism $\phi \in \mathbf{A}(x, y)$ to a morphism $F\phi \in \mathbf{B}(Fy, Fx)$.
3. F preserves identities, i.e. $F\mathbb{1}_x = \mathbb{1}_{Fx}$.
4. F reverses compositions, i.e. $F(\phi \circ \psi) = F\psi \circ F\phi$.

A contravariant functor can also be defined as a covariant functor from the opposite category and, accordingly, from now on the word "covariant" will be dropped when talking about functors.

Definition 1.2.4 (Endofunctor). A functor of the form $F : \mathbf{C} \to \mathbf{C}$.

Definition 1.2.5 (Presheaf). A functor $G : \mathbf{C}^{op} \to \mathbf{Set}$. The collection of all presheaves on a (small) category $\mathbf{C}$ forms a category $\mathbf{Psh}(\mathbf{C})$. This is sometimes also denoted by $\widehat{\mathbf{C}}$.

Example 1.2.6 (Hom-functor). Let $\mathbf{C}$ be a locally small category. Every object $x \in \mathrm{ob}(\mathbf{C})$ induces a functor $h^x : \mathbf{C} \to \mathbf{Set}$ defined as follows:

- h^x maps every object $y \in \mathrm{ob}(\mathbf{C})$ to the set $\mathbf{C}(x, y)$.
- For all $y, z \in \mathrm{ob}(\mathbf{C})$, h^x maps every morphism $f \in \mathbf{C}(y, z)$ to the fucntion

$$f \circ - : \mathbf{C}(x, y) \to \mathbf{C}(x, z) : g \mapsto f \circ g.$$

Remark 1.2.7. The contravariant hom-functor h_x is defined by replacing $\mathbf{C}(x, -)$ with $\mathbf{C}(-, x)$ and replacing postcomposition with precomposition.

Definition 1.2.8 (Faithful functor). A functor $F : \mathbf{A} \to \mathbf{B}$ for which the map

$$\mathbf{A}(x, y) \to \mathbf{B}(Fx, Fy)$$

is injective for all objects $x, y \in \mathrm{ob}(\mathbf{A})$.

Definition 1.2.9 (Full functor). A functor $F : \mathbf{A} \to \mathbf{B}$ for which the map

$$\mathbf{A}(x, y) \to \mathbf{B}(Fx, Fy)$$

is surjective for all objects $x, y \in \mathrm{ob}(\mathbf{A})$.

Definition 1.2.10 (Embedding). A fully faithful functor.

Definition 1.2.11 (Essentially surjective functor). A functor $F : \mathbf{A} \to \mathbf{B}$ such that for every object $y \in \mathrm{ob}(\mathbf{B})$, there exists an object $x \in \mathrm{ob}(\mathbf{A})$ with $Fx \cong y$.

Definition 1.2.12 (Profunctor[1]). A functor of the form $F : \mathbf{B}^{op} \times \mathbf{A} \to \mathbf{Set}$. Such a functor is often denoted by $F : \mathbf{A} \nrightarrow \mathbf{B}$.[2] Elements of the set $F(x, y)$ are called **heteromorphisms** (between x and y).

It should be noted that presheaves on $\mathbf{C}$ are profunctors of the form $1 \nrightarrow \mathbf{C}$.

Definition 1.2.13 (Reflection). A functor $F : \mathbf{A} \to \mathbf{B}$ is said to **reflect** a property if whenever the property holds for Fa, it also holds for a. (Here, a could also be a morphism.)

1.2.1 Natural transformations

Definition 1.2.14 (Natural transformation). Let $F, G : \mathbf{A} \to \mathbf{B}$ be functors. A natural transformation $\psi : F \Rightarrow G$[3] consists of a collection of morphisms satisfying the following two conditions:

1. For every object $x \in \mathrm{ob}(\mathbf{A})$, there exists a morphism $\psi_x : Fx \to Gx$ in $\hom(\mathbf{B})$. This morphism is called the **component** of ψ at x. (It is often said that ψ_x **is natural in** x.)
2. For every morphism $f \in \mathbf{A}(x, y)$, the diagram below commutes:

$$\begin{tikzcd} Fx \arrow[r, "Ff"] \arrow[d, "\psi_x"'] & Fy \arrow[d, "\psi_y"] \\ Gx \arrow[r, "Gf"'] & Gy \end{tikzcd}$$

Definition 1.2.15 (Functor category). Consider two categories $\mathbf{A}$ and $\mathbf{B}$, where $\mathbf{A}$ is small. The functors $F : \mathbf{A} \to \mathbf{B}$ form the objects of a category with the natural transformations as morphisms. This category is denoted by $[\mathbf{A}, \mathbf{B}]$ or $\mathbf{B}^{\mathbf{A}}$ (the latter is a generalization of **??**).

Definition 1.2.16 (Dinatural transformation). Consider two profunctors $F, G : \mathbf{A} \nrightarrow \mathbf{A}$ or, more generally, two functors $F, G : \mathbf{A}^{op} \times \mathbf{A} \to \mathbf{B}$. A dinatural transformation is a family of morphisms

$$\eta_x : F(x, x) \to G(x, x)$$

that make Diagram 1.1 commute for every morphism $f : y \to x$.

[1] Sometimes called a **distributor**.

[2] This is the convention by *Borceux*. Some other authors, such as [16], use the opposite convention.

[3] This notation is in analogy with the general notation for *2-morphisms*. See Section 1.6 for more information.

$$
\begin{tikzcd}
& F(x,y) \arrow[dl, "F(\mathbb{1}_x, f)"'] \arrow[dr, "F(f, \mathbb{1}_y)"] & \\
F(x,x) \arrow[d, "\eta_x"'] & & F(y,y) \arrow[d, "\eta_y"] \\
G(x,x) \arrow[dr, "G(f, \mathbb{1}_x)"'] & & G(y,y) \arrow[dl, "G(\mathbb{1}_y, f)"] \\
& G(y,x) &
\end{tikzcd}
$$

Figure 1.1: Dinatural transformation.

Definition 1.2.17 (Representable functor). Let $\mathbf{C}$ be a locally small category. A functor $F : \mathbf{C} \to \mathbf{Set}$ is said to be representable if there exists an object $x \in \mathrm{ob}(\mathbf{C})$ such that F is naturally isomorphic to h^x. The pair $(x, \psi : F \Rightarrow h^x)$ is called a **representation** of F.

Theorem 1.2.18 (Yoneda lemma). *Let $\mathbf{C}$ be a locally small category and let $F : \mathbf{C} \to \mathbf{Set}$ be a functor. For every object $x \in \mathrm{ob}(\mathbf{C})$, there exists a natural isomorphism*[4]

$$\eta_x : \mathrm{Nat}(h^x, F) \to Fx : \psi \mapsto \psi_x(\mathbb{1}_x). \tag{1.4}$$

Corollary 1.2.19 (Yoneda embedding). When F is another hom-functor h^y, the following result is obtained:

$$\mathrm{Nat}(h^x, h^y) \cong \mathbf{C}(y, x). \tag{1.5}$$

Note that y appears in the first argument on the right-hand side.

Let $\mathbf{C}(f, -)$ denote the natural transformation corresponding to the morphism $f \in \mathbf{C}(y, x)$. The functor h^-, mapping an object $x \in \mathrm{ob}(\mathbf{C})$ to its hom-functor $\mathbf{C}(x, -)$ and a morphism $f \in \mathbf{C}(y, x)$ to the natural transformation $\mathbf{C}(f, -)$, can also be interpreted as a covariant functor $G : \mathbf{C}^{op} \to \mathbf{Set}^{\mathbf{C}}$. This way the Yoneda lemma can be seen to give rise to an embedding h^- of $\mathbf{C}^{op}$ in the functor category $\mathbf{Set}^{\mathbf{C}}$.

As usual, all of this can be done for contravariant functors. This gives an embedding

$$\mathcal{Y} := h_- : \mathbf{C} \hookrightarrow \widehat{\mathbf{C}}, \tag{1.6}$$

called the Yoneda embedding.

Definition 1.2.20 (Local object). Consider a collection of morphisms $S \subseteq \mathrm{hom}(\mathbf{C})$. An object $c \in \mathrm{ob}(\mathbf{C})$ is said to be S-local if the Yoneda embedding $\mathcal{Y}c$ maps morphisms in S to isomorphisms in $\mathbf{Set}$. A morphism $f \in \mathrm{hom}(\mathbf{C})$ is said to be S-local if its image under the Yoneda embedding of every S-local object is an isomorphism in $\mathbf{Set}$.

1.2.2 Equivalences

Definition 1.2.21 (Equivalence of categories). Two categories $\mathbf{A}, \mathbf{B}$ are said to be equivalent if there exist functors $F : \mathbf{A} \to \mathbf{B}$ and $G : \mathbf{B} \to \mathbf{A}$ such that $F \circ G$ and $G \circ F$ are naturally isomorphic to the identity functors.

[4]Here, the fact that $\mathrm{Nat}(h^-, -)$ can be seen as a functor $\mathbf{Set}^{\mathbf{C}} \times \mathbf{C} \to \mathbf{Set}$ is used.

A weaker notion is that of a **weak equivalence**. Two categories $\mathbf{A}, \mathbf{B}$ are said to be weakly equivalent if there exist functors $F : \mathbf{A} \to \mathbf{B}$ and $G : \mathbf{B} \to \mathbf{A}$ that are fully faithful and essentially surjective. Assuming the axiom of choice, every weak equivalence is also a (strong) equivalence (in fact this statement is equivalent to the axiom of choice).

Definition 1.2.22 (Skeletal category). A category in which every isomorphism is necessarily an identity morphism. The **skeleton** of a category is an equivalent skeletal category (often taken to be a subcategory by choosing a representative from every isomorphism class).

If one does not assume the axiom of choice, the skeleton is merely a weakly equivalent skeletal category.

Definition 1.2.23 (Decategorification). Let $\mathbf{C}$ be an (essentially) small category. The set of isomorphism classes of $\mathbf{C}$ is called the decategorification of $\mathbf{C}$. This amounts to a functor $\mathrm{Decat} : \mathbf{Cat} \to \mathbf{Set}$.

1.2.3 Stuff, structure and property

To classify properties of objects and the *forgetfulness* of functors, it is interesting to make a distinction between stuff, structure and property. Consider for example a group. This is a set (*stuff*) equipped with a number of operations (*structure*) that obey some relations (*properties*).

Using these notions one can classify forgetful functors in the following way:

- A functor forgets nothing if it is an equivalence of categories.
- A functor forgets at most properties if it is fully faithful.
- A functor forgets at most structure if it is faithful.
- A functor forgets at most stuff if it is just a functor.

?? COMPLETE (see e.g. nLab or the paper "Why surplus structure is not superfluous" by Nicholas Teh et al.) ??

1.2.4 Adjunctions

Definition 1.2.24 (Hom-set adjunction). Let $F : \mathbf{A} \to \mathbf{B}$ and $G : \mathbf{B} \to \mathbf{A}$ be two functors. These functors form a (hom-set) adjunction $F \dashv G$ if the following isomorphism is natural in both x and y:

$$\Phi_{x,y} : \mathbf{B}(Fx, y) \cong \mathbf{A}(x, Gy). \tag{1.7}$$

The functor F (resp. G) is called the left (resp. right) adjoint and the image of a morphism under either of the natural isomoprhisms is called the adjunct of the other morphism.[5].

Notation 1.2.25. An adjunction $F \dashv G$ between categories $\mathbf{A}, \mathbf{B}$ is often denoted by

$$\begin{tikzcd} \mathbf{B} \arrow[r, bend right, "G"'] \arrow[r, phantom, "\perp"] & \mathbf{A} \arrow[l, bend right, "F"'] \end{tikzcd}.$$

Definition 1.2.26 (Unit-counit adjunction). Let $F : \mathbf{A} \to \mathbf{B}$ and $G : \mathbf{B} \to \mathbf{A}$ be two functors. These functors form a unit-counit adjunction if there exist natural transformations

$$\varepsilon : F \circ G \Rightarrow \mathbb{1}_{\mathbf{B}} \tag{1.8}$$
$$\eta : \mathbb{1}_{\mathbf{A}} \Rightarrow G \circ F \tag{1.9}$$

[5] The terms "adjunct" and "adjoint" are sometimes used interchangeably (cf. French versus Latin).

such that the following compositions are identity morphisms:

$$F \xrightarrow{F\eta} FGF \xrightarrow{\varepsilon F} F \tag{1.10}$$

$$G \xrightarrow{\eta G} GFG \xrightarrow{G\varepsilon} G. \tag{1.11}$$

These identities are sometimes called the **triangle** or **zig-zag identities** (the latter results from the shape of the associated *string diagram*). The transformations η and ε are called the **unit** and **counit** respectively.

Property 1.2.27 (Equivalence of the above definitions). Every hom-set adjunction induces a unit-counit adjunction. Let Φ be the natural isomorphism associated to the hom-set adjunction $F \dashv G$. The counit ε_y is obtained as the adjunct $\Phi^{-1}_{Gy,y}(\mathbb{1}_{Gy})$ of the identity morphism on $Gy \in \mathrm{ob}(\mathbf{A})$, and the unit η_x is analogously defined as the adjunct $\Phi_{c,Fc}(\mathbb{1}_{Fx})$ of the identity morphism at $Fx \in \mathrm{ob}(\mathbf{B})$.

Conversely, every unit-counit adjunction induces a hom-set adjunction. Consider a morphism $f : Fx \to y$. The (right) adjunct is defined as the composition

$$\widetilde{f} := Gf \circ \eta_x : x \to (G \circ F)x \to Gy.$$

To construct a (left) adjunct, consider a morphism $\widetilde{g} : x \to Gy$:

$$g := \varepsilon_y \circ F\tilde{g} : Fx \to (F \circ G)y \to y.$$

Definition 1.2.28 (Reflective subcategory). A full subcategory is said to be reflective (resp. coreflective) if the inclusion functor admits a left (resp. right) adjoint.

Property 1.2.29 (Adjoint equivalence). Any equivalence of categories is part of an adjoint equivalence, i.e. an adjunction for which the unit and counit morphisms are invertible.

Property 1.2.30. Given an adjunction $\mathbf{C} : F \dashv G : \mathbf{D}$, one obtains an adjoint equivalence by restricting to the full subcategories on which the unit and counit becomes isomorphisms.

1.3 General constructions

Definition 1.3.1 (Dagger category). A category equipped with a contravariant involutive endofunctor, this functor is often denoted by $\dagger : \mathbf{C} \to \mathbf{C}$, similar to the adjoint operator for Hermitian matrices.

Remark 1.3.2. The concept of a dagger structure allows the usual definition of **unitary** and **self-adjoint** morphisms, i.e. morphism satisfying

$$f^\dagger = f^{-1} \qquad \text{or} \qquad f^\dagger = f. \tag{1.12}$$

Definition 1.3.3 (Comma category). Let $\mathbf{A}, \mathbf{B}$ and $\mathbf{C}$ be three categories and let $F : \mathbf{A} \to \mathbf{C}$ and $G : \mathbf{B} \to \mathbf{C}$ be two functors. The comma category $F \downarrow G$ is defined as follows:

- **Objects**: The triples (x, y, γ) where $x \in \mathrm{ob}(\mathbf{A}), y \in \mathrm{ob}(\mathbf{B})$ and $\gamma : Fx \to Gy$.
- **Morphisms**: The morphisms $(x, y, \gamma) \to (k, l, \sigma)$ are pairs (f, g) with $f : x \to k \in \hom(\mathbf{A})$ and $g : y \to l \in \hom(\mathbf{B})$ such that $\sigma \circ Ff = Gg \circ \gamma$.

Composition of morphisms is defined componentwise.

Definition 1.3.4 (Arrow category). The comma category of the pair of functors $(\mathbb{1}_{\mathbf{C}}, \mathbb{1}_{\mathbf{C}})$. This is equivalently the functor category $[\mathbf{2}, \mathbf{C}]$, where $\mathbf{2}$ is the **interval category/walking arrow** $\{0 \to 1\}$.

Definition 1.3.5 (Functorial factorization). A *section* (see Definition 1.4.1) of the composition functor

$$\circ : [\mathbf{3}, \mathbf{C}] \to [\mathbf{2}, \mathbf{C}]\,,$$

where $\mathbf{3}$ is the poset $\{0 \to 1 \to 2\}$.

Definition 1.3.6 (Slice category). Let $\mathbf{C}$ be a category and consider an object $x \in \mathrm{ob}(\mathbf{C})$. The slice category $\mathbf{C}_{/x}$ of $\mathbf{C}$ over x is defined as follows:

- **Objects**: The morphisms in $\mathbf{C}$ with codomain x, and
- **Morphisms**: The morphisms $f \to g$ are morphisms h in $\mathbf{C}$ such that $g \circ h = f$.

This category is also called the **over-category** of x. By dualizing one obtains the **under-category** of x.

1.3.1 Fibred categories ♣

Definition 1.3.7 (Fibre category). Let $\Pi : \mathbf{A} \to \mathbf{B}$ be a functor. The fibre category (of Π) over $y \in \mathrm{ob}(\mathbf{B})$ is the subcategory of $\mathbf{A}$ consisting of all objects $x \in \mathrm{ob}(\mathbf{A})$ such that $\Pi x = y$ and all morphisms $f \in \hom(\mathbf{A})$ such that $\Pi f = \mathbb{1}_y$. It will be denoted by $\mathbf{A}_y$.

Morphisms in $\mathbf{A}$ that are mapped to a morphism g in $\mathbf{B}$ are called **g-morphisms** and, in particular (using the identification of objects and their identity morphisms), morphisms in $\mathbf{A}_y$ are called **y-morphisms**. Similarly, **B-categories** are defined as the categories equipped with a (covariant) functor to $\mathbf{B}$. (It is not hard to see that these form a *2-category* under composition of functors that respects the $\mathbf{B}$-category structure.)

Definition 1.3.8 (Cartesian morphism). Consider a $\mathbf{B}$-category $\Pi : \mathbf{A} \to \mathbf{B}$. A morphism f in $\mathbf{A}$ is called Π-Cartesian if every Πf-morphism factors uniquely through a y-morphism, where y is the domain of Πf.

There also exists a stronger notion. A **strongly Cartesian morphism** is a morphism $f \in \hom(\mathbf{A})$ such that for every morphism $\varphi \in \hom(\mathbf{A})$ with the same target and every factorization of $\Pi\varphi$ through Πf, there exists a unique factorization of φ through f that maps to the given factorization of $\Pi\varphi$.

The following diagram, where the triangles commute, should clarify the above (technical) definitions:

$$\begin{tikzcd} \forall x' \arrow[d, "\exists! g"'] \arrow[dr, "\forall\varphi"] & & & \Pi x' \arrow[d, "\forall\nu"'] \arrow[dr, "\Pi\varphi"] & \\ x_1 \arrow[r, "f"'] & x_2 & \overset{\Pi}{\mapsto} & \Pi x_1 \arrow[r, "\Pi f"'] & \Pi x_2 \end{tikzcd}$$

The diagram for (weak) Cartesian morphisms is obtained by identifying the objects $\Pi x'$ and Πx_1, i.e. by restricting to the case $\nu = \mathbb{1}_{\Pi x_1}$.

The Cartesian morphisms are said to be **inverse images** of their projections under Π and the object x_1 is called an **inverse image** of x_2 by Πf. The Cartesian morphisms of a fibre category are exactly the isomorphisms of that category.

Definition 1.3.9 (Fibred category). A $\mathbf{B}$-category $\Pi : \mathbf{A} \to \mathbf{B}$ is called a fibred category or **Grothendieck fibration** if the following conditions are satisfied:

1. For each morphism in $\mathbf{B}$ whose codomain lies in the range of Π and each lift of this codomain to $\mathbf{A}$, there exists at least one inverse image with the given codomain (in the weak sense).

2. The composition of two Cartesian morphisms is again Cartesian (in the weak sense).

If one instead works with strongly Cartesian morphisms, the second condition follows from the first one. However, it should be noted that in a fibred category a morphism is weakly Cartesian if and only if it is strongly Cartesian.

Definition 1.3.10 (Cleavage). Given a $\mathbf{B}$-category $\Pi : \mathbf{A} \to \mathbf{B}$, a cleavage is the choice of a Cartesian g-morphism $f : x \to y$ for every $y \in \mathrm{ob}(\mathbf{A})$ and morphism $g : b \to \Pi a'$. A $\mathbf{B}$-category equipped with a cleavage is said to be **cloven**.

The existence of cleavage is sufficient for a category to be fibred and, conversely (assuming the axiom of choice), every fibred category admits a cleavage.

The following example can be obtained as a Grothendieck fibration with discrete fibres:

Example 1.3.11 (Discrete fibration). A functor $F : \mathbf{A} \to \mathbf{B}$ such that for every object $x \in \mathrm{ob}(\mathbf{A})$ and every morphism $f : y \to Fx$ in $\mathbf{B}$ there exists a unique morphism $g : z \to x$ in $\mathbf{A}$ such that $Fg = f$.

Example 1.3.12 (Groupoidal fibration). If every morphism is required to be Cartesian, the notion of a groupoid(al) fibration or a **category fibred in groupoids** is obtained. The reason for this name is that every fibre is a groupoid. An equivalent definition is that the associated pseudofunctor (see the construction below) factors through the embedding $\mathbf{Grpd} \hookrightarrow \mathbf{Cat}$.

Property 1.3.13 (Grothendieck construction ♣). Every fibred category $\Pi : \mathbf{A} \to \mathbf{B}$ defines a *pseudofunctor*[6] $F : \mathbf{B}^{op} \to \mathbf{Cat}$ that sends objects to fibre categories and arrows $f : c \to d$ to the pullback functor $f^* : \mathbf{A}_d \to \mathbf{A}_c$ constructed from a Cartesian lift of f. This pullback functor acts as follows:

- For every object $x \in \mathbf{A}_d$, f^*x is the domain of the Cartesian lift of f through x.
- For every morphism $(\alpha : x \to y) \in \mathbf{A}_d$ there exists a diagram of the form

$$\begin{tikzcd} f^*x \arrow[r] \arrow[d, dashed, "f^*\alpha"'] & x \arrow[d, "\alpha"] \\ f^*y \arrow[r] & y \end{tikzcd}$$

  Because the horizontal morphism are both projected to f and α is projected to the identity, there exists a unique factorization of the diagram through a morphism $f^*\alpha : f^*x \to f^*y$.

Conversely, every pseudofunctor gives rise to a fibred category through the Grothendieck construction $\int : [\mathbf{C}^{op}, \mathbf{Cat}] \to \mathbf{Cat}_{/\mathbf{C}}$ as follows. (These two constructions constitute a 2-equivalence of 2-categories.). Consider a pseudofunctor $F : \mathbf{C}^{op} \to \mathbf{Cat}$. The "bundle" $\int F$ consists of the following data:

- The objects are pairs (x, y) with $x \in \mathrm{ob}(\mathbf{C})$ and $y \in \mathrm{ob}(Fx)$.
- The morphisms $(x, y) \to (x', y')$ are pairs $(f : x \to x', \alpha : y \to Ff(y'))$.

[6]See Definition 1.6.9 towards the end of this chapter.

Given a cleavage, the morphisms of the Grothendieck construction are exactly the factorizations of f-morphisms through the canonical lifting of f in the cleavage.

Property 1.3.14 (Functors). A pseudofunctor is a functor if and only if the cleavage of the associated fibred category is **split(ting)**, i.e. it contains all identities and is closed under composition.

Definition 1.3.15 (Category of elements). Consider a presheaf $F : \mathbf{C}^{op} \to \mathbf{Set}$. Its category of elements $\mathrm{El}(F)$ is defined as the comma category $(\mathcal{Y} \downarrow !_F)$, where $!_F : * \to [\mathbf{C}^{op}, \mathbf{Set}]$ sends the unique object to F itself. Equivalently, it is the category with objects the pairs $(c, x) \in \mathrm{ob}(\mathbf{C}) \times Fc$ and morphisms $f \in \mathbf{C}(c, c')$ such that $c = Ff(c')$, i.e. it is the Grothendieck construction applied to F.

This category comes equipped with a canonical forgetful functor

$$\mathbf{C}_F : \mathrm{El}(F) \to \mathbf{C} : (c, x) \mapsto c. \tag{1.13}$$

Remark 1.3.16. The category of elements is usually defined for covariant functors. To obtain that definition one should take the opposite of the category of elements (and also take the opposite of the forgetful functor).

1.3.2 Monads

Definition 1.3.17 (Monad). A monad is a triple (T, μ, η) where $T : \mathbf{C} \to \mathbf{C}$ is an endofunctor and $\mu : T^2 \to T, \eta : \mathbb{1}_\mathbf{C} \to T$ are natural transformations satisfying the following (coherence) conditions:

1. As natural transformations from T^3 to T:

$$\mu \circ T\mu = \mu \circ \mu_T. \tag{1.14}$$

2. As natural transformations from T to itself:

$$\mu \circ T\eta = \mu \circ \eta_T = \mathbb{1}. \tag{1.15}$$

These conditions say that a monad is a monoid **??** in the category $\mathbf{End}_\mathbf{C}$ of endofunctors on $\mathbf{C}$. Accordingly, η and μ are often called the **unit** and **multiplication** maps.

Example 1.3.18 (Adjunction). Every adjunction $F \dashv G$, with unit ε and counit η, induces a monad of the form $(GF, G\varepsilon F, \eta)$.

Definition 1.3.19 (Algebra over a monad[7]). Consider a monad (T, μ, η) on a category $\mathbf{C}$. An algebra over T or T-algebra is a couple (x, κ), where $x \in \mathrm{ob}(\mathbf{C})$ and $\kappa : Tx \to x$, such that the following conditions are satisfied:

1. $\kappa \circ T\kappa = \kappa \circ \mu_x$, and
2. $\kappa \circ \eta_x = \mathbb{1}_x$.

Morphisms $(x, \kappa_x) \to (y, \kappa_y)$ of T-algebras are morphisms $f : x \to y$ in $\mathbf{C}$ such that $f \circ \kappa_x = \kappa_y \circ Tf$. An algebra of the form (Tx, μ_x) is said to be **free**.

Definition 1.3.20 (Eilenberg-Moore category). Given a monad T over a category $\mathbf{C}$, the Eilenberg-Moore category $\mathbf{C}^T$ is defined as the category of T-algebras.

[7] A more suitable name would be "module over a monad", since these are modules over a monoid if monads are regarded as monoids in $\mathbf{End}_\mathbf{C}$.

Definition 1.3.21 (Kleisli category). Consider a monad T on a category $\mathbf{C}$. The Kleisli category $\mathbf{C}_T$ is defined as the full subcategory of $\mathbf{C}^T$ on the free T-algebras. This is equivalently the category with objects $\mathrm{ob}(\mathbf{C}_T) := \mathrm{ob}(\mathbf{C})$ and morphisms $\mathbf{C}_T(x,y) := \mathbf{C}(x, Ty)$.

Morphisms in the Kleisly category are composed in the "obvious way":

$$f \circ_{\mathbf{C}_T} g := \mu_Z \circ Tf \circ g \tag{1.16}$$

for all $f \in \mathbf{C}_T(Y, TZ)$ and $g \in \mathbf{C}_T(X, TY)$.

Definition 1.3.22 (Monadic adjunction). An adjunction between categories $\mathbf{A}$ and $\mathbf{B}$ is said to be monadic if there exists an equivalence between $\mathbf{B}$ and the Eilenberg-Moore category of the induced monad.

Definition 1.3.23 (Monadic functor). A functor is said to be monadic if it admits a left adjoint such that the adjunction is monadic.

The following theorem characterizes monadic functors (for more information on some of the concepts, see Section 1.4 further below):

Theorem 1.3.24 (Beck's monadicity theorem). *Consider a functor $F : \mathbf{A} \to \mathbf{B}$. This functor is monadic if and only if the following conditions are satisfied:*

- *F admits a left adjoint.*
- *F reflects isomorphisms.*
- *$\mathbf{A}$ has all coequalizers of F-split parallel pairs*[8] *and F preserves these coequalizers.*

Remark 1.3.25 (Crude monadicity theorem). A sufficient condition for monadicity is obtained by replacing the third condition above by the following weaker statement: "$\mathbf{A}$ has all coequalizers of reflexive pairs and F preserves these coequalizers."

Definition 1.3.26 (Closure operator). Consider a monad $(T : \mathbf{C} \to \mathbf{C}, \eta, \mu)$. This monad is called a closure operator or **modal operator** if the multiplication map is a natural isomorphism, i.e. if the monad is idempotent.

Given a closure operator $T : \mathbf{C} \to \mathbf{C}$, the object Tx is called the closure of $x \in \mathrm{ob}(\mathbf{C})$ and the associated morphism η_x is called the **closing map**. An object $x \in \mathrm{ob}(\mathbf{C})$ itself is said to be **T-closed** exactly if its closing map is an isomorphism.

An object $x \in \mathrm{ob}(\mathbf{C})$ is called a **modal type** if the unit $\eta_x : x \to Tx$ is an isomorphism.

Remark 1.3.27 (Bicategories ♣). A monad can be defined in any bicategory as a 1-morphism $t : x \to x$ together with two 2-morphisms that satisfy conditions similar to the ones above. The above definition is then just a specific case of this more general definition in **Cat**.

In the general setting one can then also define a **module** over a monad. First of all, one can regard any object $x \in \mathrm{ob}(\mathbf{C})$ as a functor from the terminal category $\mathbf{1}$. By replacing $\mathbf{1}$ by any other category in the ordinary definition one obtains a general algebra (or module). It is this definition that readily generalizes to bicategories, i.e. a module is a 1-morphism $a : x \to y$ together with a 2-morphism that satisfies the same conditions as an algebra over a monad in **Cat**.

[8] These are parallel pairs f, g such that the images Ff, Fg under F admit a split coequalizer.

1.4 Morphisms and diagrams

1.4.1 Morphisms

Definition 1.4.1 (Section). A section of a morphism $f : x \to y$ is a right inverse, i.e. a morphism $g : y \to x$ such that $f \circ g = \mathbb{1}_y$. f itself is called a **retraction** of g and y is called a **retract** of x.

Definition 1.4.2 (Monomorphism). Let $\mathbf{C}$ be a category. A morphism $\mu \in \mathbf{C}(x, y)$ is called a monomorphism, **mono** or **monic morphism** if for every object $z \in \mathrm{ob}(\mathbf{C})$ and every two morphisms $\alpha_1, \alpha_2 \in \mathbf{C}(z, x)$ such that $\mu \circ \alpha_1 = \mu \circ \alpha_2$, one can conclude that $\alpha_1 = \alpha_2$.

Definition 1.4.3 (Epimorphism). Let $\mathbf{C}$ be a category. A morphism $\varepsilon \in \mathbf{C}(x, y)$ is called an epimorphism, **epi** or **epic morphism** if for every object $z \in \mathrm{ob}(\mathbf{C})$ and every two morphisms $\alpha_1, \alpha_2 \in \mathbf{C}(y, z)$ such that $\alpha_1 \circ \varepsilon = \alpha_2 \circ \varepsilon$, one can conclude that $\alpha_1 = \alpha_2$.

Definition 1.4.4 (Split monomorphism). A morphism $f : x \to y$ that is a section of some other morphism $g : y \to x$. It can be shown that every split mono is in fact a mono and even an **absolute mono**, i.e. it is preserved by all functors.

The morphism g can be seen to satisfy the dual condition and, hence, is called a **split epimorphism**. It can be shown to be an absolute epi.

Definition 1.4.5 (Balanced category). A category in which every monic epi is an isomorphism.

Definition 1.4.6 (Reflexive pair). Two parallel morphisms $f, g : x \to y$ are said to form a reflexive pair if they have a common section, i.e. if there exists a morphism $\sigma : y \to x$ such that $f \circ \sigma = g \circ \sigma = \mathbb{1}_y$.

Definition 1.4.7 (Subobject). Let $\mathbf{C}$ be a category and let $x \in \mathrm{ob}(\mathbf{C})$ be any object. A subobject y of x is a mono $y \hookrightarrow x$.

In fact, one should work up to isomorphisms and, accordingly, the formal definition goes as follows: a subobject y of x is an isomorphism class of monos $i : y \hookrightarrow x$ in the slice category $\mathbf{C}/x$.

Definition 1.4.8 (Well-powered category). A category $\mathbf{C}$ such that for every object $x \in \mathrm{ob}(\mathbf{C})$ the class of subobjects $\mathrm{Sub}(x)$ is small.

1.4.2 Initial and terminal objects

Definition 1.4.9 (Initial object). An object $\emptyset$ such that for every other object x there exists a unique morphism $\iota_x : \emptyset \to x$.

Definition 1.4.10 (Terminal object). An object 1 such that for every other object x there exists a unique morphism $\tau_x : x \to 1$.

Property 1.4.11 (Uniqueness). If an initial (or terminal) object exists, it is unique (up to isomorphisms).

Definition 1.4.12 (Zero object). An object that is both initial and terminal. The zero object is often denoted by 0.

Property 1.4.13 (Zero morphism). From the definition of the zero object it follows that for any two objects x, y there exists a unique morphism $0_{xy} : x \to 0 \to y$.

Definition 1.4.14 (Pointed category). A category containing a zero object.

Definition 1.4.15 (Global element). Let $\mathbf{C}$ be a category with a terminal object 1. A global element of an object $x \in \mathrm{ob}(\mathbf{C})$ is a morphism $1 \to x$.

Property 1.4.16. Every global element is monic.

Definition 1.4.17 (Pointed object). An object x equipped with a global element $1 \to x$. This morphism is sometimes called the **basepoint**.

Remark 1.4.18. In the category **Set** the elements of a set S are in one-to-one correspondence with the global elements of S. Furthermore, there is the the important property (*axiom of functional extensionality*) that two functions $f, g : S \to S'$ coincide if their values at every element $s \in S$ coincide or, equivalently, if their precompositions with global elements coincide.

However, this way of checking equality can fail in other categories. Consider for example **Grp**, the category of groups, with its zero object $0 = \{e\}$. The only morphism from this group to any other group G is the one mapping e to the unit in G. It is obvious that precomposition with this morphism says nothing about the equality of other morphisms. To recover the extensionality property from **Set**, the notion of an "element" should be generalized:

Definition 1.4.19 (Generalized element). Let $\mathbf{C}$ be a category and consider an object $x \in \mathrm{ob}(\mathbf{C})$. For any object $y \in \mathrm{ob}(\mathbf{C})$, a morphism $y \to x$ is called a generalized element of x. These morphisms are also called **y-elements** in x or elements of **shape** y in x.

Definition 1.4.20 (Generator). Let $\mathbf{C}$ be a category. A collection of objects $\mathcal{O} \subset \mathrm{ob}(\mathbf{C})$ is called a collection of generators or **separators** for $\mathbf{C}$ if the generalized elements of shape $\mathcal{O}$ are sufficient to distinguish between all morphisms in $\mathbf{C}$:

$$\forall x, y \in \mathrm{ob}(\mathbf{C}) : \forall f, g \in \mathbf{C}(x, y) : \Big(f \neq g \implies \exists o \in \mathcal{O} : \exists h \in \mathbf{C}(o, x) : f \circ h \neq g \circ h\Big). \tag{1.17}$$

Definition 1.4.21 (Well-pointed category). A category for which the terminal object is a generator.

1.4.3 Lifts

Definition 1.4.22 (Lifts and extensions). A lift of a morphism $f : x \to y$ along an epi $e : z \to y$ is a morphism $g : x \to z$ satisfying $f = e \circ g$. Dualizing this definition gives the notion of extensions. (The epi/mono condition is often dropped in the literature.)

Definition 1.4.23 (Lifting property). A morphism $f : x \to y$ has the left lifting property with respect to a morphism $g : x' \to y'$ (or g has the right lifting property with respect to f) if for every commutative diagram

$$\begin{tikzcd}
x \arrow[r] \arrow[d, "f"'] & x' \arrow[d, "g"] \\
y \arrow[r] \arrow[ur, dashed, "\exists\psi"] & y'
\end{tikzcd}$$

there exists a morphism $\psi : y \to x'$ such that the triangles commute. If the morphism ψ is unique, then f and g are said to be **orthogonal**.

Definition 1.4.24 (Injective and projective morphisms). Consider a class of morphisms $I \subseteq \hom(\mathbf{C})$. A morphism $f \in \hom(\mathbf{C})$ is said to be I-injective (resp. I-projective) if it has the right (resp. left) lifting property with respect to all morphisms in I.

Given a set of morphisms I, the sets of I-injective and I-projective morphisms are denoted by $\mathrm{rlp}(I)$ and $\mathrm{llp}(I)$, respectively.

Definition 1.4.25 (Injective and projective objects). If $\mathbf{C}$ has a terminal object 1, an object x is called I-injective if its terminal morphism is I-injective. If $\mathbf{C}$ has an initial object, I-projective objects can be defined dually. (See Figure 1.2.)

$$
\begin{tikzcd}
x \arrow[r, "g"] \arrow[d, "\forall f \in I"'] & i \\
y \arrow[ur, dashed, "\exists\psi"'] &
\end{tikzcd}
$$

(a) Injective object i.

$$
\begin{tikzcd}
 & x \arrow[d, "\forall f \in I"] \\
p \arrow[r, "g"'] \arrow[ur, dashed, "\exists\psi"] & y
\end{tikzcd}
$$

(b) Projective object p.

Figure 1.2: Injective and projective objects.

If I is the class of monomorphisms (resp. epimorphisms), the terminology is simplified to **injective** (resp. **projective**) objects. For projective objects this is also equivalent to requiring that the (covariant) hom-functor preserves epimorphisms.

A category $\mathbf{C}$ is said to **have enough injectives** if for every object there exists a monomorphism into an injective object. The category is said to **have enough projectives** if for every object there exists an epimorphism from a projective object.

Definition 1.4.26 (Fibrations and cofibrations). Consider a category $\mathbf{C}$ together with a class $I \subseteq \hom(\mathbf{C})$ of morphisms. A morphism $f \in \hom(\mathbf{C})$ is called an I-fibration (resp. I-cofibration) if it has the right (resp. left) lifting property with respect to all I-projective (resp. I-injective) morphisms.

1.4.4 Limits and colimits

Definition 1.4.27 (Diagram). A diagram in $\mathbf{C}$ with index category $\mathbf{I}$ is a (covariant) functor $D : \mathbf{I} \to \mathbf{C}$.

Definition 1.4.28 (Cone). Let $D : \mathbf{I} \to \mathbf{C}$ be a diagram. A cone from $c \in \mathrm{ob}(\mathbf{C})$ to D consists of a family of morphisms $\psi_i : c \to Di$ indexed by $\mathbf{I}$ such that $\psi_j = Df \circ \psi_i$ for all morphisms $f : i \to j \in \hom(\mathbf{I})$. (This is depicted in Figure 1.3a.)

$$
\begin{tikzcd}
 & c \arrow[dl, "\psi_i"'] \arrow[dr, "\psi_{f(i)}"] & \\
Di \arrow[rr, "Df"'] & & Df(i)
\end{tikzcd}
$$

(a) Component of cone over D.

$$
\begin{tikzcd}
a \arrow[rr, "f"] \arrow[dr, "\psi_i"'] & & b \arrow[dl, "\phi_i"] \\
 & Di &
\end{tikzcd}
$$

(b) Morphism of cones.

Figure 1.3: Category of cones.

Alternative Definition 1.4.29. The above definition can be reformulated by defining an additional functor $\Delta_x : \mathbf{I} \to \mathbf{C}$ that maps every element $i \in \mathrm{ob}(\mathbf{I})$ to x and every morphism $g \in \hom(\mathbf{I})$ to $\mathbb{1}_x$, i.e. $\Delta : C \to [\mathbf{I}, \mathbf{C}]$ is the **diagonal functor**. The morphisms ψ_i can then be seen to be the components of a natural transformation $\psi : \Delta_x \Rightarrow D$. Hence, a cone (x, ψ) is an element of $[\mathbf{I}, \mathbf{C}](\Delta_x, D)$.

Definition 1.4.30 (Morphism of cones). Let $D : \mathbf{I} \to \mathbf{C}$ be a diagram and let (x, ψ) and (y, ϕ) be two cones over D. A morphism between these cones is a morphism of the apexes

$f : x \to y$ such that the diagrams of the form 1.3b commute for all $i \in \mathrm{ob}(\mathbf{I})$. The cones over D together with these morphisms form a category $\mathbf{Cone}(D)$. In fact this can easily be seen to be the comma category $\Delta \downarrow D$.

Definition 1.4.31 (Limit). Consider a diagram $D : \mathbf{I} \to \mathbf{C}$. The limit of this diagram, denoted by $\lim D$, is (if it exists) the terminal object of the category $\mathbf{Cone}(D)$.

Remark. In the older literature the name **projective limit** was sometimes used. The dual notion, a **colimit**, is often called an **inductive limit** in the older literature.

This definition leads to the following universal property:

Universal Property 1.4.32. Let $D : \mathbf{I} \to \mathbf{C}$ be a diagram. For every cone $(x, \psi) \in \mathbf{Cone}(D)$, there exists a unique morphism $f : x \to \lim D$. This defines a bijection

$$[\mathbf{I}, \mathbf{C}](\Delta_x, D) \cong \mathbf{C}(x, \lim D).$$

If all (small) limits exist, the limit functor $\lim : [\mathbf{I}, \mathbf{C}] \to \mathbf{C}$ can be defined. The universal property of limits then implies that it is right adjoint to the constant functor Δ.

For diagrams in **Set** one can use the fully faithfulness of the Yoneda embedding to obtain the following expression:

$$\lim D \cong [\mathbf{I}, \mathbf{Set}](\Delta_*, D). \tag{1.18}$$

Remark 1.4.33. In Section 2.2 on enriched category theory, a generalization (the so-called *weighted limits*) of the above construction will be given that is better suited to the enriched setting and allows to express a wide variety of constructions as (weighted) limits.

Example 1.4.34 (Terminal object). The terminal object 1 is the limit of the empty diagram.

Definition 1.4.35 (Finitely complete category). A category is said to be finitely complete if it has all finite limits. If all (small) limits exist, the category is said to be **complete**. The dual notion for colimits is called **(finite) cocompleteness**.

Example 1.4.36 (Presheaf categories). All presheaf categories are both complete and cocomplete.

Definition 1.4.37 (Continuous functor). A functor that preserves all small limits.

Example 1.4.38 (Hom-functors). In a locally small category every hom-functor is continuous (in fact these functors even preserve limits that are not necessarily small). This implies for example that

$$\mathbf{C}(x, \lim D) \cong \lim \mathbf{C}(x, D). \tag{1.19}$$

In the case where $\mathbf{C}$ is small, one can characterize the Yoneda embedding through a universal property:

Universal Property 1.4.39 (Free cocompletion). The Yoneda embedding $\mathbf{C} \hookrightarrow \widehat{\mathbf{C}}$ turns the presheaf category $\widehat{\mathbf{C}}$ into the **free cocompletion** of $\mathbf{C}$, i.e. there exists an equivalence of categories between the functor category of cocontinuous functors $[\widehat{\mathbf{C}}, \mathbf{D}]_{\mathrm{cont}}$ and the ordinary functor category $[\mathbf{C}, \mathbf{D}]$.

Definition 1.4.40 (Tiny object). An object in a locally small category for which the covariant hom-functor preserves small colimits. This is sometimes called a **small-projective** object since it is in particular projective[9].

[9] Epimorphisms are characterized by a *pushout* (see 1.4.61 further below).

Definition 1.4.41 (Cauchy completion). Let $\mathbf{C}$ be a small category. An important (small and full) subcategory of the free cocompletion of $\mathbf{C}$ is given by the Cauchy completion, i.e. the subcategory of $\widehat{\mathbf{C}}$ on the tiny objects.[10] It can be shown that the free cocompletion of the Cauchy completion coincides with the one on $\mathbf{C}$ (up to equivalence).

A category is said to be **Cauchy-complete** if it is equivalent to its Cauchy completion. It can be shown that a category is Cauchy-complete if and only if it has all small absolute colimits.

Definition 1.4.42 (Filtered category). A category in which every finite diagram admits a cocone. For regular cardinals κ, this notion can be generalized. A category is said to be κ-filtered if every diagram with less than κ arrows admits a cocone. (In this terminology filtered categories are the same as ω-filtered categories.)

Definition 1.4.43 (Directed limit). Consider a diagram $D : \mathbf{I} \to \mathbf{C}$. The limit (resp. colimit) of D is said to be (co)directed (resp. directed) if $\mathbf{I}$ is a downward (resp. upward) directed set **??**.

The following definition is a categorification of the previous one:

Definition 1.4.44 (Filtered limit). Consider a diagram $D : \mathbf{I} \to \mathbf{C}$. The limit (resp. colimit) of D is said to be (co)filtered (resp. filtered) if $\mathbf{I}$ is a cofiltered (resp. filtered) category.

Property 1.4.45. A category has all directed limits if and only if it has all filtered limits. (A dual statement holds for colimits.)

Definition 1.4.46 (Pro-object). A functor $F : \mathbf{I} \to \mathbf{C}$ where $\mathbf{I}$ is a small cofiltered category. The names stems from the fact that one can interpret pro-objects as formal cofiltered (projective) limits.

Definition 1.4.47 (Compact object). An object for which the covariant hom-functor preserves all filtered colimits. These objects are also said to be **finitely presentable**.[11]

Definition 1.4.48 (Product). Let $\mathbf{I}$ be a discrete category. The (co)limit over a diagram $D : \mathbf{I} \to \mathbf{C}$ is called a (co)product in $\mathbf{C}$.

Definition 1.4.49 (Equalizer). Consider a diagram of the form

$$x \overset{f}{\underset{g}{\rightrightarrows}} y.$$

The limit of this diagram is called the equalizer of f and g. It consists of an object e and a morphism $\varepsilon : e \to x$ such that the following **fork** diagram

$$e \xrightarrow{\varepsilon} x \overset{f}{\underset{g}{\rightrightarrows}} y \tag{1.20}$$

is universal with respect to (e, ε). By dualizing one obtains **cofork** diagrams $x \rightrightarrows y \to z$ and their universal versions, the **coequalizers**.

Definition 1.4.50 (Split coequalizer). A cofork diagram

$$x \overset{f}{\underset{g}{\rightrightarrows}} y \xrightarrow{\tau} z$$

together with a section φ of f and a section σ of τ such that $\sigma \circ \tau = g \circ \varphi$.

[10] A generalization in the context of enriched categories is given by the *Karoubi envelope*.

[11] This name derives from the fact that modules are finitely presented if and only if their covariant hom-functor preserves direct limits (i.e. directed colimits in the context of algebra).

Definition 1.4.51 (Regular morphisms). A mono (resp. epi) is said to be regular if it arises as an equalizer (resp. coequalizer) of two parallel morphisms.

Property 1.4.52 (Regular bimorphism). Both monic regular epimorphisms and epic regular monomorphisms are isomorphisms.

Alternative Definition 1.4.53 (Finitely complete category). A category is said to be finitely complete if it has a terminal object and if all binary equalizers and products exist.

Definition 1.4.54 (Span). A span in a category C is a diagram of the form 1.4a. By definition of a diagram, a span in C is equivalent to a functor $S : \mathbf{\Lambda} \to \mathbf{C}$, where $\mathbf{\Lambda}$ is the category with three objects $\{-1, 0, 1\}$ and two morphisms $i : 0 \to -1$ and $j : 0 \to 1$. For this reason $\mathbf{\Lambda}$ is sometimes called the walking or universal span.

$$\begin{tikzcd} & s \arrow[dl, "f"'] \arrow[dr, "g"] & \\ x & & y \end{tikzcd}$$

(a) Span (category theory).

$$\begin{tikzcd} x \arrow[dr, "f"'] & & y \arrow[dl, "g"] \\ & c & \end{tikzcd}$$

(b) Cospan.

Figure 1.4: (Co)span diagrams.

Definition 1.4.55 (Pullback). The pullback or **fibre product** of two morphisms $f : x \to z$ and $g : y \to z$ is defined as the limit of cospan 1.4b. The full diagram characterizing the pullback, which has the form of a square, is sometimes called a **Cartesian square**.

Notation 1.4.56 (Pullback). The pullback of two morphisms $f : x \to z$ and $g : y \to z$ is often denoted by $x \times_z y$. The associated pullback square is sometimes written as in Figure 1.5a.

$$\begin{tikzcd} x \times_z y \arrow[r] \arrow[d] \arrow[dr, phantom, "\text{pb}"] & y \arrow[d, "g"] \\ x \arrow[r, "f"'] & z \end{tikzcd}$$

(a) Pullback square.

$$\begin{tikzcd} z \arrow[r, "g"] \arrow[d, "f"'] \arrow[dr, phantom, "\text{po}"] & y \arrow[d] \\ x \arrow[r] & x \sqcup_z y \end{tikzcd}$$

(b) Pushout square.

Figure 1.5: Pullback and pushout diagrams.

Property 1.4.57 (Product). If a terminal object 1 exists, the pullback $x \times_1 y$ is equal to the product $x \times y$.

Definition 1.4.58 (Kernel pair). Consider a morphism $f : x \to y$. Its kernel pair is defined as the pullback of f along itself.

Definition 1.4.59 (Pushout). The dual notion of a pullback, i.e. the colimit of a span. See Figure 1.5b.

Property 1.4.60. Pullbacks preserve monos and pushouts preserve epis.

Alternative Definition 1.4.61 (Epimorphism). A morphism whose cokernel pair is the identity.

Property 1.4.62 (Span category ♣). Consider a category $\mathbf{C}$ with pullbacks. The category $\mathbf{Span}(\mathbf{C})$ is defined as the category with the same objects as $\mathbf{C}$ but with spans as morphisms. Composition of spans is given by pullbacks. By including morphisms of spans, $\mathbf{Span}(\mathbf{C})$ can be refined to a bicategory.

Definition 1.4.63 (Wedge). Consider a profunctor $F : \mathbf{C} \nrightarrow \mathbf{C}$. A wedge $e : w \rightarrow F$ is an object $w \in \mathrm{ob}(\mathbf{Set})$ together with a collection of morphisms $e_x : w \rightarrow F(x,x)$ indexed by $\mathbf{C}$ such that for every morphism $f : x \rightarrow y$ the following diagram commutes:

$$\begin{tikzcd}
 & w \arrow[dl, "e_x"'] \arrow[dr, "e_y"] & \\
F(x,x) \arrow[dr, "F(\mathbb{1}_x, f)"'] & & F(y,y) \arrow[dl, "F(f, \mathbb{1}_y)"] \\
 & F(x,y) &
\end{tikzcd}$$

As was the case for cones, this can be reformulated in terms of (di)natural transformations. A wedge (w, e) of a profunctor $F : \mathbf{C} \nrightarrow \mathbf{C}$ is a dinatural transformation from the constant profunctor Δ_w to F.

Definition 1.4.64 (End). The end of a profunctor $F : \mathbf{C} \nrightarrow \mathbf{C}$ is defined as the universal wedge of F. The components of the wedge are called the **projection maps** of the end. This stems from the fact that for a discrete category the end coincides with the product $\prod_{x \in \mathrm{ob}(\mathbf{C})} F(x,x)$.

This is equivalent to a definition in terms of equalizers. Consider the two canonical maps

$$\prod_{x \in \mathrm{ob}(\mathbf{C})} \mathbf{C}(x,x) \rightrightarrows \prod_{f : x \rightarrow y} \mathbf{C}(x,y).$$

This diagram can be interpreted as the product of all lower halves of the wedge diagrams above. It is not hard to see that its equalizer (universally) satisfies the wedge condition for all $f \in \hom(\mathbf{C})$.

Notation 1.4.65 (End). The end of a profunctor $F : \mathbf{C} \nrightarrow \mathbf{C}$ is often denoted using an integral sign with subscript:

$$\int_{x \in \mathbf{C}} F(x,x).$$

For the dual construction, called a **coend**, an integral sign with superscript is used.

Example 1.4.66 (Natural transformations). Consider two functors $F, G : \mathbf{A} \rightarrow \mathbf{B}$. The map $(x,y) \mapsto \mathbf{B}(Fx, Gy)$ gives a profunctor $H : \mathbf{A} \nrightarrow \mathbf{A}$. By looking at the wedge condition for this profunctor, the following equality for all morphisms $f : x \rightarrow y$ can be derived:

$$\tau_y \circ Ff = Gf \circ \tau_x, \tag{1.21}$$

where τ is the wedge projection. Comparing this equality to Definition 1.2.14 gives

$$\mathrm{Nat}(F,G) = \int_{x \in \mathbf{A}} \mathbf{B}(Fx, Gx). \tag{1.22}$$

Property 1.4.67. Using the continuity 1.4.37 of the hom-functor, one can prove the following equality which can be used to turn ends into coends and vice versa:

$$\mathbf{Set}\left(\int^{x \in \mathbf{C}} F(x,x), y\right) = \int_{x \in \mathbf{C}} \mathbf{Set}(F(x,x), y). \tag{1.23}$$

Using the above properties and definitions, one obtains the following two statements, called the **Yoneda reduction** and **co-Yoneda lemma**:

Property 1.4.68 (Ninja Yoneda lemma). Let $F : \mathbf{A} \to \mathbf{B}$ be a covariant functor (similar statements hold for contravariant functors).

$$\int_{x\in\mathbf{A}} \mathbf{Set}(\mathbf{A}(-,x), Fx) \cong F \tag{1.24}$$

$$\int^{x\in\mathbf{A}} \mathbf{A}(x,-) \times Fx \cong F. \tag{1.25}$$

For a generalization to the enriched setting see Definition 2.2.16.

Remark 1.4.69. A common remark at this point is the comparison with the Dirac distribution (**??**):

$$\int \delta(x-y) f(x) = f(y). \tag{1.26}$$

By interpreting the functor F as a function, the representable functors can be seen to behave as Dirac distributions.

Property 1.4.70.

$$\int_{F\in\mathbf{coPsh}(\mathbf{C})} \mathbf{Set}(Fx, Fy) \cong \mathbf{C}(x,y) \tag{1.27}$$

Definition 1.4.71 (Kan extension). Consider two functors $F : \mathbf{A} \to \mathbf{B}$ and $G : \mathbf{A} \to \mathbf{C}$. The right Kan extension of F along G is given by the universal functor $\mathrm{Ran}_G F : \mathbf{C} \to \mathbf{B}$ and natural transformation $\eta : \mathrm{Ran}_G F \circ G \Rightarrow F$:

$$\begin{tikzcd}
\mathbf{C} \arrow[dr, dashed, "\mathrm{Ran}_G F"] & \\
\mathbf{A} \arrow[u, "G"] \arrow[r, "F"'] \arrow[ur, phantom, "\Downarrow\eta"] & \mathbf{B}
\end{tikzcd}$$

The left Kan extension $\mathrm{Lan}_G F$ is obtained by dualizing this construction.

Property 1.4.72 (Complete categories). Complete (resp. cocomplete) categories admit all right (resp. left) Kan extensions.

Definition 1.4.73 (Preservation of Kan extension). A Kan extension $\mathrm{Lan}_G F$ is said to be **absolute** if every functor with the same codomain as preserves the Kan extension, i.e. a Kan extension is absolute if right whiskering it by another functor defines the Kan extension of the composition. If it is only preserved by all representable functors, the Kan extension is said to be **pointwise**.

Alternative Definition 1.4.74 (Kan extension). The construction above gives a functor Ran_G from the functor category $[\mathbf{A},\mathbf{B}]$ to the functor category $[\mathbf{C},\mathbf{B}]$. The right Kan extension Ran_G can be defined as the right adjoint to the pullback functor $G^* : F \mapsto F \circ G$. Similarly, the left Kan extension can be defined as the left adjoint to the pullback functor.

In the spirit of partial adjoints or partial limits, this definition can be used to define **local Kan extensions**. Although the left (or right) Kan extension functors do not have to exists globally, the extension of a single functor could still exist. This local version is defined by the following natural isomorphism (here given for a left extension):

$$[\mathbf{A},\mathbf{B}](F, G^*-) \cong [\mathbf{C},\mathbf{B}](\mathrm{Lan}_G F, -). \tag{1.28}$$

Remark 1.4.75. Using this equivalence of hom-spaces, Kan extensions can be generalized from **Cat** to any *2-category*.

Example 1.4.76 (Limit). Denote the terminal category by **1**. By choosing the functor G in the definition of a right Kan extension to be the unique functor $!_{\mathbf{C}} : \mathbf{C} \to \mathbf{1}$, one obtain the universal property characterizing limits 1.4.32:

$$\lim F \cong \mathrm{Ran}_{!_C} F. \tag{1.29}$$

Similarly, colimits can be obtained as left Kan extensions.

The existence of Kan extensions can also be used to determine the existence of adjoints:

Property 1.4.77 (Adjoint functors). A functor $F : \mathbf{A} \to \mathbf{B}$ admits a left (resp. right) adjoint if and only if the right (resp. left) Kan extension of the identity functor $\mathbb{1} : \mathbf{A} \to \mathbf{A}$ along F exists. If it exists as an absolute extension, the left adjoint is given exactly by this Kan extension.

Definition 1.4.78 (Codensity monad). Consider a general functor $F : \mathbf{A} \to \mathbf{B}$. If the right Kan extension $\mathrm{Ran}_F F$ exists, it defines a monad. Functors for which this monad is the identity are said to be **codense.**[12] Left Kan extensions give, by duality, rise to *density comonads*.

1.5 Internal structures

Property 1.5.1 (Eckmann-Hilton argument). A monoid internal to **Mon**, the category of monoids, is the same as a commutative monoid. (See also Property **??**.)

Definition 1.5.2 (Internal category). Let $\mathcal{E}$ be a category with pullbacks. A category $\mathbf{C}$ internal to $\mathcal{E}$ consists of the following data:

- an object $C_0 \in \mathrm{ob}(\mathcal{E})$ of objects;
- an object $C_1 \in \mathrm{ob}(\mathcal{E})$ of morphisms;
- source and target morphisms $s, t \in \mathcal{E}(C_1, C_0)$;
- an "identity-assigning" morphism $e \in \mathcal{E}(C_0, C_1)$ such that

$$s \circ e = \mathbb{1}_{C_0} \qquad\qquad t \circ e = \mathbb{1}_{C_0};$$

  and

- a composition morphism $c : C_1 \times_{C_0} C_1 \to C_1$ such that the following equations hold:

$$\begin{gathered} s \circ c = s \circ \pi_1 \qquad\qquad t \circ c = t \circ \pi_2 \\ \pi_1 = c \circ (e \times_{C_0} \mathbb{1}) \qquad\qquad c \circ (\mathbb{1} \times_{C_0} e) = \pi_2 \\ c \circ (c \times_{C_0} \mathbb{1}) = c \circ (\mathbb{1} \times_{C_0} c), \end{gathered}$$

  where π_1, π_2 are the canonical projections associated with the pullback $C_1 \times_{C_0} C_1$ of (s, t).

Morphisms between these categories, suitably called **internal functors**, are given by a pair of morphisms (in $\mathcal{E}$) between internal objects and morphisms, that preserve composition and identities. Internal natural transformations are defined in a similar way.

Notation 1.5.3. The *(bi)category* of internal categories in $\mathcal{E}$ is denoted by $\mathbf{Cat}(\mathcal{E})$. It should be noted that for $\mathcal{E} = \mathbf{Set}$, the ordinary category of small categories $\mathbf{Cat}(\mathbf{Set}) = \mathbf{Cat}$ is obtained.

Span gives source and target maps

$$\begin{tikzcd} & C_1 \arrow[dl, "s"'] \arrow[dr, "t"] & \\ C_0 & & C_0 \end{tikzcd}$$

Multiplication gives composition

$$\begin{tikzcd} & C_1 \times_{C_0} C_1 \arrow[dl, "s \circ \pi_0"'] \arrow[dr, "t \circ \pi_1"] \arrow[dd, "c"] & \\ C_0 & & C_0 \\ & C_1 \arrow[ul, "s"] \arrow[ur, "t"'] & \end{tikzcd}$$

Unit gives identity

$$\begin{tikzcd} & C_0 \arrow[dl, "\mathbb{1}_a"'] \arrow[dr, "\mathbb{1}_a"] \arrow[dd, "e"] & \\ C_0 & & C_0 \\ & C_1 \arrow[ul, "s"] \arrow[ur, "t"'] & \end{tikzcd}$$

Figure 1.6: Internal category as a monad in $\mathbf{Span}(\mathcal{E})$.

Alternative Definition 1.5.4. The above definition can be reformulated in a very elegant way. An internal category in $\mathcal{E}$ is a monad in the bicategory $\mathbf{Span}(\mathcal{E})$ of spans in $\mathcal{E}$ as shown in Figure 1.6.

Functors between internal categories are not the only relevant morphisms. However, when defining (co)presheafs such as the hom-functor, a problem occurs. In **Cat** there exist, by definition, maps to the ambient category **Set** (ordinary category theory has a set-theoretic foundation). However, for internal categories there does not necessarily exist a morphism $\mathbf{C} \to \mathcal{E}$. To solve this problem one can consider a more general structure:

Definition 1.5.5 (Internal diagram). A left module over a monad in $\mathbf{Span}(\mathcal{E})$. The dual notion is better known as an **internal presheaf**.

In fact, this is a specific instance of an even more general concept (for more information on the definitions and applications see [15, 16]):

Definition 1.5.6 (Internal profunctor). A bimodule between monads in $\mathbf{Span}(\mathcal{E})$. Together with the above definitions this gives rise to an equivalence $\mathbf{Mod}(\mathbf{Span}(\mathcal{E})) \cong \mathbf{Prof}(\mathcal{E})$.

Construction 1.5.7 (Internal Yoneda profunctor). Consider an internal functor $F : \mathbf{A} \to \mathbf{B}$. This functor induces two internal profunctors $F_* : \mathbf{B} \nrightarrow \mathbf{A}$ and $F^* : \mathbf{A} \nrightarrow \mathbf{B}$. For F_* (the profunctor F^* is defined similarly) the object span is defined as

$$A_0 \xleftarrow{\pi_0} A_0 \times_{B_0} B_1 \xrightarrow{t \circ \pi_1} B_0.$$

The action of $f \in B_1$ is given by postcomposition with f in the second factor, while the action of $g \in A_1$ is given by precomposition with Fg in the second factor and changing to the domain of g in the first factor.

It can easily be shown that the profunctors induced by an identity functor $\mathbb{1}_{\mathbf{C}}$ have an object span that corresponds to the internal category $\mathbf{C}$ with the actions given by (internal) composition. In the case of $\mathcal{E} = \mathbf{Set}$ this boils down to the hom-functor. The fact that the object span is equivalent to the category $\mathbf{C}$ is essentially the Yoneda embedding. For this reason this profunctor is in general called the (internal) Yoneda profunctor $\mathcal{Y}(\mathbf{C})$.

[12]Codense functors are usually defined in a different way, but one can show that this is an equivalent definition (hence the name).

1.6 Higher category theory ♣

1.6.1 n-categories

Definition 1.6.1 (n-category). A (strict) n-category consists of:

- objects (0-morphisms),
- 1-morphisms going between 0-morphisms,
- ...
- n-morphisms going between $(n-1)$-morphisms,

such that the composition of k-morphisms ($k \leq n$) is associative and satisfies the unit laws as required in an ordinary category. By generalizing this definition to arbitrary n one can define the notion of a (strict) ∞-category.

If one relaxes the associativity and unit laws up to higher coherent morphisms, one obtains the notion a weak n-category. Explicit definitions for such categories have been constructed up to tetracategories ($n = 4$). However, this construction by *Trimble* takes about 50 pages of diagrams.

Remark. n-morphisms are also called n**-cells**. This makes their relation to topological spaces (and in particular simplicial spaces) more visible.

Example 1.6.2. The classical examples of a 1-category and 2-category are **Set** and **Cat**, respectively.

Property 1.6.3 (Composition in 2-categories). 2-morphisms can be composed in two different ways:

- **Horizontal composition**: Consider two 2-morphisms $\alpha : f \Rightarrow g$ and $\beta : f' \Rightarrow g'$ where $f' \circ f$ and $g' \circ g$ are well-defined. These 2-morphisms can be composed as
$$\beta \circ \alpha : f' \circ f \Rightarrow g' \circ g.$$
This is sometimes called their **Godement product**.
- **Vertical composition**: Consider two 2-morphisms $\alpha : f \Rightarrow g$ and $\beta : g \Rightarrow h$ where f, g and h have the same domain and codomain. These 2-morphisms can be composed as
$$\beta \cdot \alpha : f \Rightarrow h.$$

As a consistency condition the horizontal and vertical composition are required to satisfy the following **interchange law**:

$$(\alpha \cdot \beta) \circ (\gamma \cdot \delta) = (\alpha \circ \gamma) \cdot (\beta \circ \delta). \tag{1.30}$$

Definition 1.6.4 ((n, r)-category). A higher (∞-)category for which

- all parallel k-morphisms with $k > n$ are equivalent and, hence, trivial.
- all k-morphisms with $k > r$ are invertible (or equivalences in the fully weak ∞-sense).

Definition 1.6.5 (Weak inverse). Let **C** be a 2-category. A 1-morphism $f : x \to y$ is weakly invertible if there exist a 1-morphism $g : y \to x$ and 2-isomorphisms $g \circ f \Rightarrow \mathbb{1}_x$ and $f \circ g \Rightarrow \mathbb{1}_y$.

At this point it should be obvious that the definition of a unit-counit adjunction 1.2.26 can be generalized to general 2-categories:

Definition 1.6.6 (Adjunction in 2-category). Let $\mathbf{C}$ be a 2-category. An adjunction in $\mathbf{C}$ is a pair of 1-morphisms $F : x \to y$ and $G : y \to x$ together with 2-morphisms $\varepsilon : F \circ G \Rightarrow \mathbb{1}_y$ and $\eta : \mathbb{1}_x \Rightarrow G \circ F$ that satisfy the zig-zag identities.

Remark 1.6.7 (Duals and adjunctions). By looking at the defining relations of duals in a rigid monoidal category (see Section 2.4), it should be clear that these are in fact the same as the defining relations of the unit and counit of an adjunction. This is a consequence of the fact that a 2-category with a single object can be regarded as a (strict) monoidal category where the composition in the 2-category becomes the tensor product in the monoidal category. Similarly, adjoint 1-morphisms in the 2-category become duals in the monoidal category.

Property 1.6.8 (Monoidal categories). Consider a monoidal category $(\mathbf{C}, \otimes, \mathbf{1})$. From this monoidal category one can construct the so-called **delooping** bicategory $\mathbf{BC}$ in the following way:

- There is a single object $*$.
- The 1-morphisms in $\mathbf{BC}$ are the objects in $\mathbf{C}$.
- The 2-morphisms in $\mathbf{BC}$ are the morphisms in $\mathbf{C}$.
- Horizontal composition in $\mathbf{BC}$ is the tensor product in $\mathbf{C}$.
- Vertical composition in $\mathbf{BC}$ is composition in $\mathbf{C}$.

Conversely, every 2-category with a single object comes from a monoidal category. Hence, the 2-category of (pointed) 2-categories with a single object and the 2-category of monoidal categories are equivalent. (This property and its generalizations are the content of the *delooping hypothesis*.)

In the same way one can deloop a braided monoidal category twice and find an identification with a one-object tricategory with one 1-morphism. However, this identification is not a trivial one as it makes use of the Eckmann-Hilton argument to identify different monoidal structures on this tricategory. (See also Section 2.10.)

1.6.2 n-functors

Definition 1.6.9 (2-functor). A 2-functor $F : \mathbf{A} \to \mathbf{B}$ (often called a **pseudofunctor**) is a morphism between bicategories. It consists of the following data:

- a function $F_0 : \mathrm{ob}(\mathbf{A}) \to \mathrm{ob}(\mathbf{B})$, and
- for every two objects $x, y \in \mathrm{ob}(\mathbf{A})$, a functor $F_{x,y} : \mathbf{A}(x, y) \to \mathbf{B}(Fx, Fy)$.

The function F_0 and the functors $F_{x,y}$ are also often denoted by F by abuse of notation. This data is required to satisfy some coherence conditions. These are specified by the following data:

1. **Associator**: For every pair of composable 1-morphisms $f \circ g$ in $\hom(\mathbf{A})$, a 2-isomorphism $\gamma_{f,g} : Ff \circ Fg \Rightarrow F(f \circ g)$ such that for every triple of composable morphisms $f \circ g \circ h$ in $\hom(\mathbf{A})$ the following identity holds:
$$\gamma_{f\circ g,h} \circ (\gamma_{f,g} \cdot \mathbb{1}_{Fh}) = \gamma_{f,g\circ h} \circ (\mathbb{1}_{Ff} \cdot \gamma_{g,h}). \tag{1.31}$$

2. **Unitor**: For every object $x \in \mathrm{ob}(\mathbf{A})$, a 2-isomorphism $\iota_x : \mathbb{1}_{Fx} \Rightarrow F\mathbb{1}_x$ such that for every morphism $f : x \to y$ in $\hom(\mathbf{A})$ the following identities hold:
$$\iota_y \cdot \mathbb{1}_{Ff} = \gamma_{\mathbb{1}_y,f} \tag{1.32}$$
$$\mathbb{1}_{Ff} \cdot \iota_x = \gamma_{f,\mathbb{1}_x}. \tag{1.33}$$

Note that to be completely formal one should have inserted the unitors and associators of the bicategories $\mathbf{A}, \mathbf{B}$.

Definition 1.6.10 (Lax natural transformation). Consider two 2-functors $F, G : \mathbf{A} \to \mathbf{B}$ between bicategories. A lax natural transformation $\eta : F \Rightarrow G$ consists of the following data:

1. for every object $x \in \mathrm{ob}(\mathbf{A})$, a 1-morphism $\eta_x : Fx \to Gx$, and
2. for every 1-morphism $f : x \to y$ in $\hom(\mathbf{A})$, a 2-morphism $\eta_f : Gf \circ \eta_x \Rightarrow \eta_y \circ Ff$ such that the η_f are the components of a natural transformation $(\eta_x)^* \circ G \Rightarrow (\eta_y)_* \circ F$ and such that the assignment $f \mapsto \eta_f$ satisfies the "obvious" identity and composition axioms.

Remark 1.6.11. As usual in the context of higher category theory one can speak of lax 2-functors if the associator and unitors are merely required to be 2-morphisms and of strict 2-functors if these morphisms are required to be identities. If the natural transformations between morphism categories in the definition of a lax natural transformation are all isomorphisms, this is called a **pseudonatural transformation**. If the 1-morphisms η are equivalences, they are called lax natural equivalences.

Definition 1.6.12 (Modification). Consider two bicategories $\mathbf{A}, \mathbf{B}$, two 2-functors $F, G : \mathbf{A} \to \mathbf{B}$ and two parallel (lax) natural transformations $\alpha, \beta : F \Rightarrow G$. A modification $\mathfrak{m} : \alpha \Rrightarrow \beta$ maps every object $x \in \mathrm{ob}(\mathbf{A})$ to a 2-morphism $\mathfrak{m}_x : \alpha_x \Rightarrow \beta_x$ such that $\beta_f \circ (\mathbb{1}_{Gf} \cdot \mathfrak{m}_x) = (\mathfrak{m}_y \cdot \mathbb{1}_{Ff}) \circ \alpha_f$.

This is generalized as follows:

Definition 1.6.13 (Transfor). A k-transfor[13] between two n-categories maps j-morphisms to $(j+k)$-morphisms (in a coherent way).

Example 1.6.14. The definitions for operations in bicategories above lead us to the following "explicit" expressions for k-transfors (for small k):

- $k = 0$: n-functors,
- $k = 1$: (n-)natural transformations,
- $k = 2$: modifications, and
- $k = 3$: **perturbations**.

The following definition generalizes the notion of essential surjectivity 1.2.11 to higher category theory:

Definition 1.6.15 (n-surjective functor). An ∞-functor $F : \mathbf{A} \to \mathbf{B}$ is said to be n-surjective if for any two parallel $(n-1)$-morphisms f, g in $\mathbf{A}$ and n-morphism $\alpha : Ff \to Fg$ in $\mathbf{B}$, there exists an n-morphism $\widetilde{\alpha}$ in $\mathbf{A}$ such that $F\widetilde{\alpha} \cong \alpha$.

Definition 1.6.16 (Indexed category). Consider a category $\mathbf{I}$. An $\mathbf{I}$-indexed category is a pseudofunctor $\mathbf{C} : \mathbf{I}^{op} \to \mathbf{Cat}$, i.e. a 2-presheaf on $\mathbf{I}$. Indexed functors and natural transformations are defined analogously.

[13] This name was first introduced by *Crans* in [85]. A different name that is sometimes used is (n,k)-**transformation**, but this should not be confused with the natural transformations in the context of (n,r)-categories.

1.6.3 Higher (co)limits

Definition 1.6.17 (Weighted 2-limit). Consider 2-categories $\mathbf{I}, \mathbf{C}$ together with 2-functors $W : \mathbf{I} \to \mathbf{Cat}$ and $F : \mathbf{I} \to \mathbf{C}$. By direct generalization of the ordinary definition of weighted limits, one says that $\lim^W F$ is the W-weighted (2-)limit of F if there exists a pseudonatural equivalence

$$\mathbf{C}(x, \lim^W F) \cong [\mathbf{I}, \mathbf{Cat}](W, \mathbf{C}(x, F-)). \tag{1.34}$$

By restricting to the 2-category of strict 2-categories, strict 2-functors and strict natural transformations the resulting notion of a weighted 2-limit coincides with that of an ordinary weighted limit enriched in **Cat** (since strict 2-categories are simply **Cat**-enriched 1-categories.)

?? COMPLETE ??

1.6.4 Groupoids

Definition 1.6.18 (Groupoid). A (small) groupoid $\mathcal{G}$ is a (small) category in which all morphisms are invertible.

Example 1.6.19 (Delooping). Consider a group G. Its delooping $\mathbf{B}G$ is defined as the one-object groupoid for which $\mathbf{B}G(*, *) = G$.

Property 1.6.20 (Representations). Consider a group G together with its delooping $\mathbf{B}G$. When considering *representations* as functors $\rho : \mathbf{B}G \to \mathbf{FinVect}$, one can see that the intertwiners **??** are exactly the natural transformations. More generally, all G-sets **??** can be obtained as functors $\mathbf{B}G \to \mathbf{Set}$.

Definition 1.6.21 (Core). Let $\mathbf{C}$ be a (small) category. The core $\mathrm{Core}(\mathbf{C}) \in \mathbf{Grpd}$ of $\mathbf{C}$ is defined as the maximal subgroupoid of $\mathbf{C}$.

Definition 1.6.22 (Orbit). Let $\mathcal{G}$ be a groupoid with O, M respectively the sets of objects and morphisms. On O one can define an equivalence $x \sim y \iff \exists \phi : x \to y$. The equivalence classes are called orbits and the set of orbits is denoted by O/M.

Definition 1.6.23 (Transitive component). Let $\mathcal{G}$ be a groupoid with O, M respectively the sets of objects and morphisms and let s, t denote the source and target maps on M. Given an orbit $o \in O/M$, the transitive component of M associated to o is defined as $s^{-1}(o)$, or equivalently, as $t^{-1}(o)$.

Property 1.6.24. Every groupoid is a (disjoint) union of its transitive components.

Definition 1.6.25 (Transitive groupoid). A groupoid $\mathcal{G}$ is said to be transitive if for all objects $x \neq y \in \mathrm{ob}(\mathcal{G})$, the set $\mathcal{G}(x, y)$ is not empty.

1.7 Lawvere theories ♣

Definition 1.7.1 (Lawvere theory). Let $\mathbf{F}$ denote the skeleton of **FinSet**. A Lawvere theory consists of a small category $\mathbf{L}$ and a strict (finite) product-preserving *identity-on-objects* functor $\mathcal{L} : \mathbf{F}^{op} \to \mathbf{L}$.

Equivalently, a Lawvere theory is a small category $\mathbf{L}$ with a **generic object** c_0 such that every object $c \in \mathrm{ob}(\mathbf{L})$ is a finite power of c_0.

Property 1.7.2. Lawvere theories $(\mathbf{L}, \mathcal{L})$ form a category **Law**. Morphisms between Lawvere theories are (finite) product-preserving functors.

Definition 1.7.3 (Model). A model or **algebra** over a Lawvere theory $\mathbf{L}$ is a (finite) product-preserving functor $A : \mathbf{L} \to \mathbf{Set}$.

?? COMPLETE ??

1.8 Operad theory ♣

1.8.1 Operads

Definition 1.8.1 (Plain operad[14]). Let $\mathcal{O} = \{P(n)\}_{n\in\mathbb{N}}$ be a collection of sets, called n**-ary operations** (where n is called the **arity**). The collection $\mathcal{O}$ is called a plain operad if it satisfies following axioms:

1. $P(1)$ contains an identity element $\mathbb{1}$.
2. For all positive integers $n, k_1, \ldots, k_n$ there exists a composition map

$$\begin{aligned} \circ : P(n) \times P(k_1) \times \cdots \times P(k_n) &\to P(k_1 + \cdots + k_n) \\ : (\psi, \theta_1, \ldots, \theta_n) &\mapsto \psi \circ (\theta_1, \ldots, \theta_n) \end{aligned} \tag{1.35}$$

   that satisfies two additional axioms:

   - **identity**:

$$\theta \circ (\mathbb{1}, \ldots, \mathbb{1}) = \mathbb{1} \circ \theta = \theta, \tag{1.36}$$

     and
   - **associativity**:

$$\begin{aligned} &\psi \circ \Big(\theta_1 \circ (\theta_{1,1}, \ldots, \theta_{1,k_1}), \ldots, \theta_n \circ (\theta_{n,1}, \ldots, \theta_{n,k_n})\Big) \\ &\qquad = \Big(\psi \circ (\theta_1, \ldots, \theta_n)\Big) \circ (\theta_{1,1}, \ldots, \theta_{1,k_1}, \theta_{2,1}, \ldots, \theta_{n,k_n}). \end{aligned} \tag{1.37}$$

If the operad is represented using planar tree diagrams, the associativity obtains a nice intuitive form. When combining planar tree diagrams in three layers, the associativity axiom says that one can either first glue the first two layers together or one can first glue the last two layers together.

Remark 1.8.2. Plain operads can be defined in any monoidal category. In the same way symmetric operad can be defined in any symmetric monoidal category.

Example 1.8.3 (Endomorphism operad). Consider a vector space V. For every $n \in \mathbb{N}$, one can define the endomorphism algebra $\mathrm{End}(V^{\otimes n}, V)$. The endomorphism operad $\mathcal{E}\mathrm{nd}(V)$ is defined as $\{\mathrm{End}(V^{\otimes n}, V)\}_{n\in\mathbb{N}}$.

Definition 1.8.4 (O-algebra). An object X is called an algebra over an operad O if there exist morphisms

$$O(n) \times X^n \to X$$

for every $n \in \mathbb{N}$ satisfying the usual composition and identity laws. Alternatively, this can be rephrased as the existence of a (plain) operad morphism $O(n) \to \mathcal{E}\mathrm{nd}(X)$.

Example 1.8.5 (Categorical O-algebra). An O-algebra in the category **Cat**.

[14] Also called a **nonsymmetric operad** or **non-Σ operad**.

1.8.2 Algebraic topology

Definition 1.8.6 (Stasheff operad). A topological operad $\mathcal{K}$ such that $\mathcal{K}(n)$ is given by the n^{th} *Stasheff polytope/associahedron*. Composition is given by the inclusion of faces.

Definition 1.8.7 (A_∞-space). An algebra over the Stasheff operad. This induces the structure of a multiplication that is associative up to a coherent homotopy.

Definition 1.8.8 (Little k-cubes operad). A topological operad for which every topological space $\mathcal{P}(n)$ consists of all possible configurations of n embedded k-cubes in a (unit) k-cube. Composition is given by the obvious way of inserting one unit k-cube in one of the smaller embedded k-cubes.

Property 1.8.9 (Recognition principle). If a connected topological space X forms an algebra over the little k-cubes operad, it is (weakly) homotopy equivalent to the k-fold loop space $\Omega^k Y$ of another pointed topological space Y. For $k = 1$, one should technically use the Stasheff operad, but it can be shown that this is related to the little interval operad.

?? COMPLETE ??

Chapter 2

Higher-dimensional Algebra ♣

The main reference for this chapter is the series of papers carrying the same name by *Baez et al.* [5, 68]. References for the section on Berezin calculus are [43, 100]. For Kapranov-Voevodsky 2-vector spaces the reader is referred to the original paper [31]. The section about higher Lie theory is mainly based on [93]. For fusion and modular categories the main reference is [33].

2.1 Monoidal categories

Definition 2.1.1 (Monoidal category). A category $\mathbf{C}$ equipped with a bifunctor

$$- \otimes - : \mathbf{C} \times \mathbf{C} \to \mathbf{C}$$

called the **tensor product** or **monoidal product**, a distinct object $\mathbf{1}$ called the **(monoidal) unit**, and the following three natural isomorphisms called the **coherence maps**:

- **Associator**: $\alpha_{x,y,z} : (x \otimes y) \otimes z \cong x \otimes (y \otimes z)$;
- **Left unitor**: $\lambda_x : \mathbf{1} \otimes x \cong x$; and
- **Right unitor**: $\rho_x : x \otimes \mathbf{1} \cong x$.

These natural transformations are required make the **triangle** and **pentagon** diagrams 2.1 and 2.2 commute.

$$\begin{tikzcd}
(x \otimes \mathbf{1}) \otimes y \arrow[rr, "\alpha_{x,\mathbf{1},y}"] \arrow[dr, "\rho_x \otimes \mathbb{1}_y"'] & & x \otimes (\mathbf{1} \otimes y) \arrow[dl, "\mathbb{1}_x \otimes \lambda_y"] \\
& x \otimes y &
\end{tikzcd}$$

Figure 2.1: Triangle diagram.

A monoidal category for which the associator and the unitors are identity transformations is often said to be **strict**.

Example 2.1.2 (Cartesian category). A monoidal category where the monoidal product is given by the ordinary product 1.4.48. If the monoidal product is not the ordinary product, but the monoidal unit is still terminal, the category is said to be **semicartesian**.

Definition 2.1.3 (Scalar). In a monoidal category the scalars are defined as the endomorphisms $\mathbf{1} \to \mathbf{1}$. The set of scalars forms a commutative monoid.

$$\begin{tikzcd}
 & ((w \otimes x) \otimes y) \otimes z \arrow[rr, "\alpha_{w,x,y} \otimes \mathbb{1}_z"] \arrow[dl, "\alpha_{w \otimes x, y, z}"'] & & (w \otimes (x \otimes y)) \otimes z \arrow[dr, "\alpha_{w, x \otimes y, z}"] & \\
(w \otimes x) \otimes (y \otimes z) \arrow[drr, "\alpha_{w,x,y \otimes z}"'] & & & & w \otimes ((x \otimes y) \otimes z) \arrow[dll, "\mathbb{1}_w \otimes \alpha_{x,y,z}"] \\
 & & w \otimes (x \otimes (y \otimes z)) & &
\end{tikzcd}$$

Figure 2.2: Pentagon diagram.

Property 2.1.4. Every scalar $s : \mathbf{1} \to \mathbf{1}$ induces a natural transformation $s : \mathbb{1}_{\mathbf{C}} \Rightarrow \mathbb{1}_{\mathbf{C}}$ with components

$$s_x : x \cong \mathbf{1} \otimes x \xrightarrow{s \otimes \mathbb{1}_x} \mathbf{1} \otimes x \cong x.$$

For every morphism $f \in \hom(\mathbf{C})$, the naturality square $f \circ s_x = s_y \circ f$ alo defines a morphism $s \diamond f$ that is equivalently given by $\rho_y \circ (f \otimes s) \circ \rho_x^{-1}$ (one could have used the left unitors as well). These morphisms satisfy the following well-known rules of scalar multiplication from linear algebra:

- $s \diamond (s' \diamond f) = (s \circ s') \diamond f$,
- $(s \diamond f) \circ (s' \diamond g) = (s \circ s') \diamond (f \circ g)$, and
- $(s \diamond f) \otimes (s' \diamond g) = (s \circ s') \diamond (f \otimes g)$.

Definition 2.1.5 (Weak inverse). Let $(\mathbf{C}, \otimes, \mathbf{1})$ be a monoidal category and consider an object $x \in \mathrm{ob}(\mathbf{C})$. An object $y \in \mathrm{ob}(\mathbf{C})$ is called a weak inverse of x if it satisfies $x \otimes y \cong \mathbf{1}$.

Remark 2.1.6. One can show that the existence of a one-sided weak inverse (as in the definition above) is sufficient to prove that it is in fact a two-sided weak inverse, i.e. $y \otimes x \cong \mathbf{1}$ also holds.

Theorem 2.1.7 (MacLane's coherence theorem). *Consider two functors $F, G : \mathbf{A} \to \mathbf{B}$ between two monoidal categories $\mathbf{A}, \mathbf{B}$. Any two natural transformations $\eta, \varepsilon : F \Rightarrow G$, constructed solely from the associator and the unitors, coincide.*

2.1.1 Braided categories

Definition 2.1.8 (Braided monoidal category). A monoidal category $(\mathbf{C}, \otimes, \mathbf{1})$ equipped with a natural isomorphism

$$\sigma_{x,y} : x \otimes y \cong y \otimes x$$

that makes the two **hexagon** diagrams 2.3a and 2.3b commute for all $x, y, z \in \mathrm{ob}(\mathbf{C})$. The isomorphism σ is called the **braiding** (morphism).

Property 2.1.9 (Yang-Baxter equation). The components $\sigma_{x,x}$ of a braiding satisfy the *Yang-Baxter* equation. More generally, the braiding σ satisfies the following equation for all objects $x, y, z \in \mathrm{ob}(\mathbf{C})$:

$$(\sigma_{y,z} \otimes \mathbb{1}_x) \circ (\mathbb{1}_y \otimes \sigma_{x,z}) \circ (\sigma_{x,y} \otimes \mathbb{1}_z) = (\mathbb{1}_z \otimes \sigma_{x,y}) \circ (\sigma_{x,z} \otimes \mathbb{1}_y) \circ (\mathbb{1}_x \otimes \sigma_{y,z}). \tag{2.1}$$

$$\begin{tikzcd}
& (x\otimes y)\otimes z \arrow[dl, "\sigma_{x,y}\otimes\mathbb{1}_z"'] \arrow[dr, "\alpha_{x,y,z}"] & \\
(y\otimes x)\otimes z \arrow[d, "\alpha_{y,x,z}"'] & & x\otimes(y\otimes z) \arrow[d, "\sigma_{x,y\otimes z}"] \\
y\otimes(x\otimes z) \arrow[dr, "\mathbb{1}_y\otimes\sigma_{x,z}"'] & & (y\otimes z)\otimes x \arrow[dl, "\alpha_{y,z,x}"] \\
& y\otimes(z\otimes x) &
\end{tikzcd}$$

(a) Hexagon diagram 1.

$$\begin{tikzcd}
& x\otimes(y\otimes z) \arrow[dl, "\mathbb{1}_x\otimes\sigma_{y,z}"'] \arrow[dr, "\alpha^{-1}_{x,y,z}"] & \\
x\otimes(z\otimes y) \arrow[d, "\alpha^{-1}_{x,z,y}"'] & & (x\otimes y)\otimes z \arrow[d, "\sigma_{x\otimes y,z}"] \\
(x\otimes z)\otimes y \arrow[dr, "\sigma_{x,z}\otimes\mathbb{1}_y"'] & & z\otimes(x\otimes y) \arrow[dl, "\alpha^{-1}_{z,x,y}"] \\
& (z\otimes x)\otimes y &
\end{tikzcd}$$

(b) Hexagon diagram 2.

Figure 2.3: Hexagon diagram.

Remark 2.1.10. When drawing the above equality using string diagrams, it can be seen that the Yang-Baxter equation corresponds to the invariance of string diagrams under a *Reidemeister III move.*

Definition 2.1.11 (Symmetric monoidal category). A braided monoidal category where the braiding σ satisfies

$$\sigma_{x,y}\circ\sigma_{y,x} = \mathbb{1}_{x\otimes y}\,. \tag{2.2}$$

In Chapter 2 the theory of monoidal categories is continued.

2.1.2 Monoidal functors

Definition 2.1.12 (Monoidal functor). Let $(\mathbf{A},\otimes,\mathbf{1_A}),(\mathbf{B},\circledast,\mathbf{1_B})$ be two monoidal categories. A functor $F:\mathbf{A}\to\mathbf{B}$ is said to be monoidal if there exists:

1. A natural isomorphism $\psi_{x,y}: Fx\circledast Fy\Rightarrow F(x\otimes y)$ that makes Diagram 2.4 commute.

$$\begin{tikzcd}
(Fx\circledast Fy)\circledast Fz \arrow[r, "\alpha_{\mathbf{B}}"] \arrow[d, "\psi_{x,y}\circledast\mathbb{1}_{Fz}"'] & Fx\circledast(Fy\circledast Fz) \arrow[d, "\mathbb{1}_{Fx}\circledast\psi_{y,z}"] \\
F(x\otimes y)\circledast Fz \arrow[d, "\psi_{x\otimes y,z}"'] & Fx\circledast F(y\otimes z) \arrow[d, "\psi_{ax,y\otimes z}"] \\
F\Big((x\otimes y)\otimes z\Big) \arrow[r, "F\alpha_{\mathbf{A}}"'] & F\Big(x\otimes(y\otimes z)\Big)
\end{tikzcd}$$

Figure 2.4: Monoidal functor.

2. An isomorphism $\phi:\mathbf{1_B}\to F\mathbf{1_A}$ that makes the two diagrams in Figure 2.5 commute.

Remark 2.1.13. The maps ψ and ϕ are also called **coherence maps** or **structure morphisms**.

Property 2.1.14 (Canonical unit). For every monoidal functor F there exists a canonical isomorphism $\phi:\mathbf{1_B}\to F\mathbf{1_A}$ defined by the commutative Diagram 2.6.

Definition 2.1.15 (Lax monoidal functor). A monoidal functor for which the coherence maps are merely morphisms and not isomorphisms.

$$\begin{tikzcd}
Fx \circledast \mathbf{1_B} \arrow[r, "\mathbb{1}_{Fx} \circledast \phi"] \arrow[d, "\rho_{\mathbf{B}}"'] & Fx \circledast F\mathbf{1_A} \arrow[d, "\psi_{x,\mathbf{1_A}}"] \\
Fx & F(x \otimes \mathbf{1_A}) \arrow[l, "F\rho_{\mathbf{A}}"]
\end{tikzcd}
\qquad
\begin{tikzcd}
\mathbf{1_B} \circledast Fy \arrow[r, "\phi \circledast \mathbb{1}_{Fy}"] \arrow[d, "\lambda_{\mathbf{B}}"'] & F\mathbf{1_A} \circledast Fy \arrow[d, "\psi_{\mathbf{1_A},y}"] \\
Fy & F(\mathbf{1_A} \otimes y) \arrow[l, "F\lambda_{\mathbf{A}}"]
\end{tikzcd}$$

Figure 2.5: Unitality diagrams.

$$\begin{tikzcd}
\mathbf{1_B} \circledast F\mathbf{1_A} \arrow[r, "\lambda_{\mathbf{B}}"] \arrow[d, "\phi \circledast \mathbb{1}_{F\mathbf{1_A}}"'] & F\mathbf{1_A} \arrow[d, "F\lambda_{\mathbf{A}}"] \\
F\mathbf{1_A} \circledast F\mathbf{1_A} \arrow[r, "\psi_{\mathbf{1_A},\mathbf{1_A}}"'] & F(\mathbf{1_A} \otimes \mathbf{1_A})
\end{tikzcd}$$

Figure 2.6: Canonical unit isomorphism.

Definition 2.1.16 (Monoidal natural transformation). A natural transformation η between (lax) monoidal functors (F, ψ, ϕ_F) and $(G, \widetilde{\psi}, \phi_G)$ that makes the diagrams in Figure 2.7 commute.

$$\begin{tikzcd}
& \mathbf{1_B} \arrow[dl, "\phi_F"'] \arrow[dr, "\phi_G"] & \\
F\mathbf{1_A} \arrow[rr, "\eta_{\mathbf{1_A}}"'] & & G\mathbf{1_A}
\end{tikzcd}
\qquad
\begin{tikzcd}
Fx \circledast Fy \arrow[r, "\psi_{a,b}"] \arrow[d, "\eta_a \circledast \eta_b"'] & F(x \otimes y) \arrow[d, "\eta_{a\otimes b}"] \\
Gx \circledast Gy \arrow[r, "\widetilde{\psi}_{a,b}"'] & G(x \otimes y)
\end{tikzcd}$$

Figure 2.7: Monoidal natural transformation.

Definition 2.1.17 (Monoidal equivalence). An equivalence of monoidal categories consisting of monoidal functors and monoidal natural isomorphisms.

Theorem 2.1.18 (MacLane's strictness theorem). *Every monoidal category is monoidally equivalent to a strict monoidal category.*

2.1.3 Closed categories

Definition 2.1.19 (Internal hom). Let $(\mathbf{M}, \otimes, \mathbf{1})$ be a monoidal category. In this setting one can generalize the "currying" procedure, i.e. the identification of maps $x \times y \to z$ with maps $x \to (y \to z)$. The internal hom-functor $\underline{\mathrm{Hom}}$ is defined by the following natural isomorphism:

$$\mathrm{Hom}(x \otimes y, z) \cong \mathrm{Hom}(x, \underline{\mathrm{Hom}}(y, z)). \tag{2.3}$$

The existence of all internal homs is equivalent to the existence of a right adjoint to the tensor functor.

Notation 2.1.20. The internal hom $\underline{\mathrm{Hom}}(x, y)$ is also often denoted by $[x, y]$. From now on this convention will be followed (unless otherwise specified).

Definition 2.1.21 (Closed monoidal category). A monoidal category is said to be closed monoidal if it has all internal homs. If the monoidal structure is induced by a (Cartesian) product structure, the category is often said to be **Cartesian closed**.

A category for which all slice categories are Cartesian closed is said to be **locally Cartesian closed**. A locally Cartesian closed category with a terminal object is also Cartesian closed.

Definition 2.1.22 (Exponential object). In the case of Cartesian (monoidal) categories, the internal hom $\underline{\mathrm{Hom}}(x, y)$ is called the exponential object. This object is often denoted by y^x.

In Cartesian closed categories a different, but frequently used, notation is $x \Rightarrow y$. However, this notation will not be used as it might be confusion with the notation for *2-morphisms.*

Definition 2.1.23 (Cartesian closed functor). A functor between Cartesian closed categories that preserves products and exponential objects. As such it is the natural notion of functor between Cartesian closed categories.

Property 2.1.24 (Frobenius reciprocity). A functor R between Cartesian closed categories that admits a left adjoint L is Cartesian closed if and only if the natural transformation

$$L(y \times Rx) \to Ly \times x \tag{2.4}$$

is a natural isomorphism.

Property 2.1.25 (Global elements). The following isomorphism is natural in both $x, y \in \mathrm{ob}(\mathbf{M})$:

$$\mathbf{M}(\mathbf{1}, [x, y]) \cong \mathbf{M}(x, y). \tag{2.5}$$

It is this relation that gives the best explanation for the term "internal hom". One also immediately obtains the following natural isomorphism:

$$\mathbf{M}(x, [\mathbf{1}, y]) \cong \mathbf{M}(x, y). \tag{2.6}$$

Because the Yoneda embedding is fully faithful this implies that $[\mathbf{1}, y] \cong y$. Although the global elements $\mathbf{M}(\mathbf{1}, y)$ do not fully specify an object y, this does hold internally.

Property 2.1.26 (Symmetry). Let $\mathbf{M}$ be a closed monoidal category. The definition of an internal hom can also be internalized, i.e. there exists a natural isomorphism of the form

$$[x \otimes y, z] \cong [x, [y, z]]. \tag{2.7}$$

Furthermore, if $\mathbf{M}$ is also symmetric, there exists an internal isomorphism of the form

$$[x, [y, z]] \cong [y, [x, z]]. \tag{2.8}$$

Definition 2.1.27 (Strong adjunction). Consider a monoidal category $\mathbf{M}$ together with two endofunctors $L, R : \mathbf{M} \to \mathbf{M}$. These functors are said to form a strong adjunction if there exists a natural isomorphism

$$[Lx, y] \cong [x, Ry]. \tag{2.9}$$

Property 2.1.25 above implies that every strong adjunction is in particular an adjunction in the sense of Section 1.2.4.

2.2 Enriched category theory

The following definition is due to *Bénabou*. It should represent the "ideal place in which to do category theory".

Definition 2.2.1 (Cosmos). A complete and cocomplete closed symmetric monoidal category.

Definition 2.2.2 (Enriched category). Let $(\mathcal{V}, \otimes, \mathbf{1})$ be a monoidal category. A $\mathcal{V}$-enriched category, also called a $\mathcal{V}$-category[1], consists of the following elements:

- a collection of objects $\mathrm{ob}(\mathbf{C})$, and
- for every pair of objects $x, y \in \mathrm{ob}(\mathbf{C})$, an object $\mathbf{C}(x, y) \in \mathrm{ob}(\mathcal{V})$ for which the following morphisms exist:
  1. $\mathrm{id}_x : \mathbf{1} \to \mathbf{C}(x, x)$ giving the (enriched) identity morphism, and
  2. $\circ_{xyz} : \mathbf{C}(y, z) \otimes \mathbf{C}(x, y) \to \mathbf{C}(x, z)$ replacing the usual composition.

The associativity and unity properties are given by commutative diagrams for the id and $\circ$ morphisms together with the associators and unitors in $\mathcal{V}$.

Definition 2.2.3 (Change of base). Consider a monoidal functor $F : \mathcal{V} \to \mathcal{W}$. This induces a change of base functor $F_* : \mathcal{V}\mathbf{Cat} \to \mathcal{W}\mathbf{Cat}$ by applying F to every hom-object.

Definition 2.2.4 (Underlying category). Given a $\mathcal{V}$-enriched category $\mathbf{C}$, the underlying category $\mathbf{C}_0$ is defined as follows:

- **Objects**: $\mathrm{ob}(\mathbf{C})$
- **Morphisms**: $\mathcal{V}(\mathbf{1}, \mathbf{C}(x, y))$,

where $\mathbf{1}$ is the monoidal unit in $\mathcal{V}$. This construction can be obtained as the functor $\mathcal{V}\mathbf{Cat}(\mathcal{I}, -)$ where $\mathcal{I}$ is the one-object $\mathcal{V}$-category with $\mathcal{I}(*, *) \equiv \mathbf{1}$.

Property 2.2.5 ($\mathcal{V}$ as a $\mathcal{V}$-category). Consider a closed monoidal category $\mathcal{V}$. This category can be given the structure $\widetilde{\mathcal{V}}$ of a $\mathcal{V}$-category by taking the hom-objects to be the internal homs, i.e. $\widetilde{\mathcal{V}}(x, y) := [x, y]$ for all $x, y \in \mathcal{V}$. Property 2.1.25 then implies that there exists an isomorphism between the underlying category $\widetilde{\mathcal{V}}_0$ and the original category $\mathcal{V}$.

Given two $\mathcal{V}$-enriched categories, one can define suitable functors between them:

Definition 2.2.6 (Enriched functor). A $\mathcal{V}$-enriched functor $F : \mathbf{A} \to \mathbf{B}$ consists of the following data:

- a function $F_0 : \mathrm{ob}(\mathbf{A}) \to \mathrm{ob}(\mathbf{B})$ (as for ordinary functors), and
- for every two objects $x, y \in \mathrm{ob}(\mathbf{A})$, a morphism $F_{x,y} : \mathbf{A}(x, y) \to \mathbf{B}(Fx, Fy)$ in $\mathcal{V}$.

These have to satisfy the "usual" composition and unit conditions.

By extending (1.22) using enriched ends, one obtains a definition of enriched natural transformations and, therefore, also a definition of enriched functor categories.:

$$[\mathbf{A}, \mathbf{B}](F, G) := \int_{x \in \mathbf{A}} \mathbf{B}(Fx, Gx). \tag{2.10}$$

Given two $\mathcal{V}$-enriched functors $F, G : \mathbf{A} \to \mathbf{B}$ one can also try to define $\mathcal{V}$-natural transformations by extending the usual definition of natural transformations 1.2.14:

[1] Not to be confused with the notation for fibre categories 1.3.7.

$$\begin{tikzcd}
 & \mathbf{A}(x,y) \arrow[dl, "\lambda^{-1}"'] \arrow[dr, "\rho^{-1}"] & \\
\mathbf{1}\otimes\mathbf{A}(x,y) \arrow[d, "\eta_y \otimes F_{x,y}"'] & & \mathbf{A}(x,y)\otimes\mathbf{1} \arrow[d, "G_{x,y}\otimes\eta_x"] \\
\mathbf{B}(Fy,Gy)\otimes\mathbf{B}(Fx,Fy) \arrow[dr, "\circ"'] & & \mathbf{B}(Gx,Gy)\otimes\mathbf{B}(Fx,Gx) \arrow[dl, "\circ"] \\
 & \mathbf{B}(Fx,Gy) &
\end{tikzcd}$$

Figure 2.8: $\mathcal{V}$-naturality diagram.

Definition 2.2.7 (Enriched natural transformation). An ordinary natural transformation consists of an ob($\mathbf{A}$)-indexed family of morphism $\eta_x : Fx \to Gx$. This can also be interpreted as an ob($\mathbf{A}$)-indexed family of morphisms $\eta_x : 1 \to \mathbf{B}(Fx, Gx)$ from the initial object (one-element set). By analogy, a $\mathcal{V}$-natural transformation is defined as an ob($\mathbf{A}$)-indexed family of morphisms $\eta_x : \mathbf{1} \to \mathbf{B}(Fx, Gx)$ from the monoidal unit. The usual naturality square is replaced by the naturality hexagon 2.8.

The question then becomes how these two definitions are related. The end (2.10) comes equipped with a projection $\varepsilon_x : [\mathbf{A}, \mathbf{B}](F, G) \to \mathbf{B}(Fx, Gx)$. Precomposing this morphism with a morphism in the underlying category, i.e. an element of $\mathcal{V}(\mathbf{1}, [\mathbf{A}, \mathbf{B}](F, G))$, exactly gives a $\mathcal{V}$-natural transformation. So the underlying category of $[\mathbf{A}, \mathbf{B}]$ is the ordinary category of $\mathcal{V}$-functors and $\mathcal{V}$-natural transformations.

2.2.1 Enriched constructions

Definition 2.2.8 (Functor tensor product). Consider a covariant functor $G : \mathbf{C} \to \mathcal{V}$ and a contravariant functor $F : \mathbf{C}^{op} \to \mathcal{V}$ into a monoidal category $\mathcal{V}$, where $\mathbf{C}$ does not have to be enriched over $\mathcal{V}$. The tensor product of F and G is defined as the following coend:

$$F \otimes_{\mathbf{C}} G := \int^{x\in\mathbf{C}} Fx \otimes Gx. \tag{2.11}$$

It should be noted that the above tensor product does not produce a new functor, instead it only gives an object in $\mathcal{V}$. A different type of tensor product, one that does give a functor, exists in the enriched setting (note that there is no relation between these two definitions):

Definition 2.2.9 (Day convolution). Consider a monoidally cocomplete category $\mathcal{V}$, i.e. cocomplete monoidal category for which the tensor product bifunctor is cocontinuous in each argument, together with a $\mathcal{V}$-enriched category $\mathbf{C}$. The convolution or tensor product (if it exists) of two $\mathcal{V}$-enriched functors $F, G : \mathbf{C} \to \mathcal{V}$ is defined as the following coend:

$$F \otimes_{\mathrm{Day}} G := \iint^{x,y\in\mathbf{C}} \mathbf{C}(x\otimes y, -) \otimes Fx \otimes Gy. \tag{2.12}$$

Property 2.2.10 (Monoidal structure). In the case where $\mathbf{M}$ is a closed symmetric monoidal category, the Day convolution is associative and, hence, defines a monoidal structure on the functor category $[\mathbf{C}, \mathbf{M}]$. The tensor unit is given by the functor (co)represented by the tensor unit in $\mathbf{C}$.

Definition 2.2.11 (Copower). Consider a $\mathcal{V}$-enriched category $\mathbf{C}$. The copower (or tensor) functor $\cdot : \mathcal{V} \times \mathbf{C} \to \mathbf{C}$ is defined by the following natural isomorphism:

$$\mathbf{C}(v \cdot x, y) \cong [v, \mathbf{C}(x, y)], \tag{2.13}$$

where the bracket $[-,-]$ on the right-hand side denotes the internal hom in $\mathcal{V}$. Dually, the power (or cotensor) functor $[-,-] : \mathcal{V} \times \mathbf{C} \to \mathbf{C}$ is defined by the following natural isomorphism:

$$\mathbf{C}(x, [v, y]) \cong [v, \mathbf{C}(x, y)], \tag{2.14}$$

where the bracket $[-,-]$ on the right-hand side again denotes the internal hom in $\mathcal{V}$. If an enriched category admits all (co)powers, it is said to be **(co)powered** (over its enriching category).

Remark 2.2.12. Equation (2.8) says that every (closed) symmetric monoidal category $\mathbf{M}$ is powered over itself, the power just being the internal hom. The same holds for the copower, which is just the usual tensor product functor.

Example 2.2.13 (Disjoint unions). Every (co)complete (locally) small category $\mathbf{C}$ admits the structure of a **Set**-(co)powered category:

$$x^S := \prod_{s \in S} x \tag{2.15}$$

$$S \cdot x := \bigsqcup_{s \in S} x. \tag{2.16}$$

The definition and properties of internal hom-functors and (co)powers can be formalized as follows:

Definition 2.2.14 (Two-variable adjunction). Consider three categories $\mathbf{A}, \mathbf{B}$ and $\mathbf{C}$. A two-variable adjunction $\mathbf{A} \times \mathbf{B} \to \mathbf{C}$ consists of three bifunctors:

- $- \otimes - : \mathbf{A} \times \mathbf{B} \to \mathbf{C}$,
- $\mathrm{hom}_L : \mathbf{A}^{op} \times \mathbf{C} \to \mathbf{B}$, and
- $\mathrm{hom}_R : \mathbf{B}^{op} \times \mathbf{C} \to \mathbf{A}$

admitting the following natural isomorphisms:

$$\mathbf{C}(x \otimes y, z) \cong \mathbf{A}(x, \mathrm{hom}_R(y, z)) \cong \mathbf{B}(y, \mathrm{hom}_L(x, z)). \tag{2.17}$$

It should be noted that fixing any of the variables gives rise to ordinary adjunctions in the sense of Section 1.2.4.

Property 2.2.15 (Powers and copowers). A category $\mathbf{C}$ enriched over a monoidal category $\mathcal{V}$ is powered and copowered over $\mathcal{V}$ exactly if the hom-functor $\mathbf{C}^{op} \times \mathbf{C} \to \mathcal{V}$ is the right adjoint in an enriched two-variable adjunction. The power and copower functors are then given by the other two adjoints.

The following definition constructs Kan extensions in the enriched setting (these can be shown to reduce to 1.4.71 when enriching over **Set**):

Alternative Definition 2.2.16 (Kan extension). Let $\mathbf{A}, \mathbf{B}$ and $\mathbf{C}$ be categories enriched over a monoidal category $\mathcal{V}$. If $\mathbf{B}$ is assumed to be copowered over $\mathcal{V}$, one can define the left Kan extension of $F : \mathbf{A} \to \mathbf{B}$ along $G : \mathbf{A} \to \mathbf{C}$ as a coend:

$$\mathrm{Lan}_G F := \int^{x \in \mathbf{A}} \mathbf{C}(Gx, -) \cdot Fx. \tag{2.18}$$

If $\mathbf{B}$ is assumed to be powered over $\mathcal{V}$, one can define the right Kan extension as an end:

$$\mathrm{Ran}_G F := \int_{x\in\mathbf{A}} [\mathbf{C}(-, Gx), Fx]. \tag{2.19}$$

Remark 2.2.17. By choosing $\mathcal{V} = \mathbf{Set}, \mathbf{C} = \mathbf{A}$ and $G = \mathbb{1}_\mathbf{A}$ in the previous definition, one obtains the ninja Yoneda lemma 1.4.68.

Property 2.2.18. Kan extensions computed using (co)ends as above are pointwise in the sense of Definition 1.4.73.

Alternative Definition 2.2.19 (Functor tensor product). Let $\mathbf{B}$ be a $\mathcal{V}$-enriched category. Consider a covariant functor $G : \mathbf{A} \to \mathbf{B}$ and a contravariant functor $F : \mathbf{A}^{op} \to \mathcal{V}$. The tensor product 2.2.8 can be generalized whenever $\mathbf{B}$ is copowered over $\mathcal{V}$:

$$F \otimes_\mathbf{A} G := \int^{x\in\mathbf{A}} Fx \cdot Gx. \tag{2.20}$$

2.2.2 Weighted (co)limits

In this section the definition of ordinary limits and, in particular, the defining universal property 1.4.32 is revisited. In this construction the constant functor Δ_x was one of the main ingredients. This functor can be factorized as $\mathbf{I} \to 1 \to \mathbf{C}$, where 1 denotes the terminal category. On the level of morphisms this factorization takes the form $\mathbf{I}(i, j) \to * \to \mathbf{C}(x, x)$, where $*$ denotes the terminal one-element set. However, whenever the enriching context is not $\mathbf{Set}$, one does not necessarily have access to a terminal object.

To avoid this issue, limits will first be redefined as representing objects. To this end, consider a general diagram $D : \mathbf{I} \to \mathbf{C}$. By postcomposition with the Yoneda embedding one obtains the presheaf-valued diagram $\mathbf{C}(-, D-) : \mathbf{I} \to [\mathbf{C}^{op}, \mathbf{Set}]$. Since presheaf categories are complete (Example 1.4.36), the limit of this diagram exists:

$$\mathbf{Set}(S, \lim \mathbf{C}(x, D-)) \cong [\mathbf{I}, \mathbf{Set}](\Delta_S, \mathbf{C}(x, D-)).$$

By restricting to the terminal set $S = *$, one obtains

$$\lim \mathbf{C}(x, D-) \cong [\mathbf{I}, \mathbf{Set}](\Delta_*, \mathbf{C}(x, D-)).$$

If this presheaf is representable, one can use the continuity of the hom-functor, together with the fact that the Yoneda embedding is fully faithful, to show that the representing object is (isomorphic to) $\lim D$, i.e.

$$[\mathbf{I}, \mathbf{Set}](\Delta_*, \mathbf{C}(x, D-)) \cong \mathbf{C}(x, \lim D). \tag{2.21}$$

?? CLEAN THIS UP (note that continuity and pointwise definition was already mentioned for ordinary limits) ??

Definition 2.2.20 (Weighted limit). This definition can now be generalized by replacing the constant functor Δ_* by any functor $W : \mathbf{I} \to \mathbf{Set}$. A representing object is then called the W-weighted limit of D. This object is often denoted by $\lim^W D$ or $\{W, D\}$. To distinguish weighted limits from ordinary ones, the latter are sometimes called **conical limits**.

Remark 2.2.21. A motivation for this construction is the following. As was already pointed out in Remark 1.4.18, the mere knowledge of global elements $1 \to x$ is often not enough to characterize an object x. In general one should look at the collection of generalized elements. When applying this ideology to the case of cones, one sees that replacing the functor Δ_* by a more general functor is the same as replacing the global elements $* \to Di$ by generalized elements $Wi \to Di$.

The generalization to the enriched setting is now evident. There is no reference to the terminal object left, so one can replace **Set** by any enriching category. In the enriched setting, (co)end formulas for (weighted) limits will often be used:

Formula 2.2.22 (Enriched weighted limits). By expressing the natural transformations as an end as in Equation (1.22) and by using the canonical powering in **Set**, one can express ordinary weighted limits as follows:

$$\lim^W D \cong \int_{i\in\mathbf{I}} [Wi, Di]. \tag{2.22}$$

The generalization to other enriching categories is now straightforward. Consider a diagram $D : \mathbf{I} \to \mathbf{C}$ and a weight functor $W : \mathbf{I} \to \mathcal{V}$, where $\mathbf{C}$ is $\mathcal{V}$-enriched. If $\mathbf{C}$ is powered over $\mathcal{V}$, the W-weighted limit of D is defined by the same formula as above:

$$\lim^W D := \int_{i\in\mathbf{I}} [Wi, Di]. \tag{2.23}$$

In a similar way one can define weighted colimits in copowered $\mathcal{V}$-categories as coends:

$$\operatorname{colim}^W D := \int^{i\in\mathbf{I}} Wi \cdot Di. \tag{2.24}$$

Here, the weight functor W is required to be contravariant since colimits (and cocones in general) are natural transformations between contravariant functors.

Property 2.2.23 (Weighted limits are Homs). In the case $\mathbf{C} = \mathcal{V}$, the powering functor becomes the internal hom and, therefore, one sees that weighted limits are given by (enriched) natural transformations (as was the case for ordinary conical limits).

In the following example the weighted colimit is calculated with respect to the Yoneda embedding:

Example 2.2.24 (Hom-functor). Consider a diagram $D : \mathbf{I} \to \mathbf{C}$. When using the Yoneda embedding $\mathcal{Y}i = \mathbf{I}(-, i)$ as the weight functor, one obtains the following property by virtue of the Yoneda lemma:

$$\operatorname{colim}^{\mathcal{Y}i} D \cong Di. \tag{2.25}$$

A similar statement for weighted limits can be obtained with the covariant Yoneda embedding.

Alternative Definition 2.2.25 (Weighted (co)limits). The above property can be used to axiomatize small weighted (co)limits in bicomplete categories:

1. **Yoneda**: For every object $i \in \mathrm{ob}(\mathbf{I})$ there exist isomorphisms
$$\lim^{\mathbf{I}(i,-)} D \cong Di \qquad \text{and} \qquad \operatorname{colim}^{\mathbf{I}(-,i)} D \cong Di. \tag{2.26}$$
2. **Cocontinuity**: The weighted (co)limit functors are cocontinuous in the weights.

One can also express Kan extensions as weighted limits (this simply follows from expression 2.2.16):

Property 2.2.26 (Kan extensions). Consider functors $F : \mathbf{A} \to \mathbf{B}$ and $G : \mathbf{A} \to \mathbf{C}$. If for every $x \in \mathrm{ob}(\mathbf{C})$ the weighted limit $\lim^{\mathbf{C}(x,G-)} F$ exists, these limits can be combined into a functor that can be shown to be the right Kan extension $\mathrm{Ran}_G F$. The left Kan extension can be obtained as a weighted colimit.

Property 2.2.27 (Category of elements). The weighted (co)limits of a functor (over **Set**) can also be expressed in terms of the category of elements 1.3.15 of the weight:

$$\lim^W F \cong \lim F \circ \mathbf{C}_W, \tag{2.27}$$

where the limit on the right-hand side is a conical limit.

2.3 Abelian categories

2.3.1 Additive categories

Definition 2.3.1 (Pre-additive category). A (locally small) category enriched over **Ab**, i.e. a category in which every hom-set is an Abelian group and composition is bilinear.

Property 2.3.2. Let $\mathbf{A}$ be a pre-additive category. The following statements are equivalent for an object $x \in \mathrm{ob}(\mathbf{A})$:

- x is initial,
- x is final, or
- $\mathbb{1}_x = 0$.

It follows that every initial/terminal object in a pre-additive category is automatically a zero object 1.4.12.

Property 2.3.3 (Biproducts). In a pre-additive category the following isomorphism holds for all finitely indexed sets $\{x_i\}_{i\in I}$:

$$\prod_{i\in I} x_i \cong \bigsqcup_{i\in I} x_i. \tag{2.28}$$

Finite (co)products in pre-additive categories are often called **direct sums**. In general, if a product and coproduct exist and are equal, one also speaks of a **biproduct**.

Definition 2.3.4 (Additive category). A pre-additive category in which all finite products exist.

When working with additive categories, it is generally assumed that the associated functors are of a specific type:

Definition 2.3.5 (Additive functor). Let $\mathbf{A}, \mathbf{A}'$ be additive categories. A functor $F : \mathbf{A} \to \mathbf{A}'$ is said to be additive if it preserves finite biproducts:

1. It preserves zero objects: $F\,0_{\mathbf{A}} \cong 0_{\mathbf{A}'}$.
2. There exists a natural isomorphism $F(x \oplus y) \cong Fx \oplus Fy$.

This notion can be generalized to pre-additive categories. A functor between pre-additive categories is said to be additive if it acts as a group morphism on hom-spaces.

Definition 2.3.6 (Grothendieck group). Let $\mathbf{A}$ be an additive category and consider its decategorification 1.2.23. This set carries the structure of an Abelian monoid and, hence, the Grothendieck construction **??** can be applied to obtain an Abelian group $K(\mathbf{A})$. This group is called the Grothendieck group of $\mathbf{A}$.

In a (pre-)additive category one can some classical notions from (homological) algebra such as images and kernels:

Definition 2.3.7 (Kernel). Let $f : x \to y$ be a morphism. A[2] kernel of f is a morphism $k : z \to x$ such that:

1. $f \circ k = 0$.

[2]Note the word "*a*". The kernel of a morphism is only determined up to an isomorphism.

2. **Universal property**: Every morphism $k' : z' \to x$ such that $f \circ k' = 0$ factors uniquely through k.

This implies that a kernel of f could equivalently be defined as the equalizer of f and 0.

Notation 2.3.8 (Kernel). If the kernel of $f : x \to y$ exists, it is denoted by $\ker(f)$.

Definition 2.3.9 (Cokernel). Let $f : x \to y$ be a morphism. A cokernel of f is a morphism $p : y \to z$ such that:

1. $p \circ f = 0$.
2. **Universal property**: Every morphism $p' : y \to z'$ such that $p' \circ f = 0$ factors uniquely through p.

This implies that a cokernel of f could equivalently be defined as the coequalizer of f and 0.

Notation 2.3.10 (Cokernel). If the cokernel of $f : x \to y$ exists, it is denoted by $\mathrm{coker}(f)$.

Remark 2.3.11. The name and notation of the kernel and the cokernel (in the categorical sense) is explained by remarking that $\ker(f)$ represents the functor

$$F : z \mapsto \ker\Big(\mathbf{C}(z, x) \to \mathbf{C}(z, y)\Big),$$

where ker denotes the algebraic kernel **??**, and similarly for he cokernel.

Definition 2.3.12 (Pseudo-Abelian category). An additive category in which every projection/idempotent has a kernel.

Definition 2.3.13 (Pre-Abelian category). An additive category in which every morphism has a kernel and cokernel.

Definition 2.3.14 (Abelian category). A pre-Abelian category in which every mono is a kernel and every epi is a cokernel or, equivalently, if for every morphism f there exists an isomorphism

$$\mathrm{coker}(\ker(f)) \cong \ker(\mathrm{coker}(f)). \tag{2.29}$$

Property 2.3.15 (Injectivity and surjectivity). In Abelian categories a morphism is monic if and only if it is injective, i.e. its kernel is 0. Analogously, a morphism is epic if and only if it is surjective, i.e. its cokernel is 0.

Definition 2.3.16 (Linear category). Let $\mathbf{Vect}_K$ denote the category of vector spaces over the base field K. A K-linear category is a category enriched over $\mathbf{Vect}_K$. (If the base field is clear, the subscript is often left implicit.)

2.3.2 Exact functors

Definition 2.3.17 (Exact functor). Let $F : \mathbf{A} \to \mathbf{A}'$ be an additive functor between additive categories.

- F is said to be left-exact if it preserves kernels.
- F is said to be right-exact if it preserves cokernels.
- F is said to be exact if it is both left- and right-exact.

Corollary 2.3.18. The previous definition implies the following properties (which can in fact be used as an alternative definition):

- If F is left-exact, it maps an exact sequence of the form
$$0 \longrightarrow x \longrightarrow y \longrightarrow z$$
to an exact sequence of the form
$$0 \longrightarrow Fx \longrightarrow Fy \longrightarrow Fz.$$
- If F is right-exact, it maps an exact sequence of the form
$$x \longrightarrow y \longrightarrow z \longrightarrow 0$$
to an exact sequence of the form
$$Fx \longrightarrow Fy \longrightarrow Fz \longrightarrow 0.$$
- If F is exact, it maps short exact sequences to short exact sequences.

Notation 2.3.19 (Left or right). The category of left modules ${}_R\mathbf{Mod}$ over a ring R is equivalent (as an Abelian category) to the category of right modules $\mathbf{Mod}_{R^{op}}$ over the opposite ring R. For this reason one often makes no difference between left and right modules (only bimodules are truly relevant) and "the category of R-modules" is just denoted by $R\mathbf{Mod}$.

Theorem 2.3.20 (Freyd-Mitchell embedding theorem). *Every small Abelian category admits a fully faithful, exact functor into a category of the form $R\mathbf{Mod}$ for some unital ring R.*

Theorem 2.3.21 (Eilenberg-Watts). *Let R, S be two (not necessarily unital) rings. The tensor product functor induces an equivalence between the category of R-S-bimodules and the category of cocontinuous functors $R\mathbf{Mod} \to S\mathbf{Mod}$.*

2.3.3 Finiteness

Definition 2.3.22 (Simple object). Let $\mathbf{A}$ be an Abelian category. An object $a \in \mathrm{ob}(\mathbf{A})$ is said to be simple if the only subobjects of a are 0 and a itself. An object is said to be semisimple if it is a direct sum of simple obejcts.

Definition 2.3.23 (Semisimple category). A category is said to be semisimple if every object is semisimple (where in general the direct sums are taken over finite index sets).

Definition 2.3.24 (Jordan-Hölder series). A filtration
$$0 \longrightarrow x_1 \longrightarrow x_2 \longrightarrow \cdots \longrightarrow x_n = x$$
of an object x is said to be a Jordan-Hölder series if the quotient objects x_i/x_{i-1} are simple for all $i \leq n$. If the series has finite length, the object x is said to be **finite**.

Theorem 2.3.25 (Jordan-Hölder). *If an object in an Abelian category is finite, all of its Jordan-Hölder series have the same length. In particular, the multiplicities of simple objects are the same for all such series.*

Theorem 2.3.26 (Krull-Schmidt). *Any object in an Abelian category of finite length admits a unique decomposition as a direct sum of indecomposable objects*[3]*.*

Definition 2.3.27 (Locally finite). A k-linear Abelian category is said to be locally finite if it satisfies the following conditions:

[3] An object is **indecomposable** if it cannot be written as a direct sum of its subobjects.

1. every hom-space is finite-dimensional, and
2. every object has finite length.

Definition 2.3.28 (Finite). A k-linear Abelian category is said to be finite if it satisfies the following conditions:

1. It is locally finite.
2. It has enough projectives or, equivalently, every simple object has a *projective cover*.
3. The set of isomorphism classes of simple objects is finite.

Theorem 2.3.29 (Schur's lemma). *Let* $\mathbf{A}$ *be an Abelian category. For every two simple objects* x, y*, all nonzero morphisms* $x \to y$ *are isomorphisms. In particular, if* x, y *are two non-isomorphic simple objects, then* $\mathbf{A}(x, y) = 0$*. Furthermore,* $\mathbf{A}(x, x)$ *is a division ring for every simple object* x*.*

Corollary 2.3.30. If $\mathbf{A}$ is locally finite and k is algebraically closed, then $\mathbf{A}(x, x) \cong k$ for all simple objects x. This follows from the fact that the only finite-dimensional division algebra over an algebraically closed field is the field itself.

The Freyd-Mitchell theorem 2.3.20 can be adapted to the finite linear case as follows:

Theorem 2.3.31 (Deligne). *Every finite* k*-linear Abelian category is* k*-linearly equivalent to a category of the form* $A\mathbf{Mod}^{\text{fin}}$ *for* A *a finite-dimensional* k*-algebra.*

Construction 2.3.32 (Deligne tensor product). Let $\mathbf{A}, \mathbf{B}$ be two Abelian categories. Their Deligne (tensor) product is defined (if it exists) as the category $\mathbf{A} \boxtimes \mathbf{B}$ for which there exists a bijection between right exact functors $\mathbf{A} \boxtimes \mathbf{B} \to \mathbf{C}$ and right exact functors $\mathbf{A} \times \mathbf{B} \to \mathbf{C}$ (the latter being right exact in each argument).

For finite Abelian categories it can be shown that their Deligne product always exists. By the Deligne embedding theorem one can find an explicit description. Consider two finite-dimensional k-algebras A, B. The category $A\mathbf{Mod}^{\text{fin}} \boxtimes B\mathbf{Mod}^{\text{fin}}$ is equivalent to the category $A \otimes_k B\mathbf{Mod}^{\text{fin}}$.

2.4 Monoidal categories II: Duality

The general theory of monoidal categories was introduced in Section 2.1.

Definition 2.4.1 (Dual object). Let $(\mathbf{C}, \otimes, \mathbf{1})$ be a monoidal category and consider an object $x \in \text{ob}(\mathbf{C})$. A left dual[4] of x is an object in $x^* \in \text{ob}(\mathbf{C})$ together with two morphisms $\eta : \mathbf{1} \to x \otimes x^*$ and $\varepsilon : x^* \otimes x \to \mathbf{1}$, called the **unit** and **counit** morphisms[5], such that the diagrams in Figure 2.9 commute. x is said to be **dualizable** if it admits a left dual.

Definition 2.4.2 (Rigid category). A monoidal category that has all duals. These categories are also said to be **autonomous**. If only left (resp. right) duals exist, the category is said to be left (resp. right) rigid.

Property 2.4.3 (Braided categories). In general it is not true that left and right duals coincide. However, in a braided monoidal category this is the case.

Definition 2.4.4 (Compact closed category). A symmetric rigid category.

[4] x is called the **right dual** of x^*. The right dual of y is often denoted by ${}^*y$.

[5] Also called the **coevaluation** and **evaluation** morphisms.

$$
\begin{tikzcd}
 & x & \\
\mathbf{1} \otimes x \arrow[ur, "\lambda_x"] \arrow[d, "\eta \otimes \mathbb{1}_x"'] & & x \otimes \mathbf{1} \arrow[ul, "\rho_x"'] \\
(x \otimes x^*) \otimes x \arrow[rr, "\alpha_{x,x^*,x}"] & & x \otimes (x^* \otimes x) \arrow[u, "\mathbb{1}_x \otimes \varepsilon"']
\end{tikzcd}
$$

(a) Dual object I.

$$
\begin{tikzcd}
 & x^* & \\
x^* \otimes \mathbf{1} \arrow[ur, "\rho_{x^*}"] \arrow[d, "\mathbb{1}_{x^*} \otimes \eta"'] & & \mathbf{1} \otimes x^* \arrow[ul, "\lambda_{x^*}"'] \\
x^* \otimes (x \otimes x^*) \arrow[rr, "\alpha^{-1}_{x^*,x,x^*}"] & & (x^* \otimes x) \otimes x^* \arrow[u, "\varepsilon \otimes \mathbb{1}_{x^*}"']
\end{tikzcd}
$$

(b) Dual object II.

Figure 2.9: Dualizable objects.

Example 2.4.5 (FinVect). Consider the category **FinVect** of finite-dimensional vector spaces (the ground field is assumed to be $\mathbb{R}$). The categorical dual of a vector space V is the algebraic dual V^*. The unit morphism is given by the "resolution of the identity":

$$\eta : \mathbf{1} \to V \otimes V^* : 1 \mapsto \sum_{i=1}^{\dim(V)} e_i \otimes \phi^i \,, \tag{2.30}$$

where $\{e_i\}$ and $\{\phi^i\}$ are bases of V and V^*, respectively.

It should be noted that the category **Vect** of all vector spaces is not rigid. By Property 2.4.3 above, left and right duals coincide in any braided monoidal category (such as **Vect**), but for infinite-dimensional vector spaces it is known that $A \cong (A^*)^*$ never holds and as such rigidity cannot be extended to **Vect**.

Property 2.4.6 (Tannaka duality). Consider the category $\mathcal{V} = \mathbf{FinVect}_K$. Using coends one can reconstruct the base field from its modules, i.e. the objects in $\mathcal{V}$:

$$\int^{V \in \mathcal{V}} V^* \otimes V \cong K. \tag{2.31}$$

This result can be shown to hold for all compact closed categories $\mathcal{V}$. In this context it is known as **Tannaka reconstruction**. A more general statement goes as follows:

$$\int^{V \in \mathcal{V}} \mathcal{V}(V, -) \otimes V \cong \mathrm{id}_{\mathcal{V}}. \tag{2.32}$$

For $\mathcal{V} = \mathbf{FinVect}_K$, the components $\eta_V : \mathcal{V}(V,V) \to K$ of the coend can be shown to coincide with the trace and as such the trace obtains a universal property.

Remark 2.4.7. This property can also be generalized by replacing $\mathcal{V}$ by a category of modules $A\mathbf{Mod}$ for some finite-dimensional algebra A. The end and coend give the algebra A and its dual A^*, respectively.

Definition 2.4.8 (Symmetric monoidal dagger category). A symmetric monoidal category $(\mathbf{C}, \otimes, \mathbf{1})$ that also carries the structure of a dagger category 1.3.1 such that

$$(f \otimes g)^\dagger = f^\dagger \otimes g^\dagger \tag{2.33}$$

and such that the coherence and braiding morphisms are unitary.

Definition 2.4.9 (Dagger-compact category). A symmetric monoidal dagger category that is also a compact closed category such that the following diagram commutes for all objects:

$$\begin{tikzcd} & \mathbf{1} \arrow[dl, "\eta"'] \arrow[dr, "\varepsilon^\dagger"] & \\ x^* \otimes x & & x \otimes x^* \arrow[ll, "\sigma_{x,x^*}"] \end{tikzcd}$$

The trace on **FinVect** can be generalized as follows:

Definition 2.4.10 (Trace). Let $(\mathbf{C}, \otimes, \mathbf{1})$ be a rigid category and let $f \in \mathbf{C}(x, x^{**})$. The left (**categorical** or **quantum**) trace of f is defined as the following morphism in $\mathrm{End}_{\mathbf{C}}(\mathbf{1})$:

$$\mathrm{tr}^L(f) := \varepsilon_{x^*} \circ (f \otimes \mathbb{1}_{x^*}) \circ \eta_x. \tag{2.34}$$

For $f \in \mathbf{C}(x, {}^{**}x)$, the right trace is defined similarly:

$$\mathrm{tr}^R(f) := \varepsilon_{{}^{**}x} \circ (\mathbb{1}_x \otimes f) \circ \eta_{{}^*x}. \tag{2.35}$$

Property 2.4.11. The following linear algebra-like properties hold for the categorical trace:

- $\mathrm{tr}^L(f) = \mathrm{tr}^R(f^*)$,
- $\mathrm{tr}^L(f \otimes g) = \mathrm{tr}^L(f)\,\mathrm{tr}^L(g)$, and
- for additive categories: $\mathrm{tr}^L(f \oplus g) = \mathrm{tr}^L(f) + \mathrm{tr}^L(g)$.

The second and third property can be stated analogously for the right trace.

Definition 2.4.12 (Pivotal category). Let $\mathbf{C}$ be a rigid monoidal category. A pivotal structure on $\mathbf{C}$ is a monoidal natural isomorphism $\psi : \mathrm{id}_{\mathbf{C}} \Rightarrow **$.

Definition 2.4.13 (Dimension). Let $(\mathbf{C}, \psi)$ be a pivotal category and consider an object $x \in \mathrm{ob}(\mathbf{C})$. The dimension of x is defined as follows:

$$\dim_\psi(x) := \mathrm{tr}^L(\psi_x). \tag{2.36}$$

Definition 2.4.14 (Spherical category). A pivotal category $(\mathbf{C}, \psi)$ in which the left and right traces with respect to ψ coincide: $\dim_\psi(x) = \dim_\psi(x^*)$ for all $x \in \mathrm{ob}(\mathbf{C})$.

Definition 2.4.15 (Calabi-Yau category). A **Vect**-enriched category $\mathbf{C}$ equipped with a **trace** functional

$$\mathrm{tr}_x : C(x,x) \to K \tag{2.37}$$

for each object $x \in \mathrm{ob}(\mathbf{C})$ such that the induced pairing

$$\langle \cdot, \cdot \rangle : C(x,y) \otimes C(y,x) \to K : f \otimes g \mapsto \mathrm{tr}_x(g \circ f) \tag{2.38}$$

is symmetric and nondegenerate.

Example 2.4.16. A one-object Calabi-Yau category is the (pointed) monoid delooping of a Frobenius algebra **??**.

2.5 Tensor and fusion categories

Some definitions might slightly differ from the ones in the main references and some properties might be stated less generally. K denotes an algebraically closed field (often this will be $\mathbb{C}$).

Definition 2.5.1 (Tensor category). A monoidal category with the following properties:

1. it is rigid,
2. it is Abelian,
3. it is K-linear in a way compatible with the Abelian structure,
4. $\mathrm{End}(\mathbf{1}) \cong K$, and
5. $-\otimes-$ is bilinear on morphisms.

Some authors (such as [33]) also add "locally finite" as a condition (Definition 2.3.27).

Remark 2.5.2. If K is not algebraically closed, one should replace the fourth condition by the condition that $\mathbf{1}$ is a simple object. However, if K is algebraically closed, these statements are equivalent.

Definition 2.5.3 (Pointed tensor category). A tensor category where all of the simple objects are (weakly) invertible.

Definition 2.5.4 (Fusion category). A semisimple finite tensor category.

Property 2.5.5. Let $\mathbf{C}$ be a fusion category. There exists a natural isomorphism $\mathbb{1}_{\mathbf{C}} \cong **$.

Remark 2.5.6. Although any fusion category admits a natural isomorphism between an object and its double dual, this morphism does not need to be monoidal. The fact that all fusion categories are pivotal was conjectured by *Etingof, Ostrik* and *Nikshych*. Currently the best one can do for a general fusion category is a monoidal natural transformation between the identity functor and the fourth dualization functor $\mathbb{1}_{\mathbf{C}} \cong ****$.

Definition 2.5.7 (Categorical dimension). Consider a fusion category $\mathbf{C}$ and choose a natural isomorphism $\psi : \mathbb{1}_{\mathbf{C}} \cong **$. For every simple object $x \in \mathrm{ob}(\mathbf{C})$ one can define a dimension function, sometimes called the **norm squared**, in the following way:

$$|x|^2 := \mathrm{tr}(\psi_x)\,\mathrm{tr}((\psi_x^{-1})^*). \tag{2.39}$$

If $\mathbf{C}$ is pivotal, this becomes $|x|^2 = \dim_\psi(x)\dim_\psi(x^*)$. In particular, when $\mathbf{C}$ is spherical, this becomes $|x|^2 = \dim_\psi(x)^2$.

The categorical dimension, sometimes called the **Müger dimension**, is then defined as follows:

$$\dim(\mathbf{C}) := \sum_{x\in\mathcal{O}(\mathbf{C})} |x|^2\,, \tag{2.40}$$

where $\mathcal{O}(\mathbf{C})$ denotes the set of isomorphism classes of simple objects.

Remark 2.5.8. It should be noted that the above quantities do not depend on the choice of isomorphism $\psi_x : x \cong x^{**}$ since all of them only differ by a scale factor.

Property 2.5.9 (Nonzero dimension). For any fusion category $\mathbf{C}$ one has that $\dim(\mathbf{C}) \neq 0$. In particular, if $K = \mathbb{C}$, then $\dim(\mathbf{C}) \geq 1$ (since the norm squared of any simple object is then also positive).

Definition 2.5.10 (G-graded fusion category). A semisimple linear category $\mathbf{C}$ is said to have a G**-grading**, where G is a finite group, if it can be decomposed as follows:

$$\mathbf{C} \cong \bigoplus_{g\in G} \mathbf{C}_g\,, \tag{2.41}$$

where every $\mathbf{C}_g$ is linear and semisimple. A fusion category $\mathbf{C}$ is said to be a (G-)graded fusion category if it admits a G-grading such that $\mathbf{C}_g \otimes \mathbf{C}_h \subseteq \mathbf{C}_{gh}$ for all $g, h \in G$.

Example 2.5.11 (G-graded vector spaces). Define the category $\mathbf{Vect}_G^\omega$ as having the same objects and morphisms as $\mathbf{Vect}_G$, the category of G-graded vector spaces, but with the associator given by the the 3-cocycle $\omega \in H^3(G; K^\times)$.

Property 2.5.12. Any pointed fusion category is equivalent to a category of the form $\mathbf{Vect}_G^\omega$ for some G and $\omega \in H^3(G; K^\times)$.

Theorem 2.5.13 (Tannaka duality). *The category of modules of a weak Hopf algebra has the structure of a fusion category. Conversely, any fusion category can be obtained as the category of modules of a weak Hopf algebra.*

2.6 Ribbon and modular categories

Definition 2.6.1 (Ribbon structure). Consider a braided monoidal category $(\mathbf{C}, \otimes, \mathbf{1})$ with braiding σ. A **twist** or **balancing** is a natural endomorphism θ such that the following equation is satisfied for all $x, y \in \mathrm{ob}(\mathbf{C})$:

$$\theta_{x\otimes y} = (\theta_x \otimes \theta_y) \circ \sigma_{y,x} \circ \sigma_{x,y}. \tag{2.42}$$

If in addition $\mathbf{C}$ is rigid and the twist satisfies $\theta_{x^*} = (\theta_x)^*$ for all $x \in \mathrm{ob}(\mathbf{C})$, one speaks of a ribbon or **tortile** category.

Definition 2.6.2 (Drinfel'd morphism). Let $(\mathbf{C}, \otimes, \mathbf{1})$ be a rigid braided monoidal category with braiding σ. This structure admits a canonical natural automorphism $\mathrm{id}_{\mathbf{C}} \cong **$ defined as follows:

$$x \xrightarrow{\mathbb{1}_x \otimes \eta_{x^*}} x \otimes x^* \otimes x^{**} \xrightarrow{\sigma_{x,x^*} \otimes \mathbb{1}_{x^{**}}} x^* \otimes x \otimes x^{**} \xrightarrow{\varepsilon_x \otimes \mathbb{1}_{x^{**}}} x^{**}. \tag{2.43}$$

Property 2.6.3. Let $\mathbf{C}$ be a braided monoidal category. Consider the canonical natural isomorphism $u : \mathrm{id}_{\mathbf{C}} \cong **$ defined above. Any natural isomorphism $\psi : \mathrm{id}_{\mathbf{C}} \cong **$ can be written as $u \circ \theta$ where $\theta \in \mathrm{Aut}(\mathbb{1}_{\mathbf{C}})$. It is not hard to see that this natural isomorphism is monoidal (and, hence, pivotal) exactly when θ is a twist. If $\mathbf{C}$ is a fusion category then the pivotal structure is spherical if and only if θ determines a ribbon structure.

Definition 2.6.4 (Premodular category). A ribbon fusion category. Equivalently, a spherical braided fusion category.

Definition 2.6.5 (S-matrix). Given a premodular category $\mathbf{M}$ with braiding σ, the S-matrix is defined as follows:

$$S_{xy} := \mathrm{tr}(\sigma_{y,x} \circ \sigma_{x,y})\,, \tag{2.44}$$

where $x, y \in \mathcal{O}(\mathbf{M})$ are (isomorphism classes of) simple objects. Since in a premodular category there are only finitely many isomorphism classes of simple objects (denote this number by $\mathcal{I}$), one can see that S is a $\mathcal{I} \times \mathcal{I}$-matrix.

Definition 2.6.6 (Modular category[6]). A premodular category for which the S-matrix is invertible.

Property 2.6.7. Let $\mathbf{M}$ be a modular category with S-matrix S and define the following matrix:

$$E_{xy} := \begin{cases} 1 & x = y^* \\ 0 & \text{otherwise.} \end{cases} \tag{2.45}$$

The following relation with the categorical dimension of $\mathbf{M}$ is obtained:

$$S^2 = \dim(\mathbf{M})E. \tag{2.46}$$

Formula 2.6.8 (Verlinde). Consider a modular category $\mathbf{M}$ with S-matrix S. Let $\mathcal{O}(\mathbf{M})$ denote the set of isomorphism classes of simple objects and let dim denote the dimension associated to the spherical structure on $\mathbf{M}$. Using the formula

$$S_{xy}S_{xz} = \dim(x) \sum_{w\in\mathcal{O}(\mathbf{M})} N^w_{yz} S_{xw} \tag{2.47}$$

for all $x, y, z \in \mathcal{O}(\mathbf{M})$, one obtains the following important relation:

$$\sum_{w\in\mathcal{O}(\mathbf{M})} \frac{S_{wy}S_{wz}S_{wx^*}}{\dim(w)} = \dim(\mathbf{M})N^x_{yz}. \tag{2.48}$$

This property implies that the S-matrix of a modular category determines the fusion coefficients of the underlying fusion category.

2.7 Module categories

By categorifying the definition of a module over a ring **??**, one obtains the notion of a module category:

Definition 2.7.1 (Module category). Let $\mathbf{M}$ be a monoidal category. A left $\mathbf{M}$-module (category) is a category $\mathbf{C}$ equipped with a bilinear functor $\triangleright : \mathbf{M} \times \mathbf{C} \to \mathbf{C}$ together with natural isomorphisms that categorify the associativity and unit conditions of modules (these are also required to be compatible with the associator and unitors of $\mathbf{M}$).

Remark 2.7.2. Similar to how a G-set can be defined as a functor $\mathbf{B}G \to \mathbf{Set}$ (Property 1.6.20), one can define a module category as a 2-functor $\mathbf{BM} \to \mathbf{Cat}$.

Analogous to Definition 2.1.19 one can also define internal homs for module categories:

Definition 2.7.3 (Internal hom). Consider a left $\mathbf{M}$-module $\mathbf{C}$. Given two objects $x, y \in \mathrm{ob}(\mathbf{C})$ one defines their internal hom (if it exists) as the object $\underline{\mathrm{Hom}}(x, y) \in \mathrm{ob}(\mathbf{M})$ satisfying the following condition

$$\mathbf{C}(m \triangleright x, y) \cong \mathbf{M}\big(m, \underline{\mathrm{Hom}}(x, y)\big) \tag{2.49}$$

for all $m \in \mathrm{ob}(\mathbf{M})$.

Property 2.7.4. It should be noted that for the case $\mathbf{C} \equiv \mathbf{M}$, where the action is given by the tensor product in $\mathbf{M}$, one obtains Definition 2.1.19 as a particular case.

[6] "Modular tensor category" is often abbreviated as **MTC**.

2.8 Higher vector spaces

2.8.1 Kapranov-Voevodksy 2-vector spaces

The guiding principle for the definition of 2-vector spaces in this section will be the generalization of certain observations from studying the category **Vect** of ordinary vector spaces. Linear maps between vector spaces can (at least in finite dimensions) be represented as matrices with coefficients in the ground field K. Coincidentally this ground field is also the tensor unit in **Vect**. At the same time, all finite-dimensional vector spaces are isomorphic to spaces of the form K^n, where n is given by the dimension of the vector space.

Definition 2.8.1 (2-vector space). To define 2-vector spaces, *Kapranov* and *Voevodsky* lifted these observations to categories by replacing the ground field K by the category $\mathbf{Vect}_K$. To wit, $\mathbf{2Vect}_K$ is defined as the 2-category consisting of the following data:

- **Objects**: Finite products of the form $\mathbf{Vect}_K^n$.
- **1-morphisms**: Collections $\|A_{ij}\|$ of finite-dimensional K-vector spaces, called **2-matrices**.
- **2-morphisms**: Collections (f_{ij}) of linear maps between finite-dimensional K-vector spaces.

The multiplication (or composition) of 1-morphisms is defined in analogy to the multiplication of ordinary matrices, but where the usual sum and product are replaced by the direct sum and tensor product.

A seemingly more formal definition uses the concepts of *ring* and module categories:

Alternative Definition 2.8.2. A 2-vector space is a lax module category over **Vect** that is module-equivalent to $\mathbf{Vect}^n$ for some $n \in \mathbb{N}$. The 2-category **2Vect** is then defined as the 2-category with objects these 2-vector spaces, as 1-morphisms the associated **Vect**-module functors and as 2-morphisms the module natural transformations.

2.8.2 Baez-Crans 2-vector spaces

Definition 2.8.3 (2-vector space). A category internal to **Vect**. The morphism are **linear functor**, i.e. functors internal to **Vect**.

Remark 2.8.4. The above definition should not be confused with that of categories and functors enriched over **Vect**.

Example 2.8.5 (Ground field). The ground field K can be categorified to a 2-vector spaces by taking $K_0 = K_1 := K$ and $s = t = e := \mathbb{1}_K$. This object serves as a unit for the tensor product on $\mathbf{2Vect}_K$.

Property 2.8.6 (Chain complexes). There exists an equivalence between the (2-)categories of 2-vector spaces and 2-term chain complexes.

Sketch of construction. Given a 2-vector space (V_0, V_1), one can build a chain complex $C_\bullet$ as follows:

- $C_0 := V_0$,
- $C_1 := \ker(s)$, and
- $d := t|_{C_1}$,

where s, t are the source and target morphisms.

Remark 2.8.7. The equivalence (on the level of ordinary categories) is an instance of the Dold-Kan correspondence **??**.

Definition 2.8.8 (Arrow part). Consider a 2-vector space $V = (V_0, V_1)$. For any morphism $f \in V_1$ one defines the arrow part as follows:

$$\vec{f} := f - e(s(f))\,, \tag{2.50}$$

where e, s are the identity and source morphisms in V. Any map can thus be recovered from its arrow part and its source. This allows to identify a map $f \in V_1$ with the pair $\big(s(f), \vec{f}\big)$. Using arrow parts one can rewrite the composition law of morphisms in an intuitive way:

$$g \circ f = \big(s(f), \vec{f} + \vec{g}\big). \tag{2.51}$$

Definition 2.8.9 (Antisymmetric morphism). A natural morphism between n-linear functors in **2Vect** is said to be **completely antisymmetric** if its arrow part is completely antisymmetric.

2.9 Higher Lie theory

2.9.1 Lie superalgebras

Definition 2.9.1 (Internal Lie algebra). Let $(\mathbf{C}, \otimes, \mathbf{1})$ be a linear symmetric monoidal category with braiding σ. A Lie algebra internal to $\mathbf{C}$ is an object $L \in \mathrm{ob}(\mathbf{C})$ and a morphism

$$[\cdot,\cdot] : L \otimes L \to L$$

satisyfing the following conditions:

1. **Antisymmetry**: $[\cdot,\cdot] + [\cdot,\cdot] \circ \sigma_{L,L} = 0$, and
2. **Jacobi identity**: $[\cdot,[\cdot,\cdot]] + [\cdot,[\cdot,\cdot]] \circ \tau + [\cdot,[\cdot,\cdot]] \circ \tau^2 = 0\,,$

where $\tau = (\mathbb{1}_L \otimes \sigma_{L,L}) \circ (\sigma_{L,L} \otimes \mathbb{1}_L)$ denotes cyclic permutation.

Example 2.9.2 (Lie superalgebra). When using the braiding

$$\sigma(x \otimes y) = (-1)^{\deg(x)\deg(y)} y \otimes x \tag{2.52}$$

in **sVect**, a Lie superalgebra (also called a super Lie algebra) is obtained. More generally, in $\mathbb{Z}$-**Vect**, a Lie bracket of degree k is induced by the braiding

$$\sigma(x \otimes y) = (-1)^{(\deg(x)-k)(\deg(y)-k)} y \otimes x. \tag{2.53}$$

It is simply a Lie bracket on the k-suspension $\Pi^k V$.

Example 2.9.3 (dg-Lie algebras). Lie algebras internal to $\mathbf{Ch}_\bullet(\mathbf{Vect})$ or its generalization to graded vector spaces. Sometimes these are also called strict L_∞-algebras (see further below).

The following notion is a slight modification of the idea of a (graded) Poisson algebra **??**:

Definition 2.9.4 (Gerstenhaber algebra). A graded-commutative algebra equipped with a degree-1 Lie bracket that acts as a graded derivation:

$$[x, yz] = [x, y]z + (-1)^{\deg(x)(\deg(y)-1)} y[x, z]. \tag{2.54}$$

Definition 2.9.5 (Semistrict Lie 2-algebra). A (Baez-Crans) 2-vector space $L \equiv (L_0, L_1)$ equipped with the following morphisms:

- an antisymmetric bilinear functor $[\cdot,\cdot] : L \times L \to L$ (the **bracket**), and
- a completely antisymmetric trilinear natural isomorphism

$$J_{x,y,z} : [[x,y],z] \to [x,[y,z]] + [[x,z],y]\,, \tag{2.55}$$

  called the **Jacobiator**.

These structures are required to satisfy the *Jacobiator identity* (which is just the *Zamolodchikov tetrahedron equation*). If the Jacobiator is trivial, a **strict** Lie 2-algebra is obtained. By further relaxing the antisymmetry, one can obtain the fully weak version of Lie 2-algebras (see for example the work by *Roytenberg*).

From the previous section it follows that one can define (weak) Lie 2-algebras as 2-term chain complexes equipped with a coherent Lie bracket:

Alternative Definition 2.9.6 (Lie 2-algebra). Consider a 2-term chain complex in the category **FinVect**:

$$0 \longrightarrow L_1 \longrightarrow L_0 \longrightarrow 0. \tag{2.56}$$

This complex L is called a Lie 2-algebra if it comes equipped with the following structures:

- a chain map $[\cdot,\cdot] : L \otimes L \to L$ called the **bracket**,
- a chain homotopy $S : [\cdot,\cdot] \Rightarrow -[\cdot,\cdot] \circ \sigma$ called the **alternator**, and
- a chain homotopy

$$J : [\cdot,[\cdot,\cdot]] \Rightarrow [[\cdot,\cdot],\cdot] + [\cdot,[\cdot,\cdot]] \circ (\sigma \otimes \mathbb{1})\,, \tag{2.57}$$

  called the **Jacobiator**.

These chain homotopies are again required to satisfy higher coherence relations. From the previous definition it follows that the vanishing of the alternator implies that L is semistrict. Analogously, a Lie 2-algebra for which the Jacobiator vanishes is said to be **hemistrict**. Note that this definition of weak Lie 2-algebras, when translated to the 2-vector space setting, would imply that the alternator and Jacobiator are merely natural transformations (and not isomorphisms)!

2.9.2 Lie n-algebras

Definition 2.9.7 (Semistrict Lie ω-algebra). By replacing internal categories by internal *ω-categories* and by relaxing the Jacobiator identity up to coherent homotopy, i.e. up to a completely antisymmetric quadrilinear modification which in turn satisfies an identity up to higher multilinear transfors, one obtains the definition of L_∞-algebras. Similar to A_∞-algebras, these too can be obtained as algebras over a suitable operad (however, in this case the operad is "slightly" more complex: the cofibrant replacement of the *Lie operad*).

It can be shown that these structures are equivalent to the L_∞-algebras of *Stasheff* defined below.

[7] Also called a **strong(ly) homotopy Lie algebra** (abbreviated to **sh Lie algebra**).

Definition 2.9.8 (L_∞-algebra[7]). A graded vector space V equipped with a collection of morphisms $l_n : V^{\otimes n} \to V, n \in \mathbb{N}_0$ of degree $n-2$ subject to the relations

$$l_n(v_{\sigma(1)} \dots v_{\sigma(n)}) = \chi(\sigma; v_1, \dots, v_n) l_n(v_1 \dots v_n) \tag{2.58}$$

and

$$\sum_{\substack{i+j=n+1 \\ \sigma \in \mathrm{Unshuff}(i,j-1)}} (-1)^{i(j-1)} \chi(\sigma; v_1, \dots, v_n) l_i\big(l_j(v_{\sigma(1)} \cdots v_{\sigma(j)}) v_{\sigma(j+1)} \cdots v_{\sigma(n)}\big) = 0\,, \tag{2.59}$$

where Unshuff denotes the collection of unshuffles **??**.

The l_1 map turns the L_∞-algebra into a chain complex. The l_2 map is a generalized Lie bracket since it is (graded-)antisymmetric. Higher l_n's can be identified with the Jacobiator and its generalizations. In the next section a bottom-up approach will be given.

Remark 2.9.9. The definition can be rephrased in terms of graded maps $\hat{l}_n : \mathrm{Alt}^\bullet V \to V$.

Remark 2.9.10 (Curvature). The above definition can be generalized by including a nullary bracket l_0. Such L_∞-algebras are often said to be **curved**. The reason for this is that the coherence condition for l_0 says that

$$l_1 \circ l_1 = l_2(l_0, -). \tag{2.60}$$

This terminology stems from the situation where l_1 is identified with the exterior covariant derivative on an associated vector bundle (see Formula **??**).

Example 2.9.11 (Lie algebra). It can easily be checked that the L_∞-algebra with V concentrated in degree 1 is equivalent to the structure of an ordinary Lie algebra. Similarly one obtains the notion of a Lie n-algebra by truncating an L_∞-algebra at degree n.

Property 2.9.12. 2-term L_∞-algebras, or equivalently semistrict Lie 2-algebras, are in correspondence with isomorphism classes of tuples $(\mathfrak{g}, V, \rho, l_3)$ where $\mathfrak{g}$ is a Lie algebra, (V, ρ) is Lie algebra representation of $\mathfrak{g}$ and l_3 is a V-valued Lie algebra 3-cocycle (Section **??**).

Sketch of construction. Using the representation ρ, one can extend the Lie bracket from $\mathfrak{g}$ to the complex $0 \to V \to \mathfrak{g} \to 0$ through the formulas $[g, v] := \rho(g)v$ and $[v, g] := -[g, v]$. The cocycle condition for l_3 gives rise to the Jacobiator.

Example 2.9.13. If one chooses a finite-dimensional Lie algebra $\mathfrak{g}$ with the trivial representation on $\mathbb{R}$ (or, more generally, the underlying field of $\mathfrak{g}$), one obtains

$$H^3(\mathfrak{g}; \mathbb{R}) \cong \mathbb{R}. \tag{2.61}$$

The different classes can be represented by scalar multiples of the Killing cocycle (see Example **??**). For every such scalar $\lambda \in \mathbb{R}$, one denotes the resulting Lie 2-algebra by $\mathfrak{g}_\lambda$.

Lie algebras and L_∞-algebras can also be dually characterized in terms of their Chevalley-Eilenberg algebra (see Definition **??**):

Alternative Definition 2.9.14 (Lie algebra). Consider a finite-dimensional Lie algebra $\mathfrak{g}$. The transpose/dual of the Lie bracket $[\cdot, \cdot] : \mathfrak{g} \wedge \mathfrak{g} \to \mathfrak{g}$ is a morphism $\delta : \mathfrak{g}^* \to \mathfrak{g}^* \wedge \mathfrak{g}^*$:

$$\delta\omega(g, h) := \omega([g, h]). \tag{2.62}$$

In fact, it is not hard to see that this is exactly the Chevalley-Eilenberg differential of $\mathrm{CE}(\mathfrak{g})$. Conversely, given a semifree dgca $(\mathrm{Alt}^\bullet V^*, \mathrm{d})$, for some finite-dimensional vector space V, one obtains a finite-dimensional Lie algebra by restricting the differential to V^* and taking the transpose. In fact, the nilpotency condition $\mathrm{d}^2 = 0$ is equivalent to the Jacobi identity.

More generally, by passing to graded vector spaces of finite type concentrated in positive degree, one can characterize L_∞-algebras as semifree DGCAs:

Alternative Definition 2.9.15 (L_∞-algebra). The (graded) Leibniz rule implies that the differential δ is completely defined by its restriction to the generators $V^* \leq \mathrm{Alt}^\bullet V^*$. The differential can be decomposed as follows:

$$\delta t^a := -\sum_{k=1}^{\infty} \frac{1}{k!}[t_{a_1}, \dots, t_{a_k}]^a_k\, t^{a_1} \wedge \cdots \wedge t^{a_k}\,, \tag{2.63}$$

where the basis t^a of V^* is dual to the basis t_a of V. Because δ is of degree 1, the coefficients $[\cdots]^a_k$ define a multilinear operator $[\cdots]_k : \mathrm{Alt}^k V \to V$ of degree $n-1$ (some sources rephrase these brackets as morphism from the symmetric algebra $\mathrm{Sym}^\bullet V$, in which case their degree is just -1, cf. décalage **??**).

The nilpotency condition $\delta^2 = 0$ implies a list of (quadratic) relations on the brackets $[\cdots]_k$ (with $\mathrm{d} := [\cdot]_1$):

$$\begin{aligned}
\mathrm{d}^2 &= 0\\
\mathrm{d}[\cdot,\cdot]_2 &= [\mathrm{d}\cdot,\cdot]_2 + [\cdot,\mathrm{d}\cdot]_2\\
[[v_1,v_2],v_3]_2 + \text{cyc. perm.} &= \mathrm{d}[v_1,v_2,v_3]_3 - [\mathrm{d}v_1,v_2,v_3]_3 - [v_1,\mathrm{d}v_2,v_3]_3 - [v_1,v_2,\mathrm{d}v_3]_3\\
&\ \vdots
\end{aligned}$$

These relations can be interpreted as follows:

- d is a differential.
- d acts as a derivation with respect to the binary bracket.
- The Jacobi identity holds up to a chain homotopy (given by the ternary bracket).
- The higher relations are similar to the chain homotopy for the Jacobi identity.

When written out it in full detail it can be checked that this is exactly the definition of an L_∞-algebra.

Definition 2.9.16 (Maurer-Cartan element). An element a of an L_∞-algebra V that satisfies the equation

$$\sum_{k=0}^{\infty} \frac{1}{k!}[a, \dots, a]_k = 0. \tag{2.64}$$

For dg-Lie algebras this reduces to the ordinary Maurer-Cartan equation (see **??**):

$$\mathrm{d}a + \frac{1}{2}[a,a] = 0. \tag{2.65}$$

This is no coincidence since the complex $\Omega^\bullet(M) \otimes \mathfrak{g}$ of Lie algebra-valued differential forms on a smooth manifold M carries a canonical dg-Lie algebra structure.

2.10 Monoidal n-categories

Definition 2.10.1 (Monoidal n-category). In general one can define a monoidal n-category as a one-object $(n{+}1)$-category, similar to how monoidal categories give one-object bicategories by delooping 1.6.8. For the explicit definitions of monoidal bi- and tricategories, see the papers [88] and [87] respectively.

If one would put multiple compatible monoidal products on an n-category, by a version of the Eckmann-Hilton argument 1.5.1 all of these structures will be equivalent to a "commutative" monoidal structure. By increasing the number of compatible structures the "commutativity" can be increased. This gives rise to the following definition which is stated in different terms (based on the *delooping hypothesis*):

Definition 2.10.2 (k-tuply monoidal n-categories). A pointed $(n+k)$-category (strict or weak) in which all parallel j-arrows for $j < k$ are equivalent. These categories form an $(n+k+1)$-category $\boldsymbol{k}\mathbf{Mon}\boldsymbol{n}\mathbf{Cat}$.

Example 2.10.3. For small values of k and n the resulting structures coincide with some well-known constructions:

- $n = 0$:
  - $k = 0$: pointed set,
  - $k = 1$: monoid, and
  - $k \geq 2$: Abelian monoid.
- $n = 1$:
  - $k = 0$: "pointed" category[8],
  - $k = 1$: monoidal category,
  - $k = 2$: braided monoidal category, and
  - $k \geq 3$: symmetric monoidal category.

The stabilization occurring for higher values of k is the content of the following hypothesis[9] by *Baez & Dolan*:

Theorem 2.10.4 (Stabilization hypothesis). *For values $k \geq n+2$ the structure of a k-tuply monoidal n-category becomes maximally symmetric. Formally this means that the inclusion* $\boldsymbol{k}\mathbf{Mon}\boldsymbol{n}\mathbf{Cat} \hookrightarrow (\boldsymbol{n}+\mathbf{2})\mathbf{Mon}\boldsymbol{n}\mathbf{Cat}$ *becomes an equivalence.*

2.10.1 Relation with group cohomology

See Definition **??** or Section **??** for more information on group cohomology.

Consider a finite group G. As a first step, construct the group algebra $\mathbb{C}[G]$. As a monoid one can consider this object as a G-graded monoidal 0-category. The ordinary multiplication $g * h = gh$ can be twisted to obtain a monoid $\mathbb{C}[G]^\omega$ with multiplication

$$g * h := e^{i\omega(g,h)}gh. \tag{2.66}$$

If associativity is still required to hold on the nose, one is led to the property that ω is in fact a group 2-cocycle. The equivalence classes of such twisted group algebras are then in correspondence with the second cohomology class $H^2(G; \mathrm{U}(1))$.

Before really going to higher category theory, one should first reflect on the different structures in the previous paragraph. Since the monoid is regarded as a monoidal category (call it M for convenience), one has a bifunctor $\mu : M \otimes M \to M$ (given by the twisted multiplication) that differs from the ordinary group multiplication by a phase. This phase can be viewed categorically

[8] As in category with a specified element not as in category with a zero object 1.4.14.
[9] For certain definitions of higher categories this has been proven in full generality.

as a natural isomorphism between the "tensor products" in $\mathbb{C}[G]$ and M. At the same time, all the higher coherence conditions[10] (associativity, ...) are required to hold identically.

Now, drop the restriction on the product and take this to be a more general monoidal product bifunctor. To this end, replace the monoid $\mathbb{C}[G]$ by the G-graded monoidal category $\mathbf{Vect}_G$ and relax the associativity constraint up to a natural isomorphism α. When restricted to the simple objects of $\mathbf{Vect}_G$ this is given by a phase factor $e^{i\omega(g,h,k)}$. The pentagon condition for monoidal categories then implies that the function ω is a group 3-cocycle. In analogy with the case of monoids above, the equivalence classes of (twisted) monoidal structures on $\mathbf{Vect}_G$ is in correspondence with the third cohomology group $H^3(G;\mathrm{U}(1))$.

To go yet another step higher, move up a level in the chain of coherence conditions and relax the associativity constraint even more (for simplicity the one-object n-category point of view is adopted here). Instead of a natural isomorphism it only has to be an adjoint equivalence and at the same time the pentagon condition is replaced by an invertible modification. The coherence condition of this **pentagonator** then implies a classification of (twisted) monoidal bicategories, equivalent to $\mathbf{2Vect}^{\omega}_G$, by the fourth group cohomology $H^4(G;\mathrm{U}(1))$.

In a completely analogous way one can define more and more general structures. e.g. for monoidal tricategories one can translate the K_6-*associahedron* into an equation for an invertible perturbation which by the G-graded structure is equivalent to a group 5-cocycle.

Remark 2.10.5. This section is strongly related to the twisting procedure in n-dimensional Dijkgraaf-Witten theories.

[10]These can be parametrized by the *Stasheff polytopes/associahedra*.

Chapter 3

Probability Theory

The majority of this chapter uses the language of measure theory. For an introduction see Chapter **??**.

3.1 Probability

The Kolmogorov axioms of probability state when a set admits the definition of a probability theory:

Definition 3.1.1 (Kolmogorov axioms). A probability space (Ω, Σ, P) is a measure space **??** with normalized measure $P(X) = 1$. The set Ω is called the **sample space**.

Definition 3.1.2 (Random variable). Let (Ω, Σ, P) be a probability space. A function $X : \Omega \to \mathbb{R}$ is called a random variable if $\forall a \in \mathbb{R} : X^{-1}\big([a, +\infty[\big) = \{\omega \in \Omega \mid X(\omega) \geq a\} \in \Sigma$.

Definition 3.1.3 (σ-algebra of a random variable). Let X be a random variable defined on a probability space (Ω, Σ, P) and denote the Borel σ-algebra of $\mathbb{R}$ by $\mathcal{B}$. The following family of sets is a σ-algebra:

$$X^{-1}(\mathcal{B}) := \{S \in \Sigma \mid \exists B \in \mathcal{B} : S = X^{-1}(B)\}. \tag{3.1}$$

Notation 3.1.4. The σ-algebra generated by the random variable X is often denoted by $\mathcal{F}_X$, analogous to **??**.

Definition 3.1.5 (Event). Let (Ω, Σ, P) be a probability space. An element S of the σ-algebra Σ is called an event.

From this definition it is clear that a single possible outcome of a measurement can be a part of multiple events. So, although only one outcome can occur at the same time, multiple events can occur simultaneously.

Remark. The Kolmogorov axioms use the σ-algebra **??** of events instead of the power set **??** of all events. Intuitively this seems to mean that some possible outcomes are not treated as events. However, one can make sure that the σ-algebra still contains all "useful" events by using a "nice" definition of probability spaces.

Formula 3.1.6 (Union). Let A, B be two events. The probability that at least one of them occurs is given by the following formula:

$$P(A \cup B) = P(A) + P(B) + P(A \cap B). \tag{3.2}$$

Definition 3.1.7 (Disjoint events). Two events A and B are said to be disjoint if they cannot happen at the same time:

$$P(A \cap B) = 0. \tag{3.3}$$

Corollary 3.1.8. If A and B are disjoint, the probability that both A and B occur is just the sum of their individual probabilities.

Formula 3.1.9 (Complement). Let A be an event. The probability of A being false is denoted as $P(\overline{A})$ and is given by

$$P(\overline{A}) = 1 - P(A). \tag{3.4}$$

Corollary 3.1.10. From the previous equation and de Morgan's laws (**??**) and (**??**), one can derive the following formula:

$$P(\overline{A} \cap \overline{B}) = 1 - P(A \cup B). \tag{3.5}$$

3.2 Conditional probability

Definition 3.2.1 (Conditional probability). Let A, B be two events. The probability of A given that B is true is denoted as $P(A \mid B)$:

$$P(A \mid B) = \frac{P(A \cap B)}{P(B)}. \tag{3.6}$$

By interchanging A and B in previous equation and by observing that this has no effect on the quantity $P(A \cap B)$ the following important result can be derived:

Theorem 3.2.2 (Bayes). *Let A, B be two events.*

$$P(A \mid B) = \frac{P(B \mid A)P(A)}{P(B)}. \tag{3.7}$$

Formula 3.2.3. Let $(B_n)_{n\in\mathbb{N}}$ be a sequence of pairwise disjoint events. If $\bigsqcup_{n=1}^{\infty} B_n = \Omega$, the total probability of a given event A can be calculated as follows:

$$P(A) = \sum_{n=1}^{\infty} P(A \mid B_n)P(B_n). \tag{3.8}$$

Definition 3.2.4 (Independent events). Let A, B be two events. A and B are said to be independent if they satisfy the following relation:

$$P(A \cap B) = P(A)P(B). \tag{3.9}$$

Corollary 3.2.5. If A and B are two independent events, Bayes's theorem simplifies to

$$P(A \mid B) = P(A). \tag{3.10}$$

The above definition can be generalized to multiple events:

Definition 3.2.6. The events $A_1, \dots, A_n$ are said to be independent if for each choice of k events the probability of their intersection is equal to the product of their individual probabilities.

This definition can be stated in terms of σ-algebras:

Definition 3.2.7 (Independence). The σ-algebras $\mathcal{F}_1, \dots, \mathcal{F}_n$ defined on a probability space $(\Omega, \mathcal{F}, P)$ are said to be independent if for all choices of distinct indices $i_1, \dots, i_k$ and for all choices of sets $F_{i_n} \in \mathcal{F}_{i_n}$ the following equation holds:

$$P(F_{i_1} \cap \dots \cap F_{i_k}) = P(F_{i_1}) \cdots P(F_{i_k}). \tag{3.11}$$

Corollary 3.2.8. Let X, Y be two random variables. X and Y are independent if the σ-algebras generated by them are independent.

3.3 Probability distributions

Definition 3.3.1 (Probability distribution). Let X be a random variable defined on a probability space (Ω, Σ, P). The following function is a measure on the Borel σ-algebra of $\mathbb{R}$:

$$P_X(B) = P\big(X^{-1}(B)\big). \tag{3.12}$$

This measure is called the probability distribution of X.

Example 3.3.2 (Rademacher variable). A random variable on $\Omega = \{-1, 1\}$ with probability distribution $P_X(-1) = P_X(1) = \frac{1}{2}$.

Definition 3.3.3 (Density). Let $f \geq 0$ be an integrable function and recall Property **??**. The function f is called the density of the measure $P(A) := \int_A f\,d\lambda$ (with respect to the Lebesgue measure λ). If the measure is a probability measure, i.e. is normalized to 1, f is called a **probability density function**.

More generally, by the Radon-Nikodym theorem **??**, every absolutely continuous probability distribution P is of the form

$$P(A) = \int_A f\,d\lambda \tag{3.13}$$

for some integrable function f.

In the case where P is discrete, i.e. one works with respect to the counting measure, the Radon-Nikodym derivative is called the **probability mass function**. (In this compendium this function will also often be called the density function.)

Definition 3.3.4 (Cumulative distribution function). Consider a random variable X and its associated distribution P_X. The cumulative distribution function $F_X : \mathbb{R} \to [0, 1]$ is defined as follows:

$$F_X(a) := P_X\big(\{x \in \mathbb{R} \mid x \leq a\}\big). \tag{3.14}$$

Theorem 3.3.5 (Skorokhod's representation theorem). *Let $F : \mathbb{R} \to [0, 1]$ be a function that satisfies the following three properties:*

- *F is nondecreasing.*
- *$\lim_{x\to-\infty} F(x) = 0$ and $\lim_{x\to\infty} F(x) = 1$.*
- *F is right-continuous, i.e. $\lim_{y \nearrow y_0} F(y) = F(y_0)$.*

There exists a random variable $X : [0, 1] \to \mathbb{R}$ defined on the probability space $([0, 1], \mathcal{B}_{[0,1]}, \lambda_{[0,1]})$ such that $F = F_X$, where $\mathcal{B}_{[0,1]}$ is the Borel σ-algebra of $[0, 1]$ with its Euclidean topology.

The following theorem is a specific instance of the more general change-of-variables formula:

Theorem 3.3.6 (Theorem of the unconscious statistician). *Consider a random variable X on a probability space (Ω, Σ, P). The following equality holds for every integrable function $g \in L^1(\mathbb{R})$:*

$$\int_\Omega g \circ X\,dP = \int_{\mathbb{R}} g\,dP_X. \tag{3.15}$$

Remark 3.3.7. The name of this theorem stems from the fact that many scientists take this equality to be a definition of the expectation value $\mathrm{E}[g(X)]$. However, this equality should be proven since the measure on the right-hand side is the one belonging to the random variable X and not $g(X)$.

Formula 3.3.8. Consider an absolutely continuous probability function P defined on $\mathbb{R}^n$ and let f be the associated density. Let $g : \mathbb{R}^n \to \mathbb{R}$ be integrable with respect to P.

$$\int_{\mathbb{R}^n} g\,dP = \int_{\mathbb{R}^n} f(x)g(x)\,dx \tag{3.16}$$

Corollary 3.3.9. The previous formula together with Theorem 3.3.6 gives rise to

$$\int_\Omega g \circ X\,dP = \int_{\mathbb{R}^n} f_X(x)g(x)\,dx. \tag{3.17}$$

Formula 3.3.10. Let X be a random variable with density function f_X and let $g : \mathbb{R} \to \mathbb{R}$ be smooth and strictly monotone. The random variable $g \circ X$ has an associated density f_g given by

$$f_g(y) = f\big(g^{-1}(y)\big)\left|\frac{\mathrm{d}g^{-1}}{\mathrm{d}y}(y)\right|. \tag{3.18}$$

Weak convergence of measures **??** induces a notion for convergence of random variables:

Definition 3.3.11 (Convergence in distribution). A sequence $(X_n)_{n\in\mathbb{N}}$ of random variables is said to converge in distribution to a random variable Y if the associated cumulative distribution functions F_{X_n} converge pointwise to F_Y, i.e. $\lim_{n\to\infty} F_{X_n}(x) = F_Y(x)$ for all $x \in \mathbb{R}$, where F is continuous. This is equivalent to requiring that the associated probability measures P_{X_n} converge weakly to P_X (Definition **??**).

Notation 3.3.12. If a sequence $(X_n)_{n\in\mathbb{N}}$ converges in distribution to a random variable Y, this is often denoted by $X_n \xrightarrow{d} Y$. Sometimes the d (for "distribution") is replaced by the $\mathcal{L}$ (for "law").

Theorem 3.3.13 (Slutsky). *Let $(X_n)_{n\in\mathbb{N}}, (Y_n)_{n\in\mathbb{N}}$ be two sequences of random variables converging in probability to a random variable X and a constant c, respectively. The following statements hold:*

- $X_n + Y_n \xrightarrow{d} X + c$,
- $X_nY_n \xrightarrow{d} cX$, *and*
- $X_n/Y_n \xrightarrow{d} X/c$.

Definition 3.3.14 (Convergence in probability). A sequence $(X_n)_{n\in\mathbb{N}}$ of random variables on a metric space (Ω, d) is said to converge in probability to a random variable Y if for all $\varepsilon > 0$ the following statement holds:

$$\lim_{n\to\infty} \Pr\big(d(X_n, X) > \varepsilon\big) = 0. \tag{3.19}$$

Convergence in probability implies convergence in distribution.

3.4 Moments

3.4.1 Expectation value

Definition 3.4.1 (Expectation value). Let X be random variable defined on a probability space (Ω, Σ, P).

$$\mathrm{E}[X] := \int_\Omega X\,dP \tag{3.20}$$

Notation 3.4.2. Other common notations are $\langle X\rangle$ and μ_X. However, the latter might be confused with a general measure on the space X and will, therefore, not be used here.

Property 3.4.3 (Markov's inequality). Let X be a random variable. For every constant $a > 0$ the following inequality holds:

$$\Pr(X \geq a) \leq \frac{\mathrm{E}[X]}{a}. \tag{3.21}$$

Definition 3.4.4 (Moment of order r). The moment of order r is defined as the expectation value of the r^{th} power of X. By Equation (3.17) this becomes

$$\mathrm{E}[X^r] = \int_{\mathbb{R}} x^r f_X(x)\,dx. \tag{3.22}$$

Definition 3.4.5 (Central moment of order r).

$$\mathrm{E}[(X-\mu)^r] = \int_{\mathbb{R}} (x-\mu)^r f_X(x)\,dx \tag{3.23}$$

Remark 3.4.6. Moments of order n are determined by central moments of order $k \leq n$ and, conversely, central moments of order n are determined by moments of order $k \leq n$.

Definition 3.4.7 (Variance). The central moment of order 2 is called the variance:

$$\mathrm{Var}[X] := \mathrm{E}\big[(X-\mu)^2\big]. \tag{3.24}$$

Definition 3.4.8 (Standard deviation).

$$\sigma_X := \sqrt{V[X]} \tag{3.25}$$

Property 3.4.9. If $\mathrm{E}[|X|^n]$ is finite for some $n > 0$, then $\mathrm{E}\big[X^k\big]$ exists and is finite for all $k \leq n$.

Property 3.4.10 (Chebyshev's inequality). Let X be a nonnegative random variable. For every constant $a > 0$ the following inequality holds:

$$\Pr(|X - \mathrm{E}[X]| \geq a) \leq \frac{\mathrm{Var}[X]}{a^2}. \tag{3.26}$$

Definition 3.4.11 (Moment generating function).

$$M_X(t) := \mathrm{E}\big[e^{tX}\big] = \int_{\mathbb{R}} e^{tx} f_X(x)\,dx \tag{3.27}$$

Property 3.4.12. If the moment generating function exists, the moments $\mathrm{E}[X^n]$ can be expressed in terms of M_X:

$$\mathrm{E}[X^n] = \left.\frac{\mathrm{d}^n M_X(t)}{\mathrm{d}t^n}\right|_{t=0}. \tag{3.28}$$

Method 3.4.13 (Chernoff bound). The Chernoff bound for a random variable gives a bound on the tail probabilities. For all constants $\lambda > 0$, the Markov inequality implies the following statement:

$$\Pr(X \geq a) = \Pr\Big(e^{\lambda X} \geq e^{\lambda a}\Big) \leq \frac{\mathrm{E}\big[e^{\lambda X}\big]}{e^{\lambda a}}. \tag{3.29}$$

If one has more information about the moment generating function, the Chernoff bound can be used to obtain improved concentration inequalities by optimizing over λ.

Property 3.4.14 (Hoeffding's inequalities). Consider a collection of bounded, independent random variables $X_1, \dots, X_n$. Without loss of generality one can assume that they are bounded by the unit intevral, i.e. $0 \leq X_i \leq 1$. For every constant $\lambda \geq 0$ the following inequality holds:

$$\Pr\big(\overline{X} - \mathrm{E}\big[\overline{X}\big] \geq \lambda\big) \leq \exp\big(-2n\lambda^2\big). \tag{3.30}$$

If one can sharpen the bounds for the variables such that $X_i \in [a_i, b_i]$, then

$$\Pr\big(\overline{X} - \mathrm{E}\big[\overline{X}\big] \geq \lambda\big) \leq \exp\left(-\frac{2n^2\lambda^2}{\sum_{i=1}^n (b_i - a_i)^2}\right). \tag{3.31}$$

Definition 3.4.15 (Characteristic function).

$$\varphi_X(t) := \mathrm{E}\big[e^{itX}\big] \tag{3.32}$$

Property 3.4.16. The characteristic function has the following properties:

- $\varphi_X(0) = 1$,
- $|\varphi_X(t)| \leq 1$, and
- $\varphi_{aX+b}(t) = e^{itb}\varphi_X(at)$ for all $a, b \in \mathbb{R}$.

Formula 3.4.17. If $\varphi_X(t)$ is k times continuously differentiable, then X has a finite k^{th} moment and

$$\mathrm{E}\Big[X^k\Big] = \frac{1}{i^k}\frac{\mathrm{d}^k}{\mathrm{d}t^k}\varphi_X(0). \tag{3.33}$$

Conversely, if X has a finite k^{th} moment, then $\varphi_X(t)$ is k times continuously differentiable and the above formula holds.

Formula 3.4.18 (Inversion formula). Let X be a random variable. If the CDF of X is continuous at $a, b \in \mathbb{R}$, then

$$F_X(b) - F_X(a) = \lim_{c\to\infty} \frac{1}{2\pi} \int_{-c}^{c} \frac{e^{-ita} - e^{-itb}}{it} \varphi_X(t)\, dt. \tag{3.34}$$

Formula 3.4.19. If $\varphi_X(t)$ is integrable, the CDF is given by:

$$f_X(x) = \frac{1}{2\pi}\int_{\mathbb{R}} e^{-itx}\varphi_X(t)\, dt. \tag{3.35}$$

Remark 3.4.20. This formula implies that the density function and the characteristic function form a Fourier transform pair.

3.4.2 Correlation

Property 3.4.21. Two random variables X, Y are independent if and only if $\mathrm{E}[f(X)g(Y)] = \mathrm{E}[f(X)]\mathrm{E}[g(Y)]$ holds for all measurable bounded functions f, g.

The value $\mathrm{E}[XY]$ is equal to the inner product $\langle X \mid Y\rangle$ as defined in (**??**). It follows that independence of random variables implies orthogonality. To generalize this concept, the following notions are introduced:

Definition 3.4.22 (Centred random variable). Let X be a random variable with finite expectation value $\mathrm{E}[X]$. The centred random variable X_c is defined as $X_c = X - \mathrm{E}[X]$.

Definition 3.4.23 (Covariance). The covariance of two random variables X, Y is defined as follows:

$$\mathrm{cov}(X,Y) := \langle X_c \mid Y_c\rangle = \mathrm{E}[(X-\mathrm{E}[X])(Y-\mathrm{E}[Y])]. \tag{3.36}$$

Some basic math gives

$$\mathrm{cov}(X,Y) = \mathrm{E}[XY] - \mathrm{E}[X]\mathrm{E}[Y]. \tag{3.37}$$

Definition 3.4.24 (Correlation). The correlation of two random variables X, Y is defined as the cosine of the angle between X_c and Y_c:

$$\rho_{XY} := \frac{\mathrm{cov}(X,Y)}{\sigma_X\sigma_Y}. \tag{3.38}$$

Corollary 3.4.25. From Theorem 3.4.21 it follows that independent random variables are uncorrelated.

Corollary 3.4.26. If the random variables X and Y are uncorrelated, they satisfy $\mathrm{E}[XY] = \mathrm{E}[X]\mathrm{E}[Y]$.

Formula 3.4.27 (Bienaymé formula). Let $(X_n)_{n\in\mathbb{N}}$ be a sequence of independent (or uncorrelated) random variables. Their variances satisfy the following equation:

$$\mathrm{Var}\left[\sum_{i=1}^{\infty} X_i\right] = \sum_{i=1}^{\infty}\mathrm{Var}[X_i]. \tag{3.39}$$

3.4.3 Conditional expectation

Let (Ω, Σ, P) be a probability space. Consider a random variable $X \in L^2(\Omega,\Sigma,P)$ and a sub-σ-algebra $\mathcal{G} \subset \Sigma$. Property **??** implies that the spaces $L^2(\Sigma)$ and $L^2(\mathcal{G})$ are complete and, hence, the projection theorem **??** can be applied. For every $X \in L^2(\Sigma)$ there exists a random variable $Y \in L^2(\mathcal{G})$ such that $X - Y$ is orthogonal to $L^2(\mathcal{G})$. This has the following result:

$$\forall Z \in L^2(\mathcal{G}) : \langle X - Y \mid Z\rangle \equiv \int_\Omega (X-Y)Z\,dP = 0. \tag{3.40}$$

Since $\mathbb{1}_G \in L^2(\mathcal{G})$ for every $G \in \mathcal{G}$, Equation (**??**) can be rewritten as

$$\int_G X\,dP = \int_G Y\,dP \tag{3.41}$$

for all $G \in \mathcal{G}$. This leads to the following definition:

Definition 3.4.28 (Conditional expectation). Let (Ω,Σ,P) be a probability space and let $\mathcal{G}$ be a sub-σ-algebra of Σ. For every Σ-measurable random variable $X \in L^2(\Sigma)$ there exists a unique (up to a null set) random variable $Y \in L^2(\mathcal{G})$ that satisfies Equation (3.41) for every $G \in \mathcal{G}$. This variable Y is called the conditional expectation of X given $\mathcal{G}$ and it is denoted by $\mathrm{E}[X \mid \mathcal{G}]$:

$$\int_G \mathrm{E}[X \mid \mathcal{G}]\,dP = \int_G X\,dP. \tag{3.42}$$

Remark 3.4.29. Although this construction was based on orthogonal projections, one could as well have used the (signed) Radon-Nikodym theorem **??** since $G \mapsto \int_G X\,dP$ is absolutely continuous with respect to $P|_\mathcal{G}$.

Property 3.4.30. Let (Ω,Σ,P) be a probability space and consider a sub-σ-algebra $\mathcal{G}\subset\Sigma$. If the random variable X is $\mathcal{G}$-measurable, then

$$\mathrm{E}[X \mid \mathcal{G}] = X \text{ a.s.} \tag{3.43}$$

On the other hand, if X is independent of $\mathcal{G}$, then

$$\mathrm{E}[X \mid \mathcal{G}] = \mathrm{E}[X] \text{ a.s.} \tag{3.44}$$

3.5 Joint distributions

Definition 3.5.1 (Joint distribution). Let X, Y be two random variables defined on the same probability space (Ω, Σ, P) and consider the vector random variable $(X, Y) : \Omega \to \mathbb{R}^2$. The distribution of (X, Y) isa probability measure defined on the Borel algebra of $\mathbb{R}^2$ defined by

$$P_{(X,Y)}(B) = P\big((X, Y)^{-1}(B)\big). \tag{3.45}$$

Definition 3.5.2 (Joint density). If the probability measure from the previous definition can be written as

$$P_{(X,Y)}(B) = \int_B f_{(X,Y)}(x, y)\, dx\, dy \tag{3.46}$$

for some integrable $f_{(X,Y)}$, it is said that X and Y have a joint density.

Definition 3.5.3 (Marginal distribution). The distributions of the one-dimensional random variables is determined by the joint distribution:

$$P_X(A) = P_{(X,Y)}(A \times \mathbb{R}), \tag{3.47}$$

$$P_Y(A) = P_{(X,Y)}(\mathbb{R} \times A). \tag{3.48}$$

Corollary 3.5.4. If the joint density exists, the marginal distributions are absolutely continuous and the associated density functions are given by

$$f_X(x) = \int_{\mathbb{R}} f_{(X,Y)}(x, y)\, dy, \tag{3.49}$$

$$f_Y(y) = \int_{\mathbb{R}} f_{(X,Y)}(x, y)\, dx. \tag{3.50}$$

The converse, however, is not always true. The one-dimensional distributions can be absolutely continuous without the existence of a joint density.

Property 3.5.5 (Independence). Let X, Y be two random variables with joint distribution $P_{(X,Y)}$. X and Y are independent if and only if the joint distribution coincides with the product measure:

$$P_{(X,Y)} = P_X \otimes P_Y. \tag{3.51}$$

If X and Y are absolutely continuous, the previous properties also applies to the densities instead of the distributions.

Formula 3.5.6 (Sum of random variables). Consider two independent random variables X, Y and let $Z = X + Y$ denote their sum. The density f_Z is given by the following convolution:

$$f_Z(z) := f * g(z) = \int_{\mathbb{R}} g(x)h(z - x)\, dx = \int_{\mathbb{R}} g(z - y)h(y)\, dy, \tag{3.52}$$

where g, h denote the densities of X, Y respectively.

Formula 3.5.7 (Product of random variables). Consider two independent random variables X, Y and let $Z = XY$ denote their product. The density f_Z is given by

$$f_Z(z) = \int_{\mathbb{R}} g(x)h(z/x)\, \frac{dx}{|x|} = \int_{\mathbb{R}} g(z/y)h(y)\, \frac{dy}{|y|}, \tag{3.53}$$

where g, h denote the densities of X, Y respectively.

Corollary 3.5.8. Taking the Mellin transform **??** of both the positive and negative part of the above integrand (to be able to handle the absolute value) gives the following relation:

$$\mathcal{M}\{f\} = \mathcal{M}\{g\}\mathcal{M}\{h\}. \tag{3.54}$$

Formula 3.5.9 (Conditional density). Let X, Y be two random variables with joint density $f_{(X,Y)}$. The conditional density of Y given $X \in A$ is

$$h(y \mid X \in A) = \frac{\int_A f_{(X,Y)}(x, y)\, dx}{\int_A f_X(x)\, dx}. \tag{3.55}$$

For $X = \{a\}$ this equation is ill-defined since the denominator would become 0. However, it is possible to avoid this problem by formally setting

$$h(y \mid A = a) := \frac{f_{(X,Y)}(a, y)}{f_X(a)}, \tag{3.56}$$

where $f_X(a) \neq 0$. This last condition is nonrestrictive because the probability of having a measurement $(X, Y) \in \{(x, y) \mid f_X(x) = 0\}$ is 0 (for nonsingular measures). One can thus define the conditional probability of Y given $X = a$ as follows:

$$P(Y \in B \mid X = a) := \int_B h(y \mid X = a)\, dy. \tag{3.57}$$

Formula 3.5.10 (Conditional expectation).

$$\mathrm{E}[Y \mid X](\omega) = \int_{\mathbb{R}} yh\big(y \mid X(\omega)\big)\, dy \tag{3.58}$$

Let $\mathcal{F}_X$ denote the σ-algebra generated by the random variable X as before. Using Fubini's theorem one can prove that for all sets $A \in \mathcal{F}_X$ the following equality holds:

$$\int_A \mathrm{E}[Y \mid X]\, dP = \int_A Y\, dP. \tag{3.59}$$

This implies that the conditional expectation $\mathrm{E}[Y \mid X]$ on $\mathcal{F}_X$ coincides with Definition 3.4.28.

Applying Property 3.4.30 to the case $\mathcal{G} = \mathcal{F}_X$ gives the law of total expectation:

Property 3.5.11 (Law of total expectation[1]).

$$\mathrm{E}[\mathrm{E}[Y \mid X]] = \mathrm{E}[Y] \tag{3.60}$$

Theorem 3.5.12 (Bayes's theorem). *The conditional density can be computed without prior knowledge of the joint density:*

$$g(x \mid y) = \frac{h(y \mid x) f_X(x)}{f_Y(y)}. \tag{3.61}$$

3.6 Stochastic calculus

Definition 3.6.1 (Stochastic process). A sequence of random variables $(X_t)_{t \in T}$ for some index set T. In practice T will often be a totally ordered set, e.g. $(\mathbb{R}, \leq)$ in the case of a time series. This will be assumed from here on.

[1] Also called the **tower property**.

Definition 3.6.2 (Filtered probability space). Consider a probability space (Ω, Σ, P) together with a filtration **??** of Σ, i.e. a collection of σ-algebras $\mathbb{F} \equiv (\mathbb{F}_t)_{t\in T}$, such that $i \leq j \implies \mathbb{F}_i \subseteq \mathbb{F}_j$. The quadruple $(\Omega, \Sigma, \mathbb{F}, P)$ is called a filtered probability space.

Often the filtration is required to be exhaustive and separated (where $\emptyset$ is replaced by $\mathbb{F}_0 = \{\emptyset, \Omega\}$ since any σ-algebra has to contain the total space).

Definition 3.6.3 (Adapted process). A stochastic process $(X_t)_{t\in T}$ on a filtered probability space $(\Omega, \Sigma, \mathbb{F}, P)$ is said to be adapted to the filtration $\mathbb{F}$ if X_t is $\mathbb{F}_t$-measurable for all $t \in T$.

Definition 3.6.4 (Predictable process). A stochastic process $(X_t)_{t\in T}$ on a filtered probability space $(\Omega, \Sigma, \mathbb{F}, P)$ is said to be predictable if X_{t+1} is $\mathbb{F}_t$-measurable for all $t \in T$.

Definition 3.6.5 (Stopping time). Consider a random variable τ on filtered probability space $(\Omega, \Sigma, \mathbb{F}, P)$ where the codomain of τ coincides with the index set of $\mathbb{F}$. This variable is called a stopping time for $\mathbb{F}$ if

$$\{\tau \leq t\} \in \mathbb{F}_t \tag{3.62}$$

for all t. The stopping time is a "time indicator" that only depends on the knowledge of the process up to time $t \in T$.

3.6.1 Martingales

From here on the index set T will be $\mathbb{R}_+ \equiv [0, \infty[$ so that the index t can be interpreted as a true time parameter. The discrete case $T = \mathbb{N}$ can be obtained as the restriction of most definitions or properties and, if necessary, this will be made explicit.

Definition 3.6.6 (Martingale). Consider a filtered probability space $(\Omega, \Sigma, \mathbb{F}, P)$. A stochastic process $(X_t)_{t\in T}$ is called a martingale relative to $\mathbb{F}$ if it satisfies the following conditions:

1. $(X_t)_{t\in T}$ is adapted to $\mathbb{F}$.
2. Each random variable X_t is integrable, i.e. $X_t \in L^1(P)$ for all $t \geq 0$.
3. For all $t > s \geq 0 : \mathrm{E}[X_t \mid \mathbb{F}_s] = X_s$.

If the equality in the last condition is replaced by the inequality $\leq$ (resp. $\geq$), the stochastic process is called a **supermartingale** (resp. **submartingale**).

Example 3.6.7 (Doob martingale). Consider an integrable random variable X and a filtration $\mathbb{F}$. The associated Doob martingale (a martingale with respect to $\mathbb{F}$) is given by

$$Y_t := \mathrm{E}[X \mid \mathbb{F}_t]. \tag{3.63}$$

Property 3.6.8 (Doob-Ville inequality). Consider a càdlàg submartinagle $(X_t)_{t\in T}$.

$$\Pr\left(\sup_{t\leq\tau} X_t \geq C\right) \leq \frac{\mathrm{E}[\max(0, X_\tau)]}{C} \tag{3.64}$$

for all $C \geq 1$ and $\tau \in T$.

The following property generalizes the Hoeffding inequalities 3.4.14:

Property 3.6.9 (Hoeffding-Azuma inequality). Let $(X_n)_{n\in\mathbb{N}}$ be a (super)martingale with bounded differences, i.e. there exist constants $c_k > 0$ such that

$$|X_k - X_{k-1}| \leq c_k. \tag{3.65}$$

The following inequality holds for all $\lambda \geq 0$:

$$\Pr(X_N - X_0 \geq \lambda) \leq \exp\left(-\frac{\lambda^2}{2\sum_{i=1}^N c_i^2}\right). \tag{3.66}$$

A symmetric result for the lower tail holds for (sub)martingales. Moreover, if there exist predictable processes $(A_n)_{n\in\mathbb{N}}, (B_n)_{n\in\mathbb{N}}$ such that

$$A_k \leq X_k - X_{k-1} \leq B_k \tag{3.67}$$

and

$$B_k - A_k \leq c_k \tag{3.68}$$

for all $k \in \mathbb{N}$, the inequality can be sharpened:

$$\Pr(X_N - X_0 \geq \lambda) \leq \exp\left(-\frac{2\lambda^2}{\sum_{i=1}^N c_i^2}\right). \tag{3.69}$$

Now, consider a function $f : \Omega^n \to \mathbb{R}$ such that

$$\sup_{x_1,\ldots,x_n,x'_k} |f(x_1,\ldots,x_k,\ldots,x_n) - f(x_1,\ldots,x'_k,\ldots,x_n)| \leq c_k \tag{3.70}$$

for all $k \in \mathbb{N}$. By applying the above inequalities to the Doob martingale

$$Z_m := \mathrm{E}[f(X_1,\ldots,X_n) \mid X_1,\ldots,X_m], \tag{3.71}$$

one obtains the following inequality:

$$\Pr\big(f(X_1,\ldots,X_n) - \mathrm{E}[f] \geq \lambda\big) \leq \exp\left(-\frac{2\lambda^2}{\sum_{i=1}^n c_i^2}\right). \tag{3.72}$$

This inequality is sometimes called the **McDiarmid inequality**.

Theorem 3.6.10 (Doob decomposition). *Any integrable adapted process $(X_t)_{t\in T}$ can be decomposed as $X_t = X_0 + M_t + A_t$, where $(M_t)_{t\in T}$ is a martingale and $(A_t)_{t\in T}$ is a predictable process. These two processes are constructed iteratively as follows:*

$$A_0 = 0 \qquad M_0 = 0 \tag{3.73}$$

$$\Delta A_t = \mathrm{E}[\Delta X_t \mid \mathbb{F}_{t-1}] \qquad \Delta M_t = \Delta X_t - \Delta A_t. \tag{3.74}$$

Furthermore, $(X_t)_{t\in T}$ is a submartingale if and only if $(A_t)_{t\in T}$ is (almost surely) increasing.

Corollary 3.6.11. Consider the special case $X = Y^2$ for some martingale Y. One can show the following property:

$$\Delta A_t = \mathrm{E}\big[(\Delta Y_t)^2 \mid \mathbb{F}_{t-1}\big] \qquad \forall t \in \mathbb{R}_+. \tag{3.75}$$

The process $(A_t)_{t\in T}$ is often called the **quadratic variation process** of $(X_t)_{t\in T}$ and is denoted by $([X]_t)_{t\in T}$.

Definition 3.6.12 (Discrete stochastic integral[2]). Let $(M_n)_{n\in\mathbb{N}}$ be a martingale on a filtered probability space $(\Omega, \Sigma, \mathbb{F}, P)$ and let $(X_n)_{n\in\mathbb{N}}$ be a predictable stochastic process with respect to $\mathbb{F}$. The (discrete) stochastic integral of X with respect to M is defined as follows:

$$(X \cdot M)_t(\omega) := \sum_{i=1}^t X(\omega)_i \Delta M_i(\omega), \tag{3.76}$$

where $\omega \in \Omega$. For $t = 0$ the convention $(X \cdot M)_0 = 0$ is used.

[2]Sometimes called the **martingale transform**.

Property 3.6.13. If the process $(X_n)_{n\in\mathbb{N}}$ is bounded, the stochastic integral defines a martingale.

Property 3.6.14 (Itô isometry). Consider a martingale $(M_n)_{n\in\mathbb{N}}$ and a predictable process $(X_n)_{n\in\mathbb{N}}$. Using the Doob decomposition theorem one can show the following equality for all $n \geq 0$:

$$\mathrm{E}\Big[(X \cdot M)_n^2\Big] = \mathrm{E}\big[(X^2 \cdot [M])_n\big]. \tag{3.77}$$

It is this property that allows for the definition of integrals with respect to continuous martingales, since although the martingales are not in general of bounded variation (and hence do not induce a well-defined Lebesgue-Stieltjes integral), their quadratic variations are (e.g. the Wiener process).

3.6.2 Markov processes

Definition 3.6.15 (Markov process). A Markov process (or chain) is a stochastic process $(X_t)_{t\in T}$ adapted to a filtration $(\mathbb{F}_t)_{t\in T}$ such that

$$P(X_t \mid \mathbb{F}_s) = P(X_t \mid X_s) \tag{3.78}$$

for all $t, s \in T$. For discrete processes, the first-order Markov chains are the most common. These satisfy

$$P(X_t \mid X_{t-1}, \ldots, X_{t-r}) = P(X_t \mid X_{t-1}) \tag{3.79}$$

for all $t, r \in \mathbb{N}$.

3.7 Information theory

Definition 3.7.1 (Self-information). The self-information of an event x described by a distribution P is defined as follows:

$$I(x) := -\ln P(x). \tag{3.80}$$

This definition is modeled on the following (reasonable) requirements:

- Events that are almost surely going to happen, i.e. events x such that $P(x) = 1$, contain only little information: $I(x) = 0$.[3]
- Events that are very rare contain a lot of information.
- Independent events contribute additively to the information.

Definition 3.7.2 (Shannon entropy). The amount of uncertainty in a discrete distribution P is characterized by its (Shannon) entropy

$$H(P) := \mathrm{E}[I(X)] = -\sum_i P_i \ln(P_i). \tag{3.81}$$

Definition 3.7.3 (Kullback-Leibler divergence). Let P, Q be two probability distributions. The Kullback-Leibler divergence (or **relative entropy**) of P with respect to Q is defined as follows:

$$D_{\mathrm{KL}}(P\|Q) := \int_\Omega \log\left(\frac{P}{Q}\right) dP. \tag{3.82}$$

This quantity can be interpreted as the information gained when using the distribution P instead of Q. Instead of a base-10 logarithm, any other logarithm can be used since this simply changes the result by a (positive) scaling constant.

[3] And by extension $P(x) \approx 1 \implies I(x) \approx 0$.

Property 3.7.4 (Gibbs's inequality). By noting that the logarithm is a concave function and applying Jensen's equality **??**, one can prove that the Kullback-Leibler divergence is nonnegative:

$$D_{\text{KL}}(P\|Q) \geq 0. \tag{3.83}$$

Furthermore, the Kullback-Leibler divergence is zero if and only if P and Q are equal almost everywhere.

3.8 Extreme value theory

Definition 3.8.1 (Conditional excess). Consider a random variable X with distribution P. The conditional probability that X is larger than a given threshold is given by the conditional excess distribution:

$$F_u(y) = \Pr(X - u \leq y \mid X > u) = \frac{P(u+y) - P(u)}{1 - P(u)}. \tag{3.84}$$

Definition 3.8.2 (Extreme value distribution). The extreme value distribution is given by the following formula:

$$F(x;\xi) = \exp\Big(-(1 + x\xi)^{-1/\xi}\Big). \tag{3.85}$$

For $\xi = 0$ one can use the definition of the Euler number to rewrite the definition as

$$F(x;0) = \exp(-e^{-x}). \tag{3.86}$$

The number ξ is called the **extreme value index**.

Definition 3.8.3 (Maximum domain of attraction). The (maximum) domain of attraction of a distribution function H consist of all distribution functions F for which there exist sequences $(a_n > 0)_{n\in\mathbb{N}}$ and $(b_n)_{n\in\mathbb{N}}$ such that $F^n(a_nx + b_n) \longrightarrow H(x)$.

Theorem 3.8.4 (Fischer, Tippett & Gnedenko). *Consider a sequence of i.i.d. random variables with distribution F. If F lies in the domain of attraction of G, then G has the form of an extreme value distribution.*

Theorem 3.8.5 (Pickands, Balkema & de Haan). *Consider a sequence of i.i.d. random variables with conditional excess distribution F_u. If the distribution F lies in the domain of attraction of the extreme value distribution, the conditional excess distribution F_u converges to the generalised Pareto distribution when $u \longrightarrow \infty$.*

3.9 Copulas

Property 3.9.1 (Uniformization transform). Consider a continuous random variable X and let U be the result of the probability integral transformation, i.e. $U := F_X(X)$. This transformed random variable has a uniform cumulative distribution, i.e. $F_U(u) = u$.

Definition 3.9.2 (Copula). The joint cumulative distribution function of a random variable with uniform marginal distributions.

The following alternative definition is more analytic in nature:

Alternative Definition 3.9.3 (Copula). A function $C : [0,1]^d \to [0,1]$ satisfying the following properties:

1. **Normalization** $C(x_1, \ldots, x_d) = 0$ if any of the x_i is zero.
2. **Uniformity:** $C(1, 1, \ldots, x_i, 1, \ldots) = x_i$ for all $1 \leq i \leq d$.
3. d**-nondecreasing:** For every box $B = \prod_{1\leq i\leq d}[a_i, b_i] \subseteq [0,1]^d$ the C-volume is nonnegative:

$$\int_B dC := \sum_{\mathbf{z}\in\prod_i\{a_i,b_i\}} (-1)^{N_b(\mathbf{z})} C(\mathbf{z}) \geq 0\,, \tag{3.87}$$

where $N_B(\mathbf{z}) = \mathrm{Card}(\{i \mid a_i = z_i\})$.

Theorem 3.9.4 (Sklar). *For every joint distribution function H with marginals F_i there exists a unique copula C such that*

$$H(x_1, \ldots, x_d) = C\big(F_1(x_1), \ldots, F_d(x_d)\big). \tag{3.88}$$

Property 3.9.5 (Fréchet-Hoeffding bounds). Every copula $C : [0,1]^d \to [0,1]$ is bounded in the following way:

$$\max\left(\sum_{i=1}^d u_i - d + 1, 0\right) \leq C(u_1, \ldots, u_d) \leq \min_i u_i \tag{3.89}$$

for all $(u_1, \ldots, u_d) \in [0,1]^d$. Furthermore, the upper bound is sharp, i.e. $\min_i u_i$ is itself a copula.[4]

Definition 3.9.6 (Extreme value copula). A copula C for which there exists a copula $\widetilde{C}$ such that

$$\left[\widetilde{C}(u_1^{1/n}, \ldots, u_d^{1/n})\right]^n \longrightarrow C(u_1, \ldots, u_d) \tag{3.90}$$

for all $(u_1, \ldots, u_d) \in [0,1]^d$.

Property 3.9.7. A copula C is an extreme value copula if and only if it is stable in the following sense:

$$C(u_1, \ldots, u_d) = \left[C(u_1^{1/n}, \ldots, u_d^{1/n})\right]^n \tag{3.91}$$

for all $n \geq 1$.

3.10 Randomness ♣

This section is strongly related to Section **??** on computability theory.

Definition 3.10.1 (Kolmogorov randomness). Consider a *universal Turing machine* U. The **Kolmogorov complexity** $C(\kappa)$ of a finite bit string κ (with respect to U) is defined as

$$C(\kappa) := \min\{|\sigma| \mid \sigma \text{ is finite} \wedge U(\sigma) = \kappa\}. \tag{3.92}$$

A finite bit string is said to be Kolmogorov random (with respect to U) if there exists an integer $n \in \mathbb{N}$ such that $C(\kappa) \geq |\sigma| - n$.

Property 3.10.2. For every universal Turing machine there exists at least one Kolmogorov random string. This easily follows from the pigeonhole principle since for every $n \in \mathbb{N}$ there are 2^n strings of length n but only $2^n - 1$ programs of length less than n.

[4] The lower bound is only a copula for $d = 2$. In general this bound is only pointwise sharp.

Remark 3.10.3. Note that, although universal Turing machines can emulate each other, the randomness of a string is not absolute. Its randomness depends on the chosen machine.

It would be pleasing if this notion of randomness could easily be extended to infinite bit strings, for example by giving such a string the label random if there exists a uniform choice of constant k such that all initial segments of the string are k-random. However, by a result of *Martin-Löf*, there does not exist any string satisfying this condition.

3.11 Optimal transport

In this section a new notion of atomicity of measures will be used:

Definition 3.11.1. A measure on $\mathbb{R}^n$ is said to **give mass to small sets** if there exists a subset of *Hausdorff dimension* $n-1$ (or smaller) that has nonzero measure.

3.11.1 Kantorovich duality

The problem of optimal transport constitutes the search of the most cost efficient transportation scheme that connects a set of producers to a set of consumers. Assume that these are described by the probability spaces (X, Σ_X, μ_X) and (Y, Σ_Y, μ_Y), respectively.

Definition 3.11.2 (Cost function). A measurable function $X \times Y \to \overline{\mathbb{R}}$.

Definition 3.11.3 (Transportation scheme). A transportation scheme or **transference plan** is a joint distribution $\pi \in \mathbb{P}(X \times Y)$ whose marginals coincide with μ_X and μ_Y.

Definition 3.11.4 (Monge-Kantorovich problem). The optimal transportation scheme for a given cost function according to *Kantorovich* is the solution of the following optimization problem:

$$\inf_{\pi\in\mathbb{P}(X\times Y)} \mathrm{E}_\pi[c] = \inf_{\pi\in\mathbb{P}(X\times Y)} \int_{X\times Y} c(x,y)\,d\pi(x,y). \tag{3.93}$$

The original problem of optimal transportation was considered by *Monge*. However, he studied a restricted problem, where every producer only delivers to a unique consumer. In this case the joint distributions have a specific form, namely

$$\int_{X\times Y} c(x,y)\,d\pi(x,y) = \int_X c(x,T(x))\,d\mu_X(x) \tag{3.94}$$

for some measurable function $T : X \to Y$ such that $T_*\mu_X = \mu_Y$.

Example 3.11.5 (Finite state spaces). Consider the case where both X and Y are finite of the same size and are both equipped with the uniform distribution. In this case the joint distributions π can be represented by *bistochastic matrices*, i.e. matrices with nonnegative entries such that every column and every row sums to one. This also implies that the optimization problem reduces to a linear problem on a convex, compact subset. This allows one to use Choquet's theorem **??** to restrict the attention to the extremal points, which in this case are given by permutation matrices. So, the optimal solution is given by the optimal one-to-one pairing of producers and consumers.

Property 3.11.6 (Kantorovich duality). Let X, Y be Polish spaces **??** and consider a lower semicontinuous function $c : X \times Y \to \overline{\mathbb{R}^+}$ (Definition **??**). Denote by $\mathbb{P}_{\text{Borel}}(\mu,\nu)$ the space of Borel measures on $X \times Y$ whose marginals are given by μ_X and μ_Y. Moreover, denote by $\Phi_c \subseteq \mathcal{L}^1(X) \times \mathcal{L}^1(Y)$ the space of pairs of integrable functions satisfying

$$c_X(x) + c_Y(y) \leq c(x,y) \tag{3.95}$$

for μ_X-almost all $x \in X$ and μ_Y-almost all $y \in Y$. Then

$$\inf_{\pi\in\mathbb{P}_{\text{Borel}}(\mu,\nu)} \int_{X\times Y} c\,d\pi = \sup_{(c_X,c_Y)\in\Phi_c} \int_X c_X\,d\mu_X + \int_Y c_Y\,d\mu_Y \tag{3.96}$$

and the this problem admits a solution. Moreover, one can restrict the space of would-be solutions on the right-hand side to those that are also bounded and continuous without changing the solution.

Definition 3.11.7 (Kantorovich distance). Let X be a Polish space (or even Radon space) and consider a metric d on X. The Kantorovich(-Rubinstein) distance $\mathcal{T}_d$ between two Borel probability measures μ,ν on X is defined as the optimal transport cost between them:

$$\mathcal{T}_d(\mu,\nu) := \inf_{\pi\in\mathbb{P}_{\text{Borel}}(X\times X)} \int_{X\times X} d(x,x')\,d\pi(x,x'). \tag{3.97}$$

If the metric d is the one inducing the topology on X, one obtains the definition of the **Wasserstein 1-metric**. This metric can be generalized on general complete metric spaces by restricting to the subset of Radon measures.

Theorem 3.11.8 (Kantorovich-Rubinstein). *Kantorovich duality for $X = Y$ and c equal to the metric d implies that the Kantorovich distance is equivalently given by*

$$\mathcal{T}_d(\mu,\nu) = \sup\left\{\int_X \varphi\,d\mu - \int_X \varphi\,d\nu \,\middle|\, \varphi\in\mathrm{Lip}_1(X,d)\cap\mathcal{L}^1(\mu)\cap\mathcal{L}^1(\nu)\right\}. \tag{3.98}$$

Property 3.11.9 (Translation invariance). The Kantorovich distance is invariant under translations by finite measures.

Property 3.11.10. When $X = Y = \mathbb{R}^n$ with d the Euclidean metric, the Kantorovich distance admits yet another description. In this case the Lipschitz norm is equal to the supremum norm of the gradient. This gives

$$\mathcal{T}_d(\mu,\nu) = \inf\{\|\sigma\|_1 \mid \nabla\cdot\sigma = \mu-\nu\}, \tag{3.99}$$

where the condition on σ makes sense by the Riesz-Markov theorem **??**.

3.11.2 Convex costs

In this section cost functions of the form

$$c(x,y) = h(x-y) \tag{3.100}$$

for some convex function $h:\mathbb{R}^n\to\mathbb{R}$ are considered. Moreover, the function h will be assumed to be at least differentiable with locally Lipschitz gradient.

Definition 3.11.11 (c-concave function). A function $f:\mathbb{R}^n\to[-\infty,+\infty[$, not identically $-\infty$, is said to be c-concave if there exists a set $A\subset\mathbb{R}^n\times\mathbb{R}$ such that

$$f(x) = \inf_{(x',\lambda)\in A} c(x,x')+\lambda. \tag{3.101}$$

Theorem 3.11.12 (Gangbo-McCann). *If c is strictly convex and μ does not give mass to small sets, the Monge-Kantorovich problem has a a.s. unique minimizer $\pi = (\mathbb{1}\times T)_*\mu$ with*

$$T(x) = x - (\nabla h)^{-1}\nabla\psi(x) \tag{3.102}$$

for some h-concave function $\psi:\mathbb{R}^n\to\overline{\mathbb{R}}$.

Remark 3.11.13. If h is a strictly convex function of the distance $\|x - y\|$, the theorem has to be modified:

- If $\mu \perp \nu$, the theorem still holds.
- If the measures are not singular, one has to restrict to transportation schemes that fix the shared mass. In effect, one removes the shared mass from the problem to recover the previous case.

Note that if h is sufficiently differentiable, the inverse ∇h^{-1} is equal to the gradient of the Legendre transform by Property **??**.

3.11.3 Concave costs

In this section cost functions of the form

$$c(x, y) = g(\|x - y\|) \tag{3.103}$$

for some concave function $g : \mathbb{R} \to \mathbb{R}$ are considered.

Property 3.11.14. Let c be strictly concave. If the transportation cost is not everywhere infinite and if μ does not give mass to small sets, then:

- If $\mu \perp \nu$, there exists a unique optimal transport scheme such that $\nu = T_*\mu$ with
$$T(x) = x - (\nabla g)^{-1}\nabla\varphi(x) \tag{3.104}$$
for some *c-concave function* φ.
- If the measures are not singular, there still exists a unique optimum by restricting to those schemes that fix shared mass.

3.11.4 Densities

Property 3.11.15 (Continuity equation). Let X be a complete smooth manifold and consider a family $(T_t)_{0\leq t\leq 1}$ of locally Lipschitz diffeomorphism on X such that $T_0 = \mathbb{1}_X$ with associated vector fields v_t. If μ is a probability measure on X, the family $(\mu_t := T_{t,*}\mu)_{0\leq t\leq 1}$ uniquely satisfies the **continuity equation**:

$$\frac{\partial \mu_t}{\partial t} + \nabla \cdot (\mu_t v_t) = 0\,, \tag{3.105}$$

where the divergence of a measure is defined by duality.

Let $v : \mathbb{R}^n \to \mathbb{R}^n$ be an almost everywhere smooth vector field. This induces a linear, constant velocity flow as follows:

$$T_t(x) := x - tv(x). \tag{3.106}$$

If all T_t are diffeomorphisms, the Eulerian velocity field $v_t(x) := T_t^{-1}\big(v(x)\big)$ satisfies the Eulerian continuity equation:

$$\frac{\partial v_t}{\partial t} + (v_t \cdot \nabla)v_t = 0. \tag{3.107}$$

Formula 3.11.16. Given a solution of the continuity equation, the associated flow determines an optimal transport scheme for a cost function c if and only if

$$v_0 = -(\nabla c)^{-1}\nabla\psi \tag{3.108}$$

for some c-concave function ψ. Moreover, if $v_t = (\nabla c)^{-1}\nabla u$ for some function $u(t,x)$, then u satisfies the **Hamilton-Jacobi equation** with Hamiltonian c^*:

$$\frac{\partial u}{\partial t} + c^*\nabla u = 0. \tag{3.109}$$

In this section one considers absolutely continuous measures with respect to the Lebesgue measure on $\mathbb{R}^n$:

$$d\mu_X = \rho_0 dx \qquad\qquad d\mu_Y = \rho_1 dx. \tag{3.110}$$

The transport cost in the Monge problem can then be rewritten as

$$\int_{\mathbb{R}^n} c(x, T(x))\rho_0(x)\,dx \tag{3.111}$$

with

$$\int_{T^{-1}(A)} \rho_0(x)\,dx = \int_A \rho_1(x)\,dx \tag{3.112}$$

for all measurable $A \subset \mathbb{R}^n$. By the change-of-variables formula this (weak) integral equation is equivalent to the Jacobian equation for

$$\det\big(DT(x)\big)\rho_1\big(T(x)\big) = \rho_0(x). \tag{3.113}$$

Example 3.11.17 (Euclidean metric). If the cost function c is the square of the Euclidean distance, the optimal transport mapping T, called the **Brenier map**, is given by the gradient of a convex potential:

$$T(x) = \nabla\varphi(x)\,, \tag{3.114}$$

and the optimal cost is equal to the square of the **Wasserstein 2-metric**:

$$\mathcal{T}_{\|\cdot\|_2^2}(\rho_0,\rho_1) = \inf_{\pi\in\mathbb{P}_{\text{Borel}}(\rho_0,\rho_1)} \int_{\mathbb{R}^n} \|x - x'\|\,d\pi(x,x') = W_2^2(\rho_0,\rho_1). \tag{3.115}$$

Moreover, this minimum is unique a.e.

It can also be shown that the flow acts affinely:

$$\sigma_t(x) = t\nabla\Phi(x) + (1-t)x. \tag{3.116}$$

In fact, the affinity of the flow can be shown more generally:

Property 3.11.18. Consider the time-dependent Monge-Kantorovich problem. If the differential cost $c : \mathbb{R}^n \to \mathbb{R}$ is strictly convex, the flows are given by straight lines:

$$x_t = x + t(x' - x). \tag{3.117}$$

This situation can be generalized to (complete) smooth manifolds, where the minimizers of ℓ^p-costs are geodesics with arc length parametrization.

It is possible to relate optimal transport to mechanics (Section **??**) in the following way:

Method 3.11.19 (Benamou-Brenier formulation). Let ρ_0 and ρ_1 describe the density of particles in a system at time steps $t = 0$ and $t = 1$. Assume that there exists a time-dependent velocity field $v : \mathbb{R} \times \mathbb{R}^n \to \mathbb{R}^n$. These are related by the *continuity equation* **??**:

$$\frac{\partial \rho}{\partial t} + \nabla \cdot (\rho v) = 0. \tag{3.118}$$

The optimization problem now becomes minimizing the *action* or *kinetic energy*:

$$K(\rho, v) := \frac{1}{2} \int_{\mathbb{R}^n} \int_0^T \rho(t, x) \|v(t, x)\|^2 \, dt \, dx. \tag{3.119}$$

By making the change of variables $(\rho, v) \longrightarrow (\rho, m := \rho v)$, one obtains a convex problem with a linear constraint (the continuity equation).

Property 3.11.20. The infimum of the Benamou-Brenier action is equal (up to constant factors) to the square of the Wasserstein 2-metric and, hence, gives an equivalent characterization of the Monge-Kantorovich problem for the Euclidean distance.

3.12 Probability monads ♣

Definition 3.12.1 (Giry monad). Consider the category **Meas** of measurable spaces. On this space one can define a monad 1.3.17 that sends a set X to its collection of probability distributions equipped with the σ-algebra generated by all evaluation maps ev_U, where U runs over the measurable subsets of X. Measurable functions are sent to pushforwards.

The unit of the Giry monad $\mathbb{P}$ is defined by assigning Dirac measures:

$$\eta_X(x) := \delta_x. \tag{3.120}$$

The multiplication map is given by marginalization:

$$\mu_X(Q)(U) := \int_{P \in \mathbb{P}X} \mathrm{ev}_U(P) \, dQ. \tag{3.121}$$

Now, consider the Kleisli category 1.3.21 of the Giry monad. This category has the same objects as **Meas**, but the morphisms $X \to Y$ are given by measurable functions $X \to \mathbb{P}Y$. Functions of this type are called **Markov kernels** or **stochastic maps**. This Kleisly category is often denoted by **Stoch**. Composition in **Stoch** boils down to the **Chapman-Kolmogorov equation**:

$$(f \circ_{\mathbf{Stoch}} g)(U \mid x) = \int_Y f(U \mid y) \, dg(y \mid x). \tag{3.122}$$

Global elements are simply probability distributions.

A finitary (or countable) version exists too. In this case one considers the category **FinStoch** of finite sets with **stochastic matrices** as morphisms, i.e. matrices $f \in \mathrm{M}_{m,n}(\mathbb{R})$ such that $\sum_{j \leq n} f_{ij} = 1$. In this setting the Chapman-Kolmogorov equations becomes

$$gf(k \mid i) = \sum_{j \leq n} g(k \mid j) f(j \mid i)\,, \tag{3.123}$$

where the suggestive notation $f(j \mid i) := f_{ij}$ was again used.

By restricting to Radon measures, an important submonad is found:

Definition 3.12.2 (Kantorovich monad). Consider the category of complete metric spaces and (1-)Lipschitz functions. The Kantorovich monad assigns to every space its set of Radon measures with finite first moment equipped with the Kantorovich-Wasserstein distance.

3.12.1 Markov categories

The structure of the Giry monad and its Kleisli category can be generalized as follows:

Definition 3.12.3 (Markov category). A symmetric **semicartesian** category 2.1.2 in which every object is an internal commutative comonoid. The comultiplication and counit morphisms $\mathrm{copy}_X : X \to X \otimes X$ and $\mathrm{delete}_X : X \to \mathbf{1}$ are required to be compatible with the monoidal structure in the following way:

$$\mathrm{copy}_{X\otimes Y} = (\mathbb{1}_X \otimes \sigma_{X,Y} \otimes \mathbb{1}_Y) \circ (\mathrm{copy}_X \otimes \mathrm{copy}_Y) \tag{3.124}$$

$$\mathrm{copy}_{\mathbf{1}} = \lambda_{\mathbf{1}}^{-1} \tag{3.125}$$

for all $X, Y \in \mathrm{ob}(\mathbf{M})$, where σ denotes the braiding. It can be shown that the deletion morphism is natural in X (the copy morphism is not).

Alternative Definition 3.12.4. Instead of requiring $\mathbf{M}$ to be semicartesian one can also add the following coherence conditions:

$$\mathrm{delete}_X \circ f = \mathrm{delete}_X \tag{3.126}$$

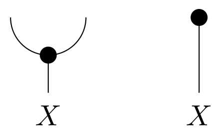
for any endomorphism $f : X \to X$ and

$$\mathrm{delete}_{X\otimes Y} = \mathrm{delete}_X \otimes \mathrm{delete}_Y \tag{3.127}$$

$$\mathrm{delete}_{\mathbf{1}} = \mathbb{1}_{\mathbf{1}} \tag{3.128}$$

for all objects $X, Y \in \mathrm{ob}(\mathbf{M})$, where λ denotes the left unitor.

Notation 3.12.5 (Graphical calculus). As is common for monoidal categories, one can also devise a graphical calculus for Markov categories. The copy and delete morphisms are denoted as follows:

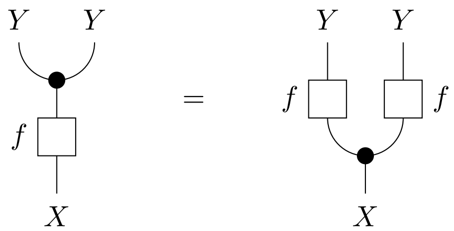


From here on, graphical calculus will be adopted.

Remark 3.12.6 (Strictness). Depending on the strictness of the monoidal structure, the above coherence conditions can be weakened. If $\mathbf{M}$ is strict monoidal, the conditions on $\mathbf{1}$ hold automatically. Moreover, in this case 3.126 follows from 3.124.

Example 3.12.7. Every Cartesian category is Markov in a unique way.

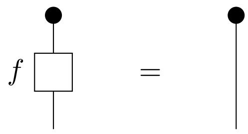
Definition 3.12.8 (Deterministic morphism). A morphism $f : X \to Y$ in a Markov category that satisfies



(3.129)

The name stems from the following intuition: if the morphism would be a random transformation of the input, the two applications on the right-hand side might act differently and, hence, the result would not coincide with the two (identical) copies on the left-hand side.

Definition 3.12.9 (Causal morphism). If the monoidal unit is not required to be terminal, and only coherence conditions (3.127) and (3.128) are required, the resulting structure is sometimes called a **CD**-category (“CD” stands for copy and discard). A morphism f in a CD-category is said to be causal or **terminal** if

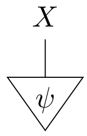


(3.130)

It follows that a Markov category is a CD-category in which all morphisms are causal.

By generalizing the situation of the Giry monad, distributions in a Markov category can be defined:

Definition 3.12.10 (Distribution). Consider a Markov category $(\mathbf{M}, \otimes, \mathbf{1})$. A distribution ψ on an object $X \in \mathrm{ob}(\mathbf{M})$ is a global element $\psi : \mathbf{1} \to X$. Graphically this is depicted as follows:



(3.131)

The category **Stoch**, the motivating example of Markov categories, is clearly not Cartesian. If it were, there would be a bijection between joint distributions $\mathbf{1} \to X \otimes Y$ and their marginals $\mathbf{1} \to X$ and $\mathbf{1} \to Y$. However, it is well-known that, in general, probability distributions are not uniquely determined by their marginals.

To take care of countable samples and asymptotic behaviour, it is necessary to extend the monoidal structure to sequences.

Definition 3.12.11 (Infinite tensor product). Consider a Markov category $\mathbf{M}$. For every two finite subsets $A \subseteq B \subset \mathbb{N}$, one obtains a morphism

$$\bigotimes_{i\in B} X_i \to \bigotimes_{j\in A} X_j \tag{3.132}$$

by applying the deletion morphism to all objects X_k for $k \in B\backslash A$. The infinite tensor product $\bigotimes_{n\in\mathbb{N}} X_n$ is defined as the inverse/filtered limit of this diagram if

1. it exists, and
2. is preserved by all $- \otimes Y$ for $Y \in \mathrm{ob}(\mathbf{M})$.

If all projection morphisms are deterministic, it is called a **Kolmogorov tensor product**.

List of Symbols

The following symbols are used throughout the summary:

Abbreviations	
PL	piecewise-linear
Operations	
e	identity element of a group
$\Gamma(E)$	set of global sections of a fibre bundle E
Par_t^γ	parallel transport map with respect to the curve γ
$X \pitchfork Y$	Transversally intersecting manifolds.
$X \times Y$	cartesian product of the sets X and Y
$\mathbb{1}_X$	identity morphism on the object X
$\approx$	is approximately equal to
$\hookrightarrow$	is included in
$\cong$	is isomorphic to
$\mapsto$	mapsto
Objects	
$C_p^\infty(M)$	ring of smooth functions $f : M \to \mathbb{R}$ on a neighbourhood of $p \in M$
Diff	category of smooth manifolds
D^n	standard n-disk
$\mathcal{D}_X$	sheaf of differential operators
$\mathrm{Hol}_p(\omega)$	holonomy group at the point p with respect to the principal connection ω
$\mathbf{Man}^p$	category of C^p-manifolds
S^n	standard n-sphere
T^n	standard n-torus (the n-fold Cartesian product of S^1)
$\mathbf{Vect}(X)$	category of vector bundles over a manifold X
$\emptyset$	empty set
$[a, b]$	closed interval
$]a, b[$	open interval
$\Omega^k(M)$	$C^\infty(M)$-module of differential k-forms on the manifold M
$\mathfrak{X}(M)$	$C^\infty(M)$-module of vector fields on the manifold M

Bibliography

[1] John C. Baez and Urs Schreiber. Higher gauge theory. 2005. arXiv:math/0511710.

[2] Urs Schreiber. *From Loop Space Mechanics to Nonabelian Strings*. PhD thesis, 2005.

[3] Richard Sanders. Commutative spectral triples & the spectral reconstruction theorem.

[4] John Baez and Alexander Hoffnung. Convenient categories of smooth spaces. *Transactions of the American Mathematical Society*, 363(11):5789--5825, 2011.

[5] John C. Baez and Aaron D. Lauda. Higher-dimensional algebra V: 2-groups. 2003. arXiv:math/0307200.

[6] Thomas Augustin, Frank PA Coolen, Gert De Cooman, and Matthias CM Troffaes. *Introduction to imprecise probabilities*. John Wiley & Sons, 2014.

[7] Tetsuji Miwa, Michio Jimbo, Michio Jimbo, and E Date. *Solitons: Differential Equations, Symmetries and Infinite-dimensional Algebras*, volume 135. Cambridge University Press, 2000.

[8] Vladimir I. Arnol'd. *Mathematical Methods of Classical Mechanics*, volume 60. Springer Science & Business Media, 2013.

[9] Edwin J. Beggs and Shahn Majid. *Quantum Riemannian Geometry*. Springer, 2020.

[10] Marc Henneaux and Claudio Teitelboim. *Quantization of Gauge Systems*. Princeton university press, 1992.

[11] Mark Hovey. *Model Categories*. Number 63. American Mathematical Soc., 2007.

[12] Gregory M. Kelly. *Basic Concepts of Enriched Category Theory*, volume 64. CUP Archive, 1982.

[13] Vladimir Vovk, Alex Gammerman, and Glenn Shafer. *Algorithmic Learning in a Random World*. Springer Science & Business Media, 2005.

[14] Mukund Rangamani and Tadashi Takayanagi. *Holographic Entanglement Entropy*. Springer, 2017.

[15] Saunders Mac Lane. *Categories for the Working Mathematician*, volume 5. Springer Science & Business Media, 2013.

[16] Peter T. Johnstone. *Topos Theory*. Dover Publications, 2014.

[17] Shun-ichi Amari. *Information Geometry and Its Applications*. Springer Publishing Company, Incorporated, 2016.

[18] Charles W. Misner, Kip S. Thorne, and John A. Wheeler. *Gravitation*. Princeton University Press, 2017.

[19] Carlo Rovelli and Francesca Vidotto. *Covariant Loop Quantum Gravity: An Elementary Introduction to Quantum Gravity and Spinfoam Theory.* Cambridge University Press, 2014.

[20] Richard W. Sharpe. *Differential Geometry: Cartan's Generalization of Klein's Erlangen Program*, volume 166. Springer Science & Business Media, 2000.

[21] John C. Baez, Irving E. Segal, and Zhengfang Zhou. *Introduction to Algebraic and Constructive Quantum Field Theory.* Princeton University Press, 2014.

[22] Raoul Bott and Loring W. Tu. *Differential Forms in Algebraic Topology.* Graduate Texts in Mathematics. Springer New York, 1995.

[23] The Univalent Foundations Program. *Homotopy Type Theory: Univalent Foundations of Mathematics.* `https://homotopytypetheory.org/book`, Institute for Advanced Study, 2013.

[24] Bruce Blackadar. *Operator Algebras: Theory of C∗-Algebras and von Neumann Algebras.* Springer, 2013.

[25] Marek Capinski and Peter E. Kopp. *Measure, Integral and Probability.* Springer Science & Business Media, 2013.

[26] Georgiev Svetlin. *Theory of Distributions.* Springer, 2015.

[27] Gerd Rudolph and Matthias Schmidt. *Differential Geometry and Mathematical Physics: Part II. Fibre Bundles, Topology and Gauge Fields.* Springer, 2017.

[28] Martin Schottenloher. *A Mathematical Introduction to Conformal Field Theory*, volume 759. 2008.

[29] Dusa McDuff and Deitmar Salamon. *Introduction to Symplectic Topology.* Oxford Graduate Texts in Mathematics. Oxford University Press, 2017.

[30] John C. Baez and Peter May. *Towards Higher Categories*, volume 152 of *IMA Volumes in Mathematics and its Applications.* Springer, 2009.

[31] Mikhail. M. Kapranov and Vladimir A. Voevodsky. *2-categories and Zamolodchikov Tetrahedra Equations*, volume 56 of *Proc. Sympos. Pure Math.* Amer. Math. Soc., Providence, RI, 1994.

[32] Geoffrey Compère. *Advanced Lectures on General Relativity*, volume 952. Springer, 2019.

[33] Pavel Etingof, Shlomo Gelaki, Dmitri Nikshych, and Victor Ostrik. *Tensor Categories*, volume 205. American Mathematical Soc., 2016.

[34] David Mumford. *The Red Book of Varieties and Schemes: Includes the Michigan Lectures (1974) on Curves and Their Jacobians*, volume 1358. Springer Science & Business Media, 1999.

[35] Charles A. Weibel. *An Introduction to Homological Algebra.* Cambridge Studies in Advanced Mathematics. Cambridge University Press, 1994.

[36] Peter J. Hilton and Urs Stammbach. *A Course in Homological Algebra.* Springer.

[37] Jean-Luc Brylinski. *Loop Spaces, Characteristic Classes and Geometric Quantization.* Birkhauser.

[38] Antoine Van Proeyen and Daniel Freedman. *Supergravity.* Cambridge University Press.

[39] William S. Massey. *A Basic Course in Algebraic Topology.* Springer.

[40] Michael E. Peskin and Daniel V. Schroeder. *An Introduction to Quantum Field Theory*. Westview Press.

[41] Nadir Jeevanjee. *An Introduction to Tensors and Group Theory for Physicists*. Birkhauser.

[42] Yvonne Choquet-Bruhat, Cecile DeWitt-Morette, and Margaret Dillard-Bleick. *Analysis, Manifolds and Physics, Part 1: Basics*. North-Holland.

[43] Yvonne Choquet-Bruhat and Cecile DeWitt-Morette. *Analysis, Manifolds and Physics, Part 2*. North-Holland.

[44] Herbet Goldstein, John L. Safko, and Charles P. Poole. *Classical Mechanics*. Pearson.

[45] Franco Cardin. *Elementary Symplectic Topology and Mechanics*. Springer.

[46] Walter Greiner and Joachim Reinhardt. *Field Quantization*. Springer.

[47] Walter Greiner. *Quantum Mechanics*. Springer.

[48] B. H. Bransden and Charles J. Joachain. *Quantum Mechanics*. Prentice Hall.

[49] Heydar Radjavi and Peter Rosenthal. *Invariant Subspaces*. Dover Publications.

[50] Max Karoubi. *K-Theory: An Introduction*. Springer.

[51] Damien Calaque and Thomas Strobl. *Mathematical Aspects of Quantum Field Theories*. Springer, 2015.

[52] Ivan Kolar, Peter W. Michor, and Jan Slovak. *Normal Operations in Differential Geometry*. Springer.

[53] Stephen B. Sontz. *Principal Bundles: The Classical Case*. Springer.

[54] Stephen B. Sontz. *Principal Bundles: The Quantum Case*. Springer.

[55] William Fulton and Joe Harris. *Representation Theory: A First Course*. Springer.

[56] Peter Petersen. *Riemannian Geometry*. Springer.

[57] Charles Nash and Siddharta Sen. *Topology and Geometry for Physicists*. Dover Publications.

[58] Ian M. Anderson. *The Variational Bicomplex*.

[59] Joel Robbin and Dietmar Salamon. The maslov index for paths. *Topology*, 32(4):827--844, 1993.

[60] Nima Moshayedi. 4-manifold topology, donaldson-witten theory, floer homology and higher gauge theory methods in the BV-BFV formalism. 2021.

[61] Edward Witten. Global anomalies in string theory. In *Symposium on Anomalies, Geometry, Topology*, 6 1985.

[62] F.A. Berezin and M.S. Marinov.

[63] Paul A. M. Dirac. Generalized Hamiltonian dynamics. *Canadian Journal of Mathematics*, 2:129–148, 1950.

[64] Angelo Vistoli. Notes on Grothendieck topologies, fibered categories and descent theory. *arXiv:math/0412512*, 2004.

[65] Emily Riehl and Dominic Verity. The theory and practice of Reedy categories. *Theory and Applications of Categories*, 29, 2013.

[66] Emily Riehl. Homotopical categories: From model categories to $(\infty, 1)$-categories. 2019. arXiv:1904.00886.

[67] Floris Takens. A global version of the inverse problem of the calculus of variations. *Journal of Differential Geometry*, 14(4):543--562, 1979.

[68] John C. Baez and Alissa S. Crans. Higher-dimensional algebra VI: Lie 2-algebras. 2003. arXiv:math/0307263.

[69] Edward Witten. Supersymmetry and Morse theory. *J. Diff. Geom*, 17(4):661--692, 1982.

[70] Jade Master. Why is homology so powerful? 2020. arXiv:2001.00314.

[71] Marcus Berg, Cécile DeWitt-Morette, Shangjr Gwo, and Eric Kramer. The Pin groups in physics: C, P and T. *Reviews in Mathematical Physics*, 13(08):953--1034, 2001.

[72] Arindam Banerjee, Srujana Merugu, Inderjit S Dhillon, and Joydeep Ghosh. Clustering with Bregman divergences. *Journal of Machine Learning Research*, 6(Oct):1705--1749, 2005.

[73] Jean-Daniel Boissonnat, Frank Nielsen, and Richard Nock. Bregman Voronoi diagrams. *Discrete & Computational Geometry*, 44(2):281--307, 2010.

[74] Richard Palais. The symmetries of solitons. *Bulletin of the American Mathematical Society*, 34(4):339--403, 1997.

[75] Michael F. Atiyah. Topological quantum field theory. *Publications Mathématiques de l'IHÉS*, 68:175--186, 1988.

[76] Jens Eisert, Christoph Simon, and Martin B Plenio. On the quantification of entanglement in infinite-dimensional quantum systems. *Journal of Physics A: Mathematical and General*, 35(17):3911--3923, 2002.

[77] Benoît Tuybens. Entanglement entropy of gauge theories. 2017.

[78] Lotfi A. Zadeh. Fuzzy sets. *Information and Control*, 8(3):338--353, 1965.

[79] John C. Baez, Alexander E. Hoffnung, and Christopher Rogers. Categorified symplectic geometry and the classical string. *Communications in Mathematical Physics*, 293:701--725, 2010.

[80] Charles Rezk. A model for the homotopy theory of homotopy theory. *Transactions of the American Mathematical Society*, 353(3):973--1007, 2001.

[81] Glenn Shafer and Vladimir Vovk. A tutorial on conformal prediction. *J. Mach. Learn. Res.*, 9:371-–421, 2008.

[82] Victor Chernozhukov, Kaspar Wüthrich, and Zhu Yinchu. Exact and robust conformal inference methods for predictive machine learning with dependent data. In *Proceedings of the 31st Conference On Learning Theory*, volume 75 of *Proceedings of Machine Learning Research*, pages 732--749. PMLR, 2018.

[83] Peter May. A note on the splitting principle. *Topology and Its Applications*, 153(4):605--609, 2005.

[84] Irina Markina. Group of diffeomorphisms of the unit circle as a principal $U(1)$-bundle.

[85] Sjoerd E. Crans. Localizations of transfors. 1998.

[86] Tom Leinster. Basic bicategories. 1998. arXiv:math/9810017.

[87] Alexander E. Hoffnung. Spans in 2-categories: A monoidal tricategory. 2011. arXiv:1112.0560.

[88] Eugenia Cheng and Nick Gurski. The periodic table of n-categories for low dimensions ii: Degenerate tricategories. 2007. arXiv:0706.2307.

[89] Mehmet B. Şahinoğlu, Dominic J. Williamson, Nick Bultinck, Michael Mariën, Jutho Haegeman, Norbert Schuch, and Frank Verstraete. Characterizing topological order with matrix product operators. 2014. arXiv:1409.2150.

[90] Dominic J. Williamson, Nick Bultinck, Michael Mariën, Mehmet B. Şahinoğlu, Jutho Haegeman, and Frank Verstraete. Matrix product operators for symmetry-protected topological phases: Gauging and edge theories. *Phys. Rev. B*, 94, 2016.

[91] Guifré Vidal. Efficient classical simulation of slightly entangled quantum computations. *Phys. Rev. Lett.*, 91, 2003.

[92] Aaron D. Lauda and Hendryk Pfeiffer. Open–closed strings: Two-dimensional extended TQFTs and Frobenius algebras. *Topology and its Applications*, 155(7):623--666, 2008.

[93] Domenico Fiorenza. An introduction to the Batalin-Vilkovisky formalism. 2004. arXiv:math/0402057v2.

[94] Stefan Cordes, Gregory Moore, and Sanjaye Ramgoolam. Lectures on 2d Yang-Mills theory, equivariant cohomology and topological field theories. arXiv:hep-th/9411210v2.

[95] Donald C. Ferguson. A theorem of Looman-Menchoff. `http://digitool.library.mcgill.ca/thesisfile111406.pdf`.

[96] Holger Lyre. Berry phase and quantum structure. arXiv:1408.6867.

[97] Florin Belgun. Gauge theory. `http://www.math.uni-hamburg.de/home/belgun/Gauge4.pdf`.

[98] Vladimir Itskov, Peter J. Olver, and Francis Valiquette. Lie completion of pseudogroups. *Transformation Groups*, 16:161--173, 2011.

[99] Richard Borcherds. Lie groups. `https://math.berkeley.edu/~reb/courses/261/`.

[100] Andrei Losev. From Berezin integral to Batalin-Vilkovisky formalism: A mathematical physicist's point of view. 2007.

[101] Edward Witten. Coadjoint orbits of the Virasoro group. *Comm. Math. Phys.*, 114(1):1--53, 1988.

[102] Sidney R. Coleman and Jeffrey E. Mandula. All possible symmetries of the S-matrix. *Phys. Rev.*, 159, 1967.

[103] Emily Riehl. Monoidal algebraic model structures. *Journal of Pure and Applied Algebra*, 217(6):1069--1104, 2013.

[104] Valter Moretti. Mathematical foundations of quantum mechanics: An advanced short course. *International Journal of Geometric Methods in Modern Physics*, 13, 2016.

[105] Antonio Michele Miti. Homotopy comomentum maps in multisymplectic geometry, 2021.

[106] Niclas Sandgren and Petre Stoica. On moving average parameter estimation. Technical Report 2006-022, Department of Information Technology, Uppsala University, 2006.

[107] John E. Roberts. Spontaneously broken gauge symmetries and superselection rules. 1974.

[108] Jean Gallier. Clifford algebras, Clifford groups, and a generalization of the quaternions, 2008. arXiv:0805.0311.

[109] Bozhidar Z. Iliev. Normal frames for general connections on differentiable fibre bundles. arXiv:math/0405004.

[110] Piotr Stachura. Short and biased introduction to groupoids. arXiv:1311.3866.

[111] Fosco Loregian. Coend calculus. arXiv:1501.02503.

[112] Frederic Schuller. Lectures on the geometric anatomy of theoretical physics. https://www.youtube.com/channel/UC6SaWe7xeOp31Vo8cQG1oXw.

[113] Nima Amini. Infinite-dimensional Lie algebras. https://people.kth.se/~namini/PartIIIEssay.pdf.

[114] Peter Selinger. Lecture notes on lambda calculus.

[115] Nigel Hitchin. Lectures on special Lagrangian submanifolds. https://arxiv.org/abs/math/9907034v1, 1999.

[116] Olivia Caramello. Lectures on topos theory at the university of Insubria. https://www.oliviacaramello.com/Teaching/Teaching.htm.

[117] Derek Sorensen. An introduction to characteristic classes. http://derekhsorensen.com/docs/sorensen-characteristic-classes.pdf, 2017.

[118] Arun Debray. Characteristic classes. https://web.ma.utexas.edu/users/a.debray/lecture_notes/u17_characteristic_classes.pdf.

[119] Jonathan R. Shewchuk. An introduction to the conjugate gradient method without the agonizing pain. Technical report, 1994.

[120] Pascal Lambrechts.

[121] Chris Tiee. Contravariance, covariance, densities, and all that: An informal discussion on tensor calculus. https://ccom.ucsd.edu/~ctiee/notes/tensors.pdf, 2006.

[122] Emily Riehl. Homotopy (limits and) colimits. http://www.math.jhu.edu/~eriehl/hocolimits.pdf.

[123] Andreas Gathmann. Algebraic geometry. https://www.mathematik.uni-kl.de/~gathmann/class/alggeom-2019/alggeom-2019.pdf.

[124] Will J. Merry. Algebraic topology. https://www.merry.io/algebraic-topology.

[125] Stacks project. https://stacks.math.columbia.edu/.

[126] The nlab. https://ncatlab.org/nlab.

[127] Wikipedia. https://www.wikipedia.org/.

[128] Joost Nuiten. Cohomological quantization of local prequantum boundary field theory. Master's thesis, 2013.

Index